



HAL
open science

Cartographie des essences forestières à partir de séries temporelles d'images satellitaires à hautes résolutions : stabilité des prédictions, autocorrélation spatiale et cohérence avec la phénologie observée in situ

Nicolas Karasiak

► **To cite this version:**

Nicolas Karasiak. Cartographie des essences forestières à partir de séries temporelles d'images satellitaires à hautes résolutions : stabilité des prédictions, autocorrélation spatiale et cohérence avec la phénologie observée in situ. Sciences agricoles. Institut National Polytechnique de Toulouse - INPT, 2020. Français. NNT : 2020INPT0115 . tel-04172021

HAL Id: tel-04172021

<https://theses.hal.science/tel-04172021>

Submitted on 27 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut National Polytechnique de Toulouse (Toulouse INP)

Discipline ou spécialité :

Agrosystèmes, Écosystèmes et Environnement

Présentée et soutenue par :

M. NICOLAS KARASIAK

le mercredi 4 novembre 2020

Titre :

Cartographie des essences forestières à partir de série temporelles d'images satellitaires à hautes résolutions: stabilité des prédictions, autocorrélation spatiale et cohérence avec la phénologie observée in situ

Ecole doctorale :

Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries (SEVAB)

Unité de recherche :

Dynamiques et écologie des paysages agriforestiers (DYNAFOR)

Directeur(s) de Thèse :

M. CLAUDE MONTEIL

M. JEAN-FRANCOIS DEJOUX

M. DAVID SHEEREN

Rapporteurs :

Mme ANNE PUISSANT, UNIVERSITE DE STRASBOURG

M. THOMAS CORPETTI, UNIVERSITE RENNES 2

Membre(s) du jury :

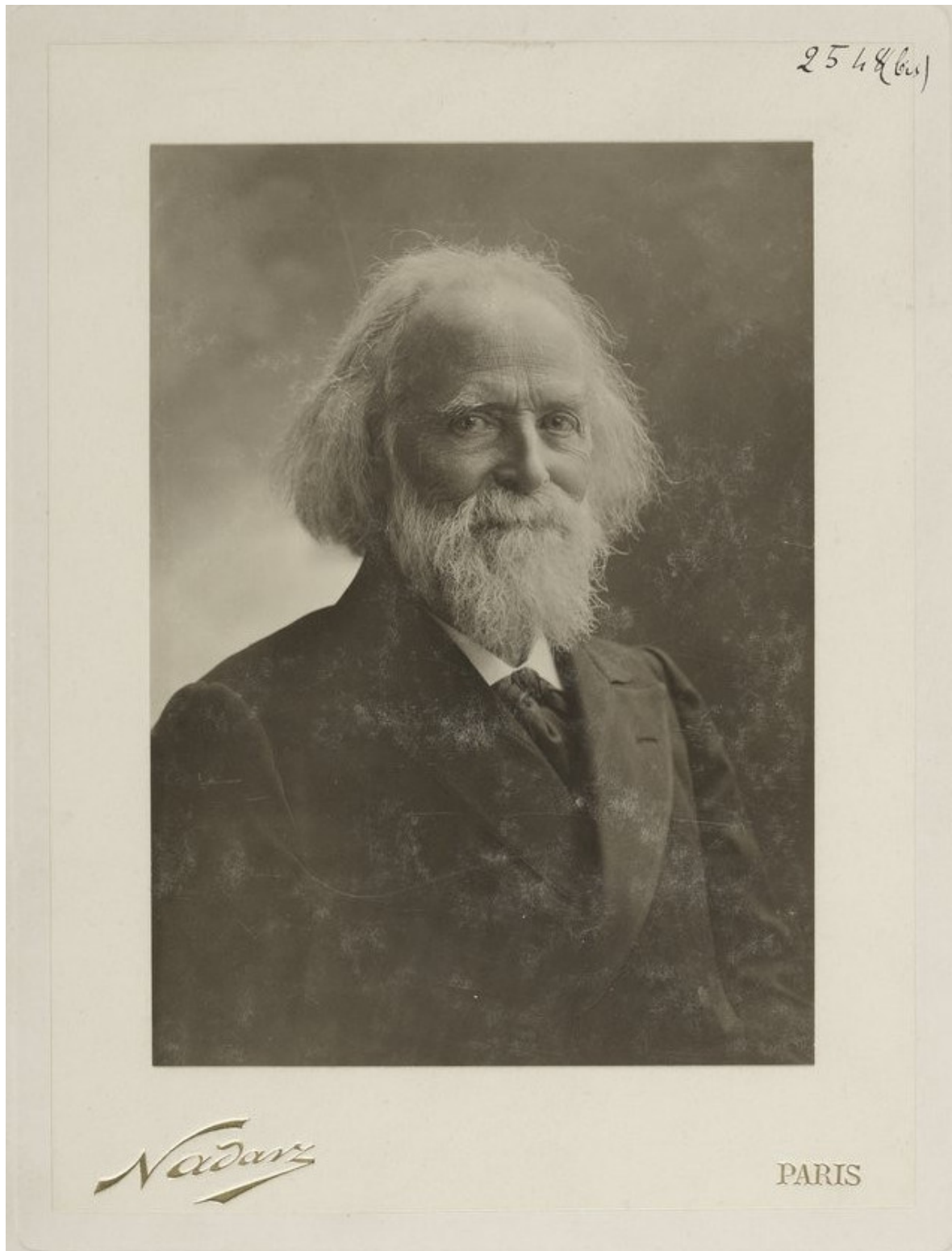
Mme CHRISTIANE WEBER, CNRS MONTPELLIER, Président

M. CLAUDE MONTEIL, TOULOUSE INP, Membre

M. CLEMENT MALLET, INSTITUT GEOGRAPHIQUE NATIONAL, Membre

M. DAVID SHEEREN, TOULOUSE INP, Membre

M. JEAN-FRANCOIS DEJOUX, CESBIO, Membre



“Devenons beaux nous mêmes et que notre vie soit belle”
Élisée Reclus photographié par Nadar en 1903. Source : BNF.

Résumé

La forêt a un rôle essentiel sur terre, que ce soit pour stocker le carbone et ainsi lutter contre le réchauffement climatique ou encore fournir un habitat à de nombreuses espèces. Or la composition de la forêt (la localisation des essences ou leur diversité) a une influence sur les services écologiques rendus.

Dans ce contexte, il est important de cartographier les forêts et les essences qui la composent. La télédétection, en particulier à partir d'images satellitaires, apparaît comme le moyen le plus adéquat pour caractériser un vaste territoire. Avec l'arrivée de constellations satellitaires comme Sentinel-2 ou Landsat-8 et leur gratuité d'acquisition pour l'utilisateur, il devient possible d'envisager l'usage de séries temporelles d'images satellites à haute résolution spatiale, spectrale et temporelle à l'aide d'algorithmes d'apprentissage automatique.

Si de nombreux travaux ont étudié le potentiel des images satellitaires pour identifier les essences, rares sont ceux qui utilisent des séries temporelles (plusieurs images par an) avec une haute résolution spatiale et en tenant compte de l'autocorrélation spatiale des références, *i.e.* la ressemblance des échantillons spatialement proches les uns des autres. Or, en ne prenant pas en compte ce phénomène, des biais d'évaluation peuvent survenir et ainsi surestimer la qualité des modèles d'apprentissage. Il s'agit aussi de mieux cerner les verrous méthodologiques afin de comprendre pourquoi il peut être facile ou compliqué pour un algorithme d'identifier une essence d'une autre.

L'objectif général de la thèse vise à étudier le potentiel et les verrous concernant la reconnaissance des essences forestières à partir des séries temporelles d'images satellite à haute résolution spatiale, spectrale, et temporelle. Le premier objectif consiste à étudier la stabilité temporelle des prédictions à partir d'une archive de neuf ans du satellite Formosat-2. Plus particulièrement, les travaux portent sur la mise en place d'une méthode de validation qui soit le plus fidèle à la qualité observée des cartographies. Le second objectif s'intéresse au lien entre les événements phénologiques *in situ* (pousse des feuilles en début de saison, ou perte et coloration des feuilles en fin de saison) et ce qu'il est observable par télédétection. Outre la capacité de détecter ces événements, il sera étudié si ce qui permet aux algorithmes de différencier les essences les unes des autres est lié à des comportements spécifiques par espèce.



Abstract

Forests have a key role on earth, whether to store carbon and so reducing the global warming or to provide a place for many species. However, the composition of the forest (the location of the tree species or their diversity) has an influence on the ecological services provided.

In this context, it seems critical to map tree species that make it up the forest. Remote sensing, especially from satellite images, appears to be the most appropriate way to map large areas. Thanks to the satellite constellations such as Sentinel-2 or Landsat-8 and their free acquisition for the user, the use of time series of satellite images with high spatial, spectral and temporal resolution using automatic learning algorithms is now easy to access.

While many works have studied the potential of satellite images to identify tree species, few use time series (several images per year) with high spatial resolution and taking into account the spatial autocorrelation of references, *i.e.* the spectral similarity of spatially close samples. However, by not taking this phenomenon into account, evaluation biases may occur and thus overestimate the quality of the learning models. It is also a question of better identifying the methodological barriers in order to understand why it can be easy or complicated for an algorithm to identify one species from another.

The general objective of the thesis is to study the potential and the obstacles concerning the identification of forest tree species from satellite images time series with high spatial, spectral and temporal resolution. The first objective is to study the temporal stability of predictions from a nine-year archive of the Formosat-2 satellite. More specifically, the work focuses on the implementation of a validation method that is as faithful as possible to the observed quality of the maps. The second objective focuses on the link between in situ phenological events (leaf growth at the beginning of the season, or leaf loss and coloration at the end of the season) and what can be observed by remote sensing. In addition to the ability to detect these events, it will be studied whether what allows the algorithms to identify tree species from each other is related to species-specific behaviors.

Sommaire

Sommaire	vii
Liste des figures	ix
Liste des tableaux	xv
1 Introduction	1
1.1 La forêt, bien plus que des arbres	2
1.2 Cartographier les forêts, une histoire ancienne	9
1.3 La télédétection comme alternative à la cartographie manuelle	12
1.4 Questions de recherche	35
2 Étude de la stabilité spatiale et statistique des prédictions	39
2.1 Introduction	41
2.2 Material	43
2.3 Classification protocol	45
2.4 Results	49
2.5 Discussion	55
2.6 Conclusion	59
3 Étude de l'impact de l'autocorrélation spatiale	61
3.1 Introduction	62
3.2 Related works	64
3.3 Experimental protocol	66
3.4 Results	71
3.5 Discussion	76
3.6 Conclusion	78
4 Étude et apport de la phénologie	79
4.1 Introduction	80
4.2 Matériel et méthode	83
4.3 Résultats des suivis phénologiques	93
4.4 Résultats de l'apport des séries temporelles pour la détection des essences	97
4.5 Discussion	102
4.6 Conclusion	108
5 Partage des outils et des données : assurer une reproductibilité des travaux	111
5.1 Introduction	112
5.2 Principes FAIR	112
5.3 Mise en œuvre des principes FAIR	114
5.4 Partage des références terrain	116
5.5 Partage les outils	118
5.6 Conclusion	136

6 Conclusion générale	139
6.1 Conclusion	140
6.2 Perspectives	141
Bibliographie	143
Acronymes	169
Annexes	173
A Compléments du chapitre 2	173
A.1 Tree species map	173
A.2 Significance tables for prediction between years	175
A.3 Effect of clouds and cloud shadows	176
A.4 Training size per species	178
B Compléments du chapitre 3	179
B.1 Moran’s I correlogram on Sentinel-2 data	179
B.2 Relative frequency distribution of NDVI values per class	180
B.3 Class-specific performances and class prevalence	181
B.4 Confusion matrices	182
C Compléments du chapitre 4	183
C.1 Protocole de suivi phénologique	184
C.2 Les meilleures dates pour les indices spectraux	188
C.3 Les paramètres des indices et de Savitzsky-Golay	192
C.4 Indice spectral avec <i>Canopy Cover</i> et BBCH	216
C.5 Évolution des spectre des feuillus	219
C.6 Observations phénologiques	222
D Liste des publications	223
D.1 Journal international à comité de lecture	223
D.2 Conférence internationale à comité de lecture	223
D.3 Conférence nationale à comité de lecture	223
D.4 Chapitre d’ouvrage	224
D.5 Revue de vulgarisation	224
D.6 Divers	224

Liste des figures

1.1	Surface minimum selon les pays pour que la forêt soit définie comme telle . . .	3
1.2	Évolution de la surface forestière entre 1830 et aujourd’hui. Source : Institut National de l’information Géographique et forestière (IGN).	4
1.3	Répartition des surfaces de plantations par espèce principale plantée en 2016 en France.	4
1.4	Répartition de la composition des peuplements en France métropolitaine. . .	5
1.5	Méthode de calcul des différents taux de couvert. En vert, les essences de feuillus, en bleu de conifères. Source : IGN.	6
1.6	Détermination du taux de couvert en tenant compte des différentes espèces par ordre décroissant de taux de couvert libre relatif	6
1.7	Répartition de la diversité des espèces d’arbres en Europe. Source : UICN. . .	7
1.8	Météo de la floraison des cerisiers du Japon (<i>Prunus serrulata</i>) en 2019. Source : WeatherMap.	8
1.9	Évolution de la date de floraison des cerisiers du Japon <i>Prunus serrulata</i> depuis le IXème siècle	9
1.10	Évolution de la cartographie de la forêt de Bouconne : de la carte de Cassini en 1740 (a), à la carte de l’État-Major en 1866 (b).	10
1.11	Les essences composant la forêt de Bouconne selon la BD Forêt v2 de l’IGN. À chaque couleur correspond une essence ou un mélange différent.	11
1.12	Différences de fonctionnement entre un capteur optique et un capteur radar .	13
1.13	Influence de la résolution spatiale sur la taille d’un échantillon du couvert forestier	14
1.14	Principaux capteurs à haute résolution spatiale (de 5 à 100 m) détaillant la longueur d’ondes de chaque bande spectrale	15
1.15	Les trois grands niveaux de traitements des images satellites	16
1.16	Différence entre une <i>Satellite Image Time Series</i> ou en français, Série Temporelle d’Images Satellites (SITS) et un ensemble d’images sans nuage	18
1.17	Exemple de la séparation de deux classes selon les algorithmes SVM et RF. Inspiré du code de Gaël Varoquaux et Andreas Müller.	19
1.18	À chaque pixel correspond une signature spectrale, soit monotemporelle soit multitemporelle.	21
1.19	Décomposition du spectre électromagnétique avec un focus sur le spectre du visible (de 400 à 700 nanomètres).	22
1.20	Illustration de l’absorption, la réflexion et de la transmission du signal	23
1.21	Spectre d’absorption de la chlorophylle <i>a</i> et <i>b</i> et des caroténoïdes.	23
1.22	Illustration des différents types de bourgeons lors du débourrement	24
1.23	Anatomie d’une feuille. Le cuticule englobe la partie interne de la feuille . . .	25
1.24	Réflectance d’une feuille de chêne fraîche (ligne épaisse) ou sèche après 24h (ligne fine)	26
1.25	Évolution de la réflectance de certains pigments foliaires lors de la sénescence.	27
1.26	Un rameau de chêne pubescent (<i>Quercus pubescens</i>) pris en photo le 27 octobre 2017	27

1.27	Mesure à l'aide d'un spectromètre de l'activité du Pin contorta, en chlorophylle et en caroténoïdes	28
1.28	Structure d'un arbre, avec et sans feuilles. Source : Nouveau Dictionnaire Visuel.	29
1.29	Exemple de la variation saisonnière du LAI dans une chênaie.	29
1.30	Exemple d'une photo hémisphérique prise le 23 septembre 2019 dans une forêt composée majoritairement de chênes.	30
1.31	Les différentes strates végétales de la forêt	31
1.32	Le score F1 est la moyenne harmonique entre la précision et le rappel	34
2.1	Location of the study area	43
2.2	Number and acquisition dates of each image in the Formosat-2 time series from 2006 to 2014.	44
2.3	Classification protocol for a single year time series, repeated for the 9 years available from 2006 to 2014	46
2.4	Spatial leave-one-out cross-validation (SLOO-CV) schema for one class	47
2.5	Moran's I correlograms of each Formosat-2 spectral band of the SITS 2013, for pixels representing forests	48
2.6	Average F1 score (in %) per species and per year based on the SLOO-CV sampling strategy using the SVM (RBF kernel) classifier	50
2.7	Confusion matrix for 13 tree species for year 2013	51
2.8	Average rate of commission and omission errors (in %) per species and per year based on the SLOO-CV sampling strategy.	52
2.9	Spatial comparison between the annual classifications of forest tree species from 2006 to 2014 (including 2006bis)	53
2.10	Annual classifications of tree species in a mixed forest composed of different conifers and broadleaf species	54
2.11	Distribution per species of the number of agreements between the annual classifications using all the predicted pixels for both the LOO-CV and SLOO-CV sampling strategies.	56
2.12	Influence of undetected clouds in the time series, an example with the reflectance evolution of a Oak tree pixel in 2006	57
3.1	Spatial and non-spatial CV strategies to investigate the effect of spatial autocorrelation between training and test sets	67
3.2	Location of the 31TDJ Sentinel-2 tile in the south of France (100 km x 100 km in extent)	69
3.3	Average overall accuracy based on the RF classifier for each cross-validation strategy (<i>k</i> -fold CV, LOO CV, SLOO CV) at pixel and object levels	72
3.4	Variability in accuracy depending on the sample size defined in pixels	73
3.5	Comparison between the average predictive performances obtained with spatial and non-spatial cross-validation from the Herault-34 <i>full model</i> fitted with the maximum training set of 1,972 forest stands and the average predictive performances obtained with the same model applied to test sets from Tarn-81 and Aveyron-12	74
3.6	Comparison of OA estimated by cross-validation in Herault-34 using the six spatial and non spatial data-splitting strategies for different sized training sets and OA estimated with the same Herault-34 model on test sets in Tarn-81	75
3.7	Comparison of OA estimated by cross-validation in Herault-34 from the six spatial and non spatial data-splitting strategies for different sized training sets and OA estimated using the same Herault-34 model on test sets in Aveyron-12	75
3.8	Learning curves calculated each from each fold according to the different cross-validation strategies applied	77

4.1	Illustration de l'extraction de métriques phénologiques	81
4.2	Zone d'étude	83
4.3	Liste des dates auxquelles un suivi phénologique a été réalisé sur le terrain.	84
4.4	Illustration des notations de BBCH pour le suivi de la phénologie des feuillus.	85
4.5	Exemple de la couverture de la canopée du peuplement de robiniers au 1er août 2019	86
4.6	Dates d'acquisition des images Sentinel-2 pour les années 2017, 2018, et 2019	87
4.7	Résumé de l'ensemble de la méthode pour valider la détection des phases phénologiques et estimer son l'apport dans la cartographie semi-automatique des espèces forestières.	93
4.8	Résumé des dates de débuts de saison (BBCH=11, 15 et 19) et des fins de saisons (BBCH=91, 95 et 99) observées sur le terrain pour les quatorze parcelles suivies en 2018 et 2019	94
4.9	Série temporelle du NDVI avec un lissage de Savitzky-Golay utilisant une taille de fenêtre de 9 avec un seuil de début de saison fixé à 80% de l'amplitude et 65% pour la fin de saison	96
4.10	Série temporelle du LCaroC avec un lissage de Savitzky-Golay utilisant une taille de fenêtre de 9 avec un seuil de début de saison fixé à 55% de l'amplitude et 75% pour la fin de saison	97
4.11	Liste des 5 meilleures dates de la série temporelle Sentinel-2 de 2018 pour classer séparément chaque essence de feuillus.	99
4.12	Liste des cinq meilleures dates de la série temporelle Sentinel-2 de 2017 pour classer séparément chaque essence de feuillus.	99
4.13	Liste des cinq meilleures dates de la série temporelle Sentinel-2 de 2019 pour classer séparément chaque essence de feuillus.	100
4.14	Photo hémisphérique de la canopée du robinier (<i>Robinia pseudoacacia</i>) et du chêne rouge (<i>Quercus rubra</i>) au 3 mai 2020, L'ensemble des feuilles était sorti (BBCH>19).	100
4.15	Photo hémisphérique de la canopée du peuplier (<i>Populus spp.</i>) au 22 novembre 2019. Les quelques feuilles restant sur les arbres étaient déjà colorées en jaune, ce qui signifie que la saison de végétation était finie.	101
4.16	Matrice de confusion des résultats obtenus avec les 10 meilleures dates pour les années 2018 et 2019. Le nombre total d'échantillons de référence par essence est de 42.	102
4.17	Série temporelle du LChloC avec un lissage de Savitzky-Golay utilisant une taille de fenêtre de 13 et un seuil de début de saison fixé à 25% de l'amplitude et 70% pour la fin de saison.	103
4.18	Série temporelle du NDVI avec un lissage de Savitzky-Golay utilisant une taille de fenêtre de 7 et un seuil de début de saison fixé à 55% de l'amplitude et de fin de saison à 75%	104
4.19	Liste des dates sélectionnées selon le meilleur score F1 moyen pour les années 2017, 2018 et 2019	106
4.20	Liste de la qualité de prédiction lorsqu'une seule date est utilisée pour prédire une essence	107
4.21	Part des feuillus prédite en 2017 et 2018 sur la zone d'étude en classant l'ensemble des essences (incluant les résineux) à partir de la meilleure sélection de caractéristiques.	108
4.22	Différence de phénologie entre deux cyprès chauves (essence non présente dans la zone d'étude au sud de Toulouse). Photo prise le 12 octobre 2019 à Blagnac dans le parc d'Andromède.	109
5.1	Que signifie l'acronyme FAIR?	113

5.2	Dessin illustrant le ressenti que l'on peut avoir à la lecture de la partie consacrée à la méthodologie dans de nombreux travaux	119
5.3	Organisation générale de Museo ToolBox.	121
5.4	Fonctionnement de RasterMath	122
5.5	Aperçu du fonctionnement des 6 principales méthodes de validation croisée et stratifiée disponibles dans Museo ToolBox.	125
5.6	Les différentes types de matrices de confusion améliorées réalisées avec la classe PlotConfusionMatrix de Museo ToolBox. Ces figures ont été écrites en utilisant le Code 5.7	127
5.7	Organisation de Museo Pheno	128
5.8	Exemple d'usage des branches dans un git. La branche <i>Master</i> est la branche principale, et chaque modification peut être effectuée sur une branche distance qui sera par la suite fusionnée sur la branche principale. Cela est particulièrement recommandé lors d'un travail collaboratif. Chaque point représente une modification historisée.	131
5.9	Un des deux badges apparaît sur les pages de Museo ToolBox afin d'avertir sur la réussite des tests de la bibliothèque. Si le badge est rouge et indique donc <i>failing</i> , cela signifie que la version en cours de développement rencontre au moins un problème à corriger.	132
5.10	La documentation python de Museo ToolBox est écrite selon le style développé par la bibliothèque Numpy. Cette convention permet de convertir de manière automatique une page internet à partir du code python ou d'afficher l'aide d'une fonction dans l'interface de développement.	134
5.11	Un des deux badges apparaît sur les pages de Museo ToolBox afin d'avertir de l'état de la documentation. Si le badge est indique <i>failing</i> , cela signifie que la version en cours rencontre un problème à corriger.	134
5.12	Liste des exemples utilisant la classe RasterMath.	135
5.13	Schéma résumant les principes FAIR auxquels cette thèse satisfait.	137
A.1	Map of the most predicted species using the 10 SITS with the Spatial Leave-One-Out cross-validation method (SLOO-CV).	173
A.2	Map of the most predicted species using the 10 SITS with the Leave-One-Out cross-validation method (LOO-CV).	174
A.3	Example of misclassified forest stand (for European ash, in the north) due to under-detected clouds in 2013-09-01. In the cloud mask, pixels in black are cloud and shadow free. Pixels in white are cloudy or shady.	177
B.1	Moran's I correlograms of each Sentinel-2 spectral band, for pixels representing forests. Each curve represents one spectral band at one date in the dataset. For a Moran's I threshold value of 0.05, spatial independence between nearby pixels was assumed. This threshold was reached at 1,964 pixels (i.e. 19.64 km) on average. The grey dashed line represents the mean distance value (in x) where Moran's I = 0.05 (in y).	179
B.2	Relative frequency distribution (in %) of NDVI values for broadleaf class. . .	180
B.3	Relative frequency distribution (in %) of NDVI values for coniferous class. . .	180
B.4	Average accuracy metrics and prevalence in the test set for broadleaf and coniferous classes (in %).	181
B.5	Confusion matrix of the test set in Aveyron-12 (left) and in Tarn-81 (right) from the Hérault-34 full model fitted with the maximum training set size of 1,972 forest stands. The average accuracy metrics have been estimated from 369,164 reference pixels for each class after undersampling repeated 30 times. The values in the cells are in pixels. The units of Producer and User's accuracy are in %	182

C.1	Liste des cinq meilleures dates pour classer chaque essence à partir de l'indice LAnthoC de l'année 2017 (a), 2018 (b) et 2019 (c).	188
C.2	Liste des cinq meilleures dates pour classer chaque essence à partir de l'indice LCaroC de l'année 2017 (a), 2018 (b) et 2019 (c).	189
C.3	Liste des cinq meilleures dates pour classer chaque essence à partir de l'indice LChloC de l'année 2017 (a), 2018 (b) et 2019 (c).	190
C.4	Liste des cinq meilleures dates pour classer chaque essence à partir de l'indice NDVI de l'année 2017 (a), 2018 (b) et 2019 (c).	191
C.5	Erreur moyenne absolue (MAE) pour le début de saison (BBCH=11) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LAnthoC.	192
C.6	Erreur moyenne absolue (MAE) pour le début de saison (BBCH=11) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LCaroC.	193
C.7	Erreur moyenne absolue (MAE) pour le début de saison (BBCH=11) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LChloC.	194
C.8	Erreur moyenne absolue (MAE) pour le début de saison (BBCH=11) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du NDVI pour l'année.	195
C.9	Erreur moyenne absolue (MAE) pour le début de saison (BBCH=11) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LAnthoC.	196
C.10	Erreur moyenne absolue (MAE) pour le début de saison (BBCH=15) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LCaroC.	197
C.11	Erreur moyenne absolue (MAE) pour le début de saison (BBCH=15) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LChloC.	198
C.12	Erreur moyenne absolue (MAE) pour le début de saison (BBCH=15) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du NDVI.	199
C.13	Erreur moyenne absolue (MAE) pour le début de saison (BBCH=19) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LAnthoC.	200
C.14	Erreur moyenne absolue (MAE) pour le début de saison (BBCH=19) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LCaroC.	201
C.15	Erreur moyenne absolue (MAE) pour le début de saison (BBCH=19) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LChloC.	202
C.16	Erreur moyenne absolue (MAE) pour le début de saison (BBCH=19) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du NDVI pour l'année.	203
C.17	Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=91) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LAnthoC.	204
C.18	Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=91) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LCaroC.	205

C.19 Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=91) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LChloC.	206
C.20 Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=91) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du NDVI.	207
C.21 Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=95) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LAnthoC.	208
C.22 Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=95) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LCaroC.	209
C.23 Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=95) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LChloC.	210
C.24 Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=95) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du NDVI.	211
C.25 Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=99) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LAnthoC.	212
C.26 Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=99) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LCaroC.	213
C.27 Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=99) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LChloC.	214
C.28 Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=99) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du NDVI.	215
C.29 Différence de signal entre le LChloC et le NDVI en comparaison avec le BBCH et le <i>Canopy Cover</i> (1/3).	216
C.29 Différence de signal entre le LChloC et le NDVI en comparaison avec le BBCH et le <i>Canopy Cover</i> (2/3).	217
C.29 Différence de signal entre le LChloC et le NDVI en comparaison avec le BBCH et le <i>Canopy Cover</i> (3/3).	218
C.30 B2, B3, B4 et B8 sont respectivement la bande bleue, verte, rouge, et proche infra-rouge à 10 m de résolution spatiale. B5 (red-edge, 705nm), B6 (red-edge, 740nm), B11 (1610 nm) et B12 (2190 nm) sont des bandes à 20 m, rééchantillonnées à 10 m pour les besoins de cette étude.	219
C.31 B2, B3, B4 et B8 sont respectivement la bande bleue, verte, rouge, et proche infra-rouge à 10 m de résolution spatiale. B5 (red-edge, 705nm), B6 (red-edge, 740nm), B11 (1610 nm) et B12 (2190 nm) sont des bandes à 20 m, rééchantillonnées à 10 m pour les besoins de cette étude.	220
C.32 B2, B3, B4 et B8 sont respectivement la bande bleue, verte, rouge, et proche infra-rouge à 10 m de résolution spatiale. B5 (red-edge, 705nm), B6 (red-edge, 740nm), B11 (1610 nm) et B12 (2190 nm) sont des bandes à 20 m, rééchantillonnées à 10 m pour les besoins de cette étude.	221
C.33 Évolution de l'ouverture de la canopée (indice <i>Canopy Cover</i> en noir) et du code BBCH (en vert). En bleu, période où a été observée la floraison.	222

Liste des tableaux

1.1	Principaux travaux de ces dernières années en télédétection des essences forestières à partir de capteur à haute résolution. Trié par ordre décroissant selon le nombre de dates utilisées. Pour les satellites les moins courants, il est spécifié leur résolution spatiale entres parenthèses.	33
2.1	List of tree species with their sample size, in pixels, collected during field surveys ($n = 1262$). The number of forest stands in which the samples were collected is also provided. Stand delimitation is based on the French National Forest Inventory database (IGN BDForêt [®] v.1)	45
2.2	Accuracy report of spatial leave-one-out cross-validation (SLOO-CV) sampling strategy and leave-one-out cross-validation (LOO-CV) for each single-year classification based on OA statistics. The 2006bis time series only includes cloud-free images of 2006. The average percentage of cloud coverage was estimated by computing for each species the number of time each reference sample was affected by clouds (detected from the MACCS processing chain).	49
3.1	Non exhaustive list of representative references.	66
3.2	Number of pixels and forest stands in each class (coniferous and broadleaves) in each administrative department in the study site.	70
4.1	Explication de la nomenclature BBCH	85
4.2	Liste des indices spectraux utilisés pour détecter la phénologie des feuillus. . .	89
4.3	Liste des essences forestières et leur effectif en pixel ($n = 1626$). Le nombre de peuplements a été calculé à partir de la donnée (IGN BDForêt [®] v.1)	90
4.4	Erreur moyenne absolue (MAE) en jours, calculée en comparant la date détectée in situ avec le meilleur seuil à partir des indices ACORVI, LAnthoC, LCaroC, LChloC et NDVI	95
4.5	Erreur moyenne absolue (MAE) en jours calculée en comparant la date détectée in situ avec le meilleur seuil à partir des indices ACORVI, LAnthoC, LCaroC, LChloC et NDVI. La date correspondant à la première fois où les stades BBCH \geq 91, BBCH \geq 95 et BBCH \geq 99 ont été notés lors des observations in situ. Les chiffres en gras correspondent au meilleur MAE par stade BBCH.	96
4.6	Score F1 maximum obtenu avec un modèle par essence en utilisant une validation croisée de type SLOO pour l'année 2017, 2018 et 2019 et à l'aide soit des 5 meilleures dates (via une sélection de caractéristiques) ou à l'aide de la série temporelle dans son ensemble. En gras sont spécifiés les meilleurs résultats par essence.	98
4.7	F1-score obtenu par essence et par indice en utilisant les 5 meilleures dates. En gras est spécifié le meilleur indice par année et par essence.	101
5.1	Liste des champs des données de référence localisées des essences. <i>int</i> signifie que le champs contient des données de type entier, <i>str</i> de type <i>string</i> (texte).	116

5.2	Liste des champs complémentaires pour chaque année. Pour 2013 qui est composée de 17 dates, il y a donc 68 colonnes <code>band_</code> (17 dates * 4 bandes spectrales) et 17 colonnes <code>cloud_</code> allant de 0 à 16. <i>int</i> signifie que le champs contient des données de type entier.	117
5.3	Liste des champs du fichier <code>plot.csv</code> qui contient l'identifiant et le détail de chaque parcelle. <i>int</i> signifie que le champs contient des données de type entier, <i>float</i> un nombre à virgule et <i>str</i> de type <i>string</i> (texte).	117
5.4	Liste des champs des relevés phénologiques de chaque année. <i>int</i> signifie que le champs contient des données de type entier, <i>float</i> un nombre à virgule et <i>str</i> de type <i>string</i> (texte).	118
A.1	Wilcoxon signed-rank test significance table for SLOO-CV computed from the overall accuracy of the 50 predictions for each single-year SITS. In bold , where $p < 0.05$	175
A.2	Wilcoxon signed-rank test significance table for LOO-CV computed from the overall accuracy of the 50 predictions for each single-year SITS. In bold , where $p < 0.05$	175
A.3	Percent of reference pixels affected by clouds or shadows detected by the MACCS algorithm for each species and each year.	176
A.4	Number of training pixels for each cross-validation method. For both methods, value per species is the mean from the 50-folds. The split between test and train set is exactly the same for each year.	178

Chapitre 1

Introduction

“ Il est certain que la destruction des forêts et la mise en culture de vastes étendues ont pour conséquence des modifications appréciables dans les diverses saisons.”

— Élisée Reclus.

De l'action humaine sur la géographie physique : l'homme et la nature (1864).

Table des matières

1.1	La forêt, bien plus que des arbres	2
1.1.1	Définition générale	2
1.1.2	Cas particulier de la forêt française	3
1.1.3	Sous la canopée	5
1.1.4	De l'intérêt de connaître la distribution des essences	6
1.2	Cartographier les forêts, une histoire ancienne	9
1.2.1	Les premières cartes topographiques	9
1.2.2	La cartographie actuelle des forêts françaises	10
1.3	La télédétection comme alternative à la cartographie manuelle	12
1.3.1	Principes de bases de la télédétection	12
1.3.2	Fonctionnement général de l'apprentissage supervisé	18
1.3.3	Les propriétés optiques des plantes	21
1.3.4	Principaux travaux de classification des essences forestières	32
1.4	Questions de recherche	35
1.4.1	Défis	35
1.4.2	Objectifs	36
1.4.3	Organisation du manuscrit	36

En ayant la possibilité de photographier l'état de notre planète plusieurs fois par an, les travaux d'observation de la terre ont su tirer profit du potentiel des images satellites dès leur déploiement dans les années 70.

Si les premiers travaux se cantonnaient à cartographier les grands types d'occupation du sol et à identifier sur plusieurs années la déforestation, l'amélioration continue des capteurs et des méthodes a permis d'envisager des cartes bien plus précises, tant sur le point de vue thématique que spatial.

Dans un premier temps, la notion de forêt sera présentée et avec elle l'intérêt de connaître sa composition et sa répartition spatiale. Nous verrons ensuite comment la forêt a été cartographiée depuis plusieurs siècles et puis comment, depuis quelques décennies, la télédétection à partir de satellites optiques a permis de débiter l'automatisation de cette cartographie. Enfin, les défis et objectifs de cette thèse seront introduits avec notamment le détail de l'organisation du manuscrit.

1.1 La forêt, bien plus que des arbres

1.1.1 Définition générale

La *Food and Agriculture Organization of the United Nations* ou, en français, l'Organisation des Nations Unies pour l'Alimentation et l'Agriculture (FAO) définit la forêt comme un espace terrestre qui occupe plus de 0,5 hectares avec un minimum de 20 mètres de large, qui est composé d'arbres pouvant atteindre à maturité une hauteur supérieure à 5 mètres et dont le couvert forestier pourra être à terme d'au moins 10% [FAO, 2001]. Si ces espaces ont pour vocation à être soit d'usage agricole comme l'est par exemple un verger, soit à prédominance urbaine, ils sont de facto exclus de cette définition. Cette définition connaît toutefois des limites concernant les plantations. Ainsi la FAO considère comme forêt une plantation d'hévéas dont on extrait la sève alors qu'une plantation d'oliviers ne sera pas comptabilisée car la récolte des fruits est directement synonyme d'agriculture.

Si la définition de la FAO prédomine au niveau mondial, elle est à la fois loin d'être la seule et loin de faire consensus. Le nombre de définitions différentes de la forêt rend compliqué leur suivi dans le temps [Putz and Redford, 2010]. Entre 1990 et 2000, la seule réduction des seuils de surface (passant de 1 à 0.5 hectare), de hauteur (de 7 à 5m) et de couverture terrestre (de 20 à 10%) a permis d'augmenter de près de 10% (soit 300 millions d'hectares) la surface de forêt à l'échelle mondiale [Chazdon et al., 2016]. Se servant de cette nouvelle définition moins restrictive, Bastin et al. [2017] ont géoréférencé par photo-interprétation les forêts des zones arides. À partir des images à très haute résolution spatiale disponibles dans Google Earth, ils ont ainsi pu répertorier plus de 467 millions d'hectares de forêt jusqu'alors jamais référencés, en particulier car le couvert forestier bien que faible était supérieur à 10%. La superficie totale mondiale des forêts serait alors aujourd'hui d'environ 4,4 milliards d'hectares [Bastin et al., 2017].

Si une définition internationale comme celle de la FAO définit les forêts, il n'en reste pas moins que les pays restent souverains et peuvent définir eux-mêmes ce qu'ils identifient comme forêt. Ainsi, un pays comme le Portugal considère qu'il suffit d'avoir 10% de couverture de canopée et 2m de hauteur à maturité pour qu'une occupation du sol soit considérée comme une forêt alors que pour le Zimbabwe il faut au minimum 15m de hauteur et 80% de couverture de canopée (Figure 1.1).

Bien que très utilisée, la définition de la FAO est toutefois contestée. En effet, il y a des différences importantes entre les forêts dites naturelles et les plantations. Bien qu'occupant moins de 10% de la surface terrestre, les forêts tropicales représentent plus des deux tiers de la biodiversité mondiale [Giam, 2017]. Or, selon Brockerhoff et al. [2008], la conversion des forêts naturelles en plantation se fait inexorablement au détriment de la biodiversité. On peut noter toutefois que, si la déforestation est inévitable, il apparaît alors que les plantations forestières sont le moindre mal pour la biodiversité selon ces mêmes auteurs. La définition de la FAO ne convient pas non plus à l'Organisation Non Gouvernementale (ONG) *World Rainforest Movement* (WRM) qui alerte sur la nécessaire redéfinition de ce qu'est la forêt. Pour cet organisme, il ne s'agit pas tant d'un problème de biodiversité que du maintien des communautés et des activités locales : *“définir les plantations comme des forêts renforce le secteur des entreprises, en particulier celui des plantations, et empêche les communautés locales de s'y opposer pour protéger leurs moyens de subsistance.”*

La forêt est aussi un territoire de lutte où les enjeux et tensions sont forts [Provost, 2019]. Ainsi, il convient de rappeler qu'en 2001 le protocole de Kyōto ne mentionnait pas la lutte contre la déforestation comme moyen de réduire les émissions de gaz à effet de serre [Moutinho et al., 2005], alors qu'on estime que la perte des forêts tropicales contribue à plus de 10% des émissions de gaz à effet de serre dans le monde [Baccini et al., 2012; Harris et al., 2012]. Quelques années plus tard, le concept de réduction compensée (2003), puis de REDD (Réduction des Émissions dues à la Déforestation et à la Dégradation forestière) et enfin de REDD+ (qui inclut la gestion durable et la conservation des forêts) est apparu

1.1. LA FORÊT, BIEN PLUS QUE DES ARBRES

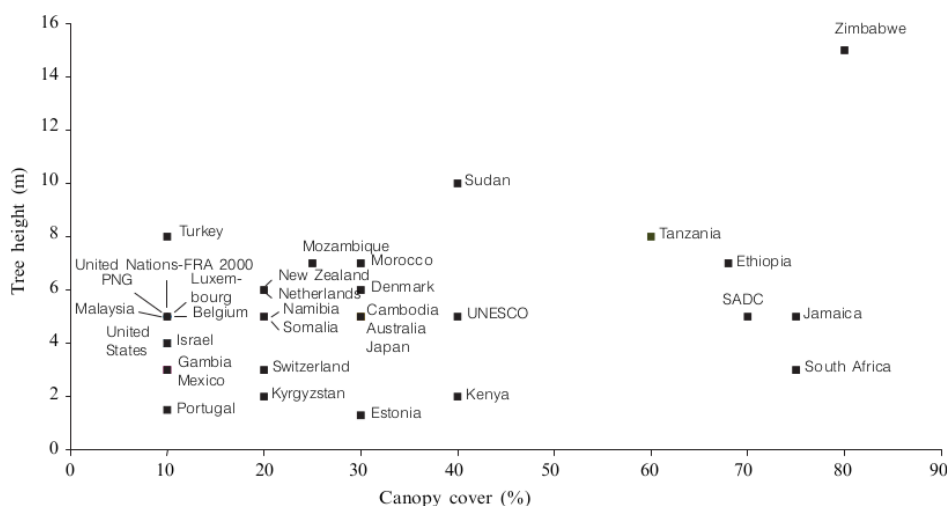


FIGURE 1.1 – Surface minimum selon les pays pour que la forêt soit définie comme telle (données de 2004 de Lund [2018] et graphique de Comber et al. [2005]). Les auteurs du graphique soulignent que la plupart des pays ne définissent pas leur forêt avec ce critère.

pour compenser ce manque. Cependant, ces mesures ont, selon Gautreau and Viard-Crétat [2020], mis la forêt sur un marché “globale, monétarisable, carbocentrée” et chaque pays se voit attribuer des quotas de prélèvements selon des inventaires pas forcément précis [Boyd, 2010]. Le résultat est donc très mitigé selon Gautreau and Viard-Crétat [2020].

Pour résumer on peut dire qu’aujourd’hui la forêt est vue à la fois comme un réservoir important de carbone et comme un lieu abritant une part importante de la biodiversité. Il convient toutefois de souligner que des courants moins anthropocentrés s’intéressent également à ce milieu, non pas pour étudier uniquement les services écosystémiques, mais pour s’intéresser aux individus sentients qui la composent [Kymlicka and Donaldson, 2014; Provoost, 2019]. Il s’agit là d’une vision qui s’éloigne de l’écologie classique qui considère que seule l’espèce compte. Alors que l’écologie ne s’intéresse globalement qu’au maintien des espèces et a comme critère la maximalisation de la biodiversité ou s’intéresse aux mécanismes du fonctionnement des écosystèmes, l’écologie sentientiste prend en compte non seulement l’espèce mais aussi les individus sentients qui la composent, c’est-à-dire les animaux qui ont la capacité à éprouver du bien-être ou de la souffrance [Lepeltier, 2020]. Les végétaux sont de facto exclus de la notion de sentience car ils ne possèdent ni requièrent de conscience [Burgat, 2020; Taiz et al., 2019]. L’accroissement de la biodiversité n’est donc plus le but ultime s’il se fait au détriment des expériences positives des habitants des écosystèmes. La préservation des écosystèmes apparaît alors comme un moyen de maximiser ces expériences et non comme une fin en soi [Fox, 1990; Mathews, 2012].

1.1.2 Cas particulier de la forêt française

En France, l’IGN, ex-IFN (Inventaire Forestier National), précise la définition de la FAO en considérant comme forêt les haies, brises-vents ou corridors s’ils mesurent au minimum 20 mètres de large.

Après des siècles de déforestation, la forêt française a atteint son minimum de couverture terrestre vers 1830 avec environ 9 millions d’hectares (Figure 1.2) [Cinotti, 1996; Rudel et al., 2005]. Grâce à la plantation du massif des Landes dès la deuxième moitié du XIX^{ème} siècle, la superficie forestière croît de manière significative en France en mettant fin au système agro-pastoral de la région. Plusieurs besoins ont poussé à la plantation massive de pins, le premier étant la demande industrielle en résine et en térébenthine [Vidalou, 2017]. La seconde était d’empêcher certains villages d’être ensevelis par les dunes. Pour Nicolas Brémontier, ingénieur des Ponts et Chaussées à l’époque, la Forêt des Landes permettait aussi d’“assainir

les marécages et d'améliorer les conditions d'hygiène". C'est pour toutes ces raisons que près d'un million d'hectares de pins ont pu être plantés massivement avec la loi du 19 juin 1857. Les bergers qui habitaient les landes se transforment en résiniers, et plantent de manière aléatoire la forêt des Landes. En 1949, un incendie vient ravager plus de la moitié des arbres. L'année d'après, une nouvelle campagne massive de plantations a eu lieu et donnera à la forêt des Landes sa forme actuelle [Deville, 2009].



FIGURE 1.2 – Évolution de la surface forestière entre 1830 et aujourd'hui. Source : IGN.

À l'instar de la forêt des Landes, 1,7 millions d'hectares des forêts plantées en France sont de type conifères, ce qui représente 80% des plantations du pays. Au total les plantations représentent selon l'IGN, 13% des forêts françaises [Du puy et al., 2017]. Le pin maritime domine donc largement avec près de 500 000 hectares plantés, suivi par le douglas, et l'épicéa. Le peuplier, la première essence de feuillus cultivée en France en terme de superficie, est présent sur près de 150 000 hectares. Les 20 essences les plus plantées en France représentent pas moins de 96% de la surface forestière cultivée (Figure 1.3).

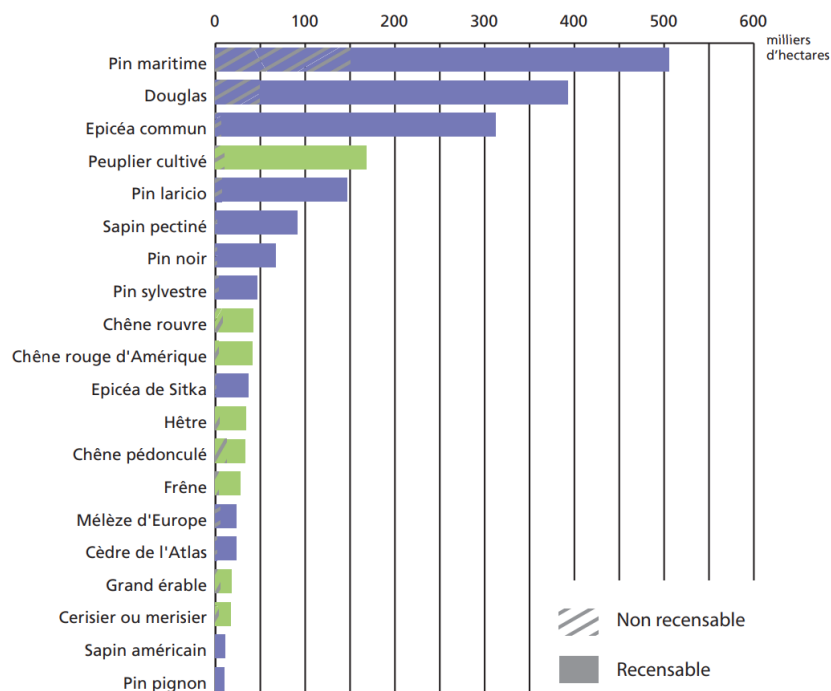


FIGURE 1.3 – Répartition des surfaces de plantations par espèce principale plantée en 2016 en France. Source : Du puy et al. [2017].

Si la surface forestière évolue positivement depuis 1830, cela n'empêche pas que certaines

1.1. LA FORÊT, BIEN PLUS QUE DES ARBRES

essences soient menacées de disparation dans notre pays. Ainsi trois sorbiers (*Sorbus latifolia*, *Sorbus remensis*, et *Sorbus legrei*), le marronnier commun (*Aesculus hippocastanum*), le laurier du Portugal (*Prunus*), et l'orme (*Ulmus glabra*) se trouvent sur la liste rouge de l'Union Internationale pour la Conservation de la Nature (UICN) [Rivers, 2019]. Pour établir cette liste des espèces risquant de disparaître, l'ONG s'est basée sur un nombre limité d'individus par essence ainsi que sur leur aire géographique restreinte.

Pour suivre également l'évolution des forêts en contexte de changement climatique, la déléguée interministérielle à la forêt et au bois recommande dans son rapport de mars 2017 [Alexandre, 2017] d'intégrer l'adaptation des forêts dans le Plan National d'Adaptation au Changement Climatique (PNACC). Le premier point de son rapport concerne l'identification de la forêt (ce qui inclut donc la localisation et la composition des forêts).

Édité chaque année par l'IGN, le Memento des forêts présente une synthèse de la forêt Française [IGN, 2019]. On y apprend que la forêt recouvre 31% du territoire métropolitain (soit 16,8 millions d'hectares), ce qui en fait la deuxième occupation du sol du pays après l'agriculture et qu'elle est majoritairement composée de feuillus (Figure 1.4). Aujourd'hui, les trois quarts de la surface forestière appartiennent à des propriétaires privés, un dixième à l'État, et le reste aux collectivités régionales, départementales ou locales.

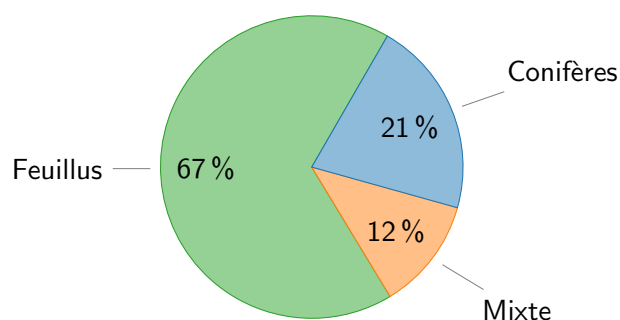


FIGURE 1.4 – Répartition de la composition des peuplements en France métropolitaine.

Chaque année l'IGN récolte des dizaines de milliers de points (55 000 points d'échantillonnage en 2018) pour qualifier et quantifier l'évolution des forêts. Cela sert de base statistique au calcul de l'évolution forestière. Ainsi on peut connaître la composition générale des peuplements : 67% de la forêt en France métropolitaine est composée de feuillus, 21% conifères et seules 12% des forêts sont dites mixtes. Pour définir une forêt comme mixte, l'IGN considère que pour recouvrir 75% du couvert forestier il faut au minimum une essence de feuillu et une de conifère. Plus le nombre d'espèces est donc important pour atteindre les 75% de couvert forestier, plus la forêt est considérée mélangée. Le taux de couvert de la forêt est déterminé à partir de l'ensemble des couverts des arbres par rapport à la superficie de la zone étudiée (Figure 1.5).

Pour classer la forêt comme mélangée, l'IGN classe les couverts relatifs par arbre de façon croissante. Si pour arriver à 75% de recouvrement un seul type d'arbre suffit (feuillu ou conifère), alors la forêt sera considérée comme pure, sinon comme mélangée (Figure 1.6). On observe dans ce dernier graphique que pour atteindre les 75% de couvert, il faut au minimum trois essences, dont deux de feuillus (en vert) et une de conifères (en bleu). Ce peuplement serait donc classé comme mixte.

1.1.3 Sous la canopée

Si depuis le début de ce manuscrit la forêt est définie par les arbres qui la dominent, il s'agit en réalité d'un milieu beaucoup plus riche et variée. Des arbres, non dominants, font aussi partie intégrante de la forêt, ainsi que nombre d'espèces végétales comme les fougères, le lierre ou le fragon, et un nombre important de fleurs sauvages [Guillot, 2014].

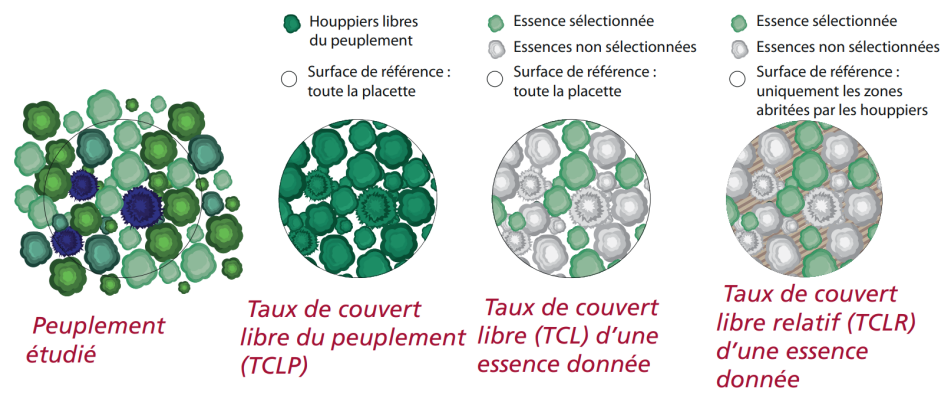


FIGURE 1.5 – Méthode de calcul des différents taux de couvert. En vert, les essences de feuillus, en bleu de conifères. Source : IGN.

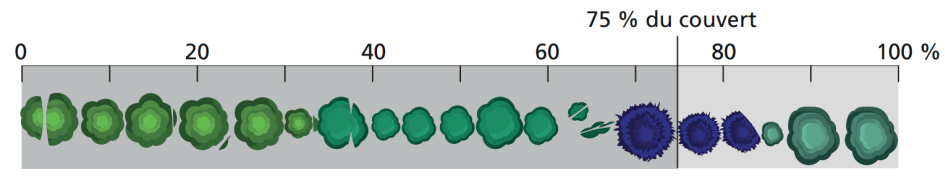


FIGURE 1.6 – Détermination du taux de couvert en tenant compte des différentes espèces par ordre décroissant de taux de couvert libre relatif. Si une espèce ne suffit pas à atteindre 75% de recouvrement, les autres espèces sont introduites. Source : IGN.

Mais la forêt ce n'est pas que des espèces végétales, c'est aussi un habitat pour les animaux, les champignons, les lichens et les algues [Vallauri, 2003]. Ainsi une forêt comme celle de Fontainebleau en France abrite plus de 6 000 espèces d'animaux dont plus de 5 700 insectes, 1 810 espèces végétales et plus de 2 000 espèces de champignons [Vallauri, 2003]. Cette diversité s'explique aussi par la taille des forêts, les plus petites (jusqu'à 300 hectares) peuvent accueillir un nombre d'espèces supérieur à 5 000 alors que les plus grandes (plusieurs milliers d'hectares) peuvent abriter plus de 10 000 espèces différentes [Vallauri, 2003].

1.1.4 De l'intérêt de connaître la distribution des essences

1.1.4.1 Pour la biodiversité

Le dernier rapport de la Plateforme Intergouvernementale scientifique et politique sur la Biodiversité et les Services Écosystémiques (IPBES) alerte sur la perte des milieux naturels "de façon globale et à des rythmes sans précédent dans l'histoire humaine", et cela concerne aussi bien la faune que la flore [IPBES, 2019]. Ainsi, depuis les années 1990, c'est plus de 290 millions d'hectares de forêts primaires qui ont disparu, et la principale cause est anthropique. La disparition des écosystèmes forestiers est majoritairement due à l'accaparement des terres pour l'agriculture et l'élevage. Ainsi, c'est plus d'un tiers de la surface terrestre du monde et près de 75% des ressources en eau douce qui destinée à cet usage.

La diminution de la forêt au niveau global s'accompagne aussi d'une baisse du nombre d'essences. Or, le nombre d'espèces est l'un des critères de base de la biodiversité [Gamfeldt et al., 2013]. Ces auteurs ont montré qu'une forêt de 5 essences différentes produisait 50% plus de biomasse et abritait 31% plus d'espèces dans son sous-bois. Afin de caractériser la richesse potentielle, l'indice Indice de Biodiversité Potentielle (IBP) a été développé par Larrieu and Gonin [2008]. L'Indice de Biodiversité Potentielle (IBP) estime indirectement la diversité d'un milieu en se basant sur les relations entre la diversité des espèces et des caractéristiques propres au peuplement (complexité structurale et continuité spatio-temporelle). Ainsi connaître le nombre d'espèces dans une forêt apparaît comme essentiel [Thompson et al., 2011]. Certains travaux ont montré que plus le nombre d'essences était élevé, plus la résistance aux pathogènes

1.1. LA FORÊT, BIEN PLUS QUE DES ARBRES

était élevée [Guo et al., 2019; Jactel et al., 2017]. Il a aussi été montré que le nombre d'essences pouvait augmenter la production de biomasse des forêts en sylviculture [Sugden, 2018].

Cependant, il se trouve qu'on assiste aujourd'hui à un risque de déclin du nombre d'espèces forestières, surtout dans les forêts tropicales. Selon Hubbell et al. [2008] ce sont, avec le scénario le plus optimiste, près de 2 000 espèces d'arbres sur les 11 210 présents en Amazonie qui sont menacés d'extinction. En Europe continentale la menace pèse aussi dans les forêts tempérées. Selon l'UICN, 37% des 454 espèces originaires du continent européen étaient menacées en 2019 [Rivers et al., 2019]. La plupart des espèces se trouve sur le pourtour méditerranéen (Figure 1.7), et plus on s'éloigne de la mer et plus on monte vers le nord, moins il y a de richesse spécifique, c'est ce qu'on appelle un effet biogéographique. Si la carte de l'UICN permet d'avoir une estimation à l'échelle européenne de la diversité des essences, elle ne permet pas de connaître précisément la localisation et encore moins leur nombre. En effet cette carte est réalisée à partir de multiples sources avec des méthodes de récolte très variées. Les données proviennent de sources divers comme l'*Atlas Floae Europaeae* Kurto et al. [2013], le *European Forest Genetic Resources Programme* (EUFORGEN), le *Global Biodiversity Information Facility* (GBIF), ou encore l'atlas européen des essences forestières [Caudullo et al., 2017; San-Miguel-Ayanz et al., 2016].

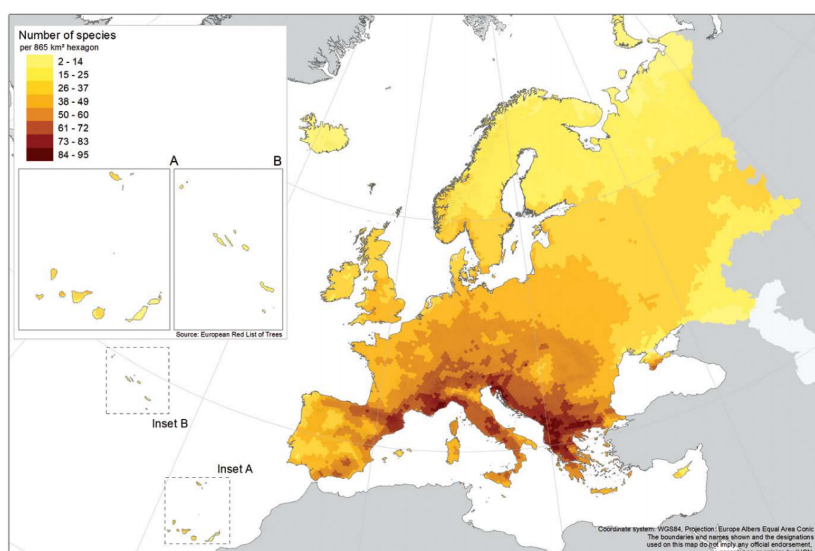


FIGURE 1.7 – Répartition de la diversité des espèces d'arbres en Europe. Source : UICN.

Avec le réchauffement climatique, la fréquence des sécheresses pourrait augmenter [Dai, 2011] mais surtout, les sécheresses pourraient être plus sévères [Trenberth et al., 2014]. Plusieurs travaux ont montré que les forêts les plus résistantes aux périodes de manque d'eau et de fortes chaleurs sont celles qui sont le plus diversifiées au niveau des essences [Andregg et al., 2018; Fichtner et al., 2020; Sousa-Silva et al., 2018; Sugden, 2020].

1.1.4.2 Pour le changement climatique

Le changement climatique a plusieurs incidences sur la végétation car il influe à la fois de manière indirecte (pathogènes par exemple) et de manière directe (température, ensoleillement ou précipitation) [Brasier and Scott, 1994; Kolb et al., 2016]. Ainsi on observe deux grandes conséquences sur les arbres :

- l'évolution du cycle annuel de la végétation (court à moyen terme)
- la capacité d'adaptation des espèces (long terme)

Selon Duputié et al. [2015], l'adaptation à moyen terme, à travers la plasticité de la phénologie, ne pourrait pas garantir l'adaptation d'une espèce sur du plus long terme.

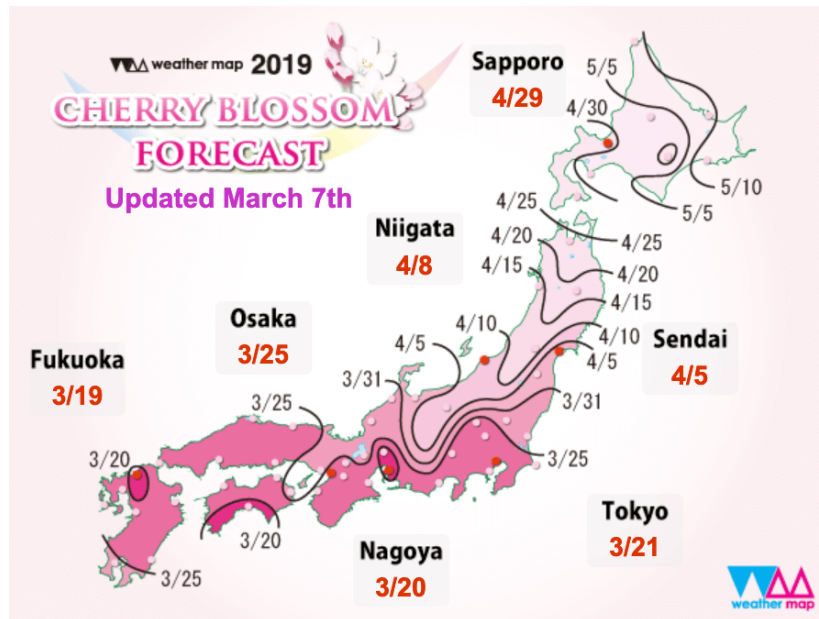


FIGURE 1.8 – Météo de la floraison des cerisiers du Japon (*Prunus serrulata*) en 2019. Source : WeatherMap.

L'évolution du cycle annuel de la végétation, nommé phénologie, est sensible à l'évolution du climat. Ainsi une espèce qui se situe à des latitudes ou des altitudes différentes n'aura pas la même date de début de saison (ouverture et pousse des feuilles), de floraison ou encore de sénescence. C'est pourquoi, pour une majorité d'espèces comme le cerisier du Japon (*Prunus serrulata*), on observe des écarts de plus d'un mois dans la période de floraison entre le sud et le nord du Japon (Figure 1.8). Il y a donc un intérêt croissant dans la recherche à étudier la phénologie de la végétation en lien avec le changement climatique [Cleland et al., 2007; Khanduri et al., 2008; Morin et al., 2009].

La phénologie permet aussi de mieux appréhender les conditions climatiques des siècles passés et de mettre en lien les changements récents avec des aléas climatiques historiques. Chuine et al. [2004] ont pu remonter jusqu'en 1370 et obtenir les dates de vendage du Pinot Noir en Bourgogne grâce à des documents historiques. Au pays du Soleil-Levant, Aono and Kazui [2008] sont remontés jusqu'en 800 et ont observé que la date de floraison du cerisier à Kyoto était beaucoup plus précoce depuis un demi siècle (Figure 1.9). De manière globale, Khanduri et al. [2008] notent une avance de 1,9 jours en moyenne par décennie dans le début de saison et un retard de 1,4 jours dans la sénescence pour des espèces de climat tempéré. Concernant le chêne, la longueur de la saison semble s'être étendue entre 6 et 13 jours [Vitasse et al., 2009].

Cependant, l'avancement du débourrement ne semble pas être uniquement lié aux températures. Des arbres de forêts tempérées mis dans des conditions tropicales n'ont jamais débuté leur saison avant mars et ce malgré de fortes chaleurs [Borchert et al., 2005], ce qui montre que dans les forêts tempérées, la température n'est pas l'unique élément déclencheur de la floraison ou de la sénescence [Körner and Basler, 2010].

Pour quantifier le changement climatique, les chercheurs et chercheuses ont besoin d'informations précises. Pour cela les *Essential Climate Variables* ou Variables Climatiques Essentielles en français (ECV) [Bojinski et al., 2014] ont été définies et nécessitent pour être calculées des informations aussi diverses que la vitesse ou la direction du vent, ou encore le type d'occupation du sol. Les ECV regroupent au total 50 variables, dont une sur le type d'occupation du sol, le Leaf Area Index (Section 1.3.3.5) ou encore la biomasse aérienne. Pour mieux estimer ces variables, le *Global Climate Observing System* (GCOS) recommande la production de trois cartes [Secretariat, 2009] dont une à haute résolution spatiale (qu'ils

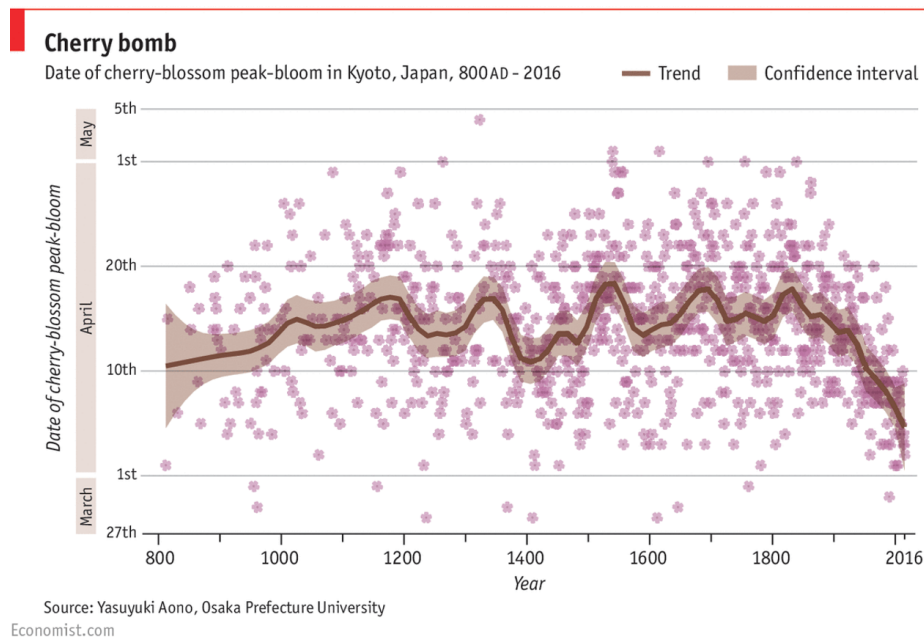


FIGURE 1.9 – Évolution de la date de floraison des cerisiers du Japon *Prunus serrulata* depuis le IX^e siècle. Source : The Economist. Données : [Aono and Kazui \[2008\]](#).

définissent comme allant de 10 à 30 m) avec 5% de taux d'incertitude par classe. Une cartographie des forêts apparaît donc comme important pour mieux définir les ECV.

Ce qu'il faut retenir

- La forêt se définit selon la FAO comme un espace d'au moins 0,5 hectare qui est composé d'arbres pouvant atteindre à maturité au moins 10% du couvert forestier et plus de 5 mètres de hauteur.
- Il s'agit d'un écosystème particulièrement riche qui peut abriter des milliers d'espèces, essentiellement animales.

Connaître la distribution spatiale des essences forestières permet :

- de mieux estimer la biodiversité des forêts car elle dépend du nombre d'espèces présentes en son sein.
- d'analyser plus finement l'impact du climat futur car certaines essences risquent de ne pas s'acclimater aux nouvelles conditions de leur habitat.

1.2 Cartographier les forêts, une histoire ancienne

1.2.1 Les premières cartes topographiques

Pour connaître son territoire Louis XIV a commandé à la famille Cassini la première carte topographique à l'échelle nationale [[Pelletier, 1990](#)]. Le niveau de détail est élevé pour le XVIII^e siècle avec une échelle de représentation de 1/86 400 (1 cm représentant 864 m sur le terrain). Pour cartographier la France entière, il a fallu pas moins de 60 ans et quatre générations de Cassini car la carte était entièrement réalisée manuellement en parcourant le territoire et en s'appuyant sur la triangulation géodésique (technique consistant à déterminer la position d'un point en mesurant les angles entre sa position et deux autres points) [[Dainville, 1955](#)]. Une vectorisation de l'ensemble des forêts de la carte de Cassini a été réalisée au début



FIGURE 1.10 – Évolution de la cartographie de la forêt de Bouconne : de la carte de Cassini en 1740 (a), à la carte de l'État-Major en 1866 (b).

des années 2010 et permet de connaître l'emplacement des forêts de feuillus et de conifères sous Louis XIV [Vallauri et al., 2012].

Après plus de 60 ans de réalisation, il s'est avéré qu'à peine quelques décennies plus tard il y avait déjà un besoin d'avoir une carte plus précise du relief et du territoire. C'est ainsi que l'ordonnance royale de 1827 confia à l'armée la conception d'une carte à l'échelle 1/80 000 avec comme condition la représentation la plus fidèle possible du relief. Les levés ayant servis à la réalisation de la carte finale sont plus précis (1/40 000).

La cartographie de l'occupation du sol en France au milieu du XIX^{ème} siècle est très utile pour les forestiers et les écologues car elle correspond au plus bas niveau de surface forestière en France. Afin de tirer profit de ce travail qui date de presque 200 ans, Lallemand et al. [2017] ont publié une méthode manuelle pour vectoriser et donc connaître facilement à l'échelle de la France l'emplacement des forêts anciennes. Afin de faciliter ce travail, une méthode de vectorisation automatique a été développée [Herrault et al., 2015] et incorporée dans le logiciel de Système d'Information Géographique (SIG) QGIS pour rendre encore plus accessible la détection automatique des forêts anciennes [Karasiak et al., 2016]. Cela permet donc de savoir si une forêt est présente depuis au moins 1850 ce qui augmente potentiellement la biodiversité. L'usage de l'histoire pour expliquer l'écologie présente se nomme l'écologie historique. Ce champs de recherche interdisciplinaire a par exemple montré que les espèces présentes dans les forêts anciennes sont plus souvent sciaphiles, acidiphiles, et moins exigeantes en azote que les espèces présentes dans les forêts récentes [Bergès and Dupouey, 2017].

1.2.2 La cartographie actuelle des forêts françaises

Si la carte de Cassini donnait une indication sur la composition des feuillus et des conifères, il n'était pas question de cartographier les espèces. La carte de l'État-Major, ayant pour but

de servir l'armée, ne contient aucune information de composition, seule la localisation des forêts est mentionnée.

Soucieux de connaître les essences peuplant notre territoire, l'IGN a lancé en 1987 une grande campagne de cartographie. Il n'était plus question cette fois-ci de passer 60 ans à parcourir la France ou de faire appel à l'armée. La nouvelle carte des forêts devait aussi recenser les principales espèces qui la composent. La partie la plus chronophage de ce travail a été faite derrière un écran en photo-interprétant des photographies aériennes à très haute résolution [Boureau \[2008\]](#). Il faudra pas moins de 17 ans (de 1987 à 2004) pour produire une cartographie de la France métropolitaine des forêts de superficie supérieure à 2,25ha. La carte numérique nommée BDForêt V1 est consultable en ligne sur le site de l'IGN¹.



FIGURE 1.11 – Les essences composant la forêt de Bouconne selon la BD Forêt v2 de l'IGN. À chaque couleur correspond une essence ou un mélange différent.

Après le succès de cette première expérience, l'IGN répéta une nouvelle fois l'opération en 2007. La nouvelle cartographie est plus ambitieuse : elle détaille au total une quinzaine d'essences, avec une surface minimale de collecte de 0,5 ha, soit un niveau de détail près de 5 fois supérieur à la précédente version [[Touzet and Lecordix, 2010](#)]. Afin de faciliter le travail des opérateurs, la détection des limites de la forêt s'est faite à travers un procédé automatique qu'on appelle la segmentation. Un algorithme, va, à partir des différences spectrales des pixels, différencier des zones hétérogènes à partir d'une approche multi-échelles [[Guigues et al., 2006](#)]. Ainsi, la nouvelle cartographie nommée BD Forêt V2 a demandé 12 ans de travail au lieu de 17, tout en améliorant significativement le niveau de détail [[Garnier, 2019](#)]. La reconnaissance des essences se fait toujours de manière manuelle à partir des BD Ortho (images à 50 cm de résolution spatiale et disponible en infrarouge couleur) et avec une étape de vérification terrain. Cependant, si la couverture est nationale, elle ne donne pas la présence et la composition de la forêt à une date précise. En effet, les images utilisées pour recenser les essences datent pour les premiers départements de 2004 (06, 10, 36, 51, 56) et vont jusqu'à 2015 pour les derniers départements traités (11, 30, 34, 66). Afin de pallier en partie à ce décalage temporel, l'IGN a décidé de produire à nouveau les premiers départements mais cette fois-ci avec des photographies aériennes à très haute résolution spatiale acquises en 2016 et 2017.

La prochaine version de la BD Forêt n'est pas encore planifiée mais les récentes avancées en apprentissage automatique laissent entrevoir une évolution aussi significative qu'entre la version 1 et la version 2. La procédure de cartographie des essences en France devrait alors davantage s'automatiser notamment grâce à l'afflux de données satellitaires qui permettent d'observer la végétation et donc leurs propriétés optiques sur tout un territoire et de manière répétée. Par exemple, afin de détecter automatiquement les parcelles homogènes, [Dechesne et al. \[2017\]](#) ont mis au point une méthode combinant des données LiDAR et des images à très haute résolution spatiale, et leurs résultats sont concluants, avec 84% d'accord global pour reconnaître l'essence dominante et 99% pour délimiter les parcelles.

1. <https://inventaire-forestier.ign.fr/spip.php?article647>

Pour globalement, pour cartographier la forêt sans distinction d'essences, une première carte d'occupation du sol à l'échelle de la France a été réalisée de manière automatique à partir de données satellites par [Inglada et al. \[2017\]](#). Cependant, cette carte se base sur la BD Forêt de l'IGN qui est réalisée de manière manuelle avec une étape de segmentation automatique pour détourner les forêts.

Ce qu'il faut retenir

La cartographie des forêts évolue rapidement :

- Les premières cartographies étaient manuelles et montraient la présence ou l'absence de forêt. Ainsi des cartes des forêts existent depuis environ deux siècles, avec en partie la distinction entre feuillus et conifères.
- L'automatisation du processus arrive petit à petit avec notamment la segmentation automatique pour délimiter les massifs ou à travers la cartographie automatique des feuillus/conifères.
- La reconnaissance des essences reste aujourd'hui entièrement du ressort des photo-interprètes.

1.3 La télédétection comme alternative à la cartographie manuelle

Initialement la télédétection se faisait par avion et était du domaine de compétence militaire. Cependant, avec le développement des satellites au début des années 70, la capacité d'observation de la terre et son suivi se sont peu à peu transférés vers le civil. Aujourd'hui les informations satellitaires sont primordiales pour de nombreux domaines mais nous allons nous intéresser aux satellites qui permettent d'observer le couvert terrestre.

Nous allons dans un premier temps voir les principes de base de la télédétection, pour ensuite étudier les méthodes et les résultats concernant la cartographie des essences forestières, et enfin nous analyserons les limites actuelles du domaine.

1.3.1 Principes de bases de la télédétection

1.3.1.1 Les grands types de capteur

La télédétection signifie littéralement la détection à distance. Elle regroupe l'ensemble des méthodes pour observer, analyser, et interpréter des phénomènes (e.g. pression atmosphérique) ou des objets comme la surface terrestre.

On distingue deux grandes familles de capteurs (Figure 1.12) :

- **Les capteurs passifs** utilisent le rayonnement naturel du soleil comme source d'énergie. Le signal reçu par ce capteur sera la radiation réfléchiée par un objet. Cela se rapproche de la vision que nous pouvons avoir des objets, mais les capteurs peuvent être plus précis dans la perception des couleurs et peuvent percevoir l'infrarouge.
- **Les capteurs actifs** émettent leur propre rayonnement électromagnétique qui est dirigé vers l'objet à étudier. On note deux grandes familles, celle du Radio Detection And Ranging (RADAR) ou encore du Light Detection And Ranging (LiDAR) qui est pour la première fois disponible par satellite pour le civil depuis 2019 [[Hancock et al., 2019](#)].

Dans le cadre de cette thèse nous nous intéresserons uniquement aux capteurs passifs qui fournissent des images optiques. Un satellite pour la télédétection se caractérise par ses différentes résolutions : radiométrique, spatiale, spectrale et temporelle.

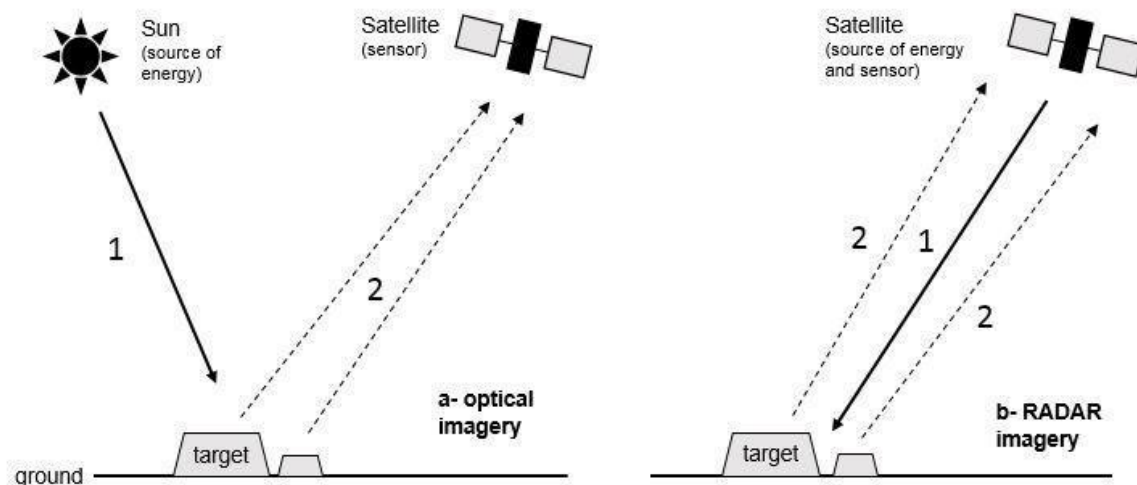


FIGURE 1.12 – Différences de fonctionnement entre un capteur optique (à gauche) et un capteur radar (à droite). Source : [Catry et al. \[2018\]](#)

1.3.1.2 Les résolutions

Résolution radiométrique

La résolution radiométrique est la capacité à reconnaître de petites différences dans l'énergie électromagnétique reçue. Plus la résolution radiométrique d'un capteur sera élevée, plus le capteur sera sensible aux différences dans l'intensité du signal. Cette résolution est exprimée en nombre de bits, allant généralement de 8 à 16 bits. Pour Sentinel-2 la résolution radiométrique est de 12 bits, ce qui correspond à une sensibilité de signal allant de 0 à 4095. Cette résolution est affectée par le rapport signal sur bruit (SNR) qui est spécifique pour chaque bande spectrale [ESA, 2015].

Résolution spatiale

Une image optique est composée d'une matrice d'éléments, appelés pixels. Ils correspondent au plus petit élément d'une image qui est la plupart du temps de taille carrée. Il y a une correspondance entre la taille du pixel et la taille que cela représente sur le terrain. Si l'élément à observer est plus petit que la taille d'un pixel, il ne sera généralement pas possible de le différencier (sauf à partir de certaines méthodes spécifiques et pour des objets particuliers).

Pour pouvoir bien différencier un objet observé à distance, il est de convenance qu'il soit au minimum égal à la résolution. Cependant, une très haute résolution spatiale inférieure à 5m va avoir tendance à capter d'autant plus d'hétérogénéité (arbre isolé, ombre etc...) alors qu'une faible résolution ne va que trop homogénéiser les objets (et ainsi inclure pour les forêts beaucoup trop des mélanges d'essences, Figure 1.13).

Selon [Shunlin Liang and Jindi Wang \[2012\]](#) on peut classer les résolutions spatiales en 4 grandes catégories :

- faible, avec une résolution inférieure au kilomètre,
- modérée allant de 100 m à 1 km,
- haute, de 5 à 100 m,
- très haute, inférieure à 5 m.

Enfin, il est à noter que la résolution spatiale du pixel peut aussi dépendre de la bande spectrale comme pour le satellite Sentinel-2 dont la résolution va de 10 m à 20 m (pour les bandes fournies au grand public).

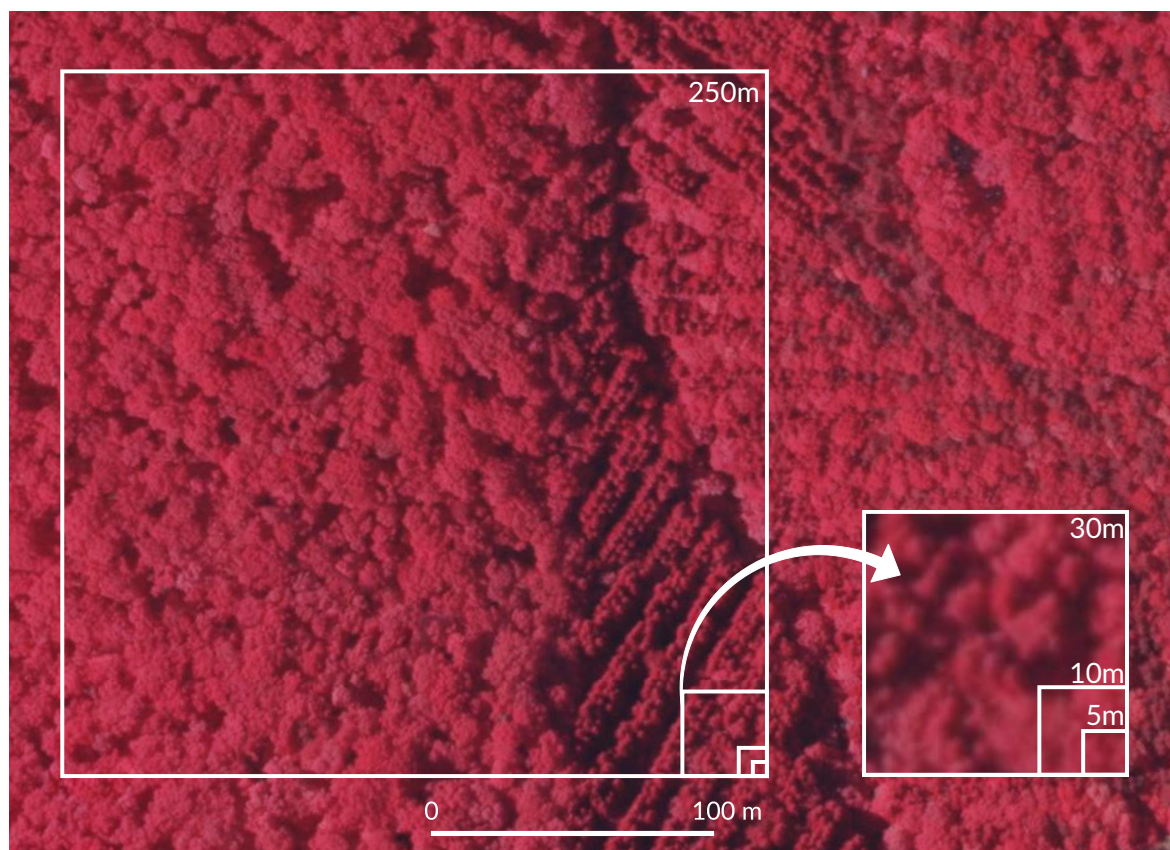


FIGURE 1.13 – Influence de la résolution spatiale sur la taille d'un échantillon du couvert forestier par des surfaces carrées de 5 m (RapidEye), 10 m (Sentinel-2), 30 m (Landsat-8) et 250 m (MODIS). Fond : BD Ortho en infrarouge couleur à la résolution spatiale native de 0,5m. Inspiré de [Guyon and Breda \[2016\]](#).

Résolution spectrale

La résolution spectrale est une résolution un peu particulière car elle peut s'interpréter de deux manières, c'est-à-dire par :

- La largeur des bandes, soit le nombre de nanomètres qu'enregistrera une bande spectrale.
- Le nombre de bandes, c'est-à-dire le nombre total de bandes spectrales que sera capable de mesurer le capteur.

Souvent ces deux paramètres sont concordants, car plus l'intervalle entre deux bandes est réduit, plus un capteur pourra acquérir de bandes spectrales.

La définition admise est le pouvoir de séparation spectrale, ainsi plus la résolution spectrale est grande, plus la largeur d'une bande sera fine, on parlera ici alors de finesse de résolution. Des satellites hyperspectraux comme EnMAP [[Guanter et al., 2015](#)] ou PRISMA [[Sacchetti et al., 2010](#)] peuvent acquérir des images avec une résolution spectrale inférieure à 12 nm de 400 à 2505 nm. Ainsi, alors qu'un capteur comme Sentinel-2 capte 3 bandes dans le visible (450 à 700 nm), EnMAP pourra en capturer 25 mais à 30 m de résolution spatiale contre 10 m pour Sentinel-2.

Résolution temporelle

Outre les caractéristiques liées à chaque image (taille de l'image, des pixels et nombre de bandes), le temps de revisite du satellite est aussi une caractéristique importante. Il représente la durée nécessaire au satellite pour imager à nouveau la même scène.

1.3. LA TÉLÉDÉTECTION COMME ALTERNATIVE À LA CARTOGRAPHIE MANUELLE

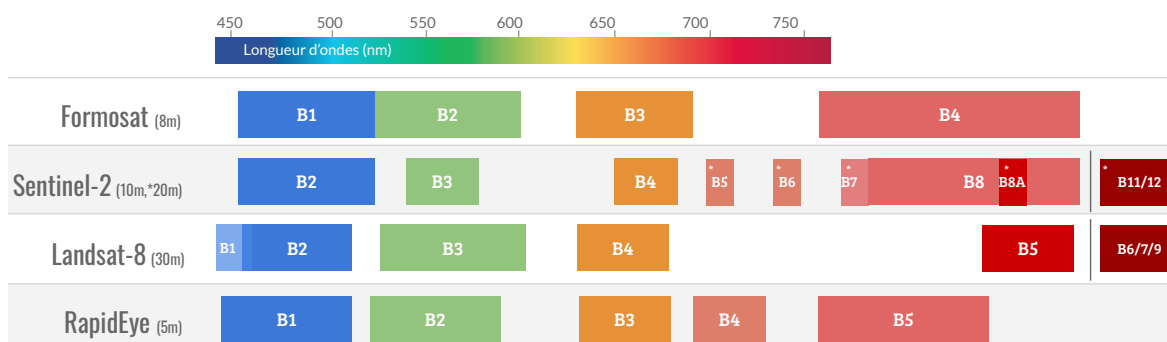


FIGURE 1.14 – Principaux capteurs à haute résolution spatiale (de 5 à 100 m) détaillant la longueur d'ondes de chaque bande spectrale. Pour Sentinel-2, la résolution spatiale dépend de la bande spectrale. Pour l'infra-rouge après 1000 nm et afin de simplifier le graphique, les bandes sont agrégées.

Certains satellites, dits géostationnaires, observent le même point de la terre systématiquement, mais ils sont généralement dédiés à la télécommunication et à la météorologie à cause de leur orbite très élevée à environ 36000 km d'altitude.

Dans l'observation de la terre, les satellites de télédétection sont souvent héliosynchrones (c'est-à-dire qu'ils suivent l'éclairage du soleil). Ils sont placés en orbite basse entre 600 et 800 km d'altitude et peuvent prendre des images de l'ensemble du territoire à la même heure solaire locale. Ces satellites capturent la plupart du temps à angle constant, ce qui signifie que le satellite ne changera pas d'angle de vue pour capturer une zone différente. Son temps de revisite sera donc le même que son cycle orbital. Pour les autres satellites qui peuvent s'orienter afin de capturer des zones précises, le temps de revisite peut être très inférieur au cycle orbital.

Malheureusement, il est difficile pour les capteurs d'observer à la fois des scènes à hautes résolutions spatiale et spectrale sur de grandes étendues avec un fort temps de revisite. En plus d'être contraints par les orbites des satellites et les capacités d'enregistrement, les capteurs sont soumis à plusieurs compromis entre la résolution spectrale et le rapport signal sur bruit, entre la résolution spatiale et le volume de données à stocker, entre la résolution spatiale et la résolution spectrale, et entre la résolution spatiale et la résolution temporelle.

Pour augmenter la résolution spatiale tout en gardant une résolution temporelle élevée voire quotidienne, il est d'usage d'augmenter le nombre de satellites. Ainsi avec deux satellites séparés de 180° sur la même orbite, on multiplie par deux la fréquence d'acquisition comme c'est le cas pour Sentinel-1 ou Sentinel-2. L'entreprise Planet a vu plus loin et s'est lancée dans la création d'une constellation de satellites nommée SkySat qui lui permet de cartographier chaque jour n'importe quel point de la terre avec une résolution spatiale de 2 m [Strauss, 2017].

Ce qu'il faut retenir

Un capteur satellite se définit par quatre éléments :

- sa **résolution radiométrique** qui est la plus petite différence d'intensité du signal qui peut être enregistrée (8, 11, 16 bits...)
- sa **résolution spatiale**. Elle dépend de l'orbite et de l'ouverture angulaire de l'instrument, et elle correspond à la taille du pixel au sol.
- sa **résolution spectrale**. Il s'agit de la largeur des bandes spectrales, plus celle-ci est fine plus la résolution spectrale est importante.
- sa **résolution temporelle**, autrement dit la fréquence d'un satellite à revisiter un même endroit.

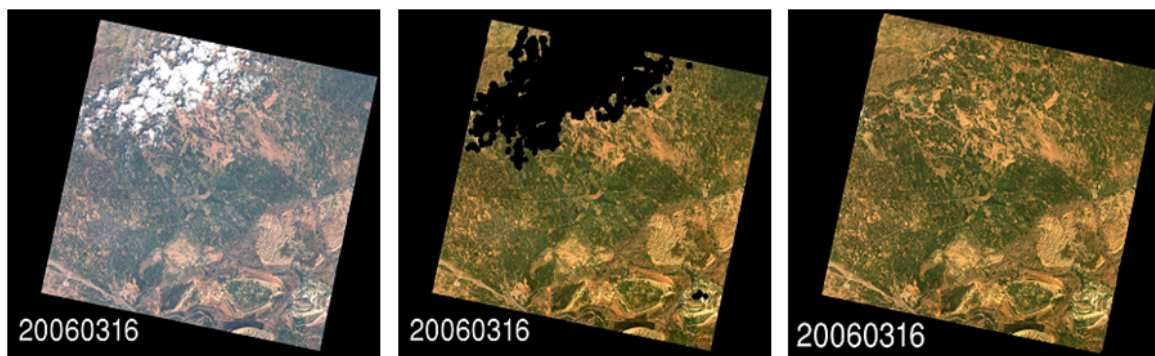
1.3.1.3 Du satellite à l'image

Entre l'acquisition brute du satellite et l'exploitation de l'image pour la télédétection, il y a beaucoup d'opérations nécessaires qui sont plus ou moins automatisées.

Les différents niveaux de pré-traitements

Cette section, bien que se référant à une nomenclature utilisée par la National Aeronautics and Space Administration (NASA), ne prévaut pas pour l'ensemble des satellites. Les niveaux de pré-traitements détaillés par la suite sont cependant conformes avec les produits développés par le pôle Theia qui a été créé fin 2012 par près d'une dizaine d'institutions publiques françaises travaillant dans l'observation de la terre et dans les sciences de l'environnement.

Le *Committee on Earth Observation Satellites* (CEOS) définit plusieurs niveaux de traitements des images, allant du niveau 0 au niveau 4. Ce comité s'est basé sur la nomenclature du Earth Observing System Data and Information System (EOSDIS) de la NASA² pour définir les niveaux de traitements. Pour plus de finesse dans la nomenclature, nous utiliserons tout au long de ce manuscrit la version définie par Oliver Hagolle³ (Figure 1.15) et qui correspond à celle utilisée par le pôle Theia. Les niveaux suivants qui seront détaillés s'appliquent aux satellites optiques dédiés à l'observation de la terre et en particulier au satellite Sentinel-2.



Level 1C:

Level 2A:

Level 3A:

FIGURE 1.15 – Les trois grands niveaux de traitements des images satellites. Pour le niveau 2A, la tache noire correspond aux nuages et ombres qui ont été détectés par un algorithme. Crédits : Oliver Hagolle.

Le niveau 0 Ce sont les données reconstruites mais non traitées des capteurs. Elles sont fournies à pleine résolution accompagnées de toutes les informations nécessaires pour les traitements des niveaux supérieurs. Ces données ne sont généralement pas communiquées à l'utilisateur en télédétection et sont réservées pour les centres de traitements qui distribuent la donnée de niveau supérieur (1 ou plus).

Le niveau 1A Pour Sentinel-2, le produit de niveau 1A est la décompression du niveau 0. Un modèle géométrique est calculé et permet de localiser chaque image à partir d'informations du capteur comme son altitude mais aussi de données externes comme une base de données mondiale d'images géoréférencées [Baillarin et al., 2012]. Chaque orbite contient environ 3500 granules de niveau 1A. Ainsi, les données 1A sont les images d'une orbite Sentinel-2 où chaque granule (ou sous-image) est géoréférencée. Là encore, il est rare que ce niveau de traitement soit accessible à l'utilisateur.

2. <https://science.nasa.gov/earth-science/earth-science-data/data-processing-levels-for-eosdis-data-products/>

3. <https://labo.obs-mip.fr/multitemp/les-noms-des-produits-comment-ca-marche/>

1.3. LA TÉLÉDÉTECTION COMME ALTERNATIVE À LA CARTOGRAPHIE MANUELLE

Les corrections radiométriques incluent la détection de pixels défaillants ou de la diaphonie (*crosstalk* en anglais), soit l'interférence d'un premier signal avec un second.

Le niveau 1B L'ensemble des images (ou granules pour Sentinel-2) est corrigé radiométriquement mais les effets de l'atmosphère (comme les aérosols) ne sont pas corrigés, le produit est donc dit *Top Of Atmosphere* (TOA).

Comme pour le niveau 1A, ces données ne sont pas accessibles à l'utilisateur final.

Le niveau 1C Il s'agit généralement du premier niveau accessible à l'utilisateur (Sentinel-2, Landsat). Dans le cas de Sentinel-2, ce niveau 1C s'appuie sur les traitements du niveau précédent (1B) pour créer des tuiles de taille 100 x 100 km. En plus du regroupement, une correction géométrique est calculée à partir d'un Modèle Numérique de Terrain (MNT). Chaque pixel de la tuile est alors traité de la même manière qu'il soit composé d'eau, de forêt, de nuages ou de routes. La valeur des pixels est exprimée en grandeur physique.

Ce type de produit est le plus populaire aujourd'hui car il est le plus couramment distribué par les producteurs de la donnée auprès du grand public. Ainsi, l'ESA ne distribuait les premières années de Sentinel-2 qu'au maximum le niveau 1C. Depuis le niveau 2A est lui aussi disponible, mais le niveau 1C reste toujours deux fois plus téléchargé en 2019 [ESA, 2020].

Le niveau 2A Le produit de niveau 2 se démarque du niveau 1 par des traitements différenciés par type de pixel. Ainsi l'on peut faire des hypothèses sur l'atmosphère, l'état de la végétation ou de la mer, et traiter les différentes classes de pixels d'une manière différente. Un masque localisant les nuages et les ombres est généralement fourni avec la donnée (comme c'est le cas avec le satellite Sentinel-2).

Afin d'avoir une information temporelle, les données sont normalisées suite au passage en réflectance et ne sont presque plus impactées par les effets de l'atmosphère. Le niveau 2A est de plus en plus populaire et distribué gratuitement, que ce soit par le pôle Theia ou plus récemment par l'European Space Agency (ESA). Il s'agit du premier niveau de traitement qui rend accessible des données dites *Top Of Canopy* (TOC).

Le niveau 3A Le produit de niveau 3 est significativement différent des niveaux précédents. La donnée produite peut être un mélange de plusieurs dates et de plusieurs emprises. Cela permet par exemple d'avoir une synthèse mensuelle, comme le fait le CESBIO, de l'état de la France. Le produit de niveau 3 est un mélange de données généralement produites à partir du niveau 2 et qui peuvent y apporter des corrections supplémentaires, comme le traitement des effets bidirectionnels (BRDF) qui peuvent apparaître lorsque les données sont acquises avec des orbites différentes⁴.

Dans le cadre de ce manuscrit et afin d'avoir les données les plus stables car normalisées suite au passage en réflectance d'une date à l'autre (en corrigeant au maximum les effets de l'atmosphère), les données utilisées seront toutes du niveau 2A.

4. <https://labo.obs-mip.fr/multitemp/les-series-temporelles-de-sentinel-2-sont-bien-lisses/>

Ce qu'il faut retenir

Parmi les nombreux niveaux de traitements des images d'un satellite, un petit nombre nous intéresse en tant qu'utilisateur final :

- Le **niveau 1C** correspond à des données géoréférencées et orthorectifiées. Cependant il n'y a pas de traitement des aérosols ou autres effets de l'atmosphère. La donnée acquise sera de type *Top Of Atmosphere* (TOA).
- Le **niveau 2A** apporte au niveau 1C un masque de nuages et enlève les effets de l'atmosphère, ce qui permet d'avoir un signal temporel plus stable. Il est ainsi possible d'obtenir des pixels en réflectance dites *Top Of Canopy* (TOC).

1.3.1.4 Série temporelle avec ou sans nuages ?

On différencie dans ce manuscrit une *Satellite Image Time Series* ou en français, Série Temporelle d'Images Satellites (SITS) d'un ensemble d'images par le fait que la première est issue d'une série d'images de grande densité temporelle comprenant des données partiellement ennuagées auxquelles on applique généralement un traitement de type *gap-filling* visant à combler les données manquantes (Figure 1.16). La résolution temporelle d'une SITS est donc systématiquement supérieure à celle d'un ensemble d'images. L'autre avantage est que la résolution temporelle peut être homogène sur tout un territoire (ce qui nécessite bien souvent une interpolation des données). Ainsi, si une zone est masquée par un nuage, la série temporelle profitera d'une donnée cohérente sur la partie de l'image non impactée et comblera le manque de données uniquement à l'endroit où un nuage a été détecté.

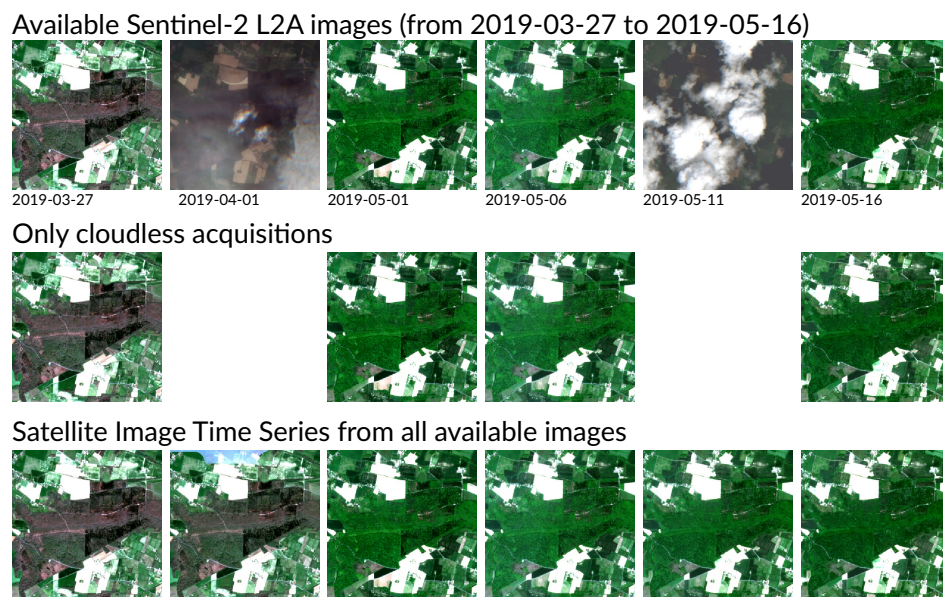


FIGURE 1.16 – Différence entre une SITS et un ensemble d'images sans nuage. Pour l'exemple ci-dessus, seules les dates entre le 27 mars et le 16 mai 2019 disponibles au niveau 2A sur le site du pôle Theia ont été gardées. Afin de générer la série temporelle, une interpolation linéaire a été réalisée lorsqu'un pixel était considéré comme masqué (ce qui a été le cas pour 1er avril et le 11 mai).

1.3.2 Fonctionnement général de l'apprentissage supervisé

L'apprentissage supervisé fonctionne avec trois éléments : tout d'abord il faut des variables qui sont dans notre cas des valeurs spectrales issues des images satellite. Puis parmi l'ensemble des variables d'une image, certaines sont annotées, c'est-à-dire que l'on a pour certaines d'entre elles la correspondance entre les valeurs spectrales et l'essence forestière. Et enfin,

vient se mettre en place un algorithme qui a pour but d'apprendre un modèle à partir des deux données précédemment citées.

1.3.2.1 Les algorithmes utilisés

Les algorithmes utilisés dans ce manuscrit appartiennent à la famille de l'apprentissage supervisé qui consiste à apprendre une fonction qui a pour but de prédire des données à partir d'exemples préalablement annotés. Les algorithmes ont pour fonction de trouver des points communs entre les variables d'échantillons d'une même classe (dans notre cas il s'agit d'essences) afin de les regrouper entre eux. Une fois le modèle appris, l'algorithme va pouvoir prédire à partir des variables la classe la plus probable d'appartenance.

Les algorithmes présentés par la suite sont de type non paramétriques, c'est-à-dire qu'ils ne supposent pas que la distribution des valeurs des données suivent une loi de distribution, à la différence des algorithmes paramétriques. Ils ne se basent donc que sur les propriétés numériques des échantillons. Le choix des deux algorithmes présentés ci-après et utilisés dans la thèse s'explique par le fait qu'ils sont très utilisés et réputés pour la cartographie des essences forestières [Fassnacht et al., 2016] et qu'ils sont robustes aux données de grande dimension comme les séries temporelles [Cortes and Vapnik, 1995; Inglada et al., 2017].

Support Vector Machines

L'algorithme *Support Vector Machines* ou en français, Machines à Vecteur de Support (SVM) [Vapnik, 1998] est le plus utilisé pour la reconnaissance des essences forestières [Fassnacht et al., 2016]. Il est reconnu pour sa robustesse et sa capacité d'apprentissage d'un modèle à partir d'un faible nombre d'échantillons par classe et d'un nombre important de variables [Melgani and Bruzzone, 2004; Mountrakis et al., 2011].

Pour apprendre un modèle, l'algorithme reprojete les données en entrée dans plusieurs hyperplans définis par des paramètres saisis par l'utilisateur. L'hyperplan où les échantillons auront la marge la plus large, c'est-à-dire où l'écart est le plus important entre les échantillons les plus proches de classes différentes, sera sélectionné pour construire le modèle.

Plusieurs noyaux peuvent être appliqués à l'algorithme (Figure 1.17). Initialement développé avec un noyau de type linéaire, d'autres noyaux ont été implémentés par la suite comme le gaussien, nommé Radial Basis Function (RBF), qui a fait ses preuves dans des problèmes similaires au nôtre [Ferreira et al., 2016; Kavzoglu and Colkesen, 2009; Sheeren et al., 2016].

De plus amples détails notamment sur le paramétrage de l'algorithme SVM sont disponibles en Section 2.3.2.

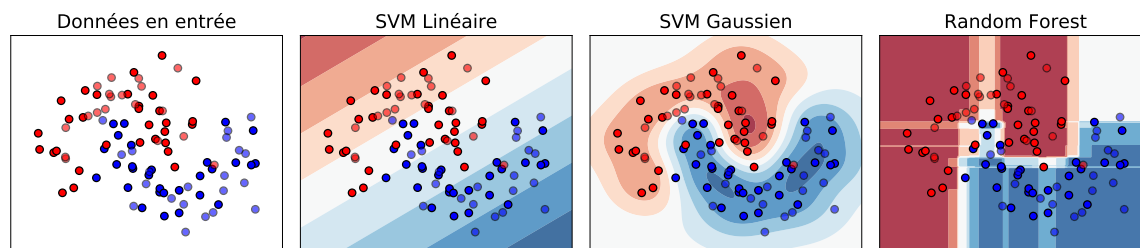


FIGURE 1.17 – Exemple de la séparation de deux classes selon les algorithmes SVM et RF. Inspiré du code de Gaël Varoquaux et Andreas Müller.

Random Forest

Créé au début des années 2000 [Breiman, 2001], l'algorithme *Random Forest* ou en français, Forêts Aléatoires (RF) est très utilisé dans notre domaine [Fassnacht et al., 2016], notamment pour sa capacité à généraliser et à être robuste au bruit dans les données ou dans les

labels [Pelletier et al., 2016, 2017]. L’algorithme est aussi assez rapide car il est peu sensible au paramétrage du nombre d’arbres. Il n’y a donc pas besoin de générer une grande grille de recherche afin de trouver les paramètres les plus optimaux [Oshiro et al., 2012; Pelletier et al., 2016; Probst and Boulesteix, 2017]

Partant du principe relativement simple d’un arbre de décision, l’algorithme en génère aléatoirement autant que l’utilisateur en a défini. En général il est fixé un minimum une centaine d’arbres et pour chaque nœud d’un arbre, la sélection de variables se fait aléatoirement à partir de l’ensemble des variables qui est par défaut défini comme la racine carré du nombre de variables. Ces arbres de décision sont appliqués aux données fournies par l’utilisateur et le critère de choix sera l’homogénéité des classes selon un critère prédéfini, le plus souvent il s’agit de la diversité de Gini [Grabska et al., 2019; Immitzer et al., 2012], plus rarement de l’entropie de Shannon.

De plus amples détails notamment sur le paramétrage de l’algorithme *Random Forest* ou en français, Forêts Aléatoires (RF) sont disponibles Section 3.3.2.

Ce qu’il faut retenir

Les algorithmes utilisés dans ce manuscrit sont de type non-paramétriques car ils ne présupposent pas d’une distribution des données selon une loi prédéfinie. Les deux algorithmes sélectionnés sont :

- **SVM** qui est réputé pour sa capacité d’apprentissage à partir d’un faible nombre d’échantillons et un nombre important de variables.
- **Random Forest** qui est réputé pour sa robustesse au bruit et ses bonnes performances en étant peu sensible paramètres.

1.3.2.2 Les références terrain des essences forestières

Pour créer un modèle capable d’apprendre ce qui différencie un chêne d’une autre espèce, il faut que l’algorithme ait la connaissance de leur signature spectral, c’est-à-dire de ce qu’est un chêne dans le spectre électromagnétique (Figure 1.18). Cette signature peut être monotemporelle (lorsqu’on utilise une seule date) ou multitemporelle (ce qui est le cas avec les séries temporelles).

C’est à partir de ces informations qui contiennent d’un côté la signature spectrale et de l’autre l’essence à laquelle elle appartient qu’un algorithme va pouvoir apprendre. Ainsi, à chaque pixel de forêt sera assigné un label, c’est-à-dire une espèce. Il est donc important d’avoir un nombre d’échantillons qui représente la variabilité de chaque essence. Or, selon Boureau [2008], il se peut que la variabilité intraspécifique soit supérieure à la variabilité interspécifique, tout du moins à une date unique.

Pour avoir suffisamment de références d’essences, certains auteurs se basent sur les inventaires forestiers nationaux, les *National Forest Inventory* (NFI) ou des campagnes de photo-interprétation. Les données forestières ont l’avantage, par rapport aux données agricoles, d’être assez pérennes dans le temps, ou du moins facilement vérifiables. En effet, il est improbable qu’un pixel change d’une année sur l’autre, par contre ce qui peut arriver est une coupe. Pour cela une validation par photo-interprétation suffit à vérifier la bonne validité de la donnée. Le problème des NFI est qu’ils sont rarement compatibles directement avec la télédétection. Leur niveau de détail est faible car ils représentent généralement des massifs ou des peuplements, et il est rare d’avoir des références pures (c’est-à-dire des endroits où l’on est sûr qu’un pixel et ses voisins ne contiendront qu’une seule essence).

D’autres auteurs ont préféré réaliser leur propre campagne terrain afin d’acquérir des références précises et compatibles directement avec la télédétection [Elatawneh et al., 2013; Persson et al., 2018; Sheeren et al., 2016]. Cependant, les effectifs sont généralement plus faibles mais le nombre d’essences est supérieur à celui des inventaires nationaux qui n’incluent

que les espèces les plus présentes. À titre d'exemple en France, la BD Forêt de l'IGN ne recense pas les bouleaux.

Entre les NFI et les références terrain créées spécifiquement pour la télédétection, de nombreuses autres sources de données peuvent être exploitées. En France il est possible de se référer aux plans simples de gestion qui sont réalisés par les propriétaires forestiers privés. Ce document précise les modalités d'entretien et de gestion de la forêt sur plusieurs années, et localise les essences présentes.

Ce qu'il faut retenir

Plusieurs types de données localisant les espèces peuvent être utilisés :

- Les *National Forest Inventory* (NFI) fournissent un grand nombre de références mais avec un taux d'erreur qui peut être important car ils ne sont pas adaptés à la télédétection. Ces données sont généralement gratuites et facilement accessibles.
- Des références récoltées pour la télédétection à travers des campagnes terrain dédiées. Ces données sont peu partagées et pas toujours accessibles.
- L'usage de plans simples de gestion ou de documents dédiés à la gestion des forêts. Ces données sont rarement accessibles à moins de faire partie des instances les ayant créées.
- De nombreuses autres sources de données existent, comme OpenTreeMap ou encore celles du GBIF.

Malheureusement il n'existe pas à l'échelle mondiale une base de données localisant les essences et qui serait compatible, ou du moins adaptable, pour la télédétection.

1.3.3 Les propriétés optiques des plantes

Une fois les images et les références localisant chaque espèce obtenues, les algorithmes vont tirer profit des propriétés optiques des plantes pour les identifier.

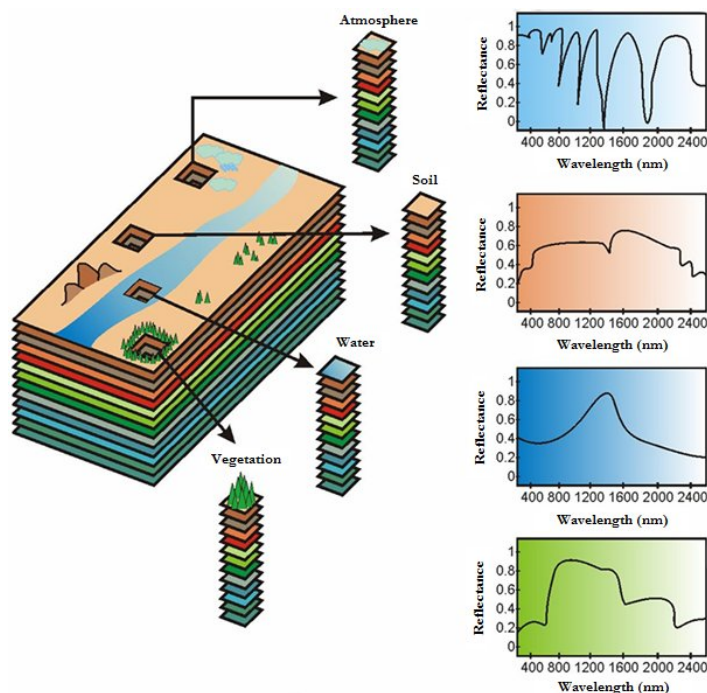


FIGURE 1.18 – À chaque pixel correspond une signature spectrale, soit monotemporelle soit multi-temporelle. Source : Plaza et al. [2009].

L'être humain ne voit qu'une petite partie du spectre électromagnétique qui s'étend de 0 nm à l'infini. Il se trouve que les plantes ne réfléchissent pas uniquement dans le spectre optique (c'est-à-dire le visible du domaine réflexif, Figure 1.19) mais elles réfléchissent aussi dans des longueurs d'ondes comme l'ultraviolet ou l'infrarouge.

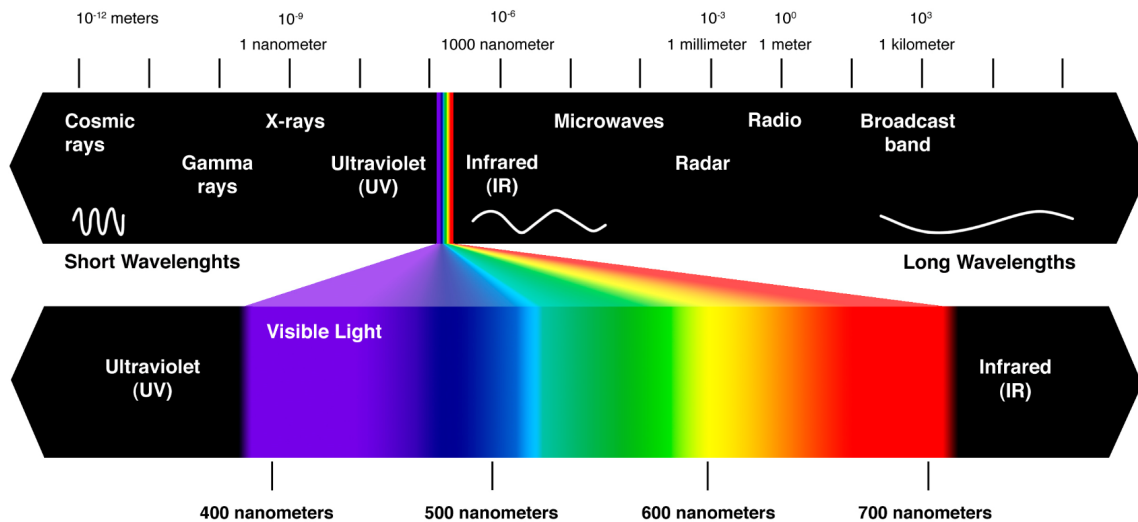


FIGURE 1.19 – Décomposition du spectre électromagnétique avec un focus sur le spectre du visible (de 400 à 700 nanomètres.).

1.3.3.1 La réponse de la feuille dans le spectre électromagnétique

La feuille est la partie de l'arbre qui nous permet de reconnaître le plus facilement une essence d'une autre. Pour cela les forestiers n'utilisent pas sa réflexion dans le spectre électromagnétique mais sa forme et sa structure. C'est d'ailleurs grâce à ces informations que l'application mobile PlantNet peut distinguer différentes essences [Goëau et al., 2013]. Les fruits, les fleurs, ou le tronc apportent aussi des informations, mais soit elles sont moins accessibles tout au long de l'année (ce qui est le cas des fruits et des fleurs par exemple), ou soit elles sont généralement moins discriminantes quand elles sont utilisées seules (le tronc).

Dans cette partie nous nous intéressons uniquement au spectre électromagnétique des feuilles qui est influencé par un nombre important d'éléments qui peuvent être directement liés à la feuille en elle-même, mais qui peuvent aussi provenir de conditions climatiques ou édaphiques (c'est-à-dire lié au sol). Nous nous concentrons ici uniquement sur ce qui influence la réflexion à l'échelle de la feuille.

Pour chacune des parties suivantes il conviendra de bien différencier les trois phénomènes qui arrivent lorsqu'un rayonnement électromagnétique atteint un objet (Figure 1.20) :

- la réflexion est la part de la radiation qui sera renvoyée,
- l'absorption est la part de la radiation qui sera bloquée par la feuillage,
- la transmission est la part de la radiation qui traversera l'objet.

1.3.3.2 Dans le spectre du visible

Entre 400 et 700 nm, ce sont les pigments foliaires qui vont influencer la réflectance de la feuille, c'est-à-dire le pourcentage de la lumière qui sera renvoyé. Ainsi les pigments comme la chlorophylle ou les caroténoïdes (Figure 1.21) sont les principaux éléments qui donnent de la couleur aux feuilles.

Le pigment le plus connu est sans conteste celui de la chlorophylle (nom provenant du grec feuille verte), c'est aussi le principal pigment de la photosynthèse. La chlorophylle s'obtient

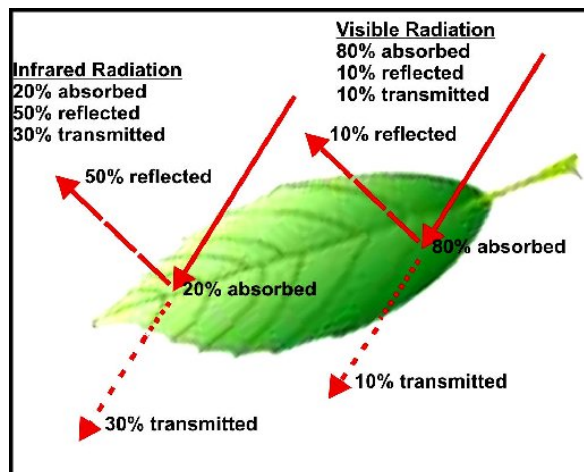


FIGURE 1.20 – Illustration de l’absorption, la réflexion et de la transmission du signal. De [Brown and Gillespie \[1995\]](#) modifié par [Shahidan and Salleh \[2007\]](#).

lors du processus qui transforme l’énergie lumineuse et le dioxyde de carbone contenu dans l’atmosphère en énergie chimique et qui restitue ensuite de l’oxygène à l’atmosphère [[Soltner, 2017](#)]. La chlorophylle est responsable de l’absorption du rayonnement des ondes bleu et rouge mais peu du vert, ce qui explique pourquoi nous voyons les feuilles de la couleur verte (Figure 1.21). Il existe différentes formes de chlorophylles, dont les seules présentes chez les végétaux sont la chlorophylle *a* et la chlorophylle *b*. Les chlorophylles *c*, *d* et *f* sont uniquement présentes chez certaines algues ou bactéries.

D’autres pigments influencent aussi la coloration des feuilles, comme les caroténoïdes (rouges orangés) [[Jacquemoud and Ustin, 2019](#)]. Comme pour la chlorophylle, certaines étapes de la synthèse des caroténoïdes sont induites par la lumière [[Sandmann et al., 2006](#)]. Son spectre d’absorption se situe principalement dans le bleu entre 400 et 550 nm. À l’opposé des caroténoïdes, les anthocyanes dont le nom signifie fleur bleue sombre en grec, absorbent fortement le spectre vert, rouge et ultraviolet. Ce pigment agit comme protection contre les ultraviolets avant que la production de chlorophylle ne commence. Il est à noter que ce composé chimique permet également aux plantes de se défendre car il est très désagréable au goût (idéal pour préserver les jeunes pousses), et agit contre certains pathogènes [[Freeman, 2008](#)].

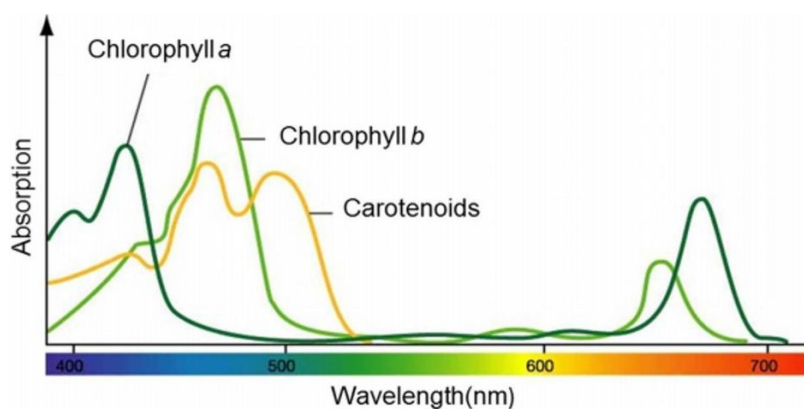


FIGURE 1.21 – Spectre d’absorption de la chlorophylle *a* et *b* et des caroténoïdes.
Source : www.cfb.unh.edu

Avant l’apparition des feuilles au printemps, les bourgeons de chaque essence ont des formes et couleurs propres qui permettent de les différencier les unes des autres (Figure 1.22). Ils sont aussi disposés selon une certaine façon le long de la tige. Cependant l’information sur

le bourgeon, bien que très utile pour reconnaître une espèce pour un forestier, ne pourra être exploitée à partir d'un capteur satellite du fait de sa trop petite taille.

Encore moins visible car de plus petite taille, le trichome (signifiant en grec la croissance de poil) peut être différent selon les espèces, mais aussi selon la maturité de la feuille (Figure 1.23). Safou and Saint-Martin [1989] ont mis en évidence que des espèces différentes de chêne (*Quercus*) peuvent être identifiées grâce à ces petits poils. Ainsi le trichome multiétoilé est uniquement présent sur le *Quercus Valentina*, et deux espèces très proches ont pu être distinguées par leur trichome. En effet un était de type étoilé *Quercus petraea* et l'autre de type unicellulaire *Quercus robur* et sa chute est totale au cours de la maturation de la feuille.



FIGURE 1.22 – Illustration des différents types de bourgeons lors du débourrement. Les différents dessins ne sont pas à l'échelle et ont uniquement vocation à illustrer les particularités de chaque essence.

Les pigments influencent grandement la réflectance de la feuille dans le visible. Selon l'échelle certains éléments seront plus déterminants que d'autres pour distinguer les espèces les unes des autres. Le trichome sera bien discriminant à l'échelle microscopique, alors que d'autres informations plus globales comme la couleur de la feuille seront mieux observables à l'échelle de l'individu. Dans notre cas l'information la plus importante est la réflectance de la feuille qui est influencée par les pigments qui évoluent au fil de l'année.

1.3.3.3 Dans le proche infra-rouge (780 nm à 3000 nm)

Le spectre de proche infrarouge est défini selon norme ISO 20473 :2007 comme allant de 780 nm à 3000 nm.

De 780 à 1400 nm

À cette longueur d'ondes, les pigments foliaires vus précédemment (comme la chlorophylle) deviennent totalement translucides et n'absorbent donc plus le rayonnement. Ainsi la part réfléchi n'est plus influencée par les pigments de la feuille, mais par ce qui est contenu sous la cuticule, à savoir le parenchyme lacuneux [Jensen, 2000] (Figure 1.23). Si l'humain était capable de voir dans le proche infrarouge, nous verrions d'une toute autre couleur la végétation. En effet c'est dans cette portion du spectre électromagnétique que la réflexion des feuilles est la plus importante. Cela s'explique par le fait que le proche infrarouge est réfléchi par le parenchyme lacuneux qui est le lieu d'échange entre l'oxygène et le dioxyde de carbone lors de la photosynthèse. Dans l'infrarouge c'est donc plutôt la structure de la feuille qui influence la réflexion que les pigments.

De manière générale, moins une feuille est rouge, plus elle contient de chlorophylle (car ce pigment absorbe une grande partie du spectre rouge) et plus la réflectance de la feuille dans l'infrarouge sera forte.

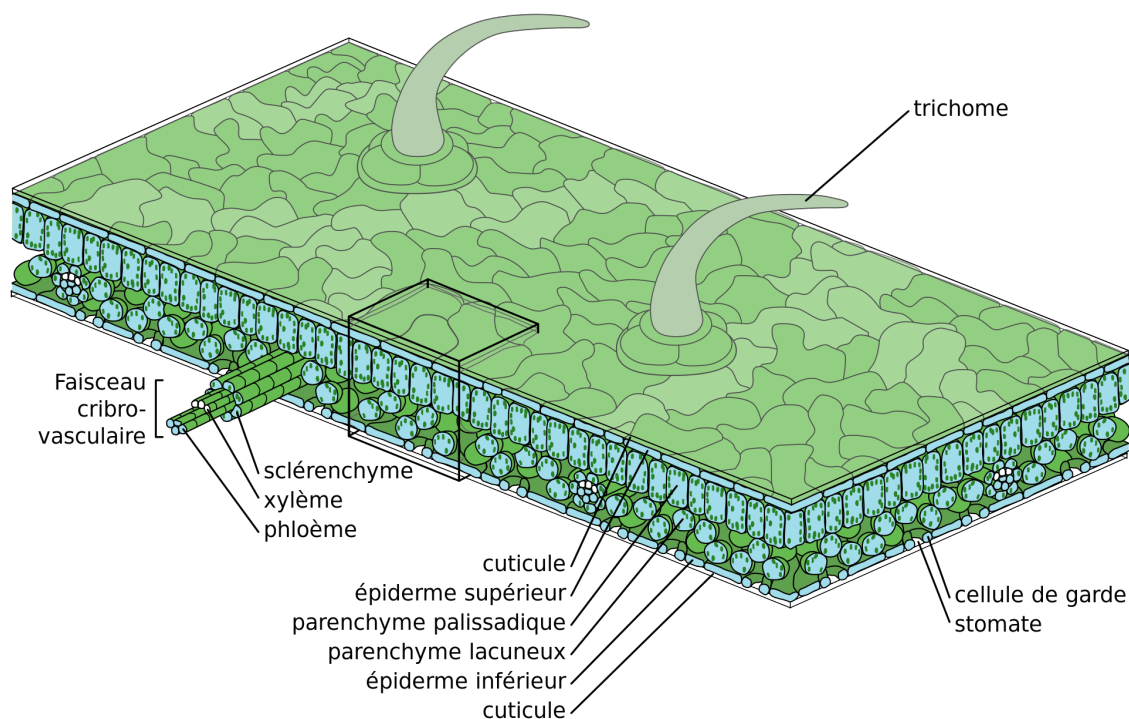


FIGURE 1.23 – Anatomie d'une feuille. Le cuticule englobe la partie interne de la feuille. Source : Wikimedia.

De 1400 nm à 3000 nm

La portion de l'infrarouge entre 1400 nm et 3000 nm n'est pas à négliger en télédétection car elle correspond à plusieurs bandes d'absorption de l'eau [Palmer and Williams, 1974], en particulier à 1450 nm, 1940 nm et 2500 nm pour les plantes [Carter, 1991]. Une corrélation très importante est observable entre la teneur en eau et la réflectance des bandes à 1450 nm et 1940 nm [Wang et al., 2009].

D'autres informations influencent aussi la réflectance du spectre infrarouge comme les protéines, la lignine, ou la cellulose, mais plutôt quand la feuille est sèche. Comme ces informations sont plus difficiles à obtenir en télédétection, nous ne les étudierons donc pas en détail. Ce qu'il faut retenir de cette portion du spectre est qu'elle est très utile pour estimer l'état d'hydratation ou de déshydratation d'une plante.

Un exemple résumant cette partie (Section 1.3.3.1) montre la réflectance d'une feuille de chêne, fraîche puis sèche, de 400 nm à 2500 nm (Fig 1.24). On remarque trois pics d'absorption de l'eau, à 1450 nm, 1940 nm, et 2500 nm. Ainsi les feuilles sèches qui ne contiennent plus d'eau ont une réflectance bien plus élevée dans les bandes de l'absorption de l'eau. On note aussi une réflexion très faible dans les bandes bleues et rouges due aux pigments de chlorophylle, et une croissance rapide de la réflectance à partir de 700 nm, ce qui correspond au spectre du proche infrarouge.

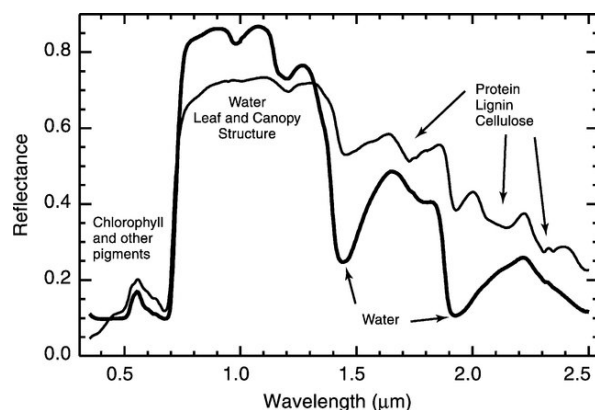


FIGURE 1.24 – Réflectance d’une feuille de chêne fraîche (ligne épaisse) ou sèche après 24h (ligne fine) observée en laboratoire par spectroscopie. Source : [Kokaly et al. \[2003\]](#).

Ce qu’il faut retenir

Si les pigments foliaires colorent les feuilles et sont donc observables dans la partie visible du spectre électromagnétique, l’infrarouge apporte de nombreuses informations indispensables concernant :

- l’état d’hydratation des plantes (avec des pics d’absorption à 1450, 1940 et 2500 nm)
- la structure des feuilles et de la canopée. Des différences non observables entre essences dans le spectre du visible apparaissent dans l’infrarouge.

1.3.3.4 L’évolution des propriétés optiques dans le temps

Les belles couleurs de l’automne chez les feuillus montrent l’évolution des pigments au fil du temps. Certaines espèces ont des feuilles qui se colorent en jaune, d’autres en orange voire en rouge. Ces comportements saisonniers peuvent varier d’une espèce à l’autre et même d’une année sur l’autre [[Jacquemoud and Ustin, 2019](#)]. Après nous être focalisé sur ce qui modifie de manière générale la réflectance d’une feuille à un instant donné, nous allons nous pencher sur le cycle annuel de la végétation, autrement dit, sa phénologie [[Badeau et al., 2017](#)]. Selon [Lebourgeois \[2018\]](#) la phénologie est “*le champ d’étude [...] qui consiste donc à enregistrer, dans le temps, le retour des étapes de croissance et de développement des plantes et à étudier les facteurs qui l’influencent.*”.

Feuillus

Les feuillus décidus (ou à feuilles caduques) ont la phénologie la plus marquée avec la chute des feuilles à l’automne et la pousse de nouvelles feuilles au printemps. Les belles couleurs orangées ou jaunes que l’on peut observer à l’automne sont dues en grande partie à des pigments qui étaient déjà présents tout au long de la phase végétative des feuilles mais qui étaient masqués par la dominance de la chlorophylle responsable de couleur verte [Jacquemoud and Ustin \[2019\]](#).

Avec l’automne qui arrive, c’est aussi la baisse du nombre d’heures d’ensoleillement et de la température. L’arbre se met en quelque sorte en veille. La photosynthèse qui a débuté dès le printemps n’a plus d’utilité, et l’arbre va commencer à décomposer la chlorophylle (Figure 1.25) pour la stocker dans les branches, racines ou encore dans le tronc. C’est à cette période qu’on observe des belles couleurs, autres que le vert des feuilles, qui trahissent la présence de certains pigments qui étaient pourtant déjà présents dès le début de la pousse de la feuille. On peut alors observer sur certaines essences la xanthophylle, le pigment responsable de la couleur jaune (Figure 1.26). Pour les couleurs oranges, il s’agit de la caroténoïde, le même

pigment responsable de la couleur orange de la carotte. Contrairement aux autres pigments présents tout au long de la vie de la feuille mais masqués dans le visible par la dominance de la chlorophylle, l'anthocyanine, responsable de la couleur rouge à la sénescence, se développe dès que la production de chlorophylle chute.

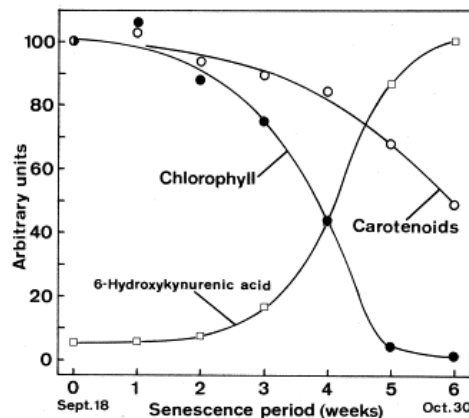


FIGURE 1.25 – Évolution de la réflectance de certains pigments foliaires lors de la sénescence. Source : Matile [2000].

Peu à peu, l'arbre n'alimente plus la feuille, et récupère toutes les substances encore présentes dans cette dernière. Quand la production de chlorophylle diminue, ce qui se traduit par une augmentation de la réflectance dans le proche infrarouge, on note que l'arbre augmente sa production d'éthylène, hormone responsable du bouchon qui se forme à la base de la feuille (pétiole). Cette barrière composée de subérine et de callose vient alors couper à tout jamais le lien entre l'arbre et la feuille. Cette cicatrice très fragile, appelée zone d'abscission, cédera au moindre souffle en faisant tomber la feuille de l'arbre.



FIGURE 1.26 – Un rameau de chêne pubescent (*Quercus pubescens*) pris en photo le 27 octobre 2017. Les quatre feuilles ont des colorations assez différentes. Les plus grandes feuilles contiennent encore de la chlorophylle, alors que les plus petites n'ont soit presque plus de pigments (marron), soit le pigment xanthophylle domine.

Conifères

Arrivés dans l'évolution bien après les feuillus, les conifères ont comme capacité d'être beaucoup moins dépendants à l'eau et plus résistants au froid que les feuillus. Cela est dû en partie à un faisceau cribro-vasculaire composé de trachéïdes de plus petite taille (parfait pour la résistance au froid) et à un système proche d'une valve pour lutter contre la sécheresse [Granier et al., 2002].

L'évolution temporelle des aiguilles est moins marquée que celle des feuillus, notamment car on n'assiste pas à leur perte totale. Cependant, chaque année, de nouvelles pousses ap-

paraissent, et allongent les rameaux déjà présents. Si le phénomène est plus discret que les feuillus, il est pourtant observable à l'échelle du rameau mais aussi du peuplement (Figure 1.27).

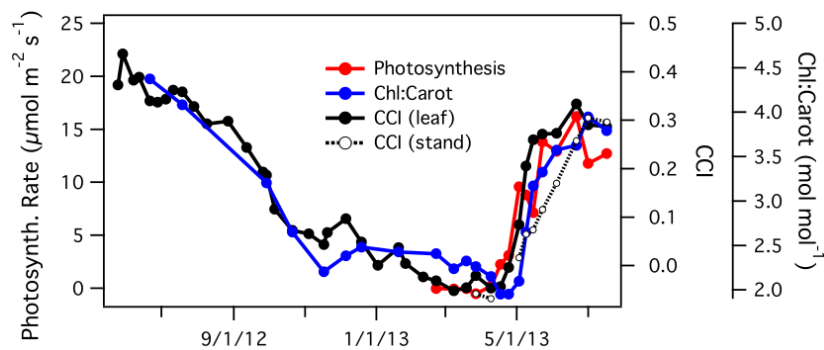


FIGURE 1.27 – Mesure à l'aide d'un spectromètre de l'activité du Pin contorta, en chlorophylle et en caroténoïdes (Chlorophyll Carotenoid Index (CCI) de décembre 2012 à janvier 2013. Source : Gamon et al. [2016]

Ce qu'il faut retenir

Le cycle annuel de la végétation, appelé phénologie, est plus marqué chez les feuillus que chez les conifères. Ainsi des différences spectrales peuvent s'observer à différentes périodes de l'année :

- lors du débourrement avec des feuilles qui se développent parfois plus vite ou plus tôt chez certaines essences.
- lors de la sénescence avec la perte et la coloration des feuilles. Certains pigments peuvent alors être davantage visibles pour certaines essences.

1.3.3.5 Un signal influencé par un changement d'échelle

La problématique du changement d'échelle est inhérente à la télédétection. Si le signal d'une feuille est plutôt homogène et pur, les différences à l'échelle d'un arbre peuvent être importantes (feuilles de taille ou d'orientation différentes par exemple). Puis à l'échelle d'un peuplement, les variations s'accroissent, il peut y avoir plusieurs essences avec pour chacune des âges, des tailles et des phénologies différentes.

On peut alors noter après la feuille, deux échelles d'analyse :

1. l'arbre
2. le peuplement

De la feuille à l'arbre

En passant à l'échelle de l'arbre, une grande variabilité des propriétés biophysiques et biochimiques peut s'observer entre les feuilles. Cela concerne la superficie, la coloration, l'orientation spatiale, ou encore l'état sanitaire des feuilles. Ainsi, le signal spectral d'un arbre va être influencé par plusieurs éléments, qu'ils soient externes (conditions du sol, maladie) ou internes (structure de l'arbre) [Serbin and Townsend, 2020].

Le feuillage de l'arbre représente la majorité de sa signature spectrale vue du ciel (Figure 1.28). Cependant, plus le feuillage sera diffus, plus on apercevra la ramure mais aussi le sous-étage qui n'a pas de lien direct avec l'essence observée.

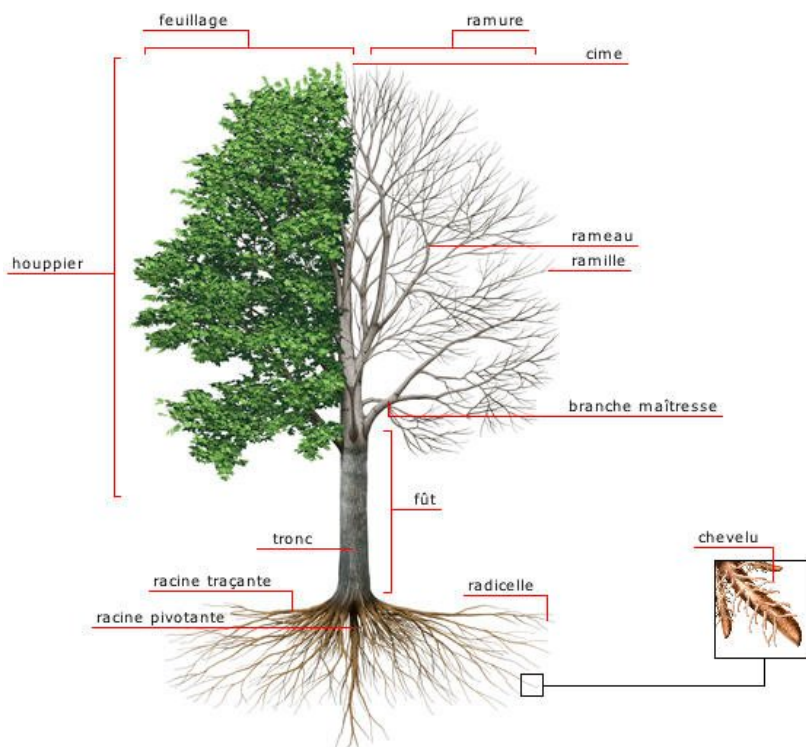


FIGURE 1.28 – Structure d'un arbre, avec et sans feuilles. Source : Nouveau Dictionnaire Visuel.

La surface foliaire La surface foliaire, qui peut être calculée avec un indice foliaire comme le *Leaf Area Index*, appelé aussi en français indice de surface foliaire (LAI), représente la surface de feuilles d'un arbre ou d'un peuplement par rapport à son occupation au sol. Si la totalité des surfaces des feuilles d'un arbre mesurent 80 m^2 , et que l'emprise du houppier mesure 40 m^2 alors le LAI sera de 2. Ainsi plus une forêt est dense, plus le LAI sera élevé.

La surface foliaire varie aussi avec la sylviculture (densité d'arbre, mode de culture, et éclaircies), les espèces, les conditions climatiques et biotiques (qui évoluent donc selon les années) [Breda, 2003].

Selon la saison, le LAI évolue, en particulier pour les feuillus. L'hiver, en l'absence de feuilles, le LAI est nulle, puis au début de saison les feuilles se développent et le LAI augmente très rapidement (Figure 1.29). La croissance des feuilles se poursuit jusqu'en août, puis décline lors de la sénescence.

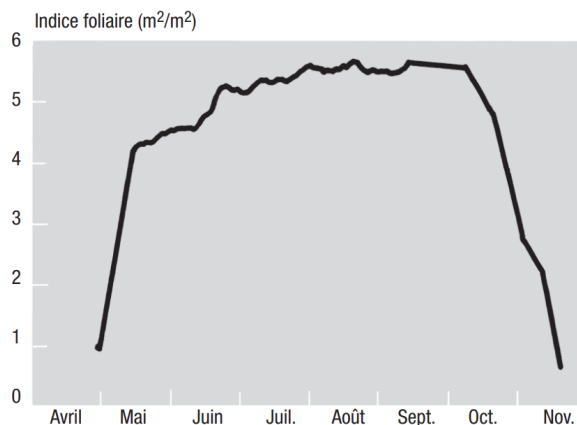


FIGURE 1.29 – Exemple de la variation saisonnière du LAI dans une chênaie. Le calcul a été fait à partir de mesures d'interception du rayonnement global par le couvert pendant la saison de végétation et par ramassage des litières à l'automne (forêt domaniale de Champenoux, Meurthe-et-Moselle) Source : Breda [1999].

L'influence de la canopée La canopée représente la partie aérienne et dominante de la forêt qui est directement exposée au rayonnement solaire. Sa densité est liée à celle du peuplement [Gill et al., 2000] et à la productivité du sous-étage [McConnell and Smith, 1970]. En effet, plus une canopée sera ouverte, plus le sous-étage pourra profiter de la lumière.

Pour mesurer le degré de fermeture de la canopée et ainsi estimer la part de la réflectance d'un pixel qui est directement assignable à l'arbre, on regarde la proportion de ciel masqué depuis un point de vue fixe au sol [Jennings, 1999] et avec un angle de vue large, souvent hémisphérique. La photographie hémisphérique sera ramenée à une projection planisphérique [Daubemire, 1959] et un indice, le *Canopy Closure* en anglais, sera utilisé pour mesurer le pourcentage de ciel masqué (Figure 1.30). D'autres techniques autres que la photographie hémisphérique sont possibles [Paletto and Tosi, 2009] mais cette méthode reste la plus utilisée. Elle convient bien pour mesurer le pourcentage de ciel masqué depuis un point fixe, mais il peut être plus délicat d'estimer la proportion de ciel masqué uniquement dû la strate arborée ou émergente en étant placé au sol.



FIGURE 1.30 – Exemple d'une photo hémisphérique prise le 23 septembre 2019 dans une forêt composée majoritairement de chênes.

Le couverture de la canopée a une signification proche, mais elle est dédiée à la strate dominante. On utilisera en anglais le terme *canopy cover* qui représente le pourcentage de la surface forestière occupée par la projection verticale des couronnes d'arbres [Korhonen et al., 2006]. Cet indice est supposé être la valeur réelle de la couverture de la strate émergente [Bonnor, 1967].

De l'arbre au peuplement

Au delà de l'arbre, l'échelle supérieure est celle du peuplement forestier, une unité de gestion avec un mode défini (futaie par exemple), une répartition par classe d'âge ou par diamètre, et par essence. Un peuplement forestier peut inclure des mélanges, que ce soit de la strate émergente ou du sous-étage (Figure 1.31). Le signal du peuplement sera donc un mélange des éléments précédemment listés auquel se rajoute des informations liées au peuplement (densité, homogénéité ou hétérogénéité des hauteurs). Si les arbres ont tendance à se ressembler au sein d'un même peuplement, cela est dû en partie aux conditions locales (maladie, humidité du sol, ensoleillement, pente...) et au mode de gestion.

À l'échelle de la forêt, c'est surtout le sol, qui sert à la fois de support (meuble et profond), mais aussi de pourvoyeur (eau et minéraux) qui va influencer les arbres. Plus précisément, la



FIGURE 1.31 – Les différentes strates végétales de la forêt. Graphique initial de Elke Freese (Wikimedia.de).

composition du sol s'explique par divers éléments comme le climat, les roches, et la topographie [Schmid et al., 1970]. La végétation repose largement sur le sol tant pour son alimentation en eau, sa nutrition minérale que pour son ancrage. En retour, les plantes rendent des services aussi sol qui sont propriétés physico-chimiques (modification du pH, agrégation, porosité) et hydriques qui constituent l'essentiel des apports de matières organiques des sols.

De très nombreux facteurs édaphiques influencent la végétation qui peuvent être regroupés en 3 catégories Freschet et al. [2018] :

- facteurs physiques (texture, densité, structure, température, etc.)
- facteurs chimiques (quantité de matière organique, potentiel redox, capacité d'échange anionique et cationique)
- de la disponibilité en ressources (eau, éléments minéraux)

La fertilité du sol a un rôle important dans la production de chlorophylle et des azotes foliaires [Larrieu, 2019], ce qui profite aussi aux strates non dominantes.

Au dessus du sol et de l'étage muscinale qui est composé de mousses et de lichen, on trouve à la fois l'étage arbustif et arborée non dominant (Figure 1.31). Ce sont ces sous-étages qui sont visibles entre les arbres et entre les feuilles. Selon la saison, et selon si le sous-étage est permanent (comme le fragon qui reste vert tout au long de l'année), la réflectance du peuplement hors strate émergente sera plus ou moins forte.

Ce qu'il faut retenir

Le signal d'un arbre ou d'un peuplement va être influencé par plusieurs facteurs :

- la taille, l'orientation, et l'état sanitaire des feuilles.
- l'ouverture de la canopée (donc la densité du couvert forestier) qui va laisser transparaître le sous-étage.
- les conditions édaphiques et climatiques.

1.3.4 Principaux travaux de classification des essences forestières

1.3.4.1 Avec un capteur à résolution spatiale modérée

La cartographie des espèces forestières sur de grands territoires est historiquement réalisée à partir de capteurs de résolution modérée, comme le satellite *Moderate-resolution Imaging Spectroradiometer* (MODIS) (250 m de résolution spatiale dans le spectre du visible). Ainsi, [Beaudoin et al. \[2014\]](#) ont cartographié les espèces dominantes de l'ensemble du Canada en utilisant des images représentant la moyenne de deux mois d'été et d'hiver. D'autres travaux ont, à partir du même satellite, étudié l'apport des séries temporelles denses (plusieurs dizaines d'images par an) dans la reconnaissance des essences. Certains sont spécifiques à un type de forêt, les conifères [[Aragones et al., 2019](#)], d'autres sont multispécifiques [[Cano et al., 2017](#)].

Les travaux à résolution spatiale modérée sont les plus développés à l'échelle globale, y compris pour calculer des variables biophysiques. Par exemple la carte mondiale du LAI (Section 1.3.3.5) est générée à partir d'images du satellite SPOT VEGETATION de résolution spatiale d'environ 1.15 km [[Baret et al., 2007](#)].

La résolution spatiale modérée n'est cependant pas adaptée aux forêts tempérées. En France par exemple, sur les 3,5 millions de propriétaires forestiers privés en France, 2,6 millions d'entre eux ont moins de 1 hectare de forêt [[Barré et al., 2015](#)]. Cet extrême morcellement empêche l'utilisation de satellites à résolution modérée. En effet la surface d'un pixel MODIS est de 6,25 hectare alors que celle de Landsat 7 ou 8 est de 0.9 hectare, ce qui rend difficile voire impossible la cartographie des forêts tempérées, et encore moins possible l'identification des essences qui la composent.

Cependant, un avantage des travaux réalisés avec cette résolution est qu'ils ont souvent la particularité d'utiliser des séries temporelles denses. Pour MODIS il est même possible d'avoir une donnée journalière directement traitée sans nuages. D'autres produits sont aussi accessibles pour l'utilisateur final comme un NDVI global actualisé tous les 16 jours, ce qui rend possible le suivi global de la phénologie de la végétation [[Zhang et al., 2003](#)]. L'écosystème qui tourne autour du satellite MODIS permet donc de faciliter le travail des chercheurs et chercheuses. Cette mise à disposition des données a aussi été rendu possible par la faible résolution spatiale de ce satellite qui permet de faire des traitements à l'échelle planétaire sans être trop gourmand niveau temps de calcul.

1.3.4.2 Avec un capteur à haute résolution spatiale

La haute résolution spatiale des satellites comme Sentinel-2 (10 à 20 m) permet de détecter des petites forêts comme celles des zones tempérées comme la France. Cependant les arbres n'ont que très rarement des houppiers excédants les 10 m, mais à part les arbres isolés, les forêts sont généralement composées de patchs d'une même essence.

Les résultats des travaux dans ce domaine sont globalement proches et le niveau de qualité est assez élevé (Tableau 1.1). La qualité de prédiction moyenne estimée avec l'accord global pour les études listées est de 84,8% ($\pm 11.1\%$), allant de 74,3% à 92,12%.

Depuis quelques années, on assiste à une évolution majeure dans ce domaine. Désormais la majorité des travaux sur la cartographie des forêts à l'échelle régionale ou supérieure se fait à partir du satellite Sentinel-2. Bien que récent, ce satellite connaît un franc succès du fait de son accessibilité et de ses hautes résolutions spatiale, spectrale et temporelle. [Fassnacht et al. \[2016\]](#) avaient déjà identifié ce futur succès en disant que "*Landsat-8 et Sentinel-2 pourraient être une bonne base pour cartographier les essences majeures présentes à l'échelle nationale dans les forêts tempérées et boréales*". Beaucoup de chercheurs utilisent ces capteurs, y compris pour la cartographie des forêts. La majorité des travaux réalisée à partir de satellite à haute résolution spatiale pour identifier les essences par télédétection se fait à partir d'images sans nuages et donc avec des séries temporelles moins denses et sans possibilité de pouvoir prédire des tuiles voisines car l'échantillonnage temporel serait différent [[Bolyn et al., 2018](#); [Grabska et al., 2019](#); [Hościło and Lewandowska, 2019](#); [Immitzer et al.,](#)

1.3. LA TÉLÉDÉTECTION COMME ALTERNATIVE À LA CARTOGRAPHIE MANUELLE

2012; Liu et al., 2018; Persson et al., 2018; Spracklen and Spracklen, 2019; Wessel et al., 2018].

TABLEAU 1.1 – Principaux travaux de ces dernières années en télédétection des essences forestières à partir de capteur à haute résolution. Trié par ordre décroissant selon le nombre de dates utilisées. Pour les satellites les moins courants, il est spécifié leur résolution spatiale entres parenthèses.

Auteurs	Capteur	Nombre de dates (↓)	Nombre de feuillus	Nombre de résineux	Essence la plus difficile	OA
Elatawneh et al. [2013]	RapidEye (5m)	20	3	5	<i>Quercus / Larix</i>	86,35
Immitzer et al. [2019]	S2	18	7	5	<i>Acer sp</i>	88,7
Grabska et al. [2019]	S2	18	5	4	<i>Betula / Alnus incana</i>	92,12
Sheeren et al. [2016]	Formosat (8m)	17	8	5	<i>Pseudotsuga menziesii</i>	92
Pasquarella et al. [2018]	Landsat	11	3	4	<i>Picea / Pinus</i>	75,48
Clark et al. [2018]	S2 simulé	8	4	12	<i>Quercus</i>	74,3
Persson et al. [2018]	S2	4	2	3	<i>Larix / Pinus Silvestris</i>	88,2
Hościło and Lewandowska [2019]	S2	4	4	4	<i>Birch / Pinus</i>	75,59
Wessel et al. [2018]	S2	4	3	5	<i>Fagus sylvatica</i>	88
Bolyn et al. [2018]	S2	2	4	6	<i>Larix</i>	88,5

Sans toutefois pouvoir estimer la stabilité de la cartographie d’une année sur l’autre, les travaux d’Immitzer sont les plus originaux sur l’utilisation des données car ils utilisent l’ensemble des dates sans nuages sur plusieurs années [Immitzer et al., 2019, 2018] comme cela se faisait aussi auparavant [Mickelson et al., 1998; Schriever and Congalton, 1995; Wolter et al., 1995] du fait de la faible résolution temporelle des images à haute résolution et du coût d’acquisition de la donnée.

À notre connaissance, un seul travail référencé utilise des séries temporelles denses avec un traitement de type *gap-filling* pour cartographier les essences forestières [Sheeren et al., 2016]. Ce travail avait aussi comparé les prédictions pour chacun des algorithmes testés, et il se trouve que les méthodes d’apprentissages prédisaient souvent différemment les unes des autres en dehors des données de validation indépendantes, et ce malgré une qualité estimée élevée (>90%) pour chacun des algorithmes, ce qui laissait sous-entendre une surestimation de la qualité de la carte produite.

Les séries temporelles denses à haute résolution spatiale sont surtout utilisées pour la cartographie de l’occupation du sol et en particulier pour aider à différencier les cultures [Inglada et al. [2017]; Khaliq et al. [2018], ou pour l’étude de la phénologie [Vrieling et al., 2018].

Ce qu’il faut retenir

La plupart des travaux de reconnaissance des essences forestières par télédétection se fait :

- avec une résolution spatiale modérée (souvent MODIS à 250 m) sur des échelles nationales.
- avec une haute ou très haute résolution spatiale jusqu’aux échelles régionales.
- en utilisant uniquement les dates sans nuages à l’échelle locale.

1.3.4.3 Une évaluation délicate à estimer

Problème de surentraînement, d’autocorrélation spatiale, nombre d’échantillons d’apprentissage variable et non homogènes entre les classes ou encore méthodes de validation différentes, la qualité prédictive des travaux précédemment cités est difficile à établir tant les différences mais aussi les biais qui peuvent affecter une prédiction sont nombreux [Fassnacht et al., 2016; Olofsson et al., 2014].

Pour estimer la qualité des travaux, la démarche est relativement simple. Dans la plupart des études la séparation entre les données d'entraînement et de validation se fait soit de manière totalement aléatoire (e.g. 70% des références pour l'entraînement et le reste pour la validation), soit en utilisant un jeu de données dédié à la validation. Cependant, il se peut que les données de validation soient très proches spatialement et donc spectralement des données d'entraînement, ce qui peut rendre l'estimation de l'erreur biaisée. Or si la méthode d'évaluation consiste à comparer le nombre de références bien classées par rapport au nombre total de références (ce qu'on appelle l'accord global), la façon dont sont séparés ces références peut être très différente d'un article à l'autre. Fasnacht et al. [2016] ont souligné l'importance de prendre la proximité des différentes références dans l'évaluation, ou à minima de souligner ce biais s'il n'était pas possible de le traiter. Il est donc aujourd'hui difficile de comparer les articles entre eux et de connaître le potentiel de reconnaissance des essences forestières par télédétection tant les biais et les méthodes de séparation des données ne sont pas similaires.

Au delà de la méthode de séparation des références, les indices permettant de mesurer la qualité d'un modèle ont été questionnés. Pour évaluer les prédictions nous n'utiliserons plus l'indice kappa qui a été mainte-fois critiqué et qui apporte une information, toute au plus redondante à l'accord global, ou une information très instable et difficilement exploitable [Delgado and Tibau, 2019; Foody, 2020a; Pontius and Millones, 2011]. Ces auteurs préconisent d'abandonner l'indice en résumant les points clés suivant :

- Si les valeurs hors diagonale de la matrice de confusion ont une entropie proche de zéro, le kappa est très instable.
- Une très bonne prédiction peut être associée à une large gamme de kappa.
- Le kappa est très corrélé à l'accord global même si parfois des écarts importants entre les deux indices peuvent être observés.
- Pour un même accord global de 95%, le kappa peut varier de -0.026 à 0.9.
- Si chaque classe a le même effectif, le kappa équivaudra à l'accord global.

Nous privilégierons alors l'accord global comme une estimation générale de notre prédiction, tout en communiquant et en fournissant la matrice de confusion ainsi que les scores F1 (moyenne harmonique entre la précision et le rappel) pour chaque classe (Figure 1.32). Ces indicateurs présentent les mêmes défauts et en particulier, celui de la dépendance à la prévalence.

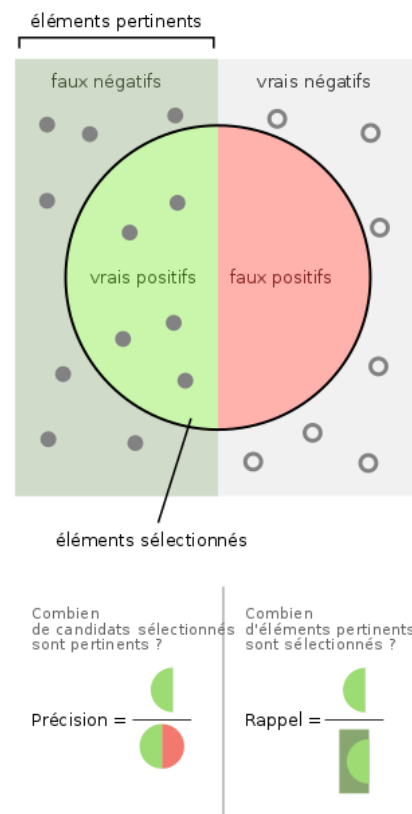


FIGURE 1.32 – Le score F1 est la moyenne harmonique entre la précision et le rappel. Source : Wikimedia.

Ce qu'il faut retenir

La plupart des travaux utilisent plusieurs images sans nuages d'une même année et cela pose plusieurs problèmes :

- Il n'est pas possible d'avoir une couverture d'un pays à une seule date (nuages et fréquence d'acquisition des satellites), ce qui freine une cartographie nationale.
- Les résultats sont difficilement comparables. Il ne s'agit jamais des mêmes essences.
- Aucune étude n'a estimé la stabilité temporelle des prédictions sur plusieurs années.

Cependant, l'ensemble des travaux suggère un véritable potentiel dans la cartographie automatique des essences même s'il est difficile de les comparer entre eux.

1.4 Questions de recherche

L'état de l'art des travaux concernant la classification supervisée des essences forestières à partir d'images satellites à haute résolution spatiale a montré que ces données semblaient prometteuses pour réaliser une cartographie à l'échelle nationale (Section 1.3.4). Cependant, s'il n'y a toujours aucune carte des essences réalisée de manière automatique pour les forêts tempérées, c'est qu'il subsiste encore des verrous méthodologiques qu'il convient d'identifier et si possible de lever.

1.4.1 Défis

Le principal défi de cette thèse est d'étudier le potentiel des *Satellite Image Time Series* ou en français, Série Temporelle d'Images Satellites (SITS) pour la reconnaissance des essences forestières sur de larges étendues en identifiant les facteurs et biais affectant la qualité des prédictions, et en évaluant si la phénologie des essences forestières aide les algorithmes à améliorer cette reconnaissance.

Un point crucial de ce travail est le faible nombre de références d'essences forestières compatibles avec la télédétection. Les algorithmes et les méthodes de validation choisis doivent être capables d'à la fois entraîner un modèle et de le valider avec un petit effectif par classe. Les progrès réalisés en intelligence artificielle permettent aujourd'hui de mettre plus facilement en œuvre des algorithmes et des méthodes de validation adaptées aux différents besoins. Cependant la méthode de validation doit :

1. Laisser un maximum de références à l'algorithme pour entraîner le modèle.
2. Prendre en compte la ressemblance spectrale des références situées dans un même peuplement.

Le défi est aussi de bien estimer la capacité prédictive des modèles et ainsi d'avoir un indice de qualité qui permettra de comprendre les erreurs mais aussi les réussites de la méthode, car des indices de qualité très élevés peuvent cacher certains biais [Sheeren et al., 2016]. L'usage d'une série temporelle avec un échantillonnage temporel dense, donc incluant des nuages sur la zone d'étude, s'explique par le fait que la finalité à plus long terme de la thèse est de rendre possible la cartographie à l'échelle nationale. Or, il est impossible d'avoir une couverture satellite nationale à une même date compte tenu des dates de passages des satellites et de l'enneigement hétérogène du territoire.

L'actuelle absence de cartographie nationale des essences dans les pays tempérés montre la présence de nombreux verrous qui empêchent la réalisation d'une telle production. L'idée de la thèse est de se servir des méthodes de télédétection qui ont fait le succès des satellites à résolution modérée comme MODIS et des avantages des approches de cartographie à haute résolution. Cette combinaison des deux approches est possible depuis l'arrivée du duo de capteurs Sentinel-2.

Bien que de nombreux travaux traitent de cartographie d'essences par satellite, la fiabilité réelle des prédictions est inconnue. Ainsi, après avoir évalué cette efficacité, il sera cherché une stratégie adaptée pour avoir l'estimation la plus proche entre la capacité estimée d'un modèle et sa réelle capacité prédictive. Enfin l'apport de la phénologie dans la cartographie des essences est inconnue tout comme l'apport des séries temporelles denses.

1.4.2 Objectifs

Le premier objectif consiste à étudier la stabilité des prédictions cartographiques à partir d'une archive du satellite Formosat-2. Pour chaque année, une SITS sera générée et contiendra l'ensemble des acquisitions du satellite, qu'il y ait ou pas des nuages. Il sera ensuite discuté de l'importance des dates et du choix de la méthode de validation du jeu de références servant à l'apprentissage de l'algorithme.

Le second objectif vise à évaluer les méthodes de validation notamment à travers l'étude de l'impact de l'autocorrélation spatiale sur un jeu de données suffisamment important pour tester différents effets d'échelle.

Le troisième objectif consiste à évaluer et à quantifier l'apport de la phénologie des essences forestières à partir d'images satellites. À l'aide de Sentinel-2 et de ses 5 jours de revisite, on pourra tout d'abord étudier si la résolution temporelle est suffisante pour suivre la phénologie des essences. Ensuite, à l'aide des méthodes validées par les précédents objectifs, il sera étudié le lien entre les dates sélectionnées pour chaque feuillu et leur signification sur le terrain.

Le dernier objectif, en se focalisant sur la reproductibilité des travaux, se dénote des objectifs précédents. L'enjeu ici est d'assurer au maximum la reproductibilité des travaux de la thèse et de faciliter aussi la transmission des bonnes pratiques identifiées dans les objectifs précédents.

Pour résumer, cette thèse va répondre aux quatre objectifs suivants :

1. évaluer le potentiel des SITS ; comprendre les facteurs influençant la qualité
2. prendre en compte l'autocorrélation spatiale dans la procédure de validation afin de retirer un biais d'évaluation
3. comprendre le réel apport de la phénologie dans les performances des classification
 - (a) identifier les dates importantes
 - (b) comparer avec la phénologie observée in-situ
4. rendre les travaux reproductibles (au regard des limites et contradictions des travaux précédents)
 - (a) création de bibliothèques python
 - (b) partage des jeux de données (références terrain et suivi phénologique)

1.4.3 Organisation du manuscrit

Le manuscrit s'organise autour de quatre chapitres thématiques répondant aux quatre objectifs.

Introduction (Chapitre 1) : Cette partie décrit le contexte des travaux de thèse. Après avoir expliqué brièvement pourquoi il y a besoin d'une cartographie des espèces forestières, la propriétés spectrales des feuilles sont étudiées en détail, que ce soit à l'échelle de la feuille, de l'arbre ou du peuplement. Puis, les approches de production des cartes des essences existantes sont décrites. Enfin, les défis, les objectifs et l'organisation du manuscrit sont détaillés.

1.4. QUESTIONS DE RECHERCHE

Chapitre 2 : Dans le chapitre 2, la stabilité temporelle et spatiale de la cartographie des essences forestières est étudiée. À l'aide de neuf années de séries temporelles du satellite Formosat-2 qui seront utilisées de manière indépendante, il est donc possible de connaître le potentiel de cartographie des essences du site d'étude grâce à la stabilité spatiale des essences années après années.

Chapitre 3 : Le troisième chapitre est dédié à l'estimation et à l'impact de l'autocorrélation spatiale dans l'évaluation des modèles d'apprentissage. Six méthodes de validation croisées, allant de la plus restrictive à la moins contraignante en terme de traitement de l'autocorrélation spatiale seront comparées en faisant varier de manière conséquente la taille des échantillons d'apprentissage, et de l'autocorrélation spatiale. L'impact des méthodes sur la qualité est évalué.

Chapitre 4 : Le chapitre 4 a pour objectif d'identifier si les dates sélectionnées par un algorithme pour cartographier une essence sont liées avec des événements phénologiques in situ. Pour cela deux années de suivi phénologique in situ ont été menées en 2018 et 2019. En plus de lier les dates avec des événements phénologiques, il sera étudié le potentiel de Sentinel-2 pour extraire les événements phénologiques clés (débourrement, pousse de feuilles, coloration et chute de feuilles) et pour différencier les essences les unes des autres.

Chapitre 5 : Le cinquième chapitre se distingue des précédents par son approche méthodologique. Tout au long de cette thèse, des outils ont été développés et des données récoltées. Afin de les valoriser, il sera présenté le grand travail de mise en valeur qui a été réalisé autour de ces contributions, le tout en respectant au maximum les principes de recherche reproductible.

Conclusion (Chapitre 6) : Ce dernier chapitre contient la conclusion de cette thèse. Il résume les principales contributions de ce manuscrit. Des perspectives sont aussi proposées pour continuer la recherche sur cette thématique.

Chapitre 2

Étude de la stabilité spatiale et statistique des prédictions

“ D’abord, la science n’est pas : elle se fait. Le savant du jour n’est que l’ignorant du lendemain.”

— Élisée Reclus.

Table des matières

2.1	Introduction	41
2.2	Material	43
2.2.1	Study area	43
2.2.2	Satellite image time series	43
2.2.3	Ancillary data	44
2.2.4	Field data	44
2.3	Classification protocol	45
2.3.1	Pre-processing	45
2.3.2	Training	45
2.3.3	Estimating prediction errors by spatial cross-validation	47
2.3.4	Accuracy assessment of one-year classifications and comparison	48
2.4	Results	49
2.4.1	Overall statistical performances	49
2.4.2	Accuracy per species	50
2.4.3	Confusion between species	50
2.4.4	Spatial agreement between years	51
2.5	Discussion	55
2.5.1	Effect of spatial autocorrelation : the SLOO-CV strategy as a standard	55
2.5.2	Effect of the size of the reference sample	55
2.5.3	Effect of clouds and cloud shadows	56
2.5.4	Effect of the available dates in the SITS	58
2.5.5	Differences between species	58
2.6	Conclusion	59

Pour cartographier à haute résolution spatiale les essences forestières à partir d’images satellites sur de larges étendues (de l’ordre de plusieurs kilomètres), il est d’usage d’utiliser des séries temporelles d’images satellites qui contiennent plusieurs acquisitions à différentes saisons sur une année.

Cependant, les travaux utilisant cette méthode se concentrent sur les images d’une seule et unique année et la stabilité temporelle (cartographie d’une année sur l’autre) et spatiale

(les endroits où aucune référence terrain n'est disponible) n'est pas étudiée. Ainsi, ce chapitre s'intéresse à la stabilité temporelle et spatiale de la cartographie des essences sur 53 km² de forêt à partir de neuf années indépendantes de séries temporelles du satellite Formosat-2.

La première partie de ce chapitre explique la problématique liée à l'estimation de la qualité des cartographies des essences forestières. Dans un deuxième temps les données utilisées pour réaliser cette étude sont présentées. Ensuite, le protocole est détaillé avec une attention particulière sur la sélection des jeux de données pour l'entraînement et la validation. Par la suite les résultats sont présentés, de manière globale, puis par espèce et par année. Une partie dédiée à discussion se focalise sur les explications des limites de la méthode et compare les résultats aux autres travaux du domaine. Enfin, les conclusions sont tirées.

Statistical Stability and Spatial Instability in Mapping Forest Tree Species by Comparing 9 Years of Satellite Image Time Series

Abstract

Mapping forest composition using multiseasonal optical time series remains a challenge. Highly contrasted results are reported from one study to another suggesting that drivers of classification errors are still under-explored. We evaluated the performances of single-year Formosat-2 time series to discriminate tree species in temperate forests in France and investigated how predictions vary statistically and spatially across multiple years. Our objective was to better estimate the impact of spatial autocorrelation in the validation data on measurement accuracy and to understand which drivers in the time series are responsible for classification errors. The experiments were based on 10 Formosat-2 image time series irregularly acquired during the seasonal vegetation cycle from 2006 to 2014. Due to lot of clouds in the year 2006, an alternative 2006 time series using only cloud-free images has been added. Thirteen tree species were classified in each single-year dataset based on the Support Vector Machine (SVM) algorithm. The performances were assessed using a spatial leave-one-out cross validation (SLOO-CV) strategy, thereby guaranteeing full independence of the validation samples, and compared with standard non-spatial leave-one-out cross-validation (LOO-CV). The results show relatively close statistical performances from one year to the next despite the differences between the annual time series. Good agreements between years were observed in monospecific tree plantations of broadleaf species versus high disparity in other forests composed of different species. A strong positive bias in the accuracy assessment (up to 0.4 of Overall Accuracy (OA)) was also found when spatial dependence in the validation data was not removed. Using the SLOO-CV approach, the average OA values per year ranged from 0.48 for 2006 to 0.60 for 2013, which satisfactorily represents the spatial instability of species prediction between years.

Authors : Nicolas Karasiak¹, Jean-François Dejoux², Mathieu Fauvel², Jérôme Willm¹, Claude Monteil¹ and David Sheeren¹.

¹ : DYNAFOR, Université de Toulouse, INRAE, 31326 Castanet-Tolosan, France.

² : CESBIO, Université de Toulouse, CNES/CNRS/INRAE/IRD/UPS, 31401 Toulouse, France.

Paper status : Published in Remote Sensing on October 26, 2019.

Paper DOI : 10.3390/rs11212512.

In situ tree species references DOI : 10.5281/zenodo.2581400.

Tree species webmap : <https://dynafor1201.github.io/publications/maps/treespeciesformosat2/>

2.1 Introduction

Forest ecosystems play a major role in global biodiversity [Thompson et al. \[2011\]](#). They provide several services to humanity including carbon sequestration (which regulates climate [Bunker et al. \[2005\]](#)), timber production [Thompson et al. \[2009\]](#), soil protection [Harris \[2009\]](#), and recreation. They also have an impact on human health and well-being. However, the provision of such ecosystem services depends on several factors including the diversity of tree species [Gamfeldt et al. \[2013\]](#). Therefore, knowing the distribution of tree species in forests is crucial to assess ecosystem functions and services. More broadly, information on tree species is required for forest management and also for long-term forest monitoring, especially in the current context of climate change and related disturbances (forest fires, windstorms, drought, pests and diseases) [Seidl et al. \[2017\]](#).

Remote sensing has long been used to collect information on forest resources including stand composition [Boyd \[2005\]](#); [Walsh \[1980\]](#). Nevertheless, accurately distinguishing tree species is still challenging [Fassnacht et al. \[2016\]](#). In the past, maps of tree species were based on field surveys completed by computer-aided analysis of aerial photographs [Meyera et al. \[1996\]](#); [Trichon and Julien \[2006\]](#). While this approach provides accurate operational results for forest managers, it is limited to small spatial extent because it is costly and time consuming, which also affects its updating. In the last few decades, various types of remotely sensed images have been used to automate the identification of forest tree species. Some authors focused on the spatial resolution using very high-resolution satellite or airborne imagery [Carleer and Wolff \[2004\]](#); [Immitzer et al. \[2012\]](#); [Lin et al. \[2015\]](#); [Waser et al. \[2011\]](#). They assumed the classification would benefit from the spatially detailed information and would therefore be accurate. Despite some successful results, this approach revealed itself to be of limited interest when only a single date was used due to the low spectral resolution of the data because of the reduced spectral and temporal information. Alternatively, as tree morphology and biochemical traits have a subtle influence on spectral reflectance [Ustin et al. \[2009\]](#), several authors explored airborne hyperspectral imagery [Aval et al. \[2019\]](#); [Féret and Asner \[2012\]](#); [Ghiyamat and Shafri \[2010\]](#). Depending on the number of classes of species, on the methodology used for classification, and on the characteristics of the images (pixel size, number of spectral bands), the accuracy of the classification varied. Nevertheless, studies based on hyperspectral imagery were typically more accurate than those based on single-date multispectral data [Fassnacht et al. \[2016\]](#).

Taking advantage of the temporal dimension of the satellite data was another way to separate tree species [Fassnacht et al. \[2016\]](#). Time series can capture the phenological behavior of the vegetation and this functional trait can be useful to discriminate the forest types. Changes in pigment contents, water and leaf morphology across seasons can vary from one species to another. Time series with images covering all phenological events from green-up to senescence (leaf-on, spring flush, autumn senescence, leaf-off) can produce detailed classification results. The use of multitemporal data for this purpose is not new. This approach has been explored from various image datasets of different spatial and temporal resolutions based on spaceborne sensors such as MODIS [Aragones et al. \[2019\]](#); [Cano et al. \[2017\]](#), Landsat [Diao and Wang \[2016\]](#); [Foody and Hill \[1996\]](#); [Pasquarella et al. \[2018\]](#); [Wolter et al. \[1995\]](#); [Zhu and Liu \[2014\]](#), RapidEye [Tigges et al. \[2013\]](#), as well as airborne sensor [Hill et al. \[2010\]](#); [Key et al. \[2001\]](#) or unmanned aerial systems [Lisein et al. \[2015\]](#). More recently, the potentialities of the new freely available high spatial resolution Sentinel-2 (S2) data have been investigated [Bolyn et al. \[2018\]](#); [Immitzer et al. \[2016\]](#); [Liu et al. \[2018\]](#); [Persson et al. \[2018\]](#); [Spracklen and Spracklen \[2019\]](#). In general, authors found a benefit to combine images acquired in spring and autumn, at the key phenological stages of temperate forests since it

influences positively the classification accuracy. Summer images are also frequently selected in features ranking procedure, especially for conifer species [Spracklen and Spracklen \[2019\]](#) but also for deciduous tree species [Hill et al. \[2010\]](#). From a spectral point of view, red-edge bands and SWIR bands were reported as important variables when S2 time series were used [Bolgen et al. \[2018\]](#); [Immitzer et al. \[2016\]](#); [Persson et al. \[2018\]](#).

More recently, the potential of the new freely available high spatial resolution Sentinel-2 (S2) data has been investigated [Bolgen et al. \[2018\]](#); [Immitzer et al. \[2016\]](#); [Liu et al. \[2018\]](#); [Persson et al. \[2018\]](#); [Spracklen and Spracklen \[2019\]](#). In general, the authors found it advantageous to combine images acquired in spring and autumn, at the key phenological stages of temperate forests, since it had a positive influence on the accuracy of the classification. Images acquired in summer are also frequently selected in features ranking procedures, particularly for conifer species [Spracklen and Spracklen \[2019\]](#), but also for deciduous species [Hill et al. \[2010\]](#). From a spectral point of view, red-edge bands and SWIR bands are reported to be important variables when S2 time series are used [Bolgen et al. \[2018\]](#); [Immitzer et al. \[2016\]](#); [Persson et al. \[2018\]](#).

Despite the increasing number of studies that use time series to identify forest types, the true predictive power of these kinds of data remain to be demonstrated. Even though it is difficult to compare studies because of the use of different methods, sensors, and classes of tree species, we observed very contrasted results from one study to another. For instance, using four dates for S2 data in 2017, [Persson et al. \[2018\]](#) obtained a kappa value of 0.83 to classify five species (Norway spruce (*Picea abies*), Scots pine (*Pinus sylvestris*), Hybrid larch (*Larix marschlinsii*), Birch (*Betula*) and Pedunculate Oak (*Quercus robur*)). This differs substantially from the Immitzer's results [Immitzer et al. \[2016\]](#) as they used only two S2 images to map seven species and so reached a kappa of 0.59 to identify seven species including Norway spruce, Scots pine, European larch (*Larix decidua*) and Oak species based on two S2 images.

There was also a difference of almost 0.2 points of OA between a study by [Persson et al. \[2018\]](#) and one by [Liu et al. \[2018\]](#) who classified eight types of forest in China with the same number of S2 images. In another study, using only two S2 images to separate 11 forest classes of broadleaf and conifers, [Bolgen et al. \[2018\]](#) obtained very accurate results (overall accuracy of 0.93) in contrast with previous works but in line with others based on dense time series acquired using different sensors [Sheeren et al. \[2016\]](#); [Tigges et al. \[2013\]](#); [Zhen et al. \[2013\]](#). The notable difference in accuracy among past studies suggests a better understanding is required of the factors that affect the classification of species, as recommended by [Fassnacht et al. \[2016\]](#). Several drivers of classification errors remain insufficiently explored, among which, spatial autocorrelation of reference data has long been identified but rarely quantified [Chen and Wei \[2009\]](#); [Hammond and Verbyla \[1996\]](#). Spatial dependence in the reference data due to an inadequate sampling strategy to split training and validation sets can wrongly increase classification accuracy [Hammond and Verbyla \[1996\]](#); [Meyer et al. \[2018\]](#); [Schratz et al. \[2019\]](#). Contamination by clouds and cloud shadows in dense image high resolution time series may also have a major impact on classification. Because the distribution of such contamination may vary over time and in space across years, a multiyear analysis is required to reliably evaluate their effect.

To our knowledge, this is the first study of variability between one-year classifications of tree species based on multiple years using dense image high spatial resolution time series. We evaluated the classification performance of single-year Formosat-2 time series in distinguishing forest types with spatially independent validation data. We also investigated how the predictions vary statistically and spatially across multiple years (from 2006 to 2014). The main contribution of this work is a better estimation of the classification accuracy of the forest maps by reducing optimistic bias due to spatial autocorrelation. The second contribution, resulting from the first, is a finer understanding of the drivers responsible for classification errors. We hypothesize that time series data improve species discrimination compared to

single-date image due to seasonal variability in spectral reflectance between species.

2.2 Material

2.2.1 Study area

The study site is located in south-western France, next to Toulouse, and covers an area of 24 km x 24 km (Fig 2.1). This delimited area was determined by a satellite acquisitions scheme by the *Centre National d'Etudes Spatiales* (CNES) who acquired a Formosat-2 Satellite Image Time Series (SITS) of the site. The Garonne river crosses the eastern part of the study area, influencing soil composition and the nearly flat topography of the area. The climate is sub-Atlantic characterized by sunny autumns, hot dry summers, and mild rainy winters (the average annual temperature is $> 13^{\circ}\text{C}$; annual precipitation = 656 mm). The landscape is dominated by arable lands (including wheat, sunflower, maize) and grasslands. Forests cover up to 10% of the landscape (53 km²).

2.2.2 Satellite image time series

We used a dense optical image dataset composed of Formosat-2 time series acquired in nine consecutive years from 2006 to 2014. This dataset was obtained during preparation for the Sentinel-2 and VEN μ S mission with cooperation between the Israeli Space Agency (ISA) and the French CNES [Dedieu et al. \[2006\]](#). A total of 156 dates was acquired with an average of 14 images per year and a maximum of 43 images in 2006. The distribution of the dates over time varied from one year to another and the number of images available during the growth season differed from the number available at the end of vegetation season (Figure 2.2).

Cloud coverage also varied considerably from one date to another, ranging from a minimum of 8 cloud-free images in 2011 to a maximum of 20 in 2006. For 2006, by visual inspection, we created an additional dataset (named 2006 bis) by selecting only the cloud-free images, resulting in a time series of 20 dates (compared to the original 46).

The Formosat-2 multispectral images are delivered in an 8-bit radiometric resolution. Each image provides 4 spectral bands ranging from the visible (Blue : 0.45–0.52 μm , Green :

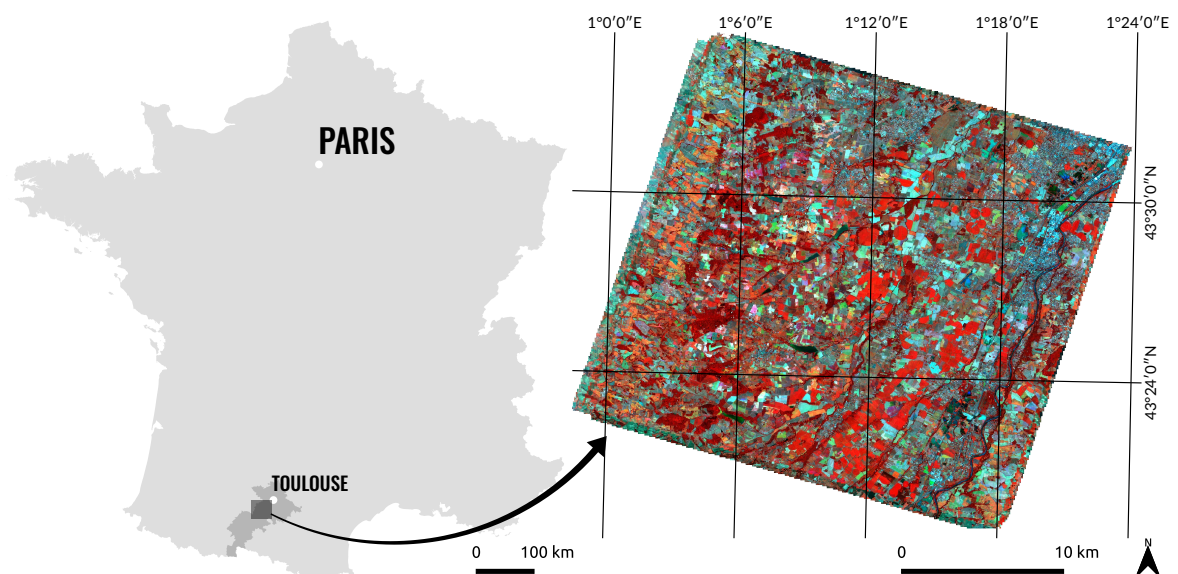


FIGURE 2.1 – The map on the left shows the location of the study area in the Haute-Garonne district (in dark grey) near Toulouse, France. On the right, a false-color image acquired on 2013-08-13 which represents the entire Formosat-2 scene covering an extent of 24 x 24 km.

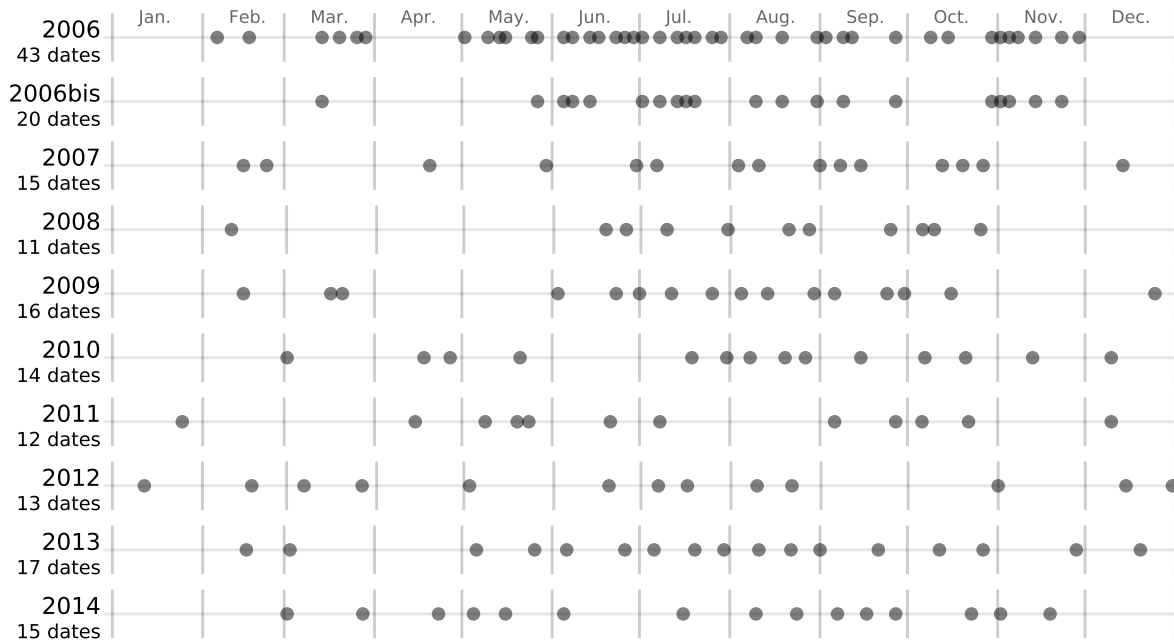


FIGURE 2.2 – Number and acquisition dates of each image in the Formosat-2 time series from 2006 to 2014.

0.53–0.60 μm , Red : 0.63–0.69 μm) to the near-infrared (NIR : 0.76–0.90 μm) with a nominal pixel size of 8 m. All the images were acquired under a constant viewing angle and a field of view of 24 km like Landsat, VEN μ S and Sentinel-2.

2.2.3 Ancillary data

A forest mask produced in 1996 by the French National Forest Inventory database (IGN BDForêt[®], v.1) was used to select forest pixels in the SITS (i.e. forest stands with a minimum area of 2.25 hectares) and to exclude non-forested areas. Based on aerial photographs taken in 2006, 2010 and 2013 (IGN BDOrtho[®]), the mask was manually updated to retain only SITS forest stands that remained stable over the nine year period

2.2.4 Field data

Four field campaigns were conducted between November 2013 and January 2017 to identify and locate reference samples of tree species in the study site. All the main forests were visited. Only the dominant broadleaf and conifer tree species were recorded. To insure tree species purity in the training samples, plots were delimited at the center of homogeneous areas covering an area of approximately 576 m² (i.e. nine contiguous 8 m \times 8 m Formosat-2 pixels). Only the pixel at the center of each area was used for the classification protocol. Plots were located using a Garmin GPSMap 62st receiver (3-5 m accuracy) and distributed over 72 distinct forest stands.

Thirteen tree species of which eight were broadleaf species and five conifer species were studied (Table 2.1). In some species, identification was limited to the genus level because of the existence of cultivars (case of Aspen) and the difficulty involved in determining the exact species in some cases Oak, Willow and Eucalyptus. We acquired a total of 1262 sample plots. Class distribution was moderately imbalanced reflecting the uneven distribution of species abundances in the forests. The number of samples varied from 50 (the minimum for Willow) to 211 (the maximum for Aspen). Conifers were less well represented with an average of 73 samples per class compared with 112 for broadleaf species.

TABLEAU 2.1 – List of tree species with their sample size, in pixels, collected during field surveys ($n = 1262$). The number of forest stands in which the samples were collected is also provided. Stand delimitation is based on the French National Forest Inventory database (IGN BDForêt[®] v.1)

Species	Sample Size	Forest stands
Broadleaf		
Silver birch (<i>Betula pendula</i>)	85	3
Oak (<i>Quercus robur/pubescens/petraea</i>)	115	12
Red Oak (<i>Quercus rubra</i>)	147	7
Aspen (<i>Populus spp.</i>)	211	6
European Ash (<i>Fraxinus excelsior</i>)	80	3
Black locust (<i>Robinia pseudoacacia</i>)	63	7
Willow (<i>Salix spp.</i>)	50	3
Eucalyptus (<i>Eucalyptus spp.</i>)	148	4
Conifer		
Corsican Pine (<i>Pinus nigra subsp. laricio</i>)	70	6
Maritime Pine (<i>Pinus pinaster</i>)	103	7
Black Pine (<i>Pinus nigra</i>)	55	2
Silver Fir (<i>Abies alba</i>)	75	5
Douglas Fir (<i>Pseudotsuga menziesii</i>)	60	7

2.3 Classification protocol

A global overview of the classification protocol applied on each Formosat-2 single-year time series is shown in Figure 2.3.

2.3.1 Pre-processing

In this step, surface reflectance time series were produced from the Formosat-2 level 1A images using the MACCS (Multisensor Atmospheric Correction and Cloud Screening) processing chain developed by the CNES Hagolle et al. [2008, 2010]. MACCS involved orthorectification, atmospheric correction, detection of clouds and cloud shadows, and reduction of topographic effects on illumination, based on multitemporal and multispectral criteria. Atmospheric correction relies on the estimation of aerosol optical thickness based on a spectro-temporal technique that minimizes (i) variations in surface reflectances between pixels acquired consecutive cloud-free images after correction and (ii) differences between the blue surface reflectance predicted from the red band (empirical relationship) and the blue surface reflectance obtained after correction Hagolle et al. [2015]. Clouds are detected using a multitemporal approach that analyzes the increase in reflectance in the blue spectral band Hagolle et al. [2010]. If high variation is observed, cloud is likely to be present. Based on this method, masks of clouds and related shadows are produced by MACCS for each image in the time series.

In the second step, SITS of each year were filtered using a linear gap-filling algorithm applied to each spectral band to remove noisy data (i.e. cloudy and shady pixels) and to retrieve their surface reflectance Grizonnet et al. [2017]. Invalid pixels were replaced by the interpolated values from the closest available valid pixels in the time series. Gap-filling was chosen for its simplicity and its previously demonstrated efficiency already demonstrated when time gaps between consecutive images are limited Kandasamy et al. [2013].

2.3.2 Training

Classification models were built using each full one-year time series as predictors with exactly the same pixels for training and testing using the supervised SVM (Support Vector

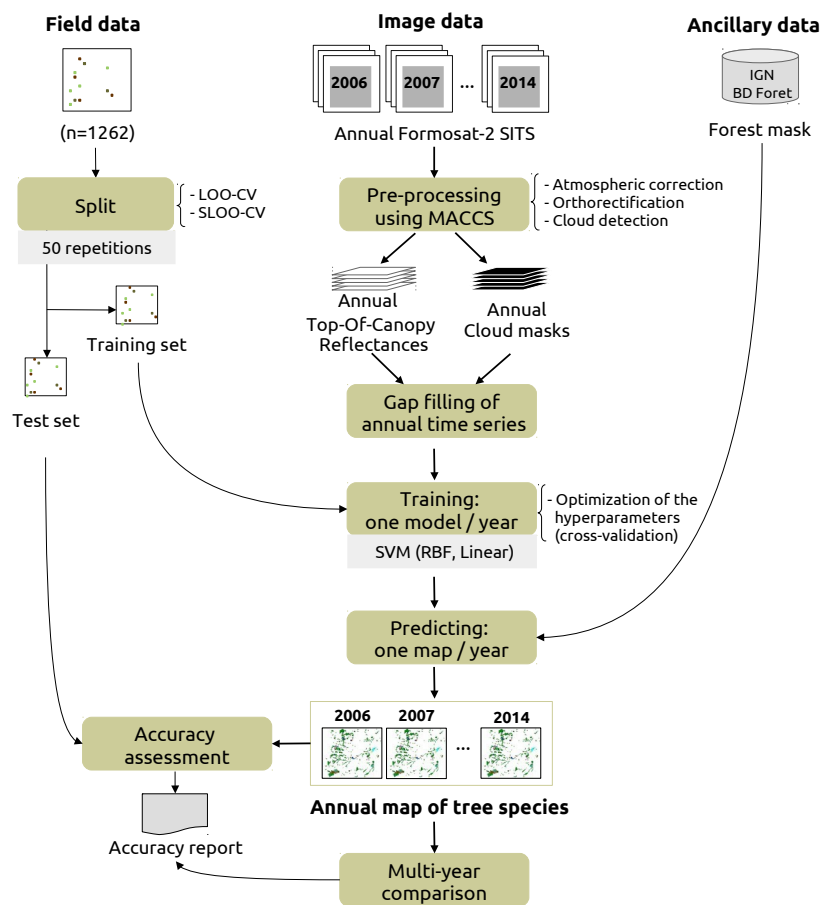


FIGURE 2.3 – Classification protocol for a single year time series, repeated for the 9 years available from 2006 to 2014. The splitting procedure to create independent training and test sets is based on a spatial and non-spatial leave-one-out cross-validation (SLOO-CV and LOO-CV respectively). The LOO-CV were trained with exactly the same number of training samples as the SLOO-CV, after random undersampling.

Machine) classifier Vapnik [1998] known to be the best approach in the case of small training data sets with respect to data dimensionality Mountrakis et al. [2011]. In this study, we selected the Radial Basis Function (RBF) kernel which is the most frequently used and has already been proven to be effective in the case of similar classification problems Kavzoglu and Colkesen [2009]. Hyperparameters including the regularization parameter (C) and the kernel bandwidth (γ) were tuned by cross-validation in a search space with the following settings : : $C = \{10^{-2}; 10^{-1}; \dots; 10^{10}\}$ and $\gamma = \{10^{-9}; 10^{-8}; \dots; 10^3\}$. A linear kernel was also tested for comparison with RBF. However, since the linear kernel performed worse, the results are not presented here. To account for imbalanced data and to prevent potential bias due to the dominant classes Graves et al. [2016], the class weights in the SVM parameters were also modified. Weights were set inversely proportional to class frequencies. SVM was computed using the scikit-learn python library Pedregosa et al. [2011]. Vector of features were standardized (i.e. centering and scaling to unit variance) prior to training.

2.3.3 Estimating prediction errors by spatial cross-validation

Because spatial autocorrelation between reference samples may produce optimistic bias in assessments of classification performance Hammond and Verbyla [1996]; Meyer et al. [2018]; Schratz et al. [2019], we used a spatial leave-one-out cross-validation (SLOO-CV) sampling strategy Le Rest et al. [2014]; Pohjankukka et al. [2017] to separate the training and test sets to guarantee full independence between them. In this approach, one reference sample is used as the test set and the remaining samples, non-spatially correlated with the test set, are used as the training set (Figure 2.4). This is repeated n times where n equals the number of reference samples. The n prediction results are then averaged to obtain an estimation of the prediction error. In our case, the test set was composed of one pixel of each class (i.e. a total of 13 pixels at each iteration) and the procedure was repeated 50 times, this being the number of reference samples of the lowest class size. We compared this splitting procedure with the classical non-spatial leave-one-out cross-validation strategy (LOO-CV) using the same training size per class as in SLOO-CV, by random undersampling. For year-to-year comparison, we also used the same training and test sets related to each sampling approach by setting the same random seed.

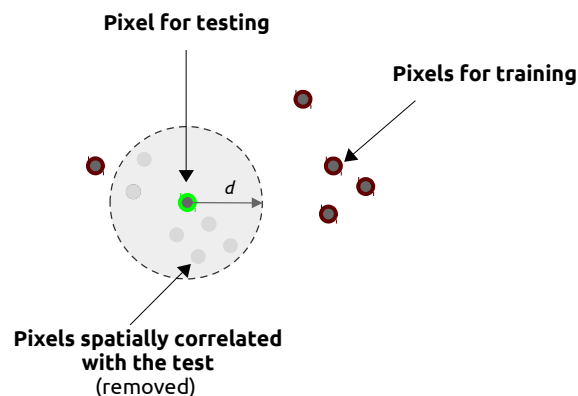


FIGURE 2.4 – Spatial leave-one-out cross-validation (SLOO-CV) schema for one class. One pixel is used for testing. The other pixels are used for training, except pixels geographically too close to the pixel selected for testing. This procedure is repeated n times where n is the number of reference samples. Spatial autocorrelation between nearby pixels is assumed up to a distance d which can be estimated using Moran's I.

The spatial autocorrelation distance was estimated by computing the Moran's Index from the pixels of forests in the SITS Moran [1950]. Moran's I estimates the correlation between the value of a variable at one location and nearby observations. The index ranges from -1

(negative spatial autocorrelation) to +1 (positive spatial autocorrelation) with a value close to 0 in the absence of spatial autocorrelation (random spatial distribution). More formally, the Moran's I is defined as the ratio of the covariance between neighborhood pixels and the variance of the entire image :

$$I(d) = \frac{n}{S_0} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{i,j} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.1)$$

where, in our case, x_i is the pixel value of x (a spectral band of the SITS for pixels of forests) at location i , x_j is the pixel value of x at location j (a nearby pixel of forest of i), \bar{x} is the average value of x , n is the number of pixels of forests in the image, $w_{i,j}$ is the weight equals to 1 if pixel j is within distance d of pixel i , otherwise $w_{i,j} = 0$, and S_0 the sum of all $w_{i,j}$'s :

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{i,j} \quad (2.2)$$

In this study, Moran's I was computed for each spectral band of each year, for neighborhoods (lags) varying from 1 to 100 pixels (i.e. from 8 m to 800 m). Based on correlograms, we evaluated the distance between nearby pixels for which Moran's I equals 0.2, considering the potential effect of spatial autocorrelation as not significantly different from the threshold value of Moran's I Dale and Fortin [2014]. Then, the median distance was calculated for each spectral band, taking all the dates of one year into account (Figure 2.5). This was done for each year. Finally, the average value of the median distance of each year was kept in the spatial cross-validation procedure to split the training and test sets. This average value was estimated to be 240 m (i.e. 49 pixels).

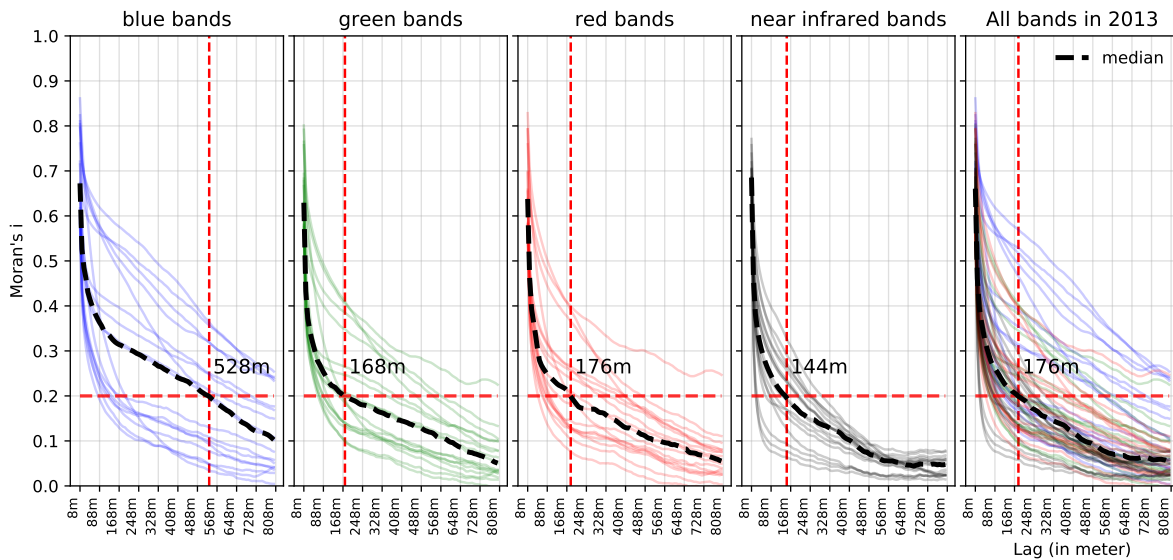


FIGURE 2.5 – Moran's I correlograms of each Formosat-2 spectral band of the SITS 2013, for pixels representing forests. Each curve represents one date of the SITS. the red dashed line represents the median distance value (in x) where Moran's I = 0.2 (in y). For a Moran's I threshold value of 0.2, spatial independence between nearby pixels was assumed. This is the case beyond to 528 m in the blue band, 168 m in the green, 176 m in the red and 144 m in the near-infrared.

2.3.4 Accuracy assessment of one-year classifications and comparison

The results of the classifications were assessed according to the confusion matrix based on Overall Accuracy (OA) and the F1 score (i.e. the harmonic mean of precision and recall

varying from 0 for the worst case to 1 for perfect classification), errors of omission and errors of commission. A Wilcoxon signed-rank test was used to determine if the difference in accuracy between annual classifications and sampling strategies (LOO-CV vs SLOO-CV) was statistically significant.

Classifications were also compared spatially to highlight instability between years. A map of uncertainty was produced by computing the number of agreements between the one-year classifications (i.e. the modal value related to the class with the highest frequency) for each pixel. Additionally, the distribution of this uncertainty was examined per class using either all the pixels or only the reference samples. Finally, the maps were visually inspected to identify problem areas and to better understand the errors with the help of field knowledge. The maps shown in results section were produced using the SLOO-CV.

2.4 Results

2.4.1 Overall statistical performances

The classification performances for each year are presented in Table 2.2. Generally speaking, the performances were similar between the years but very different between sampling strategies (SLOO-CV vs LOO-CV) in a given year.

When prediction errors were estimated by spatial cross-validation (SLOO-CV), the average OA varied from 0.48 in 2007 to 0.60 in 2013 with high variability in the results (average standard deviation of 0.12). No significant differences were observed between the years 2008-2012, 2012-2014 and between 2006 and 2007 which were the cloudiest SITS ($p < 0.05$; Wilcoxon signed-rank test statistic; see Appendix A.2 for statistical details). For the year 2006, when cloudy images were removed from the SITS (i.e. using the 2006bis dataset), the classification was improved, the performance was similar to that in the other years (average OA = 0.57). The best classification was obtained using the 2013 time series (average OA = 0.60).

When accuracy was computed using the standard leave-one-out cross-validation (LOO-CV), prediction errors were very low compared to when SLOO-CV was used, suggesting a high optimistic bias in the evaluation. The average OA varied from 0.97 in 2011 to 1.00 in 2006 and 2014 with a standard deviation close to zero. The cloudiest years (2006 and 2007) did not differ significantly in performance from the other years in most cases (Appendix A.2). These results contradict the previous ones : while the year 2007 was the worst with the SLOO-CV, with LOO-CV it had the second best score.

In the following sections, we only detail the results based on the SLOO-CV strategy since it best reflects the true performance of the classifications.

TABLEAU 2.2 – Accuracy report of spatial leave-one-out cross-validation (SLOO-CV) sampling strategy and leave-one-out cross-validation (LOO-CV) for each single-year classification based on OA statistics. The 2006bis time series only includes cloud-free images of 2006. The average percentage of cloud coverage was estimated by computing for each species the number of time each reference sample was affected by clouds (detected from the MACCS processing chain).

	2006	2006bis	2007	2008	2009	2010	2011	2012	2013	2014
Classification accuracy (average Overall Accuracy \pm standard deviation)										
SLOO-CV	0.52 \pm 0.13	0.57 \pm 0.15	0.48 \pm 0.12	0.57 \pm 0.10	0.55 \pm 0.11	0.56 \pm 0.12	0.55 \pm 0.11	0.58 \pm 0.14	0.60 \pm 0.11	0.58 \pm 0.11
LOO-CV	1.00 \pm 0.02	0.99 \pm 0.03	0.99 \pm 0.02	0.98 \pm 0.04	0.99 \pm 0.03	0.98 \pm 0.03	0.97 \pm 0.04	0.98 \pm 0.04	0.99 \pm 0.02	1.00 \pm 0.02
Characteristics of each SITS										
Number of images	43	20	15	11	16	14	12	13	17	15
Images in spring	13	4	2	1	2	3	4	3	3	5
Images in autumn	10	6	4	4	3	4	4	2	4	4
Cloud coverage	25%	0%	12%	5%	4%	3%	2%	0%	1%	0%

2.4.2 Accuracy per species

In most cases, whatever the year, broadleaf tree species were better discriminated than conifers (Figure 2.6). The highest performances were obtained for monospecific plantations of Red oak (average F1 score = 87%) and Willow (average F1 score = 86%). Aspen was also detected with good accuracy (average F1 score = 68%). Conversely, some species were difficult to identify, including European ash (average F1 score = 26%) and Silver birch (average F1 score = 36%) except in the years 2010 and 2013.

High confusion rates were obtained for conifer tree species. Black pine was the worst class with a F1 score close to zero, except in 2014 (F1 score = 62%). Maritime pines were generally better discriminated but the performances remained low (average F1 score = 40%). The best agreement was obtained for Silver fir (average F1 score = 50%) which reached its best score (average F1 score = 81%) in the year 2010.

	2006	2006bis	2007	2008	2009	2010	2011	2012	2013	2014	All years
Silver birch	24	49	16	30	10	69	33	38	72	22	36
Oak	40	44	61	70	43	50	42	50	62	58	52
Red oak	80	91	95	73	83	86	85	96	91	91	87
Aspen	82	74	62	84	73	52	65	59	71	61	68
European ash	25	35	10	40	21	15	48	5	32	30	26
Black locust	40	32	42	19	70	32	22	43	61	67	42
Willow	66	88	68	90	98	98	99	77	85	93	86
Eucalyptus	47	40	52	87	90	81	76	82	49	49	65
Corsican pine	41	36	19	19	15	27	25	25	3	62	27
Maritime pine	48	49	66	54	40	37	37	46	26	3	40
Black pine	0	0	0	0	4	4	0	24	13	62	10
Silver fir	37	53	6	37	50	81	58	69	69	43	50
Douglas fir	33	52	33	38	27	10	26	40	62	10	33
All species	43	49	40	49	48	49	47	50	53	50	

FIGURE 2.6 – Average F1 score (in %) per species and per year based on the SLOO-CV sampling strategy using the SVM (RBF kernel) classifier. Values range from white (F1 score = 0%) to dark green (F1 score = 100%). Average values of F1 score per year and per species are also provided (in the bottom row and last column on the right, respectively).

On average, the year 2013 was the best, mainly because of a high score for Silver birch compared to the other years. Year 2007 was the least accurate. Higher performance disparity was observed from one year to another for most species, except Red oak and Willow.

2.4.3 Confusion between species

Generally, when errors occurred, the broadleaf tree species were confused with each other as well as with conifers. The main source of omissions for Silver birch was mispredictions as Oak which, in turn, was confused with European ash but also with Black locust and with some pines (see the confusion matrix for the year 2013 in Figure 2.7, for example). Red oak was the subject of very little confusion. High rates of omissions were observed for European ash

with misclassifications as Oak, Aspen and Black locust. Under-detection was also observed for the evergreen Eucalyptus plantations due to confusion with Willow. In conifer species, the errors mainly appeared between species of Pine but also between Pine and Douglas fir.



FIGURE 2.7 – Confusion matrix for 13 tree species for year 2013. Each cell provides an average value of agreement or confusion (in %) based on 50 classifications (number of iterations) using the SLOO-CV sampling strategy.

Confusions between species were similar from one year to another but the commission and omission errors rates varied and accuracy was very low for some species (Figure 2.8).

2.4.4 Spatial agreement between years

As revealed by the map of the modal class values, Oak was the most representative species in the study region, especially in the small forests, which is consistent with our field observations (Appendix A.1). Conifers and plantations of broadleaf species were less frequent but pixels of the same class appear to be grouped in homogeneous stands, as expected.

When spatial uncertainty was analyzed using the map of agreements between the one-year classifications, good stability was observed in the monospecific tree plantations of broadleaf species (Figure 2.9). The stands composed of Aspen, Red oak, and Eucalyptus were clearly differentiated. In contrast, in complex forests including a mix of different species, disagreements between annual classifications were higher, as suggested by the previous statistical assessment. An example is given in Figure 2.10 showing a mix forest composed of conifers (mainly Black pine but also Douglas fir and Silver fir) and deciduous species (mainly Oak and Silver birch). There was considerable confusions between conifer species from one year to another (low agreement). The extent of Silver birch areas was also highly variable. In this forest, the dominant species were rather well-identified but their exact location was inaccurate at the pixel level.

Significant disagreements between the classifications were also observed in other contexts,

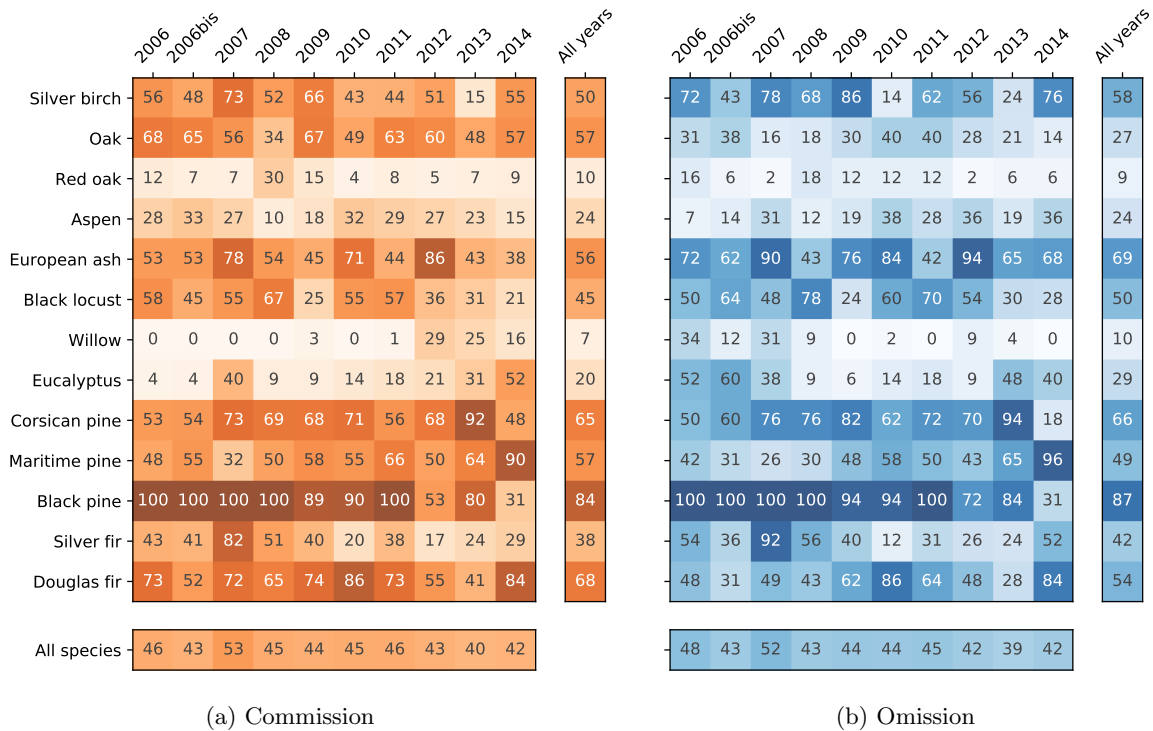


FIGURE 2.8 – Average rate of commission and omission errors (in %) per species and per year based on the SLOO-CV sampling strategy.

especially in thin riparian forests and forest edges where species composition and diversity is high, with lots of species unsampled (Figure 2.9). This was also true in low density forest stands, for which confusions appeared with the understory vegetation. Finally, disagreements were also observed in areas very affected by clouds and shadows.

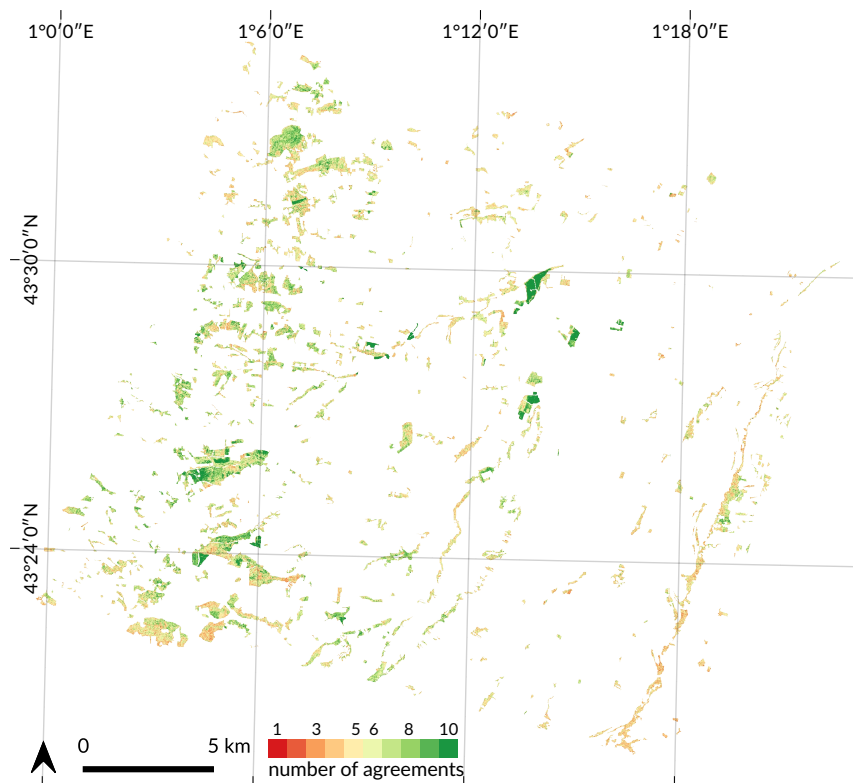


FIGURE 2.9 – Spatial comparison between the annual classifications of forest tree species from 2006 to 2014 (including 2006bis). The number of agreement is the modal value related to the class with the highest frequency. This map illustrates the stability of predictions from one year to another with a high number of disagreements in red (high uncertainty) and a low number of disagreements in green (high accuracy). In many cases, homogeneous green areas are tree plantations of broadleaf species such as Red Oak and Aspen or Eucalyptus.

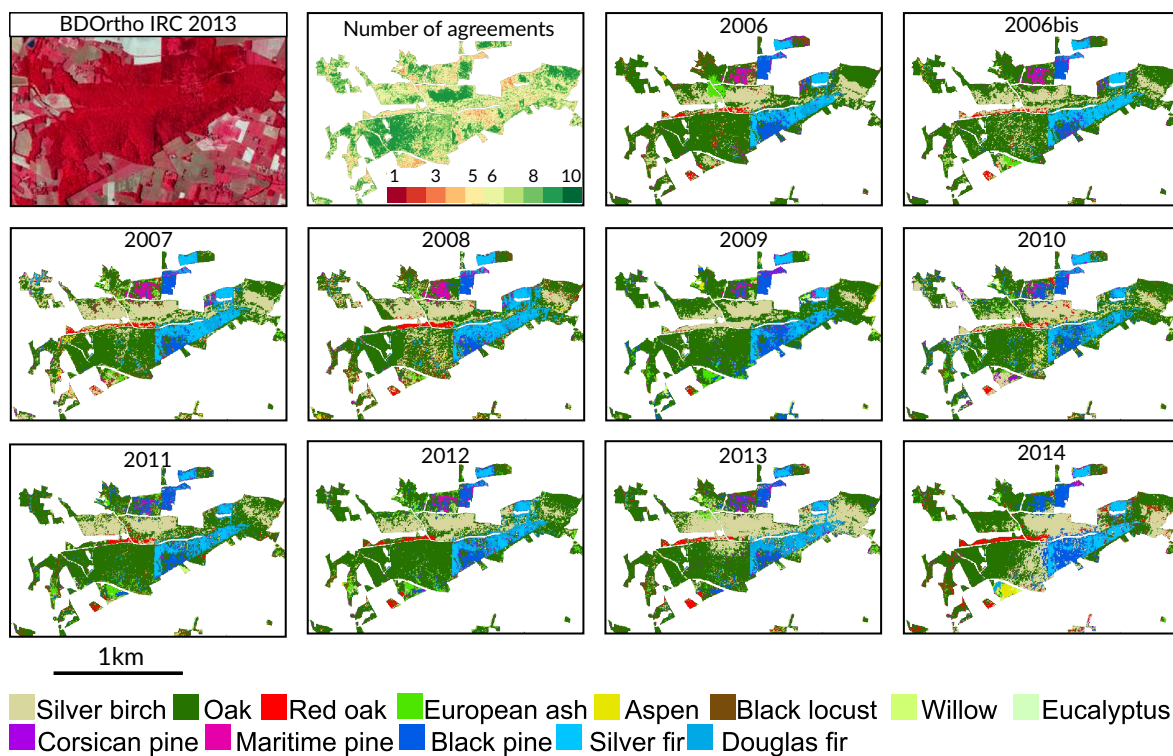


FIGURE 2.10 – Annual classifications of tree species in a mixed forest composed of different conifers and broadleaf species. Instability was observed in the conifer plantations composed of Black pine, Silver fir and Douglas fir. This was also the case for Silver birch. Part of the forest (in the north) has been excluded from the analysis because of changes during the study period (clear-cuts and reforestation).

2.5 Discussion

In this study, an archive of Formosat-2 time series was used to classify tree species in temperate forests in nine consecutive years. Each classification was validated using the same spatial leave-one-out cross-validation approach to remove the test samples that were spatially correlated with the training samples. To our knowledge, this is the first study to examine the stability of predictions from one year to another using dense SITS of high spatial resolution with spatially independent validation data. The present study is a first attempt to assess the robustness of tree species discrimination in multiple years and to better understand the drivers that affect classification performances.

2.5.1 Effect of spatial autocorrelation : the SLOO-CV strategy as a standard

Our results revealed a strong positive bias in validation based on the usual LOO-CV strategy for splitting reference data. This bias was already suspected in a previous study when we used stratified-k-fold but was not quantified Sheeren et al. [2016]. Regarding the importance of the overestimation in the classification accuracy ($\Delta OA > 0.4$ between LOO and SLOO-CV), the use of spatially independent data for validation should no longer be an option but wherever possible, a requirement, in agreement with the recommendation of Fassnacht et al. [2016].

Spatial autocorrelation in the reference data has long been known to affect the classification and accuracy assessment Chen and Wei [2009]; Congalton [1991]; Griffith and Chun [2016]; Hammond and Verbyla [1996]. Different sampling strategies for data splitting have already been studied in the literature including spatial Mu et al. [2015]; Pohjankukka et al. [2017]; Roberts et al. [2017]; Schratz et al. [2019] and aspatial approaches Lyons et al. [2018]; Mu et al. [2015]; Ramezan et al. [2019]; Stehman [2009]. Although the spatial sampling approach is recommended to reduce the spatial autocorrelation effect, an aspatial (i.e. random, systematic or stratified) sampling strategy assuming independence between samples is usually used in remote sensing analyses for the sake of simplicity Ramezan et al. [2019].

In this study, the SLOO-CV strategy was used to estimate an unbiased prediction performance, similar to used in Le Rest et al. [2014]. We measured the spatial dependence between nearby pixels of forests explicitly, using the Moran's I, as in Millard and Richardson [2015], and we separated training and validation samples that were located geographically too close to one another. In other studies, spatial partitioning was achieved differently, based on *k-means* clustering Brenning [2012]; Schratz et al. [2019] or on the definition of patches Cánovas-García et al. [2017], or blocks related to the spatial structure Roberts et al. [2017]. Whatever the spatial sampling method used, all the studies demonstrated larger errors in predictions with lower spatial autocorrelation between training and validation sets, as we observed here. The absence of independence between training and validation data thus provides an inflated estimate of classification performances as confirmed by our results.

An important point to note is that we guaranteed complete independence between the training and validation sets but not among the training samples. Thus, spatial autocorrelation still persisted in the training set. Compared to LOO-CV, the SLOO-CV strategy provides a statistical estimate of accuracy that fits the quality of the map product better (and hence predictive performance) but in terms of predictions, the results of classification results are similar. This is illustrated by the number of agreements between the annual classifications for both LOO-CV and SLOO-CV (Figure 2.11). With the exception of some classes (e.g. Maritime and Black pine), the distribution of agreements per class of species is rather similar.

2.5.2 Effect of the size of the reference sample

The size of the training sample dataset is known to be a key factor affecting both classification accuracy and predictive performance Congalton [1991]; Foody [2009b]. In practice,

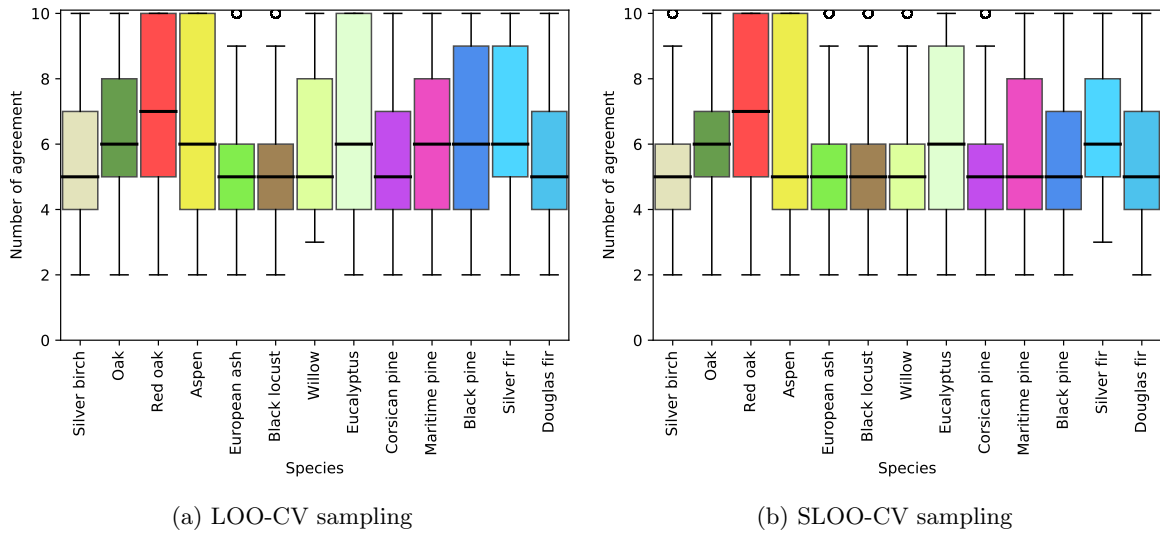


FIGURE 2.11 – Distribution per species of the number of agreements between the annual classifications using all the predicted pixels for both the LOO-CV and SLOO-CV sampling strategies.

it is hard to adequately judge the optimal training set size which depend on several factors such as the number of features, the degree of imbalance in class distribution, and the machine learning algorithm. In this study, we used the SVM classifier, which is known to be less sensitive to sample size since the decision boundaries rely on only a few support vectors. We also adjusted class weights to avoid bias due to the uneven distribution of tree species. Nevertheless, we obtained a slightly significant positive correlation ($r=0.52$; $p\text{-value}=0.06$) between the average number of pixels used for training (see Appendix A.4) and the average F1 score obtained for each species (including all the years), suggesting a potential effect of training set size on accuracy. We also observed that the least well-identified species were those with a limited number of forest stands (only three for Silver birch, European ash and Black pine). For these small classes, the presence of noise on the data (under detected clouds or cloud shadows under-detected, see below) may have a greater negative impact on their discrimination. However, Willow is an exception, as it was the least populated species with only 21 samples for training and a total of three forest stands but obtained the second highest F1 score (average = 0.86) behind Red oak (average = 0.87) with 118 samples (Appendix A.4)

More generally, due to the complexity of the learning problem (i.e. partial overlapping between spectro-temporal signatures of species in the feature space), increasing the size of the training sample and reducing the degree of imbalance in class distribution should improve the predictive performance [Sun et al. \[2009\]](#). In a recent study, [Bolgen et al. \[2018\]](#) obtained a high degree of accuracy (OA = 88.5%) when they classified 11 forest classes (including seven tree species) in the entire Belgian Ardenne ecoregion with only two Sentinel-2 dates. Although this statistical performance may be partially inflated by spatial autocorrelation, this greater accuracy could also be explained by the large sample size (from 2,589 to 7,068 pixels for each class with a minimum of 31 forest stands and a maximum of 64). An equivalent level of performance (OA = 88.2%) was obtained by [Persson et al. \[2018\]](#) when they identified five tree species in a forest in central Sweden using four S2 images acquired from spring to autumn. However, in their case study, the sample size was very close to ours (from 27 to 98 pixels per species). Spatial overfitting is suspected, as in our previous work [Sheeren et al. \[2016\]](#).

2.5.3 Effect of clouds and cloud shadows

When we compared the stability of predictions from one year to another (i.e. the map of agreements between the annual classifications) with the number of times the pixels were

affected by clouds or cloud shadows, we found no significant correlation. This indicates that disagreements between the classifications can not be attributed to this factor. We observed that the years most affected by clouds (2006 and 2007) had the lowest average OA values (52% and 48% respectively, see Table 2.2). However, for the other years, the F1 score per species was not always consistent with the extent of cloud coverage. For instance, in 2008, 11% of the reference pixels for Oak were affected by clouds (Appendix A.3) but the F1 score was the highest of all the years. This suggests that when clouds and cloud shadows are detected by the MACCS processing chain, the gap-filling approach is appropriate to correct noisy pixels. This procedure is currently used in the French production center THEIA (<https://theia.cnes.fr>) for Landsat, VEN μ S and Sentinel-2 level2A products.

An in-depth visual analysis of the map products in fact revealed that this disturbance factor (misdetection of clouds and shadows) had a major effect on classification performances. When the forests were partially affected by clouds and cloud shadows or when these were under detected (which is what happened in the case of slight fog), spectral signatures were skewed and confusion between species was likely. This issue is illustrated in Figure 2.12 which shows changes in the reflectance values of an Oak pixel in 2006. On most of the dates, the pixel was free of clouds and shadows (green dots). In some cases, clouds or cloud shadows were found and the pixel values were gap-filled by linear interpolation (orange dots). But on certain dates, clouds or cloud shadows were not detected (red dots) and this influences the spectro-temporal signature. These dates had erroneous values but also had a negative impact on nearby gap-filled values since the dates are considered to be valid (e.g., see the 10th image in Figure 2.12). Another example showing an erroneous spatial pattern in a forest stand due to undetected clouds is provided in Appendix A.3. This noise may influence the training step through the addition of inadequate support vectors, as well as the validation step, if the reference pixel to be tested is impacted by noise but the training pixels are not.

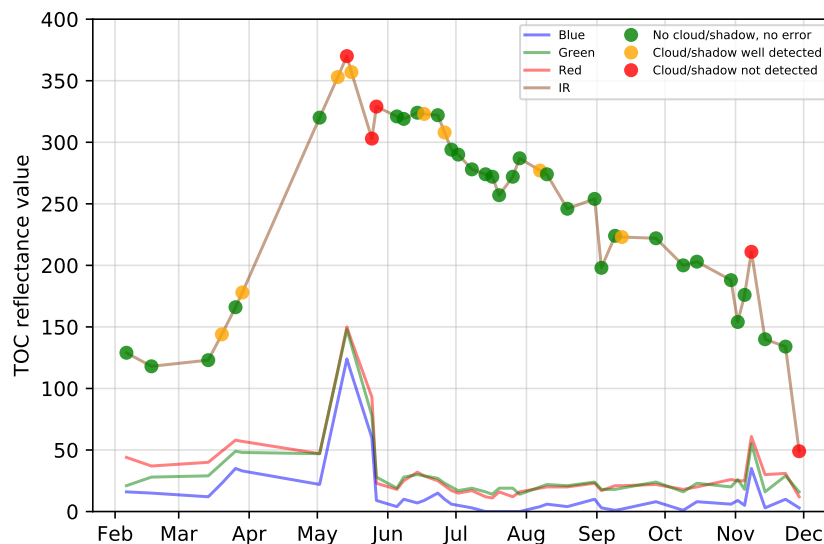


FIGURE 2.12 – Influence of undetected clouds in the time series, an example with the reflectance evolution of a Oak tree pixel in 2006. If the dot is red in the infrared time series, it means the pixel is in a cloud or in a shadow but this was not detected by the algorithm. So this pixel is taken as a valid pixel to be used to gap-fill nearest missing data.

An alternative to the gap-filling approach to reduce noise could be the use of smoothing methods applied to the whole time series and not only to a limited temporal window (i.e. the cloudy and shady pixels). Non-parametric methods such as Whittaker smoother Eilers [2003] or splines Hermance et al. [2007] may be appropriated. Another way to limit the effect of noise could be reducing the number of features Green et al. [1988]. Limiting the number of features in the classification protocol could help remove noise but also prevent the Hughes phenomenon

Hughes [1968] (i.e. a decrease in accuracy with the addition of new features after an initial increase). In theory, the Hughes effect should not be observed with the SVM classifier which is robust to the dimensionality of data (Melgani and Bruzzone [2004]). However, a previous study demonstrated a positive role of feature selection on classification results with SVM, particularly when the training set used is small Pal and Foody [2010].

2.5.4 Effect of the available dates in the SITS

We hypothesized (not verified here), that time series data improve the identification of tree species compared to the use of only one date or a couple of images due to phenological differences in spring or autumn. Using several dates should be better than using one date, whatever the year, as demonstrated in Hill et al. [2010]; Lisein et al. [2015]; Persson et al. [2018]. However, because of the variation in the dates available from one year to another, and the differences in cloud and shadow contamination, it is difficult to give a clear answer concerning the most appropriate time windows to separate species. We investigated the stability of tree species classifications over the nine consecutive years but the detailed analysis of the most informative dates was not part of the present study. Nevertheless, when we examined the relationships between the classification performances and the number of dates acquired during the key seasons, we found no evidence of a positive effect. For example, classification accuracy in the year 2008 was statistically equivalent to that in the year 2014 when there was comparable cloud coverage ($< 5\%$) whereas we only had one available image in spring 2008 versus five in 2014. Similarly, classification performance in the year 2009 was statistically identical to that in other years including in 2006bis. However, in autumn 2009, no images were available from end of October to the middle of December and only a very limited number of images were available in spring compared to 2006bis which had four dates in spring and six dates in autumn (Figure 2.2).

In a previous study, a combination of three aerial images acquired in spring (March 17th), summer (July 16th) and autumn (October 27th) provided the highest classification accuracy compared to all possible combinations based on five dates Hill et al. [2010]. When only one image was selected, autumn appeared to be the best period to distinguish between common Oak, English Oak, Field Maple, Silver birch, Aspen and small-leaved Elm. In other recent studies based on Sentinel-2 data, the optimal single date was found in May Bolyn et al. [2018]; Persson et al. [2018]. The same observation was made when discriminating deciduous tree species in a multitemporal image dataset acquired with an unmanned aerial system Lisein et al. [2015]. When more images were combined, the best datasets included data acquired in different seasons (spring, early summer and autumn) in accordance with Hill et al. [2010]. A similar analysis could be conducted in this study. However, since no ground reference data on the phenological events are available for the time period analyzed, it would be difficult to draw robust conclusions about the tangible contribution of seasonal variations in species discrimination.

2.5.5 Differences between species

We found more difficulty to separate conifers than broadleaf species as previously highlighted in other studies when multitemporal data are used Hościło and Lewandowska [2019]; Immitzer et al. [2012]; Pasquarella et al. [2018]; Persson et al. [2018]; Sheeren et al. [2016]. Seasonal changes are less pronounced which induces higher overlapping between spectro-temporal profiles. Pine species were the harder to identified (in particular Corsican and Black pines). Among broadleaf species, European Ash and Silver birch were the most confused (contrary to Immitzer et al. [2012]). The best agreements were obtained for Red Oak, Eucalyptus, Willow and Aspen, in line with Lisein et al. [2015] for the latter species.

The evergreen phenology of Eucalyptus explains its high rate of agreements among 9 years despite a medium F1-score (average of all years OA = 65%). Because of differences in

morphological and anatomical traits, optical properties also differ from those of conifers. The Red oak phenology is also specific, in particular in autumn when leaves turn red due to the production of anthocyanins. This gives a spectral characteristic which help to recognize them among the other species. For Willow, stands are located in well-suited humid areas sometimes waterlogged. The variable moisture conditions associated with a partially recovering canopy provide them a weaker reflectance in the near-infrared band compared to the other broadleaf species which may explain the good classification accuracy. For other species that cannot easily be separated, various factors may be involved, in addition to close spectral signatures. Forest managing practice is one of them. Stand age, density and the existence of understory vegetation are others. Spectral disparity for a given species (intra-species variability) may also influence the classification [Lisein et al. \[2015\]](#).

2.6 Conclusion

This study based on temperate forests in France is the first to explore the stability of tree species classification over nine consecutive years using dense high spatial resolution SITS with spatially uncorrelated validation data. The study was based on surface reflectance products derived from Formosat-2 optical time series acquired at irregular intervals from 2006 to 2014. Despite close statistical results in terms of classification accuracy, we observed high spatial disparities from one year to the next reflecting the moderate ability to predict tree species at the pixel level because of various disturbing factors.

Based on our findings, several conclusions can be drawn :

1. Spatial autocorrelation within validation data drastically overestimates the classification accuracy. In our context, an average optimistic bias of 0.4 of OA is observed when spatial dependence remained (LOO-CV strategy vs SLOO-CV). In further studies, we recommend adapting the data-splitting procedure to systematically reduce or eliminate spatial autocorrelation in the validation set in order to provide more robust conclusions about the true predictive performance.
2. Noise in the time series (i.e. undetected clouds and shadows) affects the SVM based classification performances. Despite accurate masks of clouds and shadows and a gap-filling approach to correct invalid pixels, residual noise impacts the learning and prediction processes.
3. The monospecific broadleaf plantations of Aspen, Red Oak and Eucalyptus are the easiest to classify. Conifers are the most difficult. The lowest accuracy was obtained for Silver birch, European ash and Black pines for which only a few forest stands were available.

Perspectives of this study are twofold. The first one is the use of S2 time series to confirm the results and assess the contribution of additional spectral bands such as the red-edge to separate tree species for the same in situ dataset and area. With its 5-day revisit time, S2 provides many more data in one year. These new time series should help better identify the best combination of multitemporal images and check that the combination is consistent with phenological events of the tree species concerned. Work is in progress to collect ground phenological observations on the study site. S2 also offers the possibility to work at a larger scale and will thus give us more reference pixels to reduce the bias due to the spatial autocorrelation. The second is related to the Formosat-2 time series. Annual datasets could be combined to reconstruct a synthesized multiyear time series based on all cloud-free images to combine all the phenological events of the species into a one representative year.

Author Contributions Conceptualization, N.K. and D.S. ; Data curation, N.K., D.S. and J. W. ; Methodology, N.K., D.S. and M.F. ; Investigation, N.K. ; Validation, N.K. and D.S. ;

Visualization, N.K. ; Software, N.K. ; Funding acquisition, D.S. ; Supervision, D.S., C.M. and J.-F. D. ; Writing - original draft, N.K., and D.S. ; Writing - review & editing, J.-F. D., M.F. and C.M. ;

Funding N.K. received a PhD scholarship from the French Ministry of Higher Education and Research (University of Toulouse). Field work was carried out as part of TOSCA OSO and TOSCA HyperBIO projects funded by the French Space Agency CNES.

Conflict of Interest The authors declare no conflict of interest. The funders had no role in the design of the study ; in the collection, analyses, or interpretation of data ; in the writing of the manuscript, or in the decision to publish the results.

Supplementary Materials : An open access to the tree species reference samples will be provided in the final version of the manuscript with a digital object identifier (DOI). We are also preparing an interactive web map of the classification results.

Acknowledgments : Special thanks to O. Hagolle and M. Huc from CESBIO Lab. for providing the pre-processed Formosat-2 time series with masks of clouds and cloud shadows.

Chapitre 3

Étude de l'impact de l'autocorrélation spatiale

“ Everything is related to everything else, but near things are more related than distant things.”

— Waldo Tobler.

Table des matières

3.1	Introduction	62
3.2	Related works	64
3.2.1	Spatial autocorrelation in model residuals	64
3.2.2	Spatial autocorrelation between training and test sets	65
3.3	Experimental protocol	66
3.3.1	Spatial and non-spatial data-splitting strategies	66
3.3.2	Classification algorithm	68
3.3.3	Image dataset	69
3.3.4	Reference dataset	70
3.3.5	Experimental setup	70
3.4	Results	71
3.4.1	Prediction errors according to whether spatial/non-spatial cross-validation strategies we used	71
3.4.2	Prediction errors on other spatially independent and distant sites	73
3.5	Discussion	76
3.6	Conclusion	78

Le Chapitre 2 a mis en avant le biais potentiel de l'autocorrélation spatiale dans l'estimation de la qualité des résultats des cartographies. Afin d'étudier en détail ce phénomène, la taille de l'échantillon et de la zone d'étude a été significativement augmentée. Pour cela, une tuile Sentinel-2 recouvrant plusieurs départements a été sélectionnée afin d'étudier l'impact de l'autocorrélation spatiale dans la cartographie des feuillus et des conifères.

La première partie de ce chapitre explique la problématique liée à l'autocorrélation spatiale de manière générale en listant des travaux se servant de ce phénomène (que ce soit pour améliorer les prédictions ou pour séparer les données d'entraînement ou de validation). Ensuite le protocole mis en place est détaillé et avec lui l'ensemble des six méthodes de validations croisées utilisées, allant de la moins restrictive à la plus contraignante au niveau de l'autocorrélation spatiale. Par la suite les résultats sont présentés, de manière globale, par méthode de validation mais aussi selon le nombre d'échantillons. Les résultats seront comparés avec deux jeux de données de validation indépendants spatialement et jamais vus par

l'algorithme. La discussion explique les comportements des validations et met en perspective les résultats avec d'autres travaux du domaine. Enfin, les conclusions des différentes méthodes sont tirées et avec elles des recommandations pour de futurs travaux.

Spatial dependence between training and test sets : another pitfall of classification accuracy assessment in remote sensing

Abstract

Spatial autocorrelation is inherent to remotely sensed data. Nearby pixels are more similar than distant ones. This property can help to improve the classification performance, by adding spatial or contextual features into the model. However, it can also lead to overestimation of generalisation capabilities, if the spatial dependence between training and test sets is ignored. In this paper, we review existing approaches that deal with spatial autocorrelation for image classification in remote sensing and demonstrate the importance of bias in accuracy metrics when spatial independence between the training and test sets is not respected. We compare three spatial and non-spatial cross-validation strategies at pixel and object levels and study how performances vary at different sample sizes. Experiments based on Sentinel-2 data for mapping two simple forest classes show that spatial leave-one-out cross-validation is the better strategy to provide unbiased estimates of predictive error. Its performance metrics are consistent with the real quality of the resulting map contrary to traditional non-spatial cross-validation that overestimates accuracy. This highlights the need to change practices in classification accuracy assessment. To encourage it we developed Museo ToolBox, an open-source python library that makes spatial cross-validation possible.

Authors : Nicolas Karasiak¹, Jean-François Dejoux², Claude Monteil¹ and David Sheeren¹.

¹ : DYNAFOR, Université de Toulouse, INRAE, 31326 Castanet-Tolosan, France.

² : CESBIO, Université de Toulouse, CNES/CNRS/INRAE/IRD/UPS, 31401 Toulouse, France.

Paper status : Accepted with minor revision in Machine Learning on August 13, 2020.

3.1 Introduction

The evaluation of classification accuracy has always been considered as an important issue in the remote sensing community [Congalton, 1991; Foody, 2008, 2002; Ye et al., 2018]. There is a large body of literature on this subject with precise recommendations for designing and implementing robust accuracy assessments based on reference data [Olofsson et al., 2014; Stehman and Foody, 2019]. Significant advances have been made in defining protocols that match the objective of the quality assessment. This may concern the sampling design (stratified or not, systematic or random) [Congalton, 1998; Ramezan et al., 2019; Stehman, 2009], the size of the sample [Chen and Wei, 2009; Foody, 2009a], the spatial unit of reference samples (pixel, blocks, objects) [Stehman and Wickham, 2011], or the accuracy parameters (with their pitfalls) to compute from error matrix [Foody, 2020b; Liu et al., 2007; Pontius and Millones, 2011; Stehman and Foody, 2019]. Some authors also explored how to account for and represent the spatial distribution of classification errors in order to highlight its non-uniformity and provide additional insights for map users [Comber et al., 2012; Foody, 2005;

Khatami et al., 2017; McIver and Friedl, 2001].

The standard approach used to assess the classification accuracy is to split the reference data into two subsets. The first set is used to train the classification model (learning step). The second set is used to test the model and estimate the prediction errors (validation step). The test set is never used to build the model. This ensures independence between the training and test sets and provides a generally accepted estimate of the predictive power of the model.

Most often, data-splitting is based on simple, possibly stratified, random selection. This selection is sometimes repeated (resampling by bootstrapping) to compute the sampling variability of the accuracy metrics [Lyons et al., 2018]. A common alternative to this traditional *hold-out validation* is assessing the classification accuracy by *cross-validation* (CV). In this approach, reference data are also split into subsets but the number of subsets can vary (k -fold). The model is trained iteratively on $k - 1$ subset(s) and tested on the remaining set. A accuracy is then measured by averaging the performance values computed on each subset (k models). Leave-one-out (LOO) distribution is a special case of cross-validation where $k = n$ (n being the size of the reference data set). In this case, each test set is equal to 1 and the model is trained n times.

Despite the importance attached to the accuracy assessment protocol, a gap, sometimes a serious one, is often found between the performance metrics of a model and the real quality of the resulting map. This tends to reduce the confidence placed in accuracy statistics and to discredit the true capacity of remote sensing in the opinion of the end-users. Among the factors involved in this optimistic bias [Foody, 2020b; Stehman and Foody, 2019], an important one is the spatial dependence between the training and test sets. The spatial context is often ignored in the evaluation even though it compromises the required independence of the data. Because of spatial autocorrelation, spectral values of close pixels are often more similar than those of distant ones, producing falsely high accuracy metrics if the sampling design is not used for testing [Meyer et al., 2019; Roberts et al., 2017; Schratz et al., 2019]. This dependence exists in both the spectral features (i.e. the predictor variables) and the class to predict (i.e. the response variable). A good illustration of this spatial dependence, reported by [Inglada, 2018], was found in the TiSeLaC land cover classification contest held during the 2017 ECML/PKDD conference. The winner proposed a very simple classification approach to predict the nine most important land cover classes from time-series images. No individual spectral features were used to train the model, but only the pixel geographical coordinates, leading to weighted F-scores ranging from 0.90 to 0.98. The classification was based on k-NN and the validation was carried out by cross-validation [Sergey, 2017]. The runner-up also implicitly exploited the spatial autocorrelation through convolutional neural networks. The accuracy of predictions was estimated at 0.99 of the F-score [Di Mauro et al., 2017].

The existence of spatial autocorrelation is well-known in the remote sensing community [Wulder and Boots, 1998]. Twenty years ago, Congalton already analysed the pattern of errors found in land-cover classifications and recommended correcting the sampling scheme in the case of non-random distribution [Congalton, 1998]. However, despite collective awareness, spatial dependence is often ignored in classification accuracy assessment even though a number of approaches have been proposed to prevent it. This leads to systematic overestimation of generalisation capabilities due to spatial overfitting.

The objectives of this paper are thus to (i) review existing approaches that deal with spatial autocorrelation for image classification in remote sensing, (ii) evaluate the impact of spatial autocorrelation between training and test sets on classification performances and (iii) investigate how performances vary with to the data-splitting strategy used for reference samples. The main contribution of this paper is demonstrating the importance of bias in accuracy metrics when spatial independence between the training and test sets is not respected, and hence the need to change accuracy assessment practices. We compare three spatial and non-spatial strategies at pixel and object levels using different sized training samples. The

experiments are conducted using a large number of reference samples composed of millions of pixels. Generalization capabilities of the models are evaluated through cross-validation at one site and hold-out validation at two other sites. Our assumption is that predictive power is inflated when the test set is used in the spatial domain of the training set.

3.2 Related works

One of the challenges of analysing spatial data is dealing with the interdependence between location and the value of the processes to be investigated [Anselin, 1989]. Spatial processes are distance-related, thus leading to spatial structures in the autocorrelated data. From a statistical point of view, two problems can arise with spatial autocorrelation : 1) spatial non-independence of the classification errors (or model residuals) and 2) spatial non-independence of the training and test sets used for accuracy assessment.

3.2.1 Spatial autocorrelation in model residuals

The first problem arises when the predictor variables are not able to perfectly account for the effect of the spatial structure to estimate the response variable [Roberts et al., 2017]. This is a major concern in certain disciplines like ecology, particularly in biogeographical analyses and species distribution modeling [F. Dormann et al., 2007; Kühn and Dormann, 2012; Miller et al., 2007; Roberts et al., 2017]. If parametric models are used to make inferences, the spatial autocorrelation of the prediction errors may lead to erroneous conclusions [Dormann, 2007; Kühn, 2007]. Models are built not only to predict, like in remote sensing, but also to explain the respective effect of each variable. Therefore, if the independence of the residuals (as assumed in standard regression techniques) is violated, estimation of the model parameters may be biased and the type-I error rate (i.e. incorrect rejection of the null hypothesis) may increase [F. Dormann et al., 2007; Kühn and Dormann, 2012]. This explains why in spatial ecology, testing spatial independence of residuals is becoming a standard practice.

Several approaches have been proposed to address this issue [Beale et al., 2010; F. Dormann et al., 2007; Miller et al., 2007]. The simplest is incorporating additional predictor variables to improve the model specification. If the spatial pattern of the response variable is fully reflected by the extra autocorrelated predictors, the residuals should not be spatially autocorrelated and a non-spatial model can be used [Kühn and Dormann, 2012]. Otherwise, spatial models are required. In this case, the spatial dependence is explicitly incorporated in the model, from the response variable itself or from the residuals, like in autocovariate or autoregressive models [Anselin, 1988]. Spatial eigenvector mapping is an alternative [Dray et al., 2006] like kriging and other geostatistical techniques (see [Beale et al., 2010; F. Dormann et al., 2007; Miller et al., 2007] for more details).

Spatial models have also been introduced in remote sensing. A widespread approach is the Markov random field (MRF) model that incorporates local information (the neighbors class labels) in the classification process [Magnussen et al., 2004; Shekhar et al., 2002; Solberg et al., 1996]. Spatial autologistic regression models have also been considered [Koutsias, 2003; Mallinis and Koutsias, 2008; Shekhar et al., 2002]. However, the most common approach in the field remains the use of non-spatial models. The spatial information is incorporated in contextual classifiers through additional predictors [Ghamisi et al., 2018; Wang et al., 2016]. A variety of techniques exist are available including morphological image analysis [Fauvel et al., 2013] and textural analysis [Franklin et al., 2000; Puissant et al., 2005; Sheeren et al., 2009], sometimes based on variograms [Atkinson and Lewis, 2000; Berberoglu et al., 2007] or on a local version of principal component analysis [Comber et al., 2016]. The adaptation of classifiers has also been proposed using the spatial locations of the training samples [Atkinson, 2004] or local spatial patterns [Bai et al., 2020] to estimate class probabilities. Other authors suggest using local spatial statistics [Ghimire et al., 2010; Myint et al., 2007] or interpolated spectral values and their degree of similarity with actual values to improve the classification

[Johnson et al., 2012]. Today, deep learning has become the most powerful alternative way to incorporate spectral-spatial features for classification [Ghamisi et al., 2018; Zhao and Du, 2016]. A non-exhaustive summary of these representative approaches is provided in Table 3.1.

Despite the wide range of technical options, the primary objective of these image processing methods is not to deal with spatial autocorrelation but improve the classification performance. Spatial autocorrelation is considered as an opportunity. However, the effect is the same : the spatial dependence of classification errors is reduced since the model specification is improved. Only a few works have specifically aim to control the spatial dependence between observations in the modeling process [Rocha et al., 2019].

3.2.2 Spatial autocorrelation between training and test sets

The second issue, which is the focus of this paper, is spatial dependence between the training and test sets. If the classification model is calibrated using training data that are spatially correlated with the data used for testing, predictions based on testing data will be unrealistically high and will not reflect the true predictive power of the model because of spatial overfitting [Meyer et al., 2019; Roberts et al., 2017; Schratz et al., 2019]. When the model is intended to classify new sites (assuming stationarity in the relationship between predictor and response variables across space), spatially independent data are required for the estimation of unbiased predictive performance and to test generalization capabilities within the image. While optimistic biases in the predictive power have already been discussed [Chen and Wei, 2009; Geiß et al., 2017; Hammond and Verbyla, 1996; Meyer et al., 2019; Millard and Richardson, 2015], this effect is often disregarded when image classification is evaluated, probably for the sake of simplicity [Ramezan et al., 2019] and because of its importance is underestimated.

One possible way to address this issue is to spatially segregate the training and test sets during the data-splitting procedure. This can be done by imposing a spatial stratification to select the samples, either by objects [Cánovas-García et al., 2017; Inglada et al., 2017], blocks [Lyons et al., 2018; Meyer et al., 2019; Roberts et al., 2017; Valavi et al., 2019], or clusters [Brenning, 2012; Schratz et al., 2019] (Table 3.1). Blocks can be defined arbitrarily (e.g. grid of space) or based on predefined similar characteristics. Clusters can be regular or irregular in shape (depending on the clustering techniques) and can be contiguous or disjointed, with regular or irregular spacing. Another strategy is to define distance-based buffers around the hold-out samples to be sure the learning model is only based on spatially independent data [Le Rest et al., 2014] and disjointed spectral and spatial features [Geiß et al., 2017]. The distance (related to the degree of dependency) is often fixed arbitrarily but can also be defined using the correlogram or semivariogram based on the response variable or on the predictors [Pohjankukka et al., 2017; Valavi et al., 2019]. The advantage of the latter approach over the other spatial partitioning methods is that the spatial autocorrelation is explicitly quantified. In the other cases, the spatial dependence between training and test sets is assumed to have been removed but this is not checked, meaning residual spatial dependence may persist.

In this paper, we use a distance-based buffer approach that relies on Moran's I statistics to segregate spatially referenced samples and to explicitly estimate their degree of dependence. To our knowledge, this robust approach based on correlogram has not yet been evaluated on remotely sensed data using a large reference dataset. In a previous work, we applied this buffering strategy to classify tree species in satellite image time series [Karasiak et al., 2019]. However, because the reference dataset was small, we were unable to disentangle the effects of spatial autocorrelation and training set size on classification performances.

TABLEAU 3.1 – Non exhaustive list of representative references.

Issue 1 : spatial dependence of prediction errors (model residuals)

Cause : the effect of the spatial dependent structure on the response variable is not perfectly explained by the predictor variables.

Consequence : violation of the assumption of independence of error terms in regression models, inflation of the probability of type-I error in hypotheses tests, bias in parameter estimation, limited predictive performance with *salt and pepper* noise.

Methodological solution : use spatial models that incorporates spatial dependence in classification/-prediction models; includes additional spatial-related predictors in non-spatial models.

Existing approaches**References**

- | | |
|---|--|
| <ul style="list-style-type: none"> ○ Markov random field model ○ Spatial autoregressive model ○ Spatially weighted classification ○ Spatially adapted PCA ○ Spectro-spatial classification based on : <ul style="list-style-type: none"> ○ Textural features ○ Morphological filtering ○ Local spatial statistics ○ Deep learning | <p>Magnussen et al. [2004]; Shekhar et al. [2002]; Solberg et al. [1996]</p> <p>Koutsias [2003]; Mallinis and Koutsias [2008]; Shekhar et al. [2002]</p> <p>Atkinson [2004]; Johnson et al. [2012]</p> <p>Comber et al. [2016]</p>
<p>Atkinson and Lewis [2000]; Franklin et al. [2000]; Puissant et al. [2005]</p> <p>Fauvel et al. [2013]; Ghamisi et al. [2018]</p> <p>Ghimire et al. [2010]; Myint et al. [2007]</p> <p>Ghamisi et al. [2018]; Zhao and Du [2016]</p> |
|---|--|

Issue 2 : spatial dependence between training and test sets

Cause : nearby reference samples are correlated.

Consequence : spatial overfitting leading to falsely inflated accuracy assessment, inconsistent with the true predictive power of the model.

Methodological solution : use spatial data-splitting technique to make training and test sets spatially independent

Existing approaches**References**

- | | |
|---|--|
| <ul style="list-style-type: none"> ○ Spatial k-fold/leave-one-out cross-validation : <ul style="list-style-type: none"> ○ Object-based partitioning ○ Block*-based partitioning ○ Cluster-based partitioning ○ Buffer**-based partitioning | <p>Cánovas-García et al. [2017]; Inglada et al. [2017]</p> <p>Lyons et al. [2018]; Meyer et al. [2019]; Roberts et al. [2017]; Valavi et al. [2019]</p> <p>Brenning [2012]; Schratz et al. [2019]</p> <p>Geiß et al. [2017]; Le Rest et al. [2014]; Pohjan-kukka et al. [2017]; Valavi et al. [2019]</p> |
|---|--|

*Blocks are defined arbitrarily (e.g. grid of space) or based on thematic characteristics or clusters with regular/irregular shape and regular/irregular spacing.

**A distance-based buffer is defined around hold-out testing sample. The distance (related to the degree of dependency) is fixed arbitrarily or measured on a correlogram or variogram.

3.3 Experimental protocol

3.3.1 Spatial and non-spatial data-splitting strategies

To assess the impact of spatial autocorrelation between the training and test sets, we defined a supervised classification protocol with three cross-validation (CV) sampling strategies for performance evaluation : (1) a k -fold cross-validation (k -fold-CV) based on random splitting, (2) a non-spatial leave-one-out cross-validation (LOO CV) and (3) a spatial leave-one-out

cross-validation (SLOO CV) using a distance-based buffer relying on Moran’s I statistics.

We chose 2 folds (50/50; $k = 2$) for the k -fold-CV sampling. For the LOO CV and SLOO CV, we systematically selected one sample from each class to test at each iteration (i.e. $n-1$ per class for training) in contrast to the conventional approach that selects only one sample for testing whatever the class (i.e. no stratification per class). The number of folds is equal to n_{min} , the size of the class with the fewest samples.

Data splitting was performed at both *pixel* and *object* levels (Fig. 3.1). However, at the pixel level, sampling was stratified. We imposed the object as spatial stratum to sample pixels. Thus, for the pixel-based k -fold-CV strategy, after random selection of a fixed number of objects, 50% of pixels belonging to the objects were used for training and the rest for testing (Fig. 3.1a). Pixels were sampled randomly within the selected objects, with no constraints on the sample size per object. In the object-based k -fold-CV strategy (Fig. 3.1d), all the pixels belonging to the objects were used for training or testing. This object-based sampling can be viewed as a first option to account for spatial autocorrelation, without quantify it. Test samples are forced to be more spatially distant than training ones. For the remaining LOO CV and SLOO CV strategies, only one pixel (Fig. 3.1b-c) or object with all its related pixels (Fig. 3.1e-f) was used to estimate prediction error at each iteration. In SLOO CV, pixels

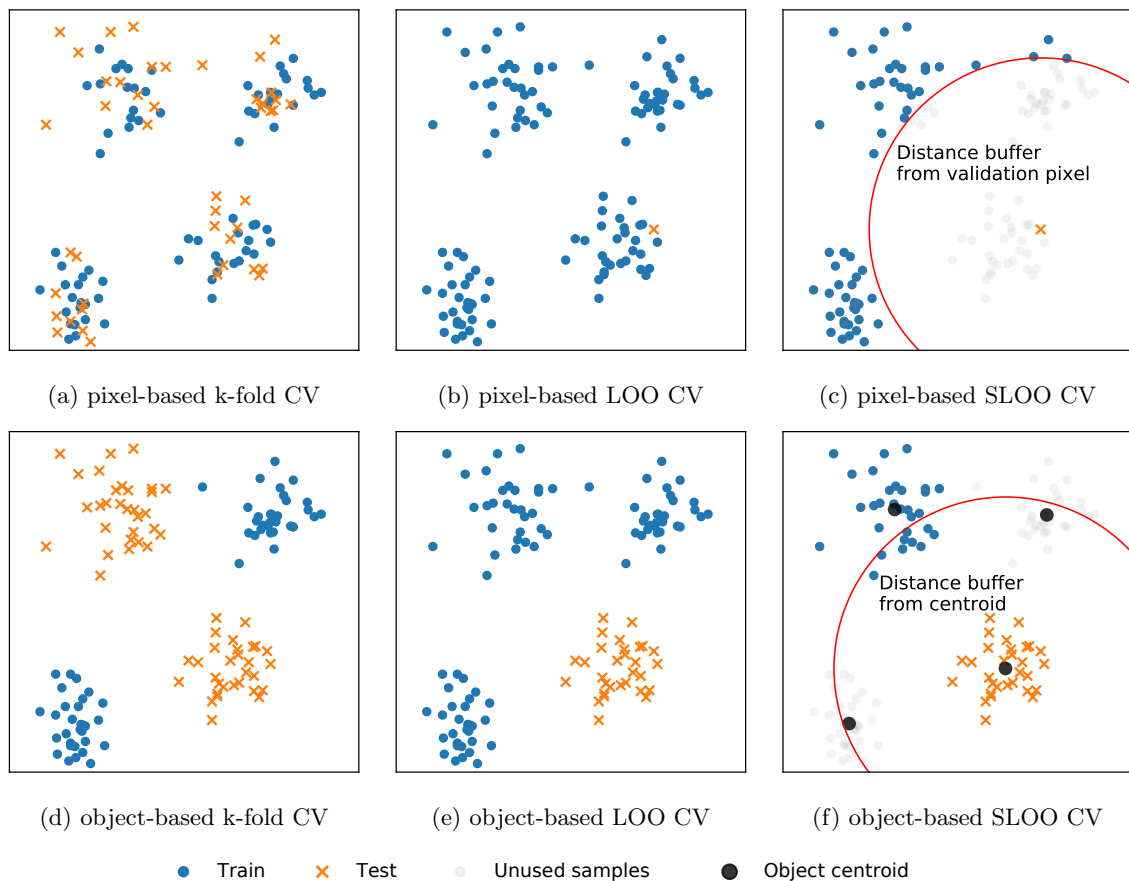


FIGURE 3.1 – Spatial and non-spatial CV strategies to investigate the effect of spatial autocorrelation between training and test sets. Clusters of points represent objects (spatial polygons in the reference data set) with their related pixels. At the pixel level (a,b,c), data-splitting is stratified using the object as a spatial stratum. At the object level (d,e,f), all the pixels belonging to the objects are either used for training or for testing. Pixels and objects are sampled randomly. For the k -fold CV strategies, $k = 2$ (i.e. split of 50/50%). In the SLOO CV (c,f), a distance-based buffer related to the degree of spatial dependency is defined around the test sample(s). Pixels that are spatially correlated with the test sample(s) are removed from the training set.

spatially correlated with the single test sample were removed from the training set (Fig. 3.1c-f). The spatial dependence between nearby pixels was estimated using Moran's Index (Moran's I) defined as the ratio of the covariance between neighboring pixels and the variance of all pixels in the image :

$$I(d) = \frac{n}{S_0} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{i,j} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (3.1)$$

where, x_i is the pixel value of x (a spectral band of the image) at location i , x_j is the pixel value of x at location j (a nearby pixel of i), \bar{x} is the average value of x , n is the number of pixels in the image, $w_{i,j}$ is the weight equals to 1 if pixel j is within distance of d of pixel i , otherwise $w_{i,j} = 0$, and S_0 the sum of all $w_{i,j}$'s :

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{i,j} \quad (3.2)$$

More simply, Moran's I expresses the correlation between the pixel value at one location and its close observations. The index ranges from -1 to +1. For positive values, the nearby pixels covary together (positive spatial autocorrelation). For negative values, the nearby pixels covary in the opposite direction (negative spatial autocorrelation). Values close to 0 indicate the absence of spatial autocorrelation (random spatial distribution).

Moran's I was computed for each spectral band available in the images we used, from the reference pixels, and for neighborhoods containing from 1 to 2700 pixels. Then, based on correlogram (a plot of Moran's I as a function of distance), we identified the threshold distance beyond which the spatial dependence between pixels is insignificant. Finally, this threshold was used to define the buffer radius in the SLOO CV strategy. At the pixel level, all the reference samples located in the buffer area centered on the test sample were excluded from the training set (Fig. 3.1c). At the object level, the exclusion of nearby pixels depends on the position of the centroids of the related objects (Fig. 3.1f).

Because the effect of spatial autocorrelation may vary with the size of the sample [Chen and Wei, 2009], we increased the number of labeled objects (with its related pixels) progressively in the learning experiments, as follows :

- from 3 to 10 labeled objects : incrementation per 1 object.
- from 10 to 100 labeled objects : incrementation per 10 objects.
- from 100 to 1000 labeled objects : incrementation per 100 objects.

The data-splitting procedure was repeated 10 times for all the strategies because we used random undersampling to investigate multiple training set sizes.

3.3.2 Classification algorithm

We used the nonparametric random forest (RF) learning algorithm to train the models [Breiman, 2001]. The ability of this algorithm to distinguish land cover classes at large scale despite the limited sensitivity of the parameter values in the classification performances has already been widely demonstrated [Pelletier et al., 2016; Rodriguez-Galiano et al., 2012]. RF is also known to be robust to outliers and faster than other classifiers such as the support vector machine algorithm.

In our experiments, we set the number of trees was at 200. The number of variables used to split a node was kept at the default value ($m = \sqrt{p}$ with p , the total number of features), as well as the stopping criteria for tree building (nodes are expanded until all leaves are pure or contain less than 2 samples). To mitigate the effect of the imbalanced distribution of the classes, we modified class weights in order to prevent bias due to the dominant class. By

default, all classes have a weight equal to one. Here, weights were adjusted inversely to the proportion of the class frequency. All the models were fitted with the same hyperparameters. Spectral features were also standardized (i.e. centering and scaling to unit variance) prior to training.

To evaluate the classification performances, we deliberately abandoned the kappa index [Foody, 2020b]. We focused our attention on producer’s accuracy per class (also known as specificity and sensitivity with two classes), in addition to overall accuracy (OA). User’s accuracy was not considered since the classifications were not intended to be used. We only report the OA as accuracy estimator in the results section. However, to control for a possible interpretation bias of OA, which is prevalence-dependent (i.e. the most prevalent class dominates the indicator value), the class-specific performances are also provided in the appendix. The average value based on 10 repetitions was computed with standard deviation as confidence interval.

We used the *scikit-learn* python library for implementation [Pedregosa et al., 2011]. The spatial and non-spatial sampling strategies were performed using *Museo ToolBox*, a python library we developed to make this research reproducible [Karasiak, 2020].

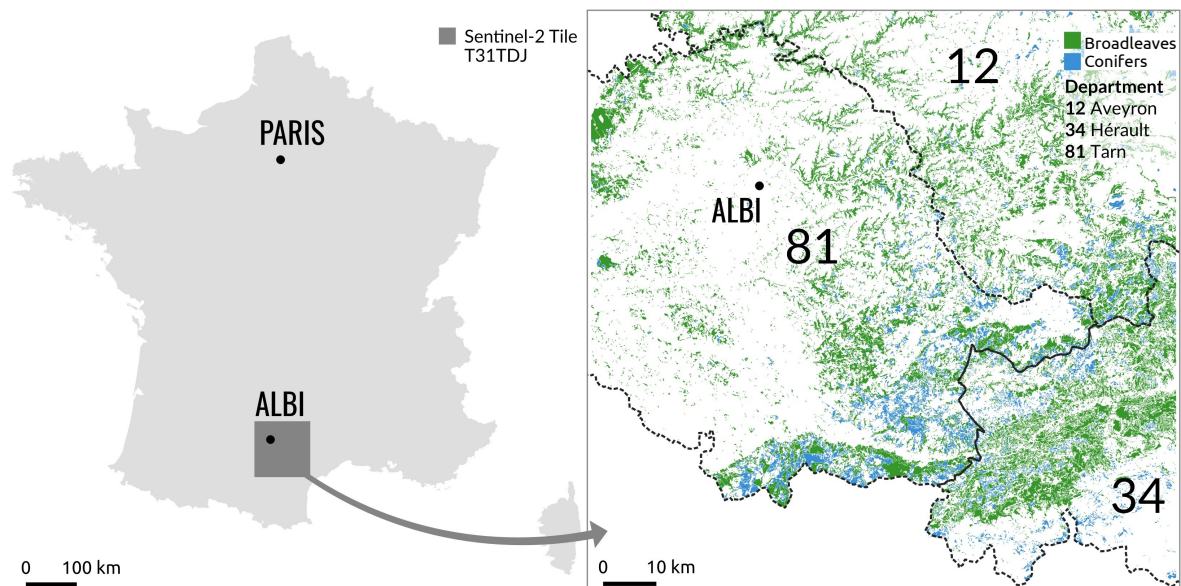


FIGURE 3.2 – Location of the 31TDJ Sentinel-2 tile in the south of France (100 km x 100 km in extent). The experiments were conducted on data originating from the three administrative departments (Hérault 34, Tarn 81, Aveyron 12).

3.3.3 Image dataset

We used Sentinel-2 (S2) optical images to test the experimental protocol and in particular, the 31TDJ S2 tile (100 km x 100 km in extent). This tile is located in the South of France and partially covers four French administrative departments, and includes the city of Albi (Fig.3.2). In practice, only three of departments were classified : Hérault-34, Tarn-81 and Aveyron-12. The effect of spatial autocorrelation was investigated by mapping two simple forest classes (coniferous and broadleaf) that can include different tree species.

The dataset includes four S2 images acquired on July 17, August 8, September 5 and October 15, 2016. We did not select images covering all four seasons, but gave priority to images with less than 5% cloud covers. The image acquisition dates are not ideal to distinguish coniferous and broadleaf stands. However, we focused on the relative differences in accuracy between the data-splitting strategies rather than on the ability to properly distinguish the two forest classes. The S2 data were downloaded from the French national Theia platform at level 2A (i.e. top-of-canopy reflectances corrected for atmospheric and topographic effects

[Hagolle et al., 2015]). We only used a subset of the available spectral bands : Blue (B2 - 490 nm), Green (B3 - 560 nm), Red (B4 - 665 nm) and Near Infra-Red (B8 - 842 nm) at 10-m spatial resolution. The cloudy pixels in the dataset were corrected using a gap-filling approach (linear interpolation) and a mask of clouds produced using the MAJA pre-processing chain of the S2 time series [Baetens et al., 2019].

3.3.4 Reference dataset

The reference dataset (labeled pixels) for the two classes of forest was derived from the French National Forest Inventory spatial database (IGN BDForet[®], v.2). This source provides a vector map of forest stands with a minimum area of 0.5 hectares. The composition of each stand is obtained from aerial stereo-image interpretation completed by field surveys. A stand is considered as pure if the proportion of the area covered by one species equals or exceeds 75% of the total area, otherwise, it is defined as a mixed stand. Here, we selected pure stands of coniferous and broadleaf species for the experiments. All mixed stands were excluded. Because of a temporal gap between the S2 images (2016) and the reference forest maps (dating from 2006 to 2015 depending on the department), we updated the forest maps to eliminate forest stands whose land cover had changed. This was done by masking S2 pixels with NDVI values lower than 0.4 using the image of July 17. We also eroded the vector forest layer with an inside buffer of 20 m to avoid selecting pixels of mixed classes at the boundaries of the masks. Finally, forest stands covering less than one hectare were removed to insure we kept a dense reference dataset that would enable us to explore different sampling strategies. The total number of forest stands and related pixels are given in Table 3.2. The forest stand is the spatial stratum used to sample pixels in pixel-based strategies and was the sample unit for the object-based ones.

TABLEAU 3.2 – Number of pixels and forest stands in each class (coniferous and broadleaves) in each administrative department in the study site.

Name of the départements		Classes	
		Broadleaf spp.	Coniferous spp.
Herault-34	pixels	4,257,112	1,119,079
	stands	2,841	1,972
	area (km ²)	425.7	111.9
Tarn-81	pixels	7,046,056	2,216,920
	stands	4,210	2,002
	area (km ²)	704.6	221.7
Aveyron-12	pixels	5,485,172	836,075
	stands	3,092	966
	area (km ²)	548.5	83.6

3.3.5 Experimental setup

The six cross-validation strategies with different sample sizes were applied to the department of Herault-34. The maximum sample size was set at 1,972 labeled objects corresponding to the number of coniferous stands in this site. However, it was not possible to investigate all the sample sizes for the pixel-based LOO CV and SLOO CV strategies. The maximum number of sampled objects was limited to 10 (i.e. $\approx 32,000$ pixels for broadleaf spp and 5,476 for conifer sp, on average) due to long computation time. Because the number of iterations increases with the sample size in cross-validation, the time required to compute the distance matrix among pixels in Moran's I rapidly becomes prohibitive.

We used reference samples in the neighboring departments, Tarn-81 and Aveyron-12, as additional independent datasets to validate the *full model* of Herault (i.e. the model with the

maximum training set size of 1972 objects per class). This extra validation on spatially distant sites was carried performed after removing all reference samples in Tarn-81 and Aveyron-12 that were spatially correlated with the Hérault-34 ones. We also removed some reference samples from the over-represented class of broadleaf spp through random under-sampling (30 repetitions). In this way, the accuracy metrics in the extra validation were computed with the same number of samples per class (369,164 pixels) in both departments Tarn-81 and Aveyron-12.

Combining the three CV strategies at both pixel and object levels for the different sized samples, means a total of 1240 classifications were trained and tested on Hérault-34 department (i.e. 27 sets of samples from 3 to 1972 stands x 2 level of analysis x 3 CV strategies x 10 repetitions, including exceptions in sample size for the pixel-based LOO and SLOO CV). Full spatial independence between the training and test sets was ensured in two configurations of the protocol : using the SLOO CV strategy on Hérault-34, and during validation on Tarn-81 and Aveyron-12. The pixel-based k -fold and LOO CV are the sampling strategies that should be most affected by spatial dependence. The object-based k -fold and LOO CV are in an intermediate position.

3.4 Results

3.4.1 Prediction errors according to whether spatial/non-spatial cross-validation strategies we used

The classification performances obtained by cross-validation on Hérault-34 are given in Fig. 3.3. Overall, the results show that ignoring dependence between training and test sets leads to very high accuracy metrics whatever the sample size. This is particularly clear at the pixel level but also at the object level with large samples. Compared to sampling strategies that account for spatial autocorrelation, the accuracy metrics are overestimated.

At the pixel level (see dotted lines with a cross in Fig. 3.3), the performances of k -fold CV and LOO CV sampling strategies (green and blue lines respectively) were close with very high average OA values, regardless of the size of the training set (e.g. OA of $97.86 \pm 0.96\%$ and $95.81 \pm 3.71\%$ respectively for 10 forest stands). We observed a gradual decrease in OA for k -fold CV with the increasing number of forest stands. With the pixel-based SLOO CV approach (orange line), the prediction errors were much higher than other pixel-based methods. Concerning the group-based SLOO CV the OA started with very low values (average OA of $68.72 \pm 11.03\%$) and then increased to reach relatively stable accuracy from 50 forest stands on (average OA of $86.08 \pm 5.73\%$).

At the object level, the general pattern of errors differed from the non-spatial pixel-based strategies (see dashed lines with the dots in Fig.3.3). Starting with a very weak performance, an improvement of OA began with the increase if the size of the training set size. The accuracy of k -fold CV and LOO CV sampling strategies (green and blue lines respectively) followed the same trend. The SLOO CV approach showed a similar pattern but differed in the magnitude and variability of prediction errors. In addition, the average OA remained relatively stable from 50 forest stands on, suggesting that the maximum performance has been reached for this side training set. In contrast, the average OA of the object-based k -fold CV and LOO CV approaches continued to grow with the increase in the size of the training set, revealing the effect of spatial dependence with large datasets (with closer objects). With the maximum sample size (1,972 forest stands), the difference in OA between SLOO CV and LOO CV was substantial (from 82.08% to 90.76% respectively) indicating an optimistic bias in the non-spatial LOO CV as well as in the k -fold CV (OA = $93.73 \pm 0.13\%$). The differences in performance between the object-based SLOO CV and the k -fold CV and LOO CV strategies were minimal between 50 and 100 forest stands.

Compared to pixel-based CV, the performance accuracy of the object-based CV was less

affected by spatial dependence (overestimation is reduced) and revealed the expected pattern with an increase in the training set. In the case of the SLOO strategy, both pixel and object levels follow the same pattern with higher OA values for the pixel-based strategy.

Object-based sampling was found to be particularly impacted by the size of the sample. Model performance was very poor when the models were trained with only a few forest stands. A higher number of pixels for training can be obtained by increasing the number of forest objects, with increased representativeness of the coniferous and broadleaf classes composed of variety of species (by reducing the sample selection bias). Since forest stands have varying extents (and are consequently composed of a different number of pixels due to their irregular shapes), different training sets that are similar in size may lead to variable prediction performances, as illustrated in Fig. 3.4. Further, between two training sets composed of the same number of pixels, the one based on more forest stands provides higher prediction accuracy. In other words, performances increase with sample size if the number of objects also increases (together with the number of pixels). Variability is particularly apparent with sample from 10^3 to 10^4 pixels (Fig. 3.4). Beyond 50 forest stands (i.e. $>10^4$ pixels), the variability gradually disappears with the k -fold CV and LOO CV strategies and to a lesser extent, with the SLOO CV approach.

An analysis was conducted of the number of species included in the coniferous and broadleaf classes according to the number of forest stands. The maximum number of dominant broadleaf species existing in Hérault-34 (i.e. existing in pure forest stands as defined in the IGN BDForet[®] reference database) was systematically reached with 30 forest stands. Five dominant tree species were found in this class. For conifers, more forest stands of pure species exist with a less even distribution both in membership and space. Ten dominant conifer tree species were found. Eight of them were systematically sampled in 50 forest stands. Beyond 50,

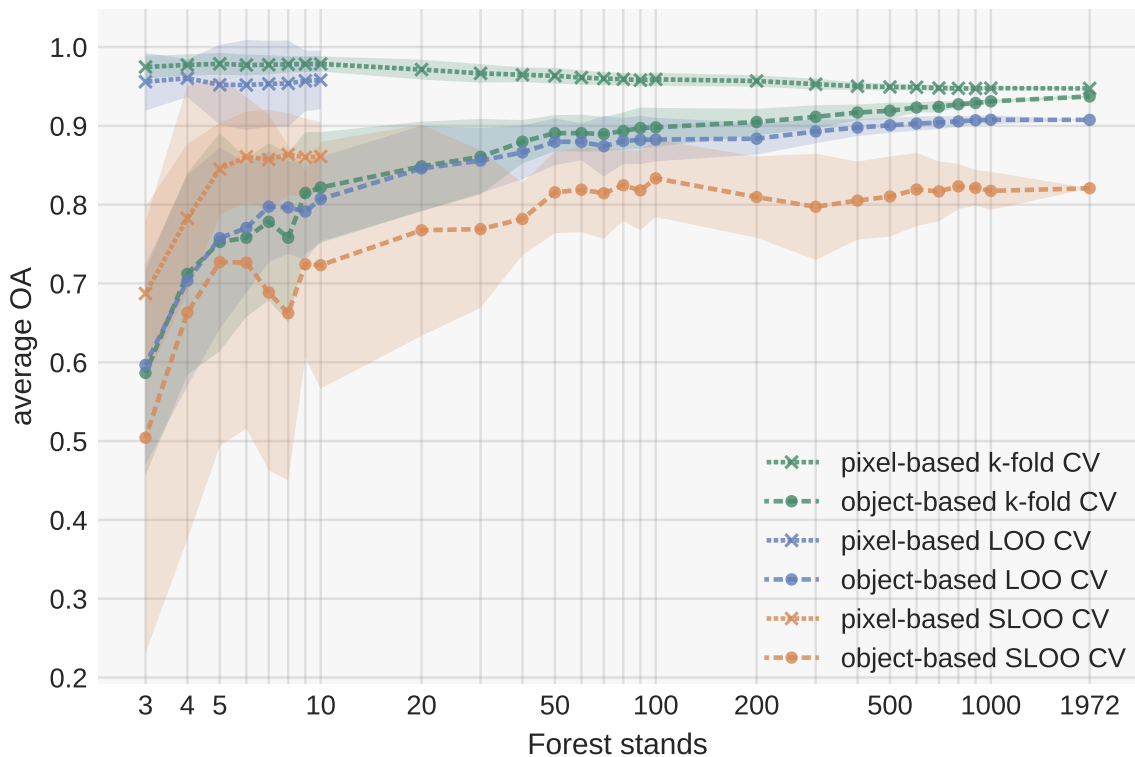


FIGURE 3.3 – Average overall accuracy based on the RF classifier for each cross-validation strategy (k -fold CV, LOO CV, SLOO CV) at pixel and object levels. Models were fitted with reference samples of Hérault-34 and repeated 10 times (i.e. the y-axis provides the average OA value \pm standard deviation). The premature stopping of the pixel-based LOO and SLOO CV approaches was due to excessive computational time.

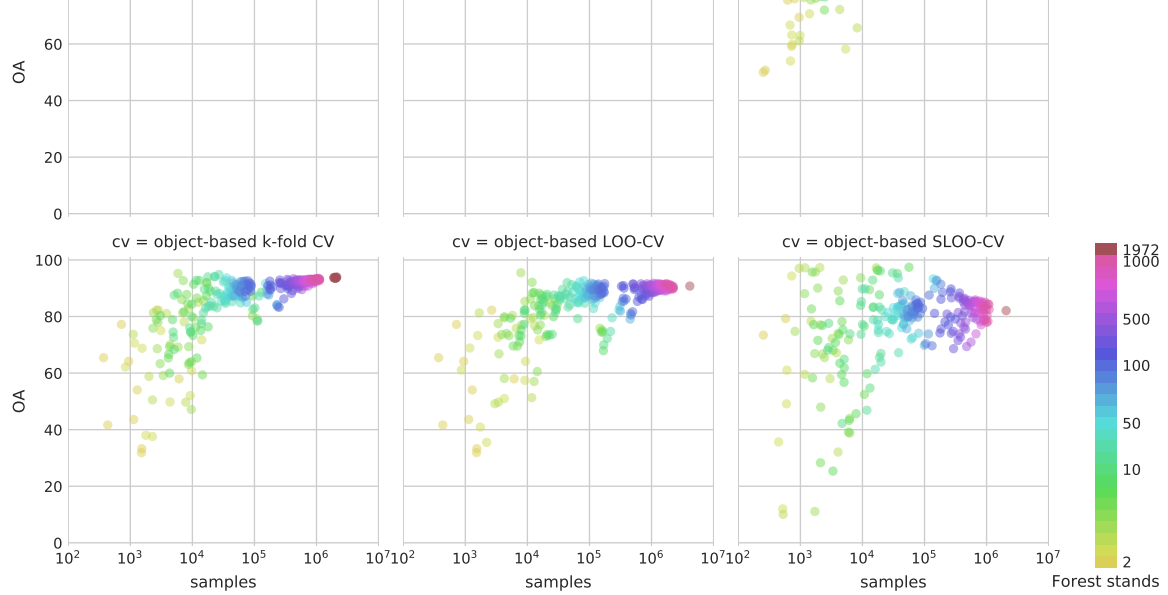


FIGURE 3.4 – Variability in accuracy depending on the sample size defined in pixels. For each prediction (i.e. each point), the number of forest stands related to the number of samples is given by the color variable.

the number of average sampled species oscillated continuously between 8 and 10 up to 1000 forest stands. Thus, because of larger number and unequal frequency of species composing the conifer class, and the lower number of available conifer samples compared to broadleaf samples, the error rate for conifers was higher (see producer’s accuracy metric in Appendix B.4). This was particularly clear when accuracy was assessed with the SLOO CV approach but less apparent when non-spatial strategies were used.

For the spatial CV, the distance threshold with a negligible spatial dependence between all the reference data was estimated at 19.74 km based on Moran’s I correlogram (see Appendix B.1). This threshold value was used in the pixel and object SLOO CV strategies.

3.4.2 Prediction errors on other spatially independent and distant sites

Prediction errors were also estimated for the neighboring departments (Tarn-81 and Aveyron-12) from the Herault-34 *full model*, i.e. the model fitted using all the available training data (1972 forest stands with related pixels). We assumed there was no change in the dataset between Herault-34 and the neighboring departments since the dominant composition of species in the coniferous and broadleaf classes of Tarn-81 and Aveyron-12 is included in Herault-34, according to the IGN BDForet[®] reference database (and confirmed by the relative frequency distribution of NDVI values of each class ; see Appendices B.2 and B.3).

Interestingly, we found OA performed similarly in predictions for Tarn-81 ($80.4 \pm 0.03\%$), Aveyron-12 ($82.2 \pm 0.01\%$) and the comparable object-based SLOO CV strategy of Herault-34 based on 1972 forest stands ($82.1 \pm 0.03\%$). Whether by cross-validation or by validation based on distant sites, predictions made outside the spatial domain of training set produced equivalent error rates (Fig. 3.5) The confusion matrices are provided in the Appendices B.5. When spatial dependence between training and test sets is ignored using standard CV, it tends to overestimate the accuracy (from +8% to +14% of OA according to the strategy compared to the OA values of distant sites).

Additional predictions were estimated for distant sites, based on Herault-34 *full models* with different sized training sets. We compared the OA values of these models evaluated using the spatial and non-spatial cross-validation strategies with the OA values based on predictions made with these models in Tarn-81 (Fig. 3.6) and Aveyron-12 (Fig. 3.7). These results show that the OA values are consistent among themselves. For an OA value estimated for a distant site, we observed a marked variability in OA estimated by cross-validation, with an optimistic bias with the non-spatial CV strategy. On average, the OA estimated from SLOO CV matches the OA estimated from spatially distant sites better, especially for large samples size (more than 50 stands). The true predictive performances of the Herault-34 *full models* of each strategy do not differ fundamentally, for a given training set size. Apparent

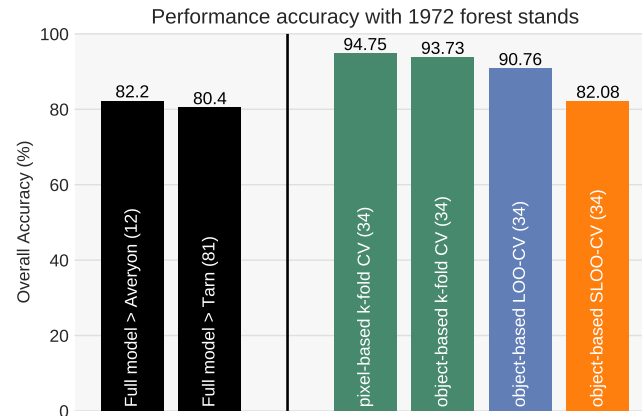


FIGURE 3.5 – Comparison between the average predictive performances obtained with spatial and non-spatial cross-validation from the Herault-34 *full model* fitted with the maximum training set of 1,972 forest stands (on the right) and the average predictive performances obtained with the same model applied to test sets from Tarn-81 and Aveyron-12 (in black, on the left). The number of test samples in Tarn-81 and Aveyron-12 is exactly the same with balanced class distributions. Pixel-based LOO CV and SLOO CV are not shown because of the excessive computation time required for this big training set.

differences in performance are only due to differences in the way they are assessed using CV strategies. This explains why, with large training sets size (more than 500 forest stands), the predictive performance for distant sites are all equivalent (see the x-axis values of the purple dots in Fig.3.6 and 3.7; OA is approximately to 80% with 1,972 stands).

3.4. RESULTS

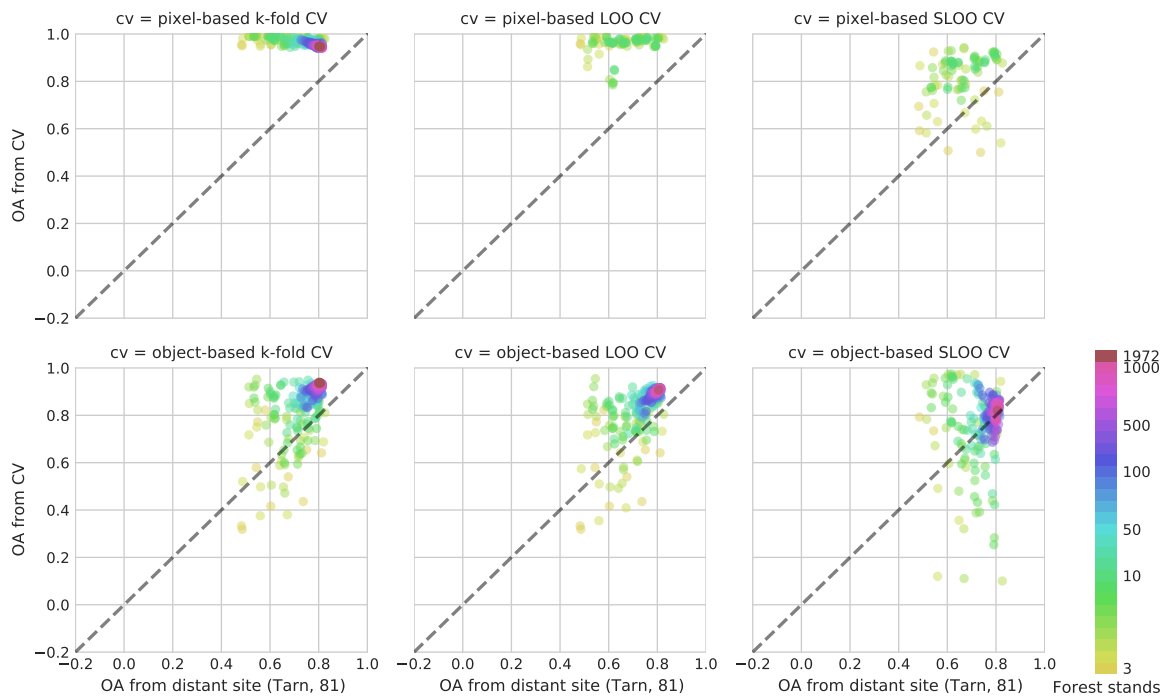


FIGURE 3.6 – Comparison of OA estimated by cross-validation in Herault-34 using the six spatial and non spatial data-splitting strategies for different sized training sets and OA estimated with the same Herault-34 model on test sets in Tarn-81. The number of training set sizes is given by the number of forest stands represented by the graphical color variable. Results for pixel-based LOO CV and SLOO CV are not shown for more than 10 forest stands because of excessive computation time.

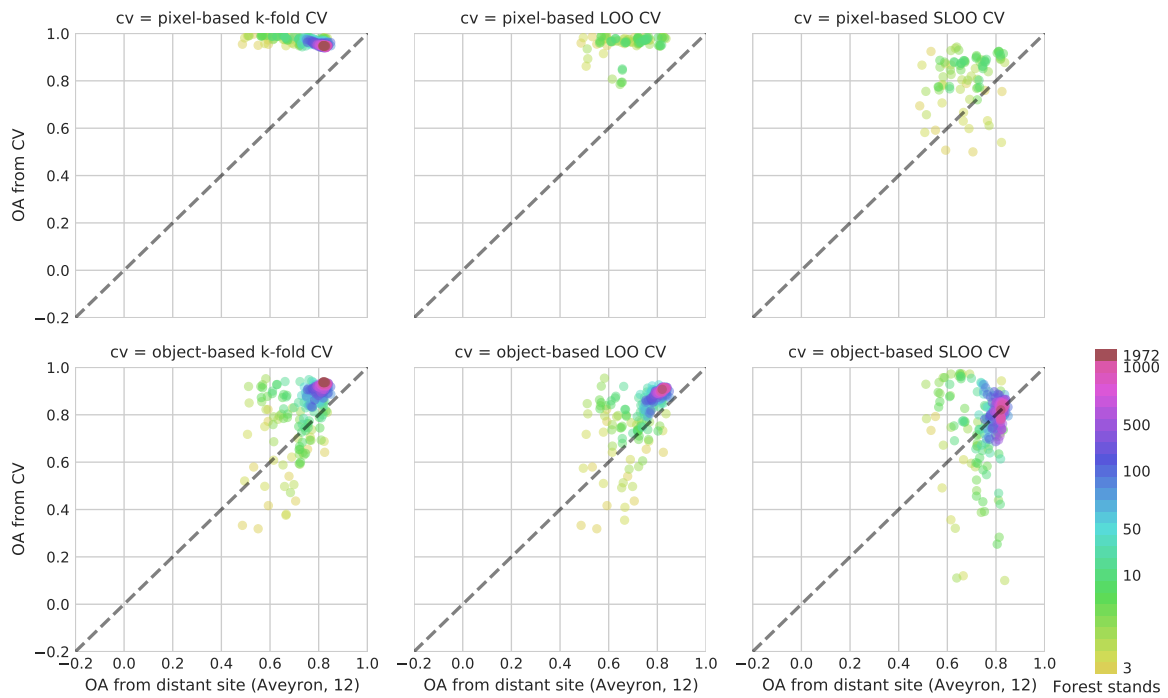


FIGURE 3.7 – Comparison of OA estimated by cross-validation in Herault-34 from the six spatial and non spatial data-splitting strategies for different sized training sets and OA estimated using the same Herault-34 model on test sets in Aveyron-12. The number of training set sizes is given by the number of forest stands represented through the graphical color variable. Results for pixel-based LOO CV and SLOO CV are not shown for more than 10 forest stands because of excessive computation time.

3.5 Discussion

Our results revealed notable underestimation of generalization errors when traditional non-spatial approaches were used to assess the accuracy. Pixel-based samplings were the most affected. Object-based strategies mitigate the effect of spatial dependence since the pixels used for training and testing never belong to the same forest stands. Nonetheless, non-spatial data-splitting at the object level also leads to overestimation of predictive performance.

Three distinct performance trends were observed with the increase in the size of the training set : (1) a slight gradual decrease in the non-spatial pixel-based CV, (2) a marked and continuous increase in the non-spatial object-based CV, and (3) a marked increase up to an optimal training set size in SLOO CV with no improvement beyond. To our knowledge, these contrasting patterns have never previously been demonstrated.

Learning curves can help interpret these trends. The curves describe how the model's error rates on training and test sets vary as the training set size increases. This makes it possible to diagnose model bias and variance. Theoretically, with only a few training samples, the model error rate on the training set would be expected to be very low (if not zero). The model can fit the training samples perfectly but does not have the ability to generalize new data. Thus, the model error rate on the test set is expected to be high. With larger datasets, the error on the training set should increase because of less overfitting. Conversely, the model should perform better on the test set. The generalization capabilities should be improved, reducing the error rate on the test set with more data but keeping an irreducible error. This analysis was conducted using three contrasted CV strategies : the pixel-based k -fold CV, the object-based LOO CV, and the object-based SLOO CV.

However, the expected behavior of learning curves was not observed with the pixel-based k -fold CV most affected by spatial autocorrelation. With both small and large training sets, the error rate in training set was zero with a convergence of the learning curve on the test set from the first sample sizes onward (Fig. 3.8). The narrow gap between the two curves falsely suggests a model with low variance irrespective of the size of the training set. As it also shows limited bias, the model appears to be perfect with no difficulty expected with generalization. This unrealistic behavior is mainly due to spatial dependence. The pixels in the test sets belong to the same forest stands as those in the training sets. The model consequently overfits the training samples but overfitting is masked by the overly optimistic estimates of accuracy using similar and correlated test sets. The generalization capabilities of the model are better with large samples. The spatial autocorrelation between training and test sets does not change, but the model slightly reduces overfitting on test sets, thereby increasing the error rate.

A more realistic pattern of errors was found with the object-based SLOO CV that was not affected by spatial autocorrelation. Notable model variance was observed when the training set size was small but variance tended to be reduced by adding more training samples, as expected (Fig. 3.8). The learning curves of the training and test converged from 50 forest stands on, suggesting that adding more samples beyond this threshold is not necessary, but there is no substantial improvement. This sample size is to be related to the number of species sampled in the coniferous class which affect the classification accuracy, as previously mentioned (see section 3.4.1). Beyond 50 forest stands, eight out of the ten dominant coniferous species are sampled systematically, reducing confusion with broadleaf species.

The behavior of the non-spatial object-based LOO CV is a blend of the two previous trends. When the training sets are small, accuracy is less affected by spatial autocorrelation. Compared to pixel-based CV, the spatial dependence between training and test sets is lower because of object-based sampling. In addition, the generalization capabilities of the model are limited because of the high variability of species composition between the training and test sets. This leads to considerable model variance. With the addition of new samples, the model performs better on test sets but with progressive overestimation of performance. The increase

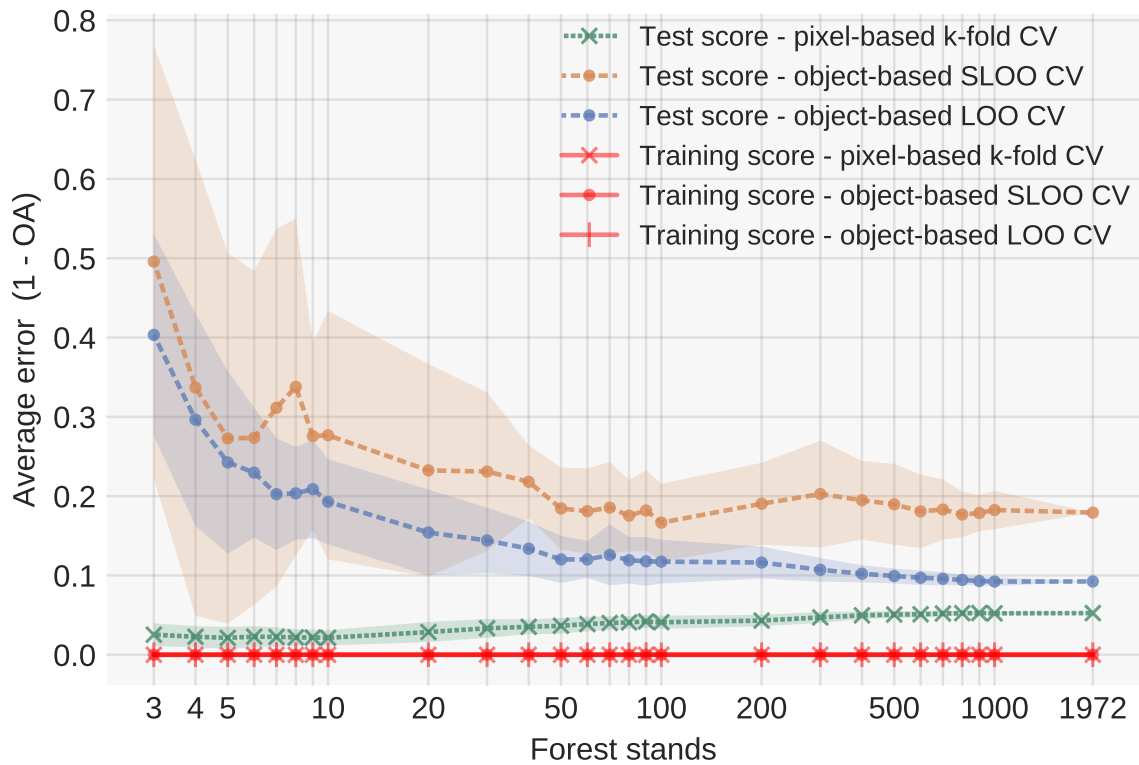


FIGURE 3.8 – Learning curves calculated each from each fold according to the different cross-validation strategies applied. Pixel-based SLOO CV and pixel-based LOO CV are not shown as they stop at 10 forest stands due to excessive computation time. The three curves related to the training scores (in red) are equal and thus, superimposed.

in performance is not due to the improvement of generalization capabilities of the model. With reference to the SLOO CV, the predictive power of the model might be reached beyond approximately 50 forest stands. Large sample sizes tend to reduce the distance between forest stands thereby increasing the effect of spatial dependence between training and test sets and hence, estimated accuracy. This explains why the error rates between non-spatial pixel-based and object-based CV are similar with large samples.

With all these strategies, whatever the sample size, the average zero error rate in the training set was unexpected, suggesting constant model overfitting with a reduction in the error rate in the test sets (in particular, in object-based CVs). Since, to avoid overfitting, we included a large number of trees in the RF algorithm, we attribute this deceptive behavior to spatial autocorrelation too but in this case, between the samples used in the training set. And this is another fundamental point that should be taken into consideration. All spatial and non-spatial CV strategies guarantee spatial independence among training samples. Spatial data splitting removed autocorrelation between the training and test sets but this autocorrelation persisted in the training set. Thus, learning curves need to be interpreted with caution.

Our findings support evidence from previous studies that spatial CV is required to estimate unbiased predictive error [Meyer et al., 2019; Schratz et al., 2019]. Our results also show that the non-spatial object-based sampling is less affected by spatial autocorrelation than pixel-based sampling, in line with [Cánovas-García et al., 2017]. When an optimal sample size is used for training (i.e. with no undersampling or oversampling) this strategy is a possible alternative to the SLOO CV if spatial dependence is ignored in the accuracy assessment. However, this non-spatial strategy only mitigates overfitting, it does not prevent it. Thus, we recommend to choosing a spatial CV whenever possible.

We used a distance-based buffer approach for spatial CV [Le Rest et al., 2014] instead of block or cluster-based partitioning [Meyer et al., 2019; Roberts et al., 2017; Schratz et al.,

2019]. Data-splitting with blocks also produces better estimates of predictive performances than random sampling. However, with no explicit quantification and no control of the degree of dependence, residual correlation may exist between training and test sets when the spatial block approach is used (i.e. where blocks consist of geographical units). Objects belonging to contiguous blocks and located close to borders may be very almost identical. Finding the ideal block size does not solve the problem, but the distance-based buffer approach avoids it. Systematic quantification of the spatial autocorrelation range should also be performed using Moran's I or empirical variogram.

3.6 Conclusion

The take-home message of this paper is that we need to change practices in classification accuracy assessment using spatial imagery. A data splitting design ensuring spatial independence between the training and test sets should be the standard approach for validation. Non-spatial LOO or k -fold CV at the object level (i.e. cross-validation leaving all the pixels belonging to one object for testing and the rest for training) is a absolute minimum required to mitigate overfitting. Spatial LOO CV is a better way to provide unbiased estimates of predictive error and to reduce the gap between the accuracy statistics given to users and the real quality of the maps produced.

To facilitate this change, we assembled the *Museo ToolBox*, an open-source python library that makes it possible to validate classification results with a range of spatial and non-spatial CV approaches (<https://museotoolbox.readthedocs.io/>). We hope this library will help the community leave its traditional approach behind to instead to routinely use spatial CV. We also expect the library to open the way for new prospects on this topic.

Acknowledgments : N.K. received a PhD scholarship from the French Ministry of Higher Education and Research (University of Toulouse). The authors are grateful to the European Space Agency (ESA) and the French national Theia infrastructure for access to Copernicus Sentinel data.

Author Contributions : Conceptualization, N.K. and D.S. ; Methodology, N.K. and D.S. ; Investigation, N.K. and D.S. ; Software, N.K. ; Visualization, N.K. ; Validation, N.K. and D.S. ; Funding acquisition, D.S. ; Supervision, J.-F.D., C.M. and D.S. ; Writing–original draft preparation, N.K. and D.S. ; Writing–review and editing, J.-F.D. and C.M.

Conflicts of Interest : The authors declare no conflict of interest.

Chapitre 4

Étude et apport de la phénologie

“ Aucune description, aussi belle qu’elle soit, ne peut être vraie, car elle ne peut reproduire la vie du paysage, la fuite de l’eau, le frémissement des feuilles, le chant des oiseaux, le parfum des fleurs, les formes changeantes des nuages ; pour connaître, il faut voir.”

— Élisée Reclus.

Table des matières

4.1	Introduction	80
4.1.1	Détection de la phénologie	80
4.1.2	Lien entre phénologie et reconnaissance des feuillus par télédétection	82
4.2	Matériel et méthode	83
4.2.1	Zone d’étude	83
4.2.2	Suivi phénologique in situ	84
4.2.3	Images et références terrain	87
4.2.4	Apprentissage des modèles	90
4.3	Résultats des suivis phénologiques	93
4.3.1	Les stades phénologiques observés sur le terrain	93
4.3.2	Longueur de saison de croissance	95
4.3.3	Les stades phénologiques par satellite	95
4.4	Résultats de l’apport des séries temporelles pour la détection des essences	97
4.4.1	Classification : Un feuillu contre l’ensemble des essences	97
4.4.2	Classification de l’ensemble des essences	102
4.5	Discussion	102
4.5.1	Stabilité des seuils pour détecter les stades phénologiques	103
4.5.2	Influence du sous-étage	105
4.5.3	Explication des dates sélectionnées pour l’ensemble des feuillus	105
4.5.4	Explication des dates par feuillu	106
4.6	Conclusion	108

De nombreux articles utilisent plusieurs images satellitaires obtenues à plusieurs dates tout au long d’une année pour cartographier les essences forestières. Généralement, le jeu de données est composé au maximum d’une dizaine de dates et inclut les différentes saisons. Ces articles utilisent généralement pour estimer les dates les plus importantes une sortie de l’algorithme Random Forest, mais comme les auteurs se basent sur une seule année il est difficile de connaître la stabilité temporelle de ces choix. Comme il a été vu dans le Chapitre 2, en s’aidant de neuf années d’archives du satellite Formosat-2, il est difficile d’estimer si une

date ou une saison en particulière aide à bien classer l'ensemble des essences présentes sur la zone d'étude.

Après une campagne de suivi phénologique complète sur deux années (de 2018 à 2019) ayant eu pour but de noter les différents stades phénologiques sur des parcelles homogènes d'essences de feuillus, les dates qui seront alors sélectionnées par un algorithme avec une sélection de caractéristiques pourront être interprétées selon des connaissances in situ.

Afin de savoir si les événements phénologiques par essence sont liés avec le choix des dates sélectionnées par l'algorithme, un cadre méthodologique a été développé. Ce dernier consiste à classer l'ensemble des essences date par date et à garder pour chaque essence la date qui obtient le meilleur score F1, et d'y ajouter une à une, le reste des dates toujours selon ce même critère.

Ce chapitre en plus de se focaliser sur l'apport des dates par essence, interprétera à l'aide de l'algorithme de lissage Savitzky-Golay le potentiel de détection du début et de la fin de saison de végétation pour les feuillus.

4.1 Introduction

La phénologie est un domaine avec un regain d'intérêt depuis environ une décennie dû au développement des études pour mieux connaître le changement climatique. Ce terme provient du grec *phainō* qui veut dire apparaître, et de *logos* qui signifie l'étude. On pourrait donc traduire littéralement ce mot par l'étude de ce qui apparaît. Ce domaine, bien que souvent associé au suivi de la végétation, étudie l'ensemble des événements récurrents qu'ils concernent la végétation (e.g. pousse des feuilles, des fruits, chute des feuilles) ou les animaux (e.g. migrations ou périodes de reproduction) [Lieth, 1974].

Si le réchauffement climatique a remis sur le devant de la scène la phénologie de la végétation notamment auprès du grand public, c'est que l'apparition des événements comme la pousse des feuilles, leur chute ou leur coloration sont en grande partie liés au climat. On estime d'ailleurs que depuis 1960 la saison de croissance a été allongée de 11 jours en Europe [Way, 2011]. Cependant, si l'augmentation des températures peut retarder la sénescence, les essences forestières sont aussi fortement influencées par la durée du jour, appelée photopériode [Wareing, 1956]. Si des espèces sont uniquement sensibles à la photopériode (rapport entre la durée du jour et de la nuit), alors une augmentation des températures n'aura aucun effet sur la date de leur sénescence puisque par définition la photopériode n'a pas évolué depuis bien longtemps.

Dans ce manuscrit de thèse, nous avons opté pour l'utilisation de *Satellite Image Time Series* ou en français, Série Temporelle d'Images Satellites (SITS), ce qui nous donne l'opportunité d'avoir un nombre important d'acquisitions dans une année et de pouvoir tenter de suivre la phénologie de la végétation. Cependant, nous ne savons pas aujourd'hui si ce qui permet de différencier les essences les unes des autres est lié à des stades phénologiques spécifiques (e.g. différences dans les dates de débournement, coloration des feuilles) ou à d'autres facteurs (structure, surface foliaire, sous-étage...). L'hypothèse avancée dans ce chapitre est que la phénologie nous aide à différencier certaines essences.

4.1.1 Détection de la phénologie

Historiquement, comme la résolution spatiale des satellites ne permettait pas d'identifier de manière suffisamment précise les objets terrestres, le terme de *Land Surface Phenology* (LSP) est apparu pour désigner l'évolution spatio-temporelle des surfaces végétalisées telle que révélée à travers l'observation des capteurs satellites [De Beurs and Henebry, 2005]. La plupart des auteurs qui ont étudié la phénologie par télédétection s'appuie sur des observations à large échelle à partir de capteurs à résolution spatiale modérée comme MODIS (250 m) [Ahl et al., 2006; Garrity et al., 2011; Jönsson et al., 2010; Xiao et al., 2006; Zhang et al.,

2003], voire SPOT VEGETATION ou PROBA-V (1 km) [Bórnez et al., 2020; Delbart et al., 2005, 2006]. Pour compenser la faible résolution spatiale de capteurs comme MODIS ou les récents produits PROBA-V à 100 m de résolution spatiale, certains auteurs comme Fisher and Mustard [2007]; Isaacson et al. [2012]; Rivas et al. [2019] ont comparé en validant ou fusionnant les résultats avec des satellites à plus haute résolution spatiale comme Sentinel-2 (10 m) ou Landsat (30 m). Cependant ces résultats ne peuvent s'appliquer que pour des forêts assez homogènes en terme d'essences. Les résultats des précédents travaux sont assez hétérogènes tant sur les méthodes que sur les données, mais l'ensemble des travaux montre qu'il est courant d'observer avec moins de 10 jours d'écart la phénologie in situ depuis des capteurs satellites. Cependant beaucoup de travaux ne valident pas avec des données terrain et la définition des termes que sont le début et la fin de saison de la végétation n'est pas toujours la même selon les auteurs, ce qui rend difficile les comparaisons [Zeng et al., 2020].

Pour détecter la date des stades phénologiques qui représentent des événements clés comme le début du débourrement ou la chute des feuilles, il est d'usage d'utiliser un indice spectral comme le *Normalized Difference Vegetation Index* (NDVI), sauf dans des conditions de fortes biomasses où cet indice a tendance à saturer, ce qui n'est pas le cas dans nos forêts tempérées Zeng et al. [2020]. Généralement les séries temporelles de NDVI sont issues du satellite MODIS. Ce succès s'explique par plusieurs raisons : (1) la disponibilité de produits prêts à l'emploi comme les séries temporelles de NDVI (les nuages ont été effacés à l'aide d'algorithmes), (2) une disponibilité mondiale de la donnée, (3) la légèreté des images due à la résolution spatiale de 250 m du satellite, (4) et parce que le NDVI est un bon proxy de l'activité photosynthétique de la végétation.

L'extraction des métriques phénologiques (i.e. les informations relatives à ces événements phénologiques comme la date de l'évènement) se fait généralement en lissant le signal à l'aide d'un algorithme dans le but de débruiter au maximum la série temporelle et en fixant des seuils par rapport à l'amplitude minimale et maximale de l'indice tout au long de l'année [Zeng et al., 2020]. La Figure 4.1 illustre les différentes métriques phénologiques qu'il est possible d'extraire à partir d'une série temporelle. Ces métriques permettent de suivre année après année l'évolution de la phénologie. Elles peuvent aussi s'utiliser en apprentissage semi-supervisé comme variables seules [Clerici et al., 2012] ou en supplément aux valeurs spectrales des images [Aragones et al., 2019; Pasquarella et al., 2018; Yan et al., 2015] pour cartographier les grandes catégories de forêt voire même les essences.

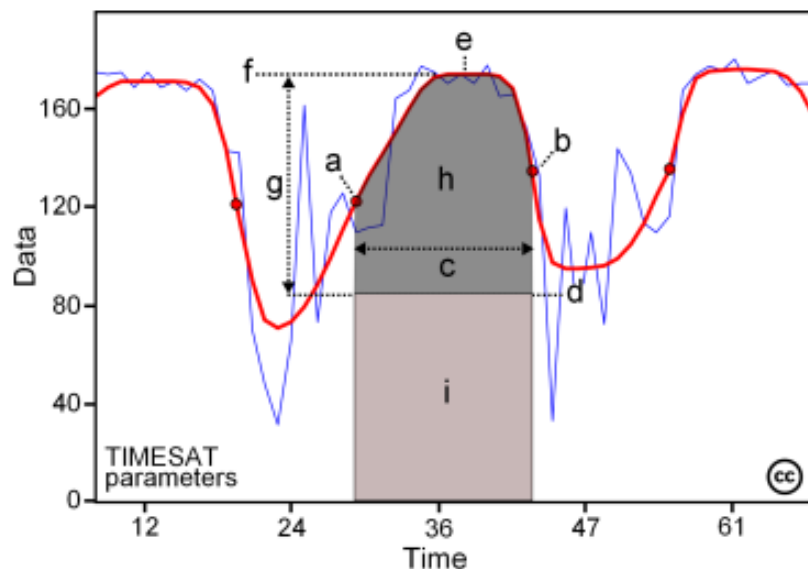


FIGURE 4.1 – Illustration de l'extraction de métriques phénologiques : (a) début de saison, (b) fin de saison, (c) longueur de la saison, (d) valeur minimale, (e) date du milieu de saison, (f) valeur maximale, (g) amplitude, (h) petite intégrale, (h+i) grande intégrale. Source Jönsson and Eklundh [2004].

Pour extraire les métriques phénologiques, de nombreuses méthodes existent mais la plus utilisée est celle du seuillage [Zeng et al., 2020]. Cela consiste à établir un seuil soit en absolu (e.g. début de saison lorsque le NDVI dépasse une valeur de 0,7) soit en relatif (e.g. début de saison lorsque l'amplitude du NDVI sur l'année dépasse les 30%). Certains auteurs ont fait varier de façon empirique ces seuils afin de trouver les meilleurs paramètres [Myneni et al., 1997].

Concernant les essences forestières, de nombreuses études sont consacrées à la phénologie mais essentiellement à partir du capteur à résolution spatiale modérée MODIS ou SPOT VEGETATION que ce soit pour les feuillus décidus [Ahl et al., 2006; Delbart et al., 2005], les feuillus sempervirent [Xiao et al., 2006] ou les conifères [Jönsson et al., 2010]. Dans leur revue de la littérature, Zeng et al. [2020] notent que les faibles variations des indices de végétation des essences sempervirentes rendent difficiles l'extraction des métriques phénologiques, contrairement aux feuillus qui ont un cycle bien marqué et observable à travers ces mêmes indices. Ce présent chapitre se focalisera donc sur la détection de la phénologie des feuillus et sur son apport dans la reconnaissance des essences par télédétection.

Cependant notre optique n'est pas de quantifier dans ce chapitre l'impact du réchauffement climatique sur la végétation ou d'étudier les changements de phénologie au fil du temps comme la plupart des travaux précédemment cités, mais plutôt de comprendre si les potentiels décalages phénologiques entre essences permettent de mieux différencier les espèces les unes des autres.

4.1.2 Lien entre phénologie et reconnaissance des feuillus par télédétection

Pour savoir si les événements phénologiques majeurs comme la floraison, le débourrement ou la sénescence permettent de différencier les essences les unes des autres, il faut déjà comprendre ce que l'on peut observer depuis des centaines de kilomètres d'altitude. En 2009, Reed et al. [2009] mentionnaient que trop peu de travaux étaient validés avec des données in situ. Hufkens et al. [2012] signalaient aussi le besoin de développer des méthodes de validation et ce en définissant notamment des protocoles communs. Sans cela, il est difficile d'évaluer la capacité de détection des événements annuels de la végétation. Pour pallier au manque de données de validation, il s'est mis en place depuis environ une décennie un réseau de caméra nommé PhenoCam [Richardson et al., 2011] qui fournit gratuitement des images au moins quotidiennes d'un point précis sur la végétation.

Des travaux exploitant Sentinel-2 ont émergé pour, soit étudier le potentiel de détection des stades phénologiques, soit estimer la contribution des dates pour prédire les essences forestières. Bénéficiant d'une résolution temporelle plus élevée que Landsat, ces travaux sont encore peu nombreux car les deux satellites fonctionnent en coordination depuis l'été 2017.

Dans le cadre de la cartographie des essences forestières par télédétection à une échelle d'au minimum plusieurs massifs, il est d'usage d'utiliser plusieurs images satellites sans nuages obtenues lors d'une même année [Fassnacht et al., 2016]. Conformément aux résultats de Hill et al. [2010], les travaux les plus récents confirment qu'il vaut mieux privilégier peu d'images à des dates clés que beaucoup d'images [Grabska et al., 2019; Immitzer et al., 2019]. Cependant, il peut être difficile de connaître a priori ces dates, et dans le cadre d'une cartographie à large échelle il ne sera pas possible d'avoir la même date d'acquisition pour tout un territoire. Il y a donc un intérêt non négligeable à utiliser des SITS dans l'optique d'une cartographie à large échelle comme il a été fait pour la carte d'occupation du sol de la France [Inglada et al., 2017].

À notre connaissance, l'apport de la phénologie dans la classification des feuillus et la compréhension de l'interaction entre les phénomènes observés en forêt et l'observation satellite n'a jamais fait l'objet d'une étude en détail par espèce à partir d'un capteur de haute résolution spatiale, spectrale et temporelle. Afin d'estimer au mieux l'apport de la phénologie, et comme le nombre important de dates et donc de variables spectrales peut réduire la

qualité des modèles (le fléau de la dimension selon Bellman [1984]), un protocole de sélection de caractéristiques [Whitney, 1971] a été réalisé de manière systématique pour estimer non seulement les meilleures dates, mais pour aussi comprendre si ce choix est en lien avec des évènements phénologiques majeurs.

Dans ce chapitre, nous allons utiliser les séries temporelles d'images Sentinel-2 avec pour objectif en s'appuyant sur des relevés phénologiques de terrain (1) d'estimer la précision des métriques phénologiques obtenues à l'aide d'un satellite optique, et (2) d'estimer l'apport de la phénologie en analysant si les dates sélectionnées lors de la sélection de caractéristiques pour apprendre un modèle par essence sont liées à des évènements phénologiques sur les essences observées.

À notre connaissance, les nouveautés scientifiques présentées dans ce chapitre sont :

1. La détection et validation des évènements phénologiques des essences forestières obtenues à partir d'images satellites Sentinel-2 et de données terrain avec l'aide de la nomenclature BBCH.
2. La stabilité interannuelle des résultats des métriques phénologiques et de leurs paramètres.
3. le lien entre les dates sélectionnées par un algorithme pour prédire les essences et les évènements phénologiques.

Les deux capteurs Sentinel-2 seront utilisés pour évaluer et valider la phénologie des essences forestières à partir de deux années de séries temporelles d'images satellites à haute résolution temporelle et avec une validation terrain compatible avec la résolution spatiale et temporelle du satellite.

4.2 Matériel et méthode

4.2.1 Zone d'étude

La zone d'étude se situe au sud ouest de Toulouse. Elle couvre un territoire de 24 km x 24 km (Figure 4.2). Cette limitation correspond à l'emprise du satellite Formosat-2 sur laquelle le CNES a un suivi régulier avec des acquisitions privilégiées à cet endroit grâce à l'Observatoire Spatial Régional fondé en 2006 par le CESBIO.

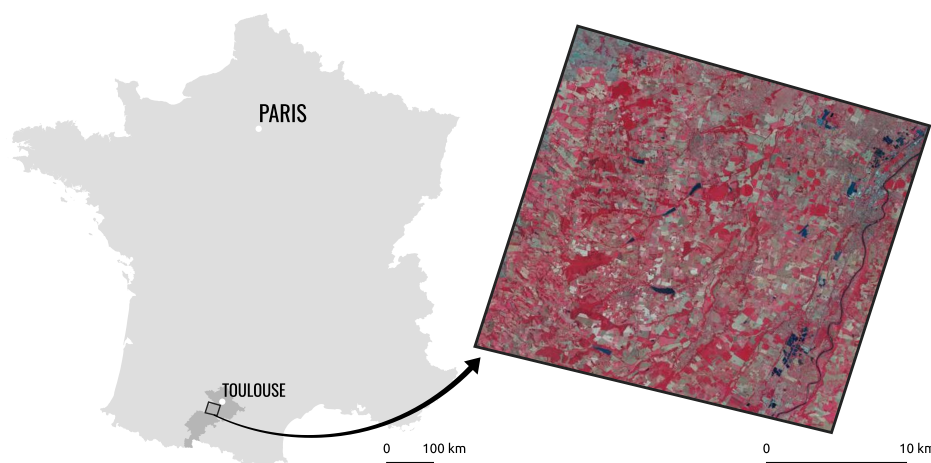


FIGURE 4.2 – Zone d'étude

La Garonne traverse la zone d'étude à l'est du nord au sud et influence la composition du sol et la topographie. Le climat est typique du subaltantique avec un automne ensoleillé, un été chaud et sec et un hiver doux et pluvieux (température annuelle moyenne $> 13^{\circ}\text{C}$; précipitation = 656 mm). Le paysage est dominé par les cultures (blé, maïs, colza, tournesol). La forêt recouvre un peu plus de 10% de la scène soit 53 km².

4.2.2 Suivi phénologique in situ

Afin d'étudier les interactions possibles entre l'évolution du couvert végétal et les observations satellitaires, une campagne de terrain a commencé à l'automne 2017 avec comme objectif de suivre régulièrement l'évolution la végétation, en particulier l'évolution de la canopée et les stades phénologiques des feuillus. Comme l'apparition des évènements phénologiques peut être très rapide, il est important d'avoir une fréquence de revisite assez élevée et régulière afin de ne pas les manquer.

Pour cela nous avons donc décidé de suivre, pour chacune des sept espèces de feuillu décidu dont nous avons les références terrain, deux parcelles sélectionnées aléatoirement dans nos références terrain. Les parcelles ont été contrôlées afin que la strate émergente appartienne bien à la même espèce et que les essences n'étaient pas affectées par une maladie. La localisation précise des parcelles correspond à leur centroïde Sentinel-2 qui a été matérialisé par un piquet sur le terrain à l'aide d'un GPS différentiel Trimble Geo 7X. La précision du piquet par rapport au centroïde est de l'ordre du mètre.

Des différences intra et extra-spécifiques, comme la structure de la forêt ou la composition du sous-étage sont observables entre les différentes parcelles. Pour qualifier et quantifier les évènements phénologiques de chaque parcelle, le protocole se déroule exactement de la même manière pour chaque placette. La fiche rédigée pour suivre le protocole est disponible en annexe C.1.

Toutes les données récoltées sont directement enregistrées sur le terrain à l'aide de l'application android Memento. Cette application nous a permis de facilement mettre au point un formulaire à remplir à chaque visite. L'usage d'un formulaire facilite grandement la transition entre différentes personnes et surtout permet de ne pas oublier certains éléments à remplir lors de la visite. Le tout est directement synchronisé en ligne sur google sheets (le tableur de Google). Il est toutefois possible d'extraire de l'application le tableau sous format csv.

4.2.2.1 Fréquence des visites

La campagne de terrain a commencé en septembre 2017, mais comme le protocole était en cours d'expérimentation et de validation, nous analyserons uniquement les résultats des deux années entières où un suivi a pu être fait à la fois au printemps et à l'automne, c'est-à-dire 2018 et 2019.

Pour chaque année, il y a eu 17 sorties qui ont été réalisées de février à fin mai et de septembre à fin novembre (Figure 4.3). Pour le printemps 2018, certaines sorties ont été doublées car nous avons manqué de temps sur le terrain, mais une fois le protocole rodé, une seule journée suffisait à visiter et noter les informations de l'ensemble des quatorze parcelles (deux parcelles par essence).



FIGURE 4.3 – Liste des dates auxquelles un suivi phénologique a été réalisé sur le terrain.

4.2.2.2 BBCH

Afin d'estimer le stade phénologique moyen des feuillus localisés à l'intérieur de chacun des quatorze pixels étudiés, nous avons opté, sous les conseils avisés de Jean-François Lebourgeois et comme spécifié dans l'ouvrage "*Les plantes au rythme des saisons : guide d'observation phénologique*" [Badeau et al., 2017] pour la nomenclature BBCH [Meier et al., 2009]. Initialement développée pour le suivi des cultures [Meier, 1985], la nomenclature BBCH est un code composé d'un nombre à deux chiffres. Le premier chiffre correspond au stade phénologique et

le deuxième au pourcentage de complétion de ce stade. Ainsi, pour le stade 1 qui correspond à la sortie des feuilles, si 50% des feuilles sont sorties pour l'individu considéré, le code BBCH sera de 15.

Pour aider à mieux appréhender le code BBCH, une illustration qui synthétise la notation des principaux stades phénologiques a été réalisée (Figure 4.4), ainsi qu'un tableau détaillant les principaux stades (Table 4.1).

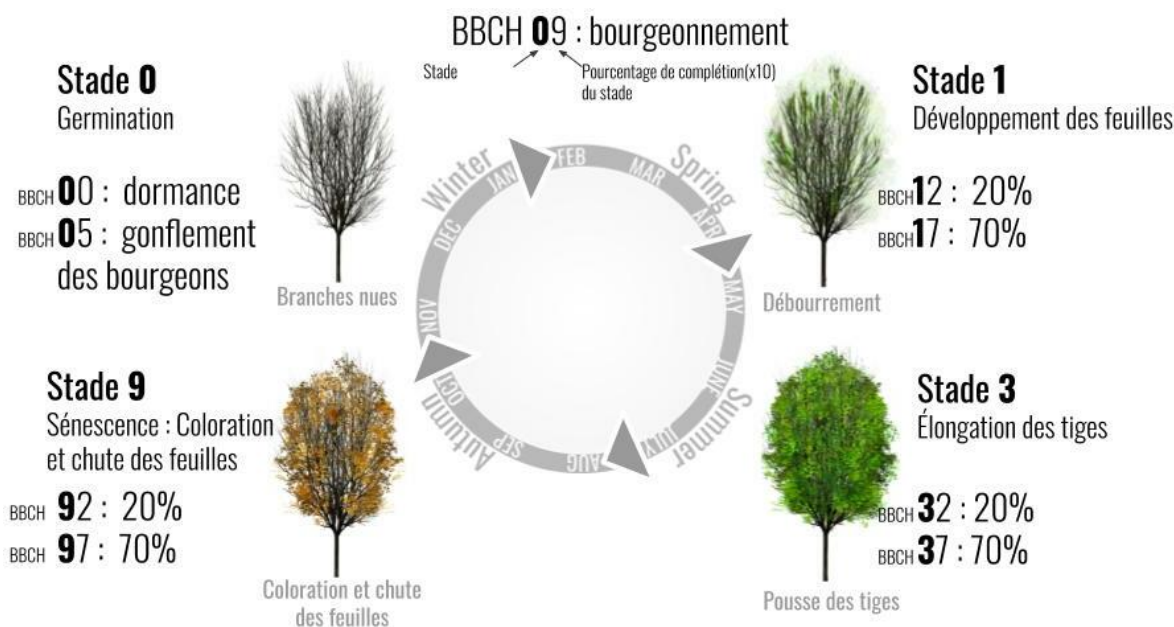


FIGURE 4.4 – Illustration des notations de BBCH pour le suivi de la phénologie des feuillus.

L'intérêt d'utiliser la nomenclature BBCH est qu'elle est reconnue dans le milieu scientifique et très utilisée, notamment par l'Observatoire des Saisons¹ ou encore par le réseau RENECOFOR [Ulrich, 1995]. Cependant, pour certains stades comme la sénescence, la notation peut mélanger plusieurs éléments comme la coloration des feuilles avec la chute des feuilles. Pour pallier à cette limite en télédétection, il a été suivi l'ouverture de la canopée à chaque sortie terrain.

TABLEAU 4.1 – Explication de la nomenclature BBCH (d'après Badeau et al. [2017]). Les pourcentages ci-dessous sont donnés à titre indicatif, la nomenclature BBCH permet de noter tous les 10% d'un même stade.

Stade phénologique (chiffre du stade)	Pourcentage du stade (%)	Code BBCH	Commentaire
Germination (0)	10	01	Dormance (hiver)
	50	05	Les bourgeons gonflent
	90	09	Les bourgeons sont en train d'exploser
Développement des feuilles (1)	10	11	10% des feuilles sont sorties
	50	15	50% des feuilles sont sorties
	90	19	90% des feuilles sont sorties
Floraison (7)	50	75	La moitié des fleurs sont apparues
Maturation (8)	50	85	La moitié des fleurs sont tombées/sèches
Sénescence (9)	10	91	10% des feuilles sont colorées ou tombées
	50	95	50% des feuilles sont colorées ou tombées
	90	99	90% des feuilles sont colorées ou tombées

1. <https://www.obs-saisons.fr/>

4.2.2.3 Ouverture de la canopée

Le signal reçu par le satellite va être très dépendant de la densité de la canopée. En effet, plus la canopée sera ouverte, plus le signal reçu par le satellite sera influencé par les strates inférieures. Comme le code BBCH mélange pour un même pourcentage la perte des feuilles et leur coloration, mesurer le taux de couverture de la canopée permet de compenser cette limite dans le cadre d'un usage en télédétection. Grâce au travail de Tichý [2016], nous avons utilisé l'application *Gap Light Analysis Mobile Application* (GLAMA) qui permet à l'aide d'un objectif fisheye qui peut être placé sur un smartphone de réaliser des photos hémisphériques à partir desquelles il est possible d'en extraire différentes métriques comme le *Canopy Cover*, *Canopy Closure*, le *Canopy Cover Modified*. Pour rappel, ces termes ont été évoqués dans la partie 1.3.3.5. La *Canopy Cover* représente le pourcentage de canopée une fois que l'image a été mise à plat (et non sur la sphère) [Gonsamo et al., 2013]. Afin de rendre le protocole le plus léger possible, les photos sont prises directement avec la tablette à partir de l'application *Level Camera* qui permet de prendre une photo en étant à l'horizontal (un niveau à bulle apparaît et un bip retentit lorsque l'appareil est à l'horizontal).



FIGURE 4.5 – Exemple de la couverture de la canopée du peuplement de robiniers au 1er août 2019. La photo a été prise à l'aide de l'application *Level Camera* avec un objectif fisheye compatible pour smartphone.

Pour chaque parcelle (qui équivaut à un centroïde Sentinel-2), cinq photos hémisphériques sont prises, la première à l'endroit même du centroïde du pixel, et les quatre autres aux quatre points cardinaux à 2,5 m du centroïde afin d'avoir une valeur moyenne de la densité de la canopée.

4.2.2.4 Estimation de la chlorophylle

Des mesures non destructives de chlorophylle sur un unique individu ont été effectuées à chaque parcelle quand les branches étaient accessibles à l'aide d'une perche de 5 m de long. Les rameaux en hauteur ont été privilégiés mais il n'a pas toujours été possible de suivre strictement un protocole de saisie adapté (avec un prélèvement qui conserve notamment l'angle d'exposition des feuilles, etc). Par conséquent, les informations recueillies apportent une indication sur la teneur en chlorophylle mais elles n'ont pas été directement exploitées.

Pour mesurer la teneur en chlorophylle nmol/cm^2 nous avons utilisé le Chlorophyll meter SPAD-502 Plus (Konica Minolta). Cinq mesures pour chacune des six feuilles récoltées ont été réalisées. Les résultats, qui s'échelonnent entre 0 à $50 \text{ nmol}/\text{cm}^2$ ont été moyennés à chaque

parcelle.

4.2.2.5 Observation générale

Toutes les données disponibles sont partagées sur Zenodo afin d'assurer la pérennité et la reproductibilité des travaux (Chapitre 5).

Pour résumer, à chaque parcelle, l'ensemble des étapes ci-dessous est effectué :

1. Notation du BBCH [Badeau et al., 2017] calculé en observant le stade moyen des individus de l'essence dominante présents dans un carré de 10 m x 10 m.
2. 5 photos prises à l'aide d'un objectif fisheye en utilisant l'application *Level Camera* pour être à l'horizontal sans trépied. Une photo est prise au centroïde même, les 4 autres sont prises à 2,5 m du centroïde en direction de chaque point cardinal.
3. Le calcul des indices d'ouverture de la canopée est réalisé à partir de l'application GLAMA [Tichý, 2016].
4. L'acquisition d'un rameau au nord de l'arbre à la hauteur maximale possible (environ 5 m) dans le but d'estimer la chlorophylle des feuilles.
5. Photographie de la parcelle et du sous-étage.
6. Commentaires spécifiques.

4.2.3 Images et références terrain

Les images utilisées proviennent des satellites Sentinel-2 et sont composées de quatre bandes à 10 m de résolution spatiale (bleu, vert, rouge et proche infra-rouge), et de six bandes à 20 m allant du red-edge (portion du spectre juste après le rouge mais avant le proche infrarouge) au proche et moyen infrarouge. La résolution radiométrique est de 12 bits. Les images ont été téléchargées via la plateforme THEIA. Ces données sont disponibles directement au niveau 2A (i.e. produit ortho-rectifié et en réflectance de surface) et incluent un masque de nuages et d'ombres. L'ensemble des images traitées au niveau 2A a été téléchargé, et l'échantillonnage temporel est irrégulier compte tenu de la couverture nuageuse aléatoire d'une année sur l'autre qui rend indisponible la donnée (Figure 4.6).

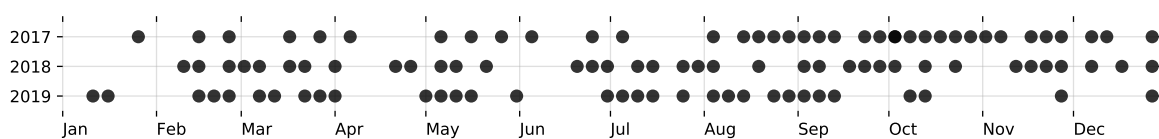


FIGURE 4.6 – Dates d'acquisition des images Sentinel-2 pour les années 2017, 2018, et 2019. L'échantillonnage temporel des données Sentinel-2 n'est pas régulier car à certaines dates la scène était masquée entièrement ou presque par des nuages. Cela signifie que ces images de niveau 1C n'ont pas été traitées vers le niveau 2A par le pôle Théia.

4.2.3.1 Lissage des séries temporelles

La série temporelle générée à partir des données Sentinel-2 contient toutefois des imperfections voire des erreurs dans la signature temporelle de l'occupation du sol. Lorsqu'un nuage est détecté par l'agorithme MAJA développé par le Centre d'Études Spatiales de la BIOSphère (CESBIO) en collaboration avec le Deutsches Zentrum für Luft- und Raumfahrt, ou, en français, le centre allemand pour l'aéronautique et l'astronautique (DLR) [Hagolle et al., 2015], une interpolation linéaire est effectuée afin de combler la donnée manquante. Cependant, il arrive que des nuages ou des ombres ne soient pas détectées, et dans ce cas aucun traitement n'était initialement réalisé. Afin de diminuer les effets de ces bruits, il est donc d'usage d'appliquer des filtres afin de réduire les anomalies.

Différentes méthodes ont été proposées pour lisser les SITS, comme la médiane mobile [Reed et al., 1994], l'algorithme de Savitzky-Golay [Chen et al., 2004] ou encore celui de Whittaker [Eilers, 2003].

En comparant quatre méthodes de lissage de séries temporelles pour étudier la phénologie de la végétation, Atkinson et al. [2012] ont conclu que Whittaker était généralement la solution la plus adaptée lorsqu'il n'y avait que peu voire pas de bruits. Quand les auteurs ont ajouté des anomalies dans la série temporelle, l'application d'une fonction de type double logistique s'est montrée plus performante. Comme il n'est pas rare d'avoir des données bruitées, notamment à cause de nuages ou d'ombres mal détectés, la double logistique semble donner de bons résultats pour estimer la phénologie de la végétation à partir de capteurs à moyenne résolution spatiale [Beck et al., 2006; Julien and Sobrino, 2009]. Certains auteurs ont opté pour une approche multi-filtres, en combinant plusieurs méthodes de lissage. Par exemple, en lissant tout d'abord la série temporelle avec l'algorithme de Savitzky-Golay, puis en lui appliquant un filtre de médiane mobile [Zhang et al., 2017].

Pour les besoins de cette étude nous utiliserons pour la partie détection des métriques une série temporelle lissée avec l'algorithme Savitzky-Golay car cette méthode a déjà fait ses preuves sur de nombreux travaux et qu'elle s'applique à toute forme de spectre [Cao et al., 2018; Chen et al., 2004]. Comme la taille de la fenêtre d'application de l'algorithme de Savitzky-Golay influence grandement les résultats, la taille optimale de la fenêtre sera recherchée dans une plage de valeur allant de trois dates à 19 avec un pas de deux dates (soit tous les nombres impairs de trois à 19). La validation de ces paramètres sera faite à partir de la qualité et de la stabilité des résultats des métriques phénologiques obtenus.

Cependant, dans le cadre d'une application à un modèle d'apprentissage automatique, l'ensemble des méthodes de lissage a diminué la variabilité intra-classe [Shao et al., 2016] et c'est pour cette raison que la partie dédiée à l'apprentissage semi-supervisé des essences forestières utilisera uniquement la série temporelle brute et non lissée.

4.2.3.2 Indices spectraux

Les indices spectraux ont plusieurs usages : ils permettent tout d'abord de réduire le nombre de variables dans l'apprentissage donc ils réduisent la dimension de la donnée (si une date contient dix bandes spectrales, l'indice n'aura plus qu'une bande). Cela permet de limiter, dans les cas de données avec un grand nombre de dimensions, l'effet de Hughes [Hughes, 1968] qui se traduit par une perte de la qualité du modèle d'apprentissage lorsque le nombre de variables augmente. Aussi les indices permettent de donner une interprétation biophysique ou biochimique à la forme du spectre. Par exemple pour l'indice NDVI qui s'échelonne entre -1 et 1, si le résultat est supérieur à 0,5 alors il s'agira très certainement de végétation photosynthétiquement active. À la différence des bandes utilisées seules (verte ou proche infrarouge par exemple) qui, lorsqu'elles sont utilisées seules, peuvent pour une même valeur, correspondre à différents types d'occupation du sol.

La liste des indices utilisés est présenté dans le Tableau 4.2. Outre le traditionnel NDVI et son homologue ACORVI qui est un NDVI modifié pour ne pas saturer sur des données corrigées des effets atmosphériques [Hagolle, 2019], nous allons tester des indices spécifiques à Sentinel-2 qui utilisent les bandes du red-edge, c'est-à-dire la partie du spectre entre 670 nm et 780 nm où les feuilles vertes ont une brusque augmentation de réflectance [Féret, 2009]. Trois indices développés pour refléter les pigments foliaires de chlorophylle, d'anthocyanine ou de caroténoïdes ont été étudiés [Wulf and Stuhler, 2015]. Il existe toutefois un grand nombre d'autres indices pour Sentinel-2, mais notre choix s'est porté sur des indices qui représentent les pigments foliaires. L'objectif recherché est de voir si les observations in situ à certaines dates clés, comme la couleur rouge liée à l'anthocyanine lors de la sénescence chez le chêne rouge (*Quercus rubra*) [Lee and Gould, 2002] peuvent être directement observables à travers l'évolution des indices décrivant ces pigments foliaires.

TABLEAU 4.2 – Liste des indices spectraux utilisés pour détecter la phénologie des feuillus.

Nom de l'indice	Formule	Signification
ACORVI	$\frac{B8-(B4+0.05)}{B8+(B4+0.05)}$	Indice corrigé et normalisé d'activité de la végétation (NDVI modifié pour la correction atmosphérique)
LAnthoC	$\frac{B7}{B3-B5}$	Contenu des feuilles en anthocyanine
LCaroC	$\frac{B7}{B2-B5}$	Contenu des feuilles en caroténoïde
LChloC	$\frac{B7}{B5}$	Contenu des feuilles en chlorophylle
NDVI	$\frac{B8-B4}{B8+B4}$	Indice normalisé d'activité de la végétation

4.2.3.3 Extraction des métriques phénologiques

L'une des méthodes les plus utilisées pour détecter les dates de début et de fin de saison est d'établir un seuil par rapport à l'amplitude de la saison ou de l'année [Zeng et al., 2020]. L'indice de végétation NDVI est le plus populaire pour extraire ces métriques, si ce n'est dans les régions où la couverture nuageuse peu être significative, il est alors préférable d'utiliser l'indice *Normalized Difference Vegetation Index* (NDVI) [Delbart et al., 2005, 2006]. C'est une méthode très simple et facile à mettre en place, mais qui n'a aucune signification biophysique quand au choix du seuil, et qui peut être influencée par du bruit dans la série temporelle [Zeng et al., 2020].

Pour nous aider dans le choix des méthodes et des seuils, Zeng et al. [2020] ont écrit une revue de la littérature sur l'étude de la phénologie à partir de séries temporelles d'images satellites. Ils ont estimé qu'il était difficile de dire quels sont les meilleurs seuils et quelle est la meilleure méthode de lissage. En se focalisant sur le lissage avec le filtre de Savitzky-Golay, nous allons étudier l'impact des différents seuils et les comparer à nos relevés terrain afin de trouver les paramètres optimaux et les plus stables entre les deux années de relevés pour détecter les stades phénologiques.

À la façon de Myneni et al. [1997], nous optons pour une méthode empirique basée sur l'amplitude d'une saison, en faisant varier les seuils de détection tous les 0,05 de 0,15 à 0,85 (soit de 15 à 85%). L'amplitude est calculée en se focalisant sur le début de saison (maximum de l'année - minimum du début de saison) et sur la fin de saison (maximum de l'année - minimum de la fin de saison).

Afin de détecter à partir d'indices de végétation adaptés (ACORVI, LAnthoC, LCaroC, LChloC et NDVI) le moment où les essences atteignent les stades BBCH sélectionnés (11, 15, 19, 91, 95 et 99), nous avons choisi d'utiliser un seuil sur l'amplitude du signal. Pour cela, nous regardons la différence entre la valeur minimale et maximale de la saison, et selon un pourcentage de l'amplitude (les seuils de 15 et 85% tous les 5% ont été testés). Pour mesurer la précision d'un seuil, il a été comparé la date trouvée à partir de la série temporelle avec la date où le stade BBCH a été atteint sur le terrain. L'erreur moyenne absolue (*mean absolute error*) a permis de comparer les méthodes et d'estimer les meilleurs paramètres, c'est-à-dire la combinaison entre la meilleure taille de la fenêtre de Savitzki-Golay avec le meilleur seuil pour l'amplitude.

Les deux stades phénologiques qui seront analysés sont le début et la fin de saison, et les pourcentages d'accomplissement étudiés sont soit au moins 10%, soit 50% soit au moins 90% du stade (BBCH \geq 11, BBCH \geq 15 et BBCH \geq 19 pour le développement des feuilles et BBCH \geq 91, BBCH \geq 95 et BBCH \geq 99 pour la sénescence).

Pour calculer à la volée à partir d'une série temporelle brute (i.e. bandes spectrales Sentinel-2) des séries temporelles d'indices spectraux et les lisser selon un algorithme au choix, la bibliothèque Museo Pheno a été développée lors de cette thèse (Chapitre 5).

4.2.3.4 Références terrain

Quatre campagnes de relevés ont été effectuées entre novembre 2013 et janvier 2017 afin d'identifier et de référencer les essences forestières de la zone d'étude. Les 72 principaux massifs forestiers ont été visités. Seules les espèces dominantes de feuillus ou de résineux ont été sélectionnées. Pour s'assurer de l'homogénéité de la donnée, les références ont été acquises à partir d'un récepteur GPS Garmin GPSMap 62st (3 à 5 m de précision) au centre d'une zone couvrant approximativement 576 m² (i.e. neuf pixels contigus du satellite Formosat-2, soit par pixel 8 m × 8 m, voir Sheeren et al. [2016]). Seul les pixels au centre de ces zones ont été utilisés pour la classification.

Au total il a été récolté 1626 références terrain pour quatorze espèces, dont huit feuillus et six conifères (Tableau 4.3). Pour certaines essences, l'identification a été limitée au genre en raison de l'existence de nombreux cultivars (comme c'est le cas pour le peuplier) ou de la difficulté à déterminer l'espèce exacte (chêne, saule ou eucalyptus). La distribution par classe est déséquilibrée mais reflète l'abondance des essences dans la forêt. Le nombre d'échantillons varie de 42 (cyprès) à 247 (chêne).

TABLEAU 4.3 – Liste des essences forestières et leur effectif en pixel ($n = 1626$). Le nombre de peuplements a été calculé à partir de la donnée (IGN BDForêt[®] v.1)

Espèce	Effectif	Peuplement
Feuillus		
Bouleau (<i>Betula pendula</i>)	67	3
Chêne (<i>Quercus robur/pubescens/petraea</i>)	247	12
Chêne rouge (<i>Quercus rubra</i>)	205	9
Peuplier (<i>Populus spp.</i>)	139	7
Frêne (<i>Fraxinus excelsior</i>)	130	5
Robinier (<i>Robinia pseudoacacia</i>)	88	9
Saule (<i>Salix alba</i>)	52	4
Eucalyptus* (<i>Eucalyptus spp.</i>)	143	5
Conifères		
Pin laricio (<i>Pinus nigra subsp. laricio</i>)	201	9
Pin maritime (<i>Pinus pinaster</i>)	157	78
Pin noir (<i>Pinus nigra</i>)	52	2
Sapin blanc (<i>Abies alba</i>)	53	4
Sapin douglas (<i>Pseudotsuga menziesii</i>)	50	8
Cyprès (<i>Cupressus</i>)	42	1
*Seul feuillus à feuilles persistantes présent dans la zone d'étude.		

4.2.4 Apprentissage des modèles

4.2.4.1 Algorithme

Pour entraîner le modèle en vue de reconnaître les essences à partir de chaque SITS, l'algorithme SVM (Support Vecteur Machine) a été choisi [Vapnik, 1998]. Il est réputé pour être efficace avec un faible nombre de données d'entraînement et un nombre important de caractéristiques [Mountrakis et al., 2011]. Pour cette étude, nous avons fait le choix d'un noyau de type gaussien (Radial Basis Function, RBF) qui a déjà fait ses preuves sur des problématiques similaires [Kavzoglu and Colkesen, 2009]. Le réglage des paramètres de régularisation de l'algorithme (C) et du noyau (γ) a été fixé par validation croisée parmi ces valeurs : $C = \{10^{-2}; 10^{-1}; \dots; 10^{10}\}$ et $\gamma = \{10^{-9}; 10^{-8}; \dots; 10^3\}$.

L'apprentissage s'est fait via la bibliothèque python *scikit-learn* [Pedregosa et al., 2011]. Pour la série temporelle et pour chaque indice spectral, les variables ont été standardisées en amont de l'apprentissage.

4.2.4.2 Validation des modèles

Les différents modèles ont été évalués et appris en utilisant une validation croisée. Étant donné le biais optimiste possible lié à l'autocorrélation spatiale dans les jeux de données comme nous l'avons montré dans les chapitres 3 et 4, nous avons opté pour une validation croisée de type spatialisée [Le Rest et al., 2014; Pohjankukka et al., 2017].

Une validation croisée spatialisée consiste à sélectionner des échantillons de validation qui soient distants d'au moins d mètres des échantillons d'entraînement. Pour calculer cette distance d , nous avons utilisé l'indice de Moran [Moran, 1950] sur chaque bande spectrale et moyenné le résultat, cela nous donne un indice allant de -1 à 1, où 1 correspond au maximum d'autocorrélation spatiale, et 0 le cas idéal sans autocorrélation.

Nous avons fait varier de manière linéaire la distance de 1 pixel (10 m) à 100 pixels (1000 m) pour estimer la distance à laquelle la variabilité entre les valeurs spectrales d'un voisinage défini (de 1 à 100 pixels) est égale à la variabilité totale du jeu de données. Dale and Fortin [2014] considère qu'en dessous de Moran's $i=0,2$, l'autocorrélation spatiale peut être considérée comme négligeable.

Nous obtenons alors une distance de 360 m en 2017, 250 m en 2018 et 340 m en 2019. Afin de comparer les résultats de ce travail avec les précédents travaux obtenus à l'aide du satellite Formosat-2, nous avons choisi de garder une distance de $d=340$ m.

Concrètement, la validation croisée sera composée de 42 plis (soit 42 jeux de validation et d'entraînement) car la classe ayant le plus faible effectif, celle du cyprès, est composée de 42 échantillons. Chacun des 42 plis est composé de quatorze échantillons de validation (un pour chaque classe) et d'un nombre variable d'échantillons d'entraînement. Pour chaque essence, après avoir sélectionné de manière aléatoire un premier échantillon, une zone de non-sélection de 340 m est appliquée, les échantillons plus loin et donc peu corrélés spatialement sont utilisés pour l'apprentissage (pour plus de détails, voir 2.3.3). Le résultat des 42 prédictions est moyenné pour avoir une estimation globale de l'erreur.

Afin d'estimer la qualité des modèles et la qualité de chaque caractéristique (i.e. date), nous avons choisi d'utiliser le score F1 car il a la particularité de pouvoir se calculer à la fois par classe, mais aussi en le moyennant pour les classes désirés il peut donner une estimation globale de la qualité (il s'agit alors de la moyenne des F1, pour rappel, voir Section 1.3.4.3). Cet indice se calcul en faisant la moyenne harmonique entre la précision et le rappel :

$$F1 = 2 \cdot \frac{\text{précision} \cdot \text{rappel}}{\text{précision} + \text{rappel}} \quad (4.1)$$

Pour rappel, dans l'équation 4.1 la précision (ou "*précision utilisateur*") se calcule comme le ratio entre le nombre d'échantillons bien classés par rapport à l'ensemble des prédits de la classe. Le rappel (ou "*précision producteur*") correspond au ratio entre le nombre d'échantillons bien classés par rapport à tous les observés de la classe. Le F1-score varie donc de 0 (le score le plus faible possible) à 1, le score le plus élevé.

4.2.4.3 Équilibre des effectifs par classe

Comme notre jeu de données représente la réalité des forêts du site d'étude, le nombre de références par espèce est donc inégal. Les deux essences les plus présentes que sont le chêne (*Quercus petraea/pubescens/petraea*) et le pin laricio (*Pinus nigra subsp. laricio*) ont plus de 200 échantillons chacun, à la différence du saule (*Salix alba*) et du cyprès (*Cupressus*) qui en ont quatre fois moins, soit une cinquantaine chacun. Pour chaque espèce, nous avons alors décidé de sous-échantillonner l'effectif pour qu'il corresponde à celui de l'essence la moins peuplée et ce en utilisant leur valeur spectrale [Jain et al., 1999]. Cela évite une interprétation potentiellement erronée des différences de F1 car cet indice est dépendant de la prévalence (tout comme l'accord global ou l'indice Kappa). Une approche très simple aurait été de diminuer l'effectif de l'entraînement en sélectionnant aléatoirement le même nombre

d'échantillons que celui du cyprès qui est la classe la plus faiblement peuplée. Compte tenu de la richesse de l'information et du faible nombre d'échantillons par classe, il nous est apparu crucial de ne pas perdre certaines données qui pouvaient être faiblement représentées mais très importantes pour que le modèle puisse bien apprendre. Nous avons alors opté pour une approche utilisant un K-means pour regrouper les échantillons ayant des signatures spectrales similaires [Hartigan and Wong, 1979]. Cette technique est recommandée et fonctionne très bien avec des jeux d'entraînement de petite taille [Lin et al., 2017].

Concrètement, le frêne (*Fraxinus excelsior*) avec ses 130 échantillons sera réduit à 42 échantillons, soit autant que l'essence la moins bien représentée qu'est le cyprès (*Cupressus*) ($n=42$). L'algorithme de sous-échantillonnage va alors générer un clustering avec 42 groupes pour cette essence ($k = 42$) et utilisera les valeurs spectrales des 42 clusters comme nouvelles valeurs d'échantillons pour le frêne. Cette opération est répétée pour chaque essence. Cette procédure a été réalisée à partir de la bibliothèque python *imbalanced-learn* [Lemaître et al., 2017].

4.2.4.4 Sélection des meilleures caractéristiques

Pour estimer l'apport de chaque date à la qualité du modèle, nous avons opté pour une sélection de caractéristiques de type *Sequential Feature Selection* (SFS) [Chouaib, 2011; Whitney, 1971]. Afin de faciliter la compréhension des algorithmes, lors de la sélection de caractéristiques à partir de la SITS, une caractéristique correspond à l'ensemble des 10 bandes spectrales d'une même date. Tout au long de ce document, le terme caractéristiques sera donc synonyme de dates. L'ajout des caractéristiques se fait de manière itérative, une à une, et la meilleure caractéristique (définie selon le score F1) est ajoutée à un ensemble initialement vide. Ce protocole a été utilisé de manière systématique pour évaluer les performances du modèle à prédire les différentes essences selon une difficulté d'apprentissage croissante :

1. **Chaque essence de feuillu est classée contre l'ensemble des essences.** La date qui obtient le meilleur score F1 pour l'essence en question sera sélectionnée. L'enjeu est d'identifier les meilleures dates pour chaque essence est de faire le lien avec les suivis phénologiques in situ.
2. **Les quatorze essences sont classées simultanément** et le meilleur score F1 est gardé comme critère de sélection de la date. L'enjeu est de cartographier l'ensemble des essences du mieux possible et de voir si les dates correspondent à des événements phénologiques majeurs sont choisies en priorité.

La Figure 4.7 résume l'ensemble de la méthode utilisée pour réaliser ce chapitre.

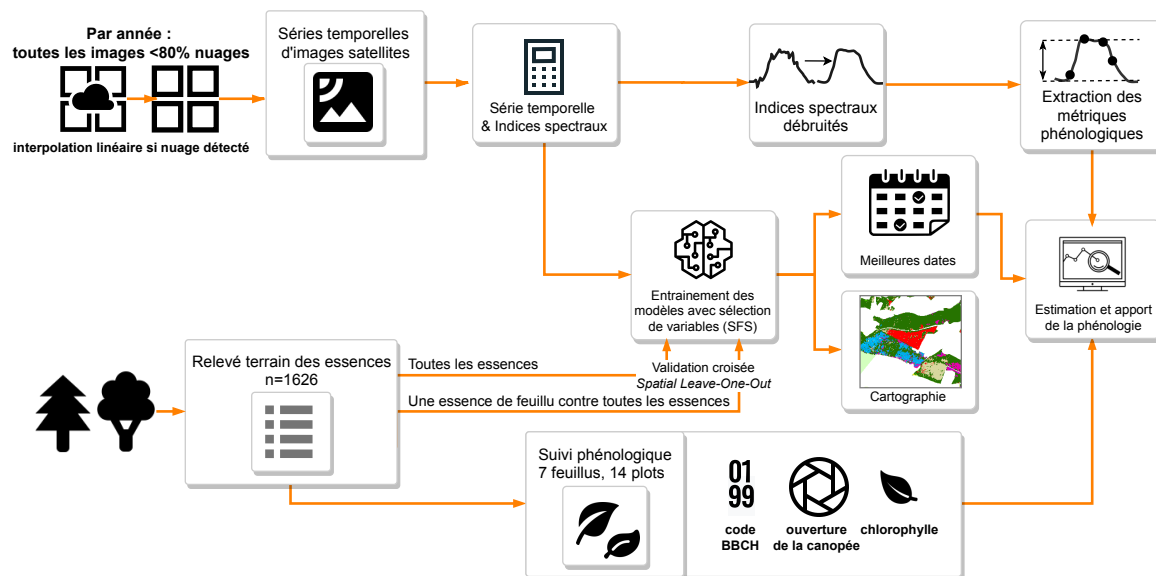


FIGURE 4.7 – Résumé de l'ensemble de la méthode pour valider la détection des phases phénologiques et estimer son l'apport dans la cartographie semi-automatique des espèces forestières.

4.3 Résultats des suivis phénologiques

Pour rappel, les deux stades phénologiques analysés sont le début et la fin de saison. Afin de détailler au maximum les analyses, nous nous sommes concentrés sur trois pourcentages d'accomplissement, à savoir 10%, 50% et 90% (BBCH \geq 11, BBCH \geq 15 et BBCH \geq 19 pour le développement des feuilles et BBCH \geq 91, BBCH \geq 95 et BBCH \geq 99 pour la sénescence). Les résultats des suivis phénologiques se basent sur les quatorze parcelles observées.

4.3.1 Les stades phénologiques observés sur le terrain

4.3.1.1 Début de saison

De manière générale le développement des feuilles est très rapide et les différences intraspécifiques peuvent être nulles. On note une grande stabilité entre 2018 et 2019 dans les dates de débuts de début de saison (Figure 4.8a,4.8b). Que ce soit pour l'apparition des premières feuilles (BBCH=11) ou lorsqu'au moins 50% des feuilles sont déployées (BBCH=15), l'écart maximale entre essences est de 20 jours (Figure 4.8).

Les essences les plus précoces sont le peuplier (*Populus spp.*), le saule (*Salix alba*) et le bouleau (*Betula pendula*) avec un début de saison mi-avril. Le frêne (*Fraxinus excelsior*) fait aussi partie des essences les plus précoces mais en 2018 uniquement, en 2019 on observe un retard de 10 jours.

Les essences les plus tardives sont le chêne rouge (*Quercus rubra*) et le robinier (*Robinia pseudoacacia*) avec un début de saison (BBCH=11) mi-avril en 2018 et un milieu de début de saison (BBCH=15) tout début mai pour les deux années. Le chêne rouge (*Quercus rubra*) est systématiquement la dernière essence à sortir ses feuilles (BBCH=11), et passe en moins de dix jours de quelques feuilles sorties à plus de la moitié des feuilles déjà déployées.

Globalement on note une certaine stabilité entre les deux années observées, avec un début de saison pour les essences les plus précoces estimé vers mi-avril et pour les essences les plus tardives comme le chêne rouge (*Quercus rubra*) et le robinier (*Robinia pseudoacacia*) vers fin avril ou tout début mai.

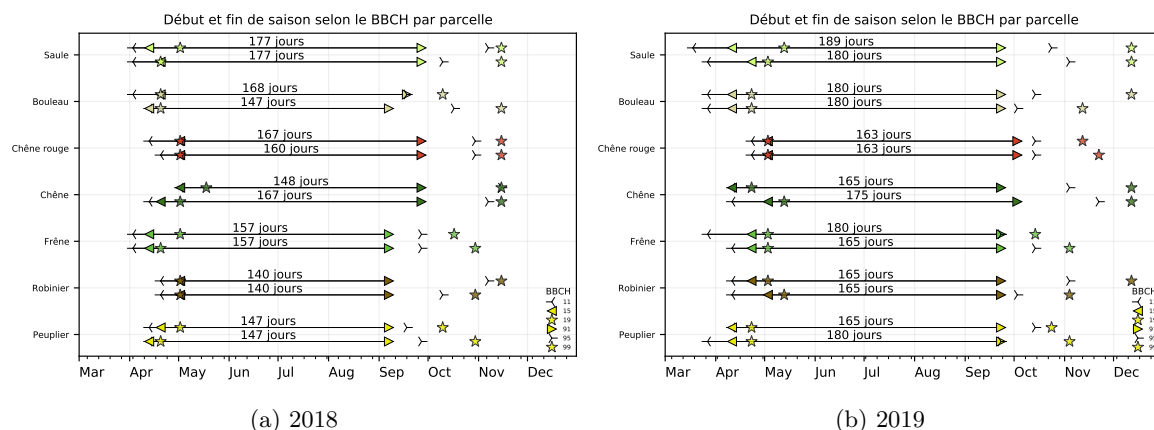


FIGURE 4.8 – Résumé des dates de débuts de saison (BBCH=11, 15 et 19) et des fins de saisons (BBCH=91, 95 et 99) observées sur le terrain pour les quatorze parcelles suivies en 2018 et 2019. La longueur de la saison a été calculée entre la date où le BBCH ≥ 91 a été observée pour la première fois dans l'année et la date du BBCH ≥ 11 [Kalvane et al., 2009; Meier, 1997].

4.3.1.2 Fin de saison

Du début de sénescence

Le tout début de la fin de saison (BBCH=91) se passe de manière assez synchrone entre les essences et il n'y a globalement que 20 jours qui séparent les essences les plus précoces des plus tardives en 2018, et même pas 10 jours pour 2019. Le chêne rouge (*Quercus rubra*) est, comme pour le début de saison, la dernière essence à entamer ce stade phénologique. On peut noter que le peuplier (*Populus spp.*), le robinier (*Robinia pseudoacacia*) et le frêne (*Fraxinus excelsior*) ont vu leur début de sénescence passer de début septembre en 2018 à fin septembre en 2019.

Vers la fin de la sénescence

Si pour le tout début de la fin de saison il y avait peu d'écart entre les essences, les écarts sont beaucoup plus importants quand au moins la moitié des feuilles est tombée ou s'est colorée (BBCH=95).

En effet certaines espèces ont déjà au début du mois de septembre bien entamé leur fin de saison (*Populus spp.*, *Fraxinus excelsior* et *Betula pendula*), et ce pour les deux années, alors que d'autres espèces comme le saule (*Salix alba*) ou le chêne (*Quercus petraea*) la commenceront seulement à partir d'octobre voire novembre pour cette dernière espèce. Il est à noter que les deux parcelles de chêne sessile (*Quercus petraea*) ont un comportement différent des autres essences car cette espèce est marcescente, ce qui signifie que la plupart des feuilles mortes restent sur l'arbre pendant l'hiver, et ne tombent que quand les nouvelles pousses arrivent au printemps.

Le peuplier (*Populus spp.*), le frêne (*Fraxinus excelsior*) et le bouleau (*Betula pendula*) sont pour les deux années 2018 et 2019 les essences qui ont terminé leur saison le plus tôt relativement aux autres feuillus (Figure 4.8).

Globalement, on peut noter que des différences importantes existent pour les mêmes peuplements lors des relevés effectués sur ces deux années. Par exemple, quand une parcelle de peuplier (*Populus spp.*) était notée pour le milieu de la fin de saison (BBCH ≥ 95) mi-septembre en 2018, cette même parcelle arrivait au même stade l'année suivante avec un mois supplémentaire (mi-octobre). Il en est aussi de même pour un saule (*Salix alba*) qui est passé pour ce même stade de début octobre en 2018 à début novembre en 2019.

4.3.2 Longueur de saison de croissance

En calculant la différence entre la date de fin de saison (BBCH=91) et la date de début (BBCH=11) nous obtenons pour chaque peuplement la longueur de sa saison de croissance [Kalvane et al., 2009; Meier, 1997]. Pour 2018 (Figure 4.8a) la saison la plus courte concerne les robiniers (*Robinia pseudoacacia*, 140 jours), suivi des peupliers (*Populus spp.*, 147 jours). Le saule (*Salix alba*) a incontestablement la saison la plus longue avec 177 jours en 2018, suivi de près par un peuplement de bouleaux (*Betula pendula*), chênes rouge (*Quercus rubra*) et chênes (*Quercus petraea*) qui pour certains atteignent les 167 jours.

Si des changements sont notables en 2019 par rapport à l'année 2018, le saule (*Salix alba*) a toujours la saison la plus longue (180 à 189 jours). Le robinier (*Robinia pseudoacacia*) qui n'avait que 140 jours de saison est toujours l'essence dont l'écart entre le début et la fin de saison est le plus court mais avec 165 jours en 2019.

À l'exception des bouleaux (*Betula pendula*) et des chênes (*Quercus petraea*), il n'a pas été observé plus de 10 voire 15 jours d'écart dans la longueur de saison entre les peuplements d'une même essence. Ces différences importantes ont toutefois été observées pour l'année 2018 pour ces deux essences (Figure 4.8a).

4.3.3 Les stades phénologiques par satellite

4.3.3.1 Début de saison

L'ensemble des figures détaillant les résultats selon la taille de fenêtre de Savitzky-Golay et selon l'amplitude des indices est disponible en annexe C.3. Le Tableau 4.4 récapitule par indice et pour chacun des trois pourcentages d'accomplissement du début de saison (BBCH=11, BBCH=15 et BBCH=19) la meilleure combinaison pour chaque indice spectral pour détecter le début de saison en 2018 et 2019 pour les quatorze parcelles étudiées.

Sur les trois pourcentages d'accomplissement du stade de développement des feuilles du BBCH à deviner pour les années 2018 et 2019, le NDVI a les meilleures performances pour la moitié des résultats (le stade BBCH=11 de 2018 est prédit avec 5,1 jour d'erreurs par rapport au LChloC qui l'estime 5 jours, ce qui est non significatif). L'indice NDVI a de très bonnes performances pour prédire l'ensemble des stades du début de saison, et une grande stabilité entre les deux années pour prédire la fin du début de saison (BBCH>=19) avec une erreur moyenne absolue de l'ordre de 6,5 jours (Figure 4.9). Ces résultats s'obtiennent avec une taille de fenêtre de 3 de Savitzky-Golay et une amplitude de 85%, et ce que ce soit pour 2018 et 2019, ce qui montre une bonne stabilité des paramètres (Annexe C.16a et C.16b). Il est à noter que comme l'indice NDVI ne sature pas (il n'atteint jamais la valeur 1), il n'y a donc aucune différence de résultats avec l'indice ACORVI.

TABLEAU 4.4 – Erreur moyenne absolue (MAE) en jours, calculée en comparant la date détectée in situ avec le meilleur seuil à partir des indices ACORVI, LAnthoC, LCaroC, LChloC et NDVI. La date correspond à la première fois où les stades BBCH>=11, BBCH >=15 et BBCH>=19 ont été observés lors des observations in situ. Les chiffres en gras correspondent au meilleur MAE par stade BBCH.

Indice	2018			2019		
	BBCH=11	BBCH=15	BBCH=19	BBCH=11	BBCH=15	BBCH=19
ACORVI	5.1	3.9	6.5	10,2	8.4	6.4
LAnthoC	9.9	8	11.1	6.9	4.6	8.1
LCaroC	9.5	7.6	9.9	8.1	3.9	8.2
LChloC	5	3.4	6.8	8.5	4.9	7.1
NDVI	5.1	3.9	6.5	10,2	8.4	6.4

Parmi les nouveaux indices utilisant les bandes du red-edge de Sentinel-2 (i.e. LAnthoC, LCaroC et LChloC), le LChloC (*Leaf Chlorophyll Content*) tire son épingle du jeu pour prédire le stade BBCH=15 avec moins de 5 jours d'erreur que ce soit pour 2018 ou 2019.

Avec une taille de fenêtre de Savitzky-Golay de 13 et 40% de l'amplitude du début de saison, le LChloC prédit à 5,5 jours près le code BBCH=15 (Annexe C.11a et C.11b). Le LCaroC et LAnthoC peuvent aussi obtenir sur un stade particulier de très bons résultats mais ils sont plus instables que le NDVI, surtout au regard des paramètres à fixer entre l'amplitude et la taille de la fenêtre du lissage.

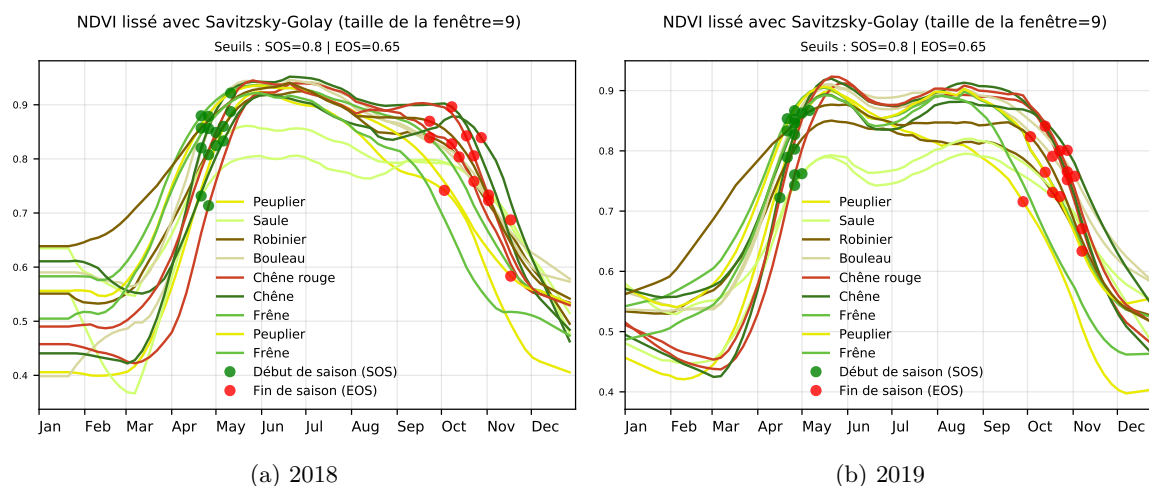


FIGURE 4.9 – Série temporelle du NDVI avec un lissage de Savitzky-Golay utilisant une taille de fenêtre de 9 avec un seuil de début de saison fixé à 80% de l'amplitude et 65% pour la fin de saison. Ce paramètre permet de détecter le début de saison (BBCH=19) avec une erreur de l'ordre de 6,5 jours.

4.3.3.2 Fin de saison

Le Tableau 4.5 récapitule les performances des indices spectraux pour prédire les stades de la sénescence observés sur le terrain (BBCH=91, BBCH=95 et BBCH=99). Cela signifie que l'on a observé soit au minimum 10%, 50% ou au moins 90% des feuilles qui étaient colorées ou tombées lors de la visite de la parcelle.

TABLEAU 4.5 – Erreur moyenne absolue (MAE) en jours calculée en comparant la date détectée in situ avec le meilleur seuil à partir des indices ACORVI, LAnthoC, LCaroC, LChloC et NDVI. La date correspondant à la première fois où les stades BBCH>=91, BBCH>=95 et BBCH>=99 ont été notés lors des observations in situ. Les chiffres en gras correspondent au meilleur MAE par stade BBCH.

Indice	2018			2019		
	BBCH=91	BBCH=95	BBCH=99	BBCH=91	BBCH=95	BBCH=99
ACORVI	11.1	9.8	8.2	9.3	8.8	11.6
LAnthoC	6.8	10,1	9.7	10,4	7.3	9.5
LCaroC	8.1	10,7	10,6	8.6	7.9	7.4
LChloC	11.8	10,6	9.2	12.1	7.8	8.9
NDVI	11.1	9.8	8.2	9.3	8.8	11.6

En utilisant les meilleurs paramètres (combinaison d'une taille de la fenêtre de Savitzky-Golay et d'un seuil entre 15 et 85% de l'amplitude de la fin de saison), nous notons que les résultats sont légèrement moins précis qu'avec le début de saison. Il n'y a pas un indice qui reste sous les 10 jours d'erreur moyenne absolue lors de deux années et des trois pourcentages d'accomplissement de la fin de saison.

Le NDVI ou l'ACORVI, qui ont les mêmes performances, prédisent avec environ 10 jours d'erreurs les stades de sénescence des deux années. Ces indices sont toujours aussi stables et ne nécessitent pas un paramétrage très précis (Annexe C.3).

Si l'on se focalise sur les meilleures performances, l'indice de caroténoïdes (LCaroC) est le seul à prédire avec moins de dix jours d'erreur le tout début de la sénescence (BBCH=91, Figure 4.10). Cependant, comme les autres indices à l'exception du NDVI, il faut une fenêtre très précise de paramètres (Figure C.18a et C.18b).

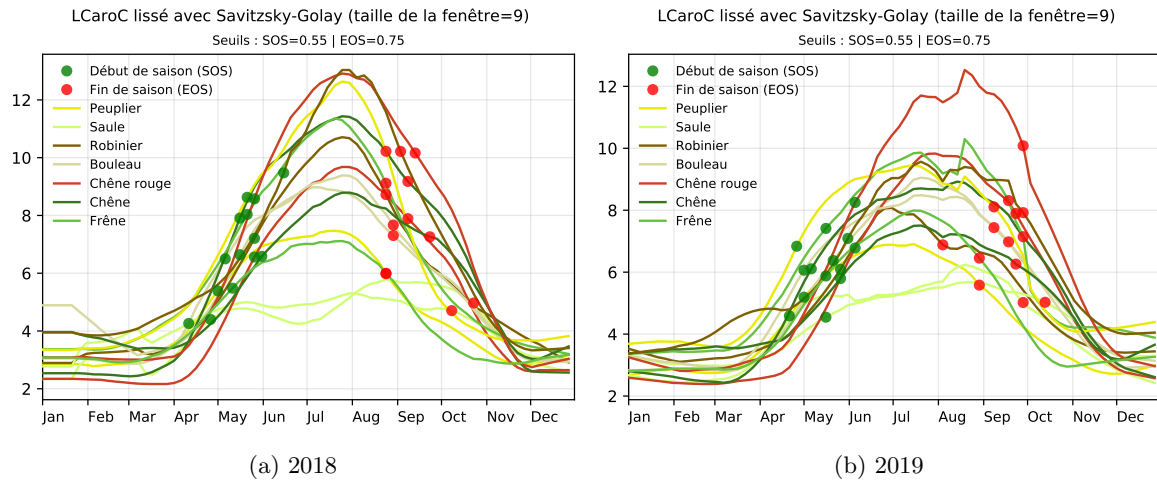


FIGURE 4.10 – Série temporelle du LCaroC avec un lissage de Savitzky-Golay utilisant une taille de fenêtre de 9 avec un seuil de début de saison fixé à 55% de l'amplitude et 75% pour la fin de saison. Cette amplitude permet de prédire le BBCH=91 avec 8,6 jours d'erreurs en 2018 et 8,9 jours en 2019.

4.4 Résultats de l'apport des séries temporelles pour la détection des essences

Pour rappel, l'apport de séries temporelles d'images satellites dans les modèles d'apprentissage automatique a été évalué en utilisant la série brute mais aussi les indices spectraux qui découlent des années 2017, 2018 et 2019. Nous détaillerons tout d'abord l'apport de la sélection des dates pour chaque feuillu, puis nous analyserons si les dates sélectionnées pour prédire l'ensemble des espèces (feuillus et résineux) ont un lien avec des événements phénologiques observés in situ.

4.4.1 Classification : Un feuillu contre l'ensemble des essences

4.4.1.1 Analyse générale

Le Tableau 4.6 résume la meilleure qualité qu'il a été possible d'obtenir pour identifier chaque feuillu à partir des 5 meilleures dates (Tableau 4.6a), ou en utilisant l'entièreté de la série temporelle (Tableau 4.6b). Le choix des 5 meilleures dates s'est effectuée car que ce soit pour classer une seule essence comme dans cette partie ou pour classer l'ensemble des essences il n'y a pas d'amélioration significative au-delà.

Avec sélection des meilleures dates

Afin de bien comprendre l'apport des dates dans la classification des feuillus, chaque essence a été classée à l'aide d'une sélection de caractéristiques. Pour ne pas perdre la complexité des modèles d'apprentissage, le critère de choix de la meilleure date est le F1 de l'essence désirée. Cela signifie que l'apprentissage se fait toujours avec les quatorze essences mais que seul le résultat de l'essence à prédire comptera lors de la sélection de la date.

Les espèces les mieux prédites sont tout d'abord le chêne rouge (*Quercus rubra*) avec un score F1 moyen de 0,99, puis le saule (*Salix alba*) avec un score de 0,96.

TABLEAU 4.6 – Score F1 maximum obtenu avec un modèle par essence en utilisant une validation croisée de type SLOO pour l'année 2017, 2018 et 2019 et à l'aide soit des 5 meilleures dates (via une sélection de caractéristiques) ou à l'aide de la série temporelle dans son ensemble. En gras sont spécifiés les meilleurs résultats par essence.

(a) Avec sélection de caractéristiques (5 meilleures dates)								
Espèce	Peuplier	Robinier	Frêne	Chêne	Chêne rouge	Bouleau	Saule	Moyenne
2017	0,86	0,63	0,79	0,86	0,97	0,8	0,99	0,84
2018	0,93	0,78	0,81	0,87	1	0,94	0,93	0,89
2019	0,83	0,67	0,78	0,8	1	0,88	0,95	0,84
Moyenne	0,87	0,69	0,79	0,85	0,99	0,87	0,96	

(b) Sans sélection de caractéristiques (SITS entière)								
Espèce	Peuplier	Robinier	Frêne	Chêne	Chêne rouge	Bouleau	Saule	Moyenne
2017	0,78	0,28	0,61	0,63	0,96	0,18	0,93	0,62
2018	0,73	0,32	0,62	0,55	0,9	0,46	0,75	0,61
2019	0,71	0,42	0,77	0,56	0,96	0,65	0,76	0,69
Moyenne	0,74	0,34	0,66	0,58	0,94	0,43	0,81	

Ensuite les espèces comme le peuplier (*Populus spp.*), le bouleau (*Betula pendula*), le frêne (*Fraxinus excelsior*) ou le chêne (*Quercus petraea*) sont en moyenne vers 0,87 pour les deux premiers puis 0,85 pour les deux derniers.

Enfin l'essence qui pose le plus de complications est le robinier (*Robinia pseudoacacia*) avec un score F1 moyen de 0,69 sur trois années.

Si l'on se focalise sur les années, 2018 semble être la meilleure année avec 0,05 points de plus que 2017 et 2019, soit une moyenne des F1 des feuillus de 0,89 contre 0,84 pour les autres années.

Sans sélection des meilleures dates

L'utilisation de la sélection de caractéristiques a permis d'augmenter significativement la qualité des modèles comparé à l'utilisation de la série temporelle dans son ensemble. Alors que le score moyen des F1 des feuillus par année oscillait entre 0,84 et 0,89 avec les 5 meilleures dates, les performances ont baissé entre 0,15 et 0,28 points selon les années (Tableau 4.6).

Les performances du chêne rouge (*Quercus rubra*) restent très élevées avec une moyenne de F1 de 0,94 sur les 3 années. Le saule (*Salix alba*) qui avait des performances très proches du chêne rouge (*Quercus rubra*) avec la sélection de dates subit une perte de F1 de 0,15 points en moyenne.

Le robinier (*Robinia pseudoacacia*) a comme avec la précédente méthode le plus mauvais score F1 (F1=0,34) lorsque l'ensemble de la série temporelle était utilisée pour apprendre leur modèle dédié (Tableau 4.6). Le bouleau (*Betula pendula*) qui a été très bien reconnu avec les 5 meilleures dates (F1=0,87) est la deuxième essence la moins bien détectée après le robinier en utilisant toute la série temporelle (F1=0,43). Pour ces deux essences, réduire la dimension de la donnée et sélectionner les dates les plus significatives a permis d'augmenter la qualité de prédiction de plus de 0,35 points.

4.4.1.2 Importance des dates et des saisons

Parmi les 37 dates disponibles en 2017 et 2018 (soit 370 bandes spectrales), et les 32 dates disponibles en 2019, il faut en moyenne entre deux et maximum cinq dates pour prédire avec le niveau maximal de qualité chaque espèce.

En 2018, meilleure année pour prédire les feuillus avec une sélection de caractéristiques, on observe que les dates les plus importantes pour chaque espèce sont essentiellement soit au printemps (fin avril à début mai), soit en novembre (Figure 4.11). Par exemple pour le

4.4. RÉSULTATS DE L'APPORT DES SÉRIES TEMPORELLES POUR LA DÉTECTION DES ESSENCES

début de saison, le 11 mai est la date la plus importante pour classer le peuplier (*Populus spp.*), et pour le robinier (*Robinia pseudoacacia*) il s'agit du 21 avril. En fin de saison, le mois de novembre est très prisé comme premier choix avec le 12 novembre pour le chêne rouge (*Quercus rubra*) et pour le frêne (*Fraxinus excelsior*), le 22 novembre pour le bouleau (*Betula pendula*) et le 27 novembre pour le chêne (*Quercus petraea*).

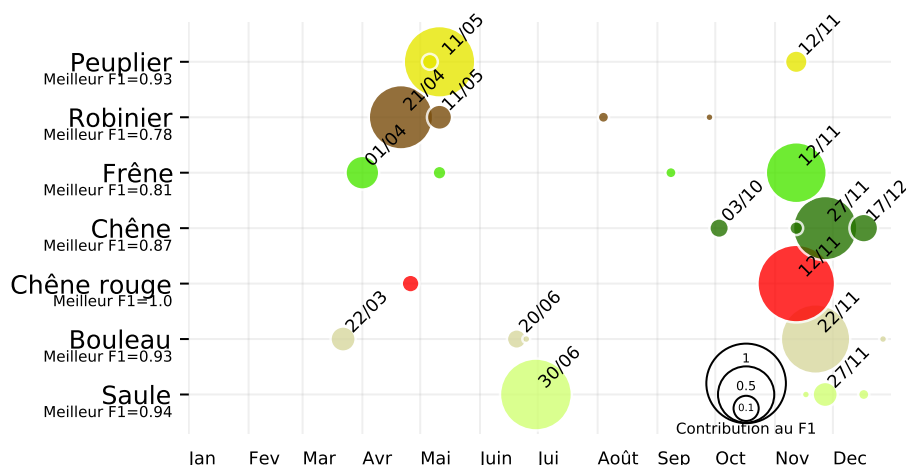


FIGURE 4.11 – Liste des 5 meilleures dates de la série temporelle Sentinel-2 de 2018 pour classer séparément chaque essence de feuillus.

Les dates sélectionnées en 2018 sont très semblables à celles de 2017 (Figure 4.12) avec début novembre (2 et 7 novembre) comme dates les plus importantes pour la totalité des essences à l'exception du robinier (*Robinia pseudoacacia*) au 8 octobre et du saule (*Salix alba*) au 5 juillet.

Le saule (*Salix alba*) est la seule essence étudiée qui a la même période sélectionnée pour les trois années étudiées (voir les résultats de l'année 2019 à la figure 4.13). En effet fin juin ou tout début juillet est l'unique date nécessaire pour classer cette espèce avec un score F1 de plus de 0,8.

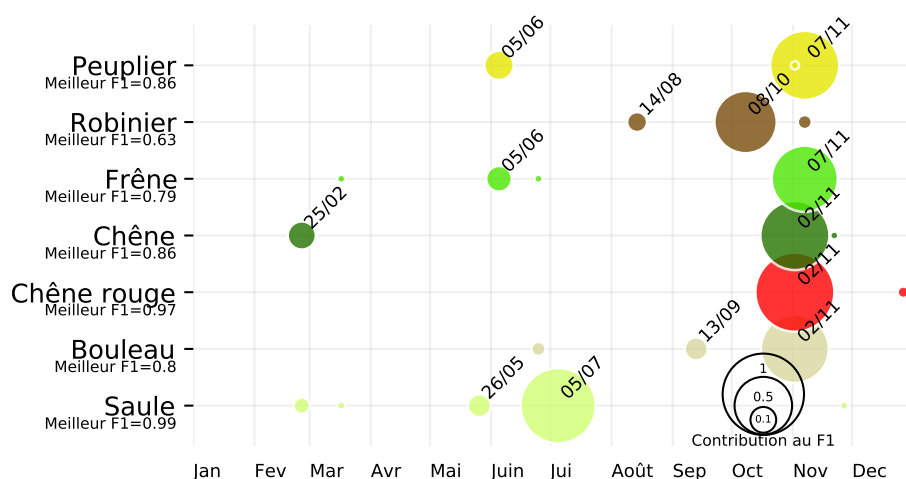


FIGURE 4.12 – Liste des cinq meilleures dates de la série temporelle Sentinel-2 de 2017 pour classer séparément chaque essence de feuillus.

L'année 2019 est singulière car elle contient moins de dates que les autres années (32 contre 37) et que quasiment aucune date n'a été sélectionnée en novembre, si ce n'est le 27 novembre pour détecter le peuplier (*Populus spp.*). À cette date, nos relevés terrain montrent que la saison était déjà terminée depuis plus d'une semaine et qu'il ne restait presque aucune feuille

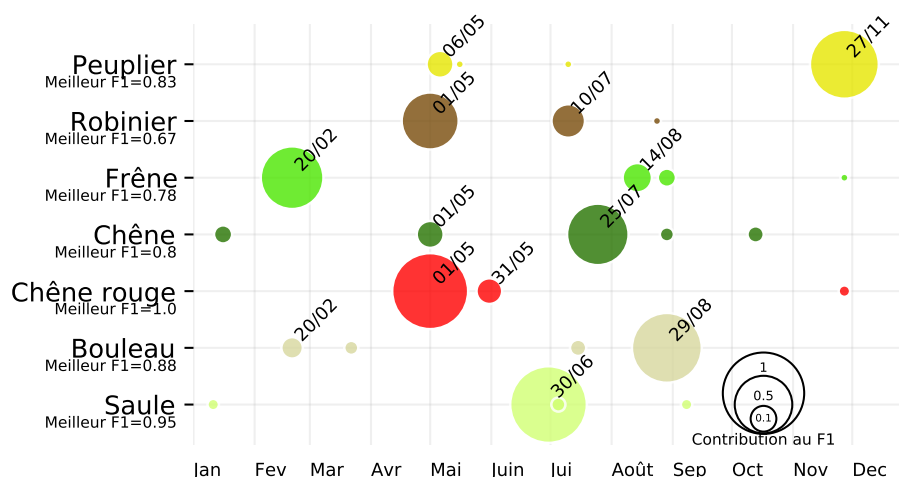
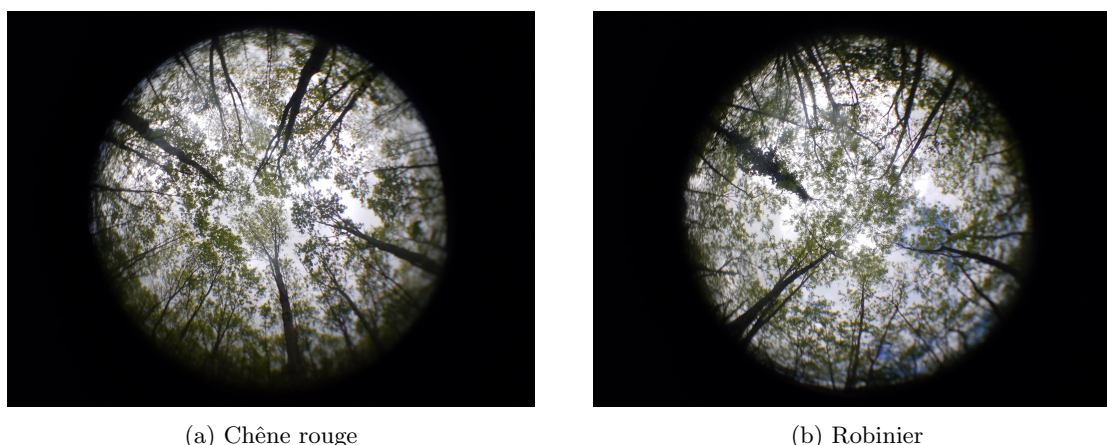


FIGURE 4.13 – Liste des cinq meilleures dates de la série temporelle Sentinel-2 de 2019 pour classer séparément chaque essence de feuillus.

sur les arbres des parcelles étudiées cinq jours avant (Figure 4.15), ce qui montre que ce n'est pas le signal de la canopée qui a servi à l'algorithme pour différencier les peupliers des autres essences. Concernant le printemps 2019, la date du 1er mai est la plus importante pour le chêne rouge (*Quercus rubra*) ou le robinier (*Robinia pseudoacacia*). Les feuilles étaient toutes dépliées pour ces deux essences (Figure 4.14), ce qui semble confirmer que la phénologie a joué un rôle important dans la sélection de la date. Pour le bouleau (*Betula pendula*) ou le chêne (*Quercus petraea*) il s'agit désormais de dates estivales (respectivement le 29 août et 25 juillet pour la date la plus importante).

Enfin, le cas du frêne (*Fraxinus excelsior*) est intéressant, car les dates les plus importantes sont systématiquement pour 2018 et 2019 des dates avant le début de la saison de végétation (respectivement février ou tout début avril) ou après la fin de saison (12 novembre 2018), ce qui semble signifier que ce qui permet à l'algorithme de différencier cette essence n'est pas lié au signal de l'arbre et de ses feuilles mais à d'autres éléments annexes (sous-étage, structure...).



(a) Chêne rouge

(b) Robinier

FIGURE 4.14 – Photo hémisphérique de la canopée du robinier (*Robinia pseudoacacia*) et du chêne rouge (*Quercus rubra*) au 3 mai 2020, L'ensemble des feuilles était sorti (BBCH>19).

4.4.1.3 Classification à partir d'indices spectraux

Pour comparer la performance des indices, il a été classé chaque essence à l'aide d'un modèle dédié et en utilisant les cinq meilleures dates obtenues à partir d'une sélection de



FIGURE 4.15 – Photo hémisphérique de la canopée du peuplier (*Populus spp.*) au 22 novembre 2019. Les quelques feuilles restant sur les arbres étaient déjà colorées en jaune, ce qui signifie que la saison de végétation était finie.

caractéristiques (Tableau 4.7). Comparativement à la sélection de caractéristiques sur l'ensemble de la série temporelle, les cinq meilleures dates des indices sont moins performantes que les cinq meilleures dates utilisant les 10 bandes spectrales. La part de F1 se situe en moyenne aux alentours de 0,1 selon les indices et selon les années.

TABLEAU 4.7 – F1-score obtenu par essence et par indice en utilisant les 5 meilleures dates. En gras est spécifié le meilleur indice par année et par essence.

	LAnthoC			LCaroC			LChloC			NDVI		
	2017	2018	2019	2017	2018	2019	2017	2018	2019	2017	2018	2019
Peuplier	0.7	0.57	0.54	0.63	0.58	0.54	0.69	0.44	0.63	0.69	0.58	0.67
Robinier	0.61	0.57	0.4	0.6	0.52	0.51	0.55	0.52	0.6	0.57	0.64	0.57
Frêne	0.82	0.59	0.55	0.73	0.48	0.53	0.66	0.66	0.43	0.65	0.63	0.5
Chêne	0.67	0.67	0.57	0.8	0.65	0.58	0.76	0.65	0.56	0.71	0.52	0.56
Chêne rouge	0.96	0.89	0.83	0.86	0.97	0.86	0.9	0.93	0.9	0.96	0.92	0.9
Bouleau	0.93	0.76	0.67	0.8	0.68	0.67	0.69	0.73	0.8	0.82	0.78	0.85
Saule	0.95	0.91	0.85	1	0.96	0.95	1	0.94	0.91	0.97	0.94	0.85
Moyenne	0.81	0.71	0.63	0.77	0.69	0.66	0.75	0.7	0.69	0.77	0.72	0.7

En se focalisant sur les indices spectraux, à l'exception de l'indice LAnthoC qui est le plus efficace pour l'année 2017 (F1 moyen des feuillus = 0,81), le NDVI obtient systématiquement les meilleurs scores en moyenne, et se classe devant l'indice LChloC avec 0,02 points pour chaque année, ce qui est très proche. Les écarts sont globalement assez faibles entre indices.

Concernant les essences, le saule (*Salix alba*) et le chêne rouge (*Quercus rubra*) sont toujours les deux espèces les mieux identifiées avec des scores F1 supérieurs en moyenne à 0,9.

Les figures montrant le détail des cinq meilleures dates pour chaque année et chaque feuillu est disponible en annexe C.2. On observe, contrairement à la sélection de dates de la série temporelle, des changements importants selon les années et selon les essences, il est donc difficile de conclure qu'une date précise ou qu'un indice en particulier est plus stable qu'un autre, si ce n'est que le saule (*Salix alba*) est toujours très bien classé avec des dates sélectionnées vers juillet.

4.4.2 Classification de l'ensemble des essences

Comme pour le premier chapitre consacré aux archives du satellite Formosat-2 (Chapitre 3), la Figure 4.16 montre que les résultats avec sélection de caractéristiques permettent d'obtenir une qualité avoisinant les 0,70 d'accord global. Concernant les confusions entre essences, elles se font très majoritairement entre types, les feuillus se confondant avec les feuillus et les conifères avec les conifères. À noter une exception avec l'eucalyptus (*Eucalyptus spp.*) qui est le seul feuillu à garder ses feuilles toute l'année et qui se confond avec les conifères et non les feuillus.

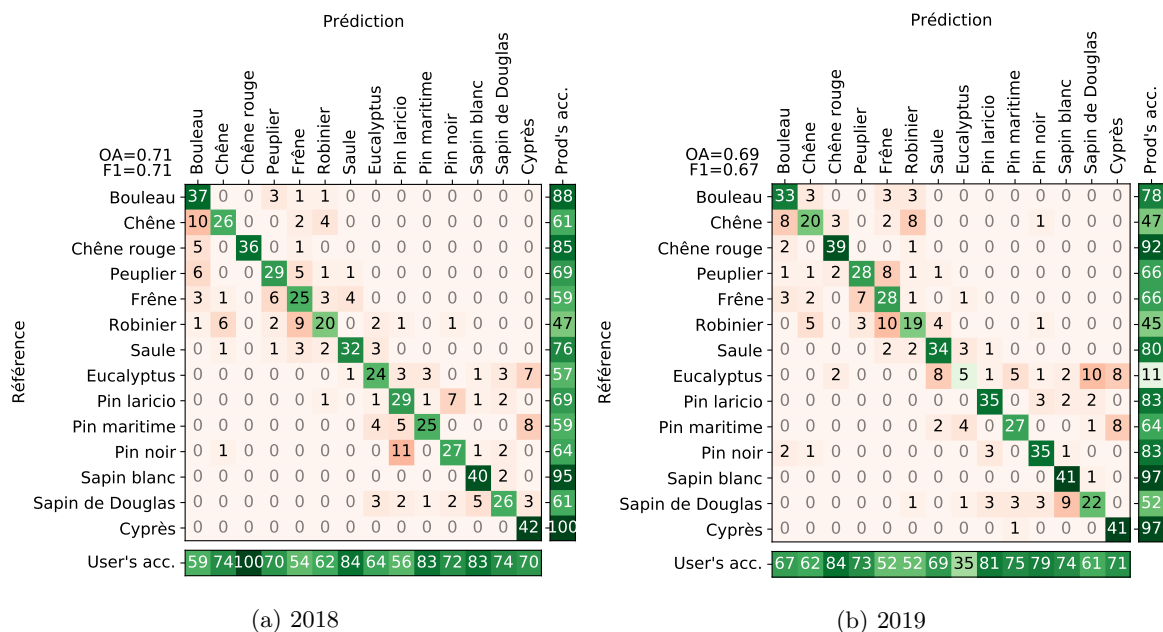


FIGURE 4.16 – Matrice de confusion des résultats obtenus avec les 10 meilleures dates pour les années 2018 et 2019. Le nombre total d'échantillons de référence par essence est de 42.

Parmi les essences les mieux identifiées, on trouve le chêne rouge (*Quercus rubra*) qui se confond légèrement avec le bouleau (*Betula pendula*), alors que le saule (*Salix alba*) se confond avec le frêne (*Fraxinus excelsior*) ou le robinier (*Robinia pseudoacacia*). Du côté des conifères, le sapin blanc (*Abies alba*) est très bien détectée et il se confond légèrement avec le sapin de douglas (*Pseudotsuga menziesii*).

Certaines essences sont surprédites comme le bouleau (*Betula pendula*) qui se confond avec le chêne (*Quercus petraea/pubescens/petraea*), ou le frêne (*Fraxinus excelsior*) qui est prédit comme robinier (*Robinia pseudoacacia*) ou peuplier (*Populus spp.*). L'eucalyptus (*Eucalyptus spp.*), seule feuillu à être à feuillage persistant est souvent omis et sa précision producteur est très faible pour les deux années (57% en 2018 et 11% en 2019).

4.5 Discussion

Il est à noter que les visites sur le terrain ne se faisant que tous les 10 jours, nous ne pouvons pas dire si un stade aurait été atteint un jour ou 9 jours après notre visite. Avoir une précision d'environ 10 jours est donc déjà un très bon résultat.

Plusieurs auteurs ont validé la phénologie détectée par satellite avec des suivis in situ [Beck et al., 2006; Delbart et al., 2006]. Avec 108 références de bouleaux sur la péninsule scandinave, [Beck et al., 2006] ont pu avec 12 jours de décalage au printemps et 10 jours à l'automne détecter le début et la fin de saison. Ces auteurs ont signalé que le signal du NDVI pouvait être influencé plus par le sous-étage que par la canopée, en particulier en altitude où les forêts de bouleaux sont plus éparées.

Dans notre cas, il est difficile de conclure avec peu d'échantillons par essence ($n = 2$ pour chaque espèce). Il serait hasardeux de comparer l'influence de la canopée en mélangeant les essences tant des différences de structure, d'orientation des feuilles, d'âge des plantations peuvent expliquer de grands écarts déjà au sein d'une même essence.

4.5.1 Stabilité des seuils pour détecter les stades phénologiques

4.5.1.1 Début de saison

Le LChloC comme nouvelle référence de début de saison ?

L'indice LChloC a montré qu'il était presque systématiquement l'un des plus performants mais surtout l'un des plus stables, c'est-à-dire que ses performances par année et par stade BBCH (11, 15) sont presque systématiquement inférieures à cinq jours d'erreurs en moyenne (précision de 7 jours pour le stade BBCH=19). Cela peut s'expliquer par une non saturation du signal qui augmente jusqu'en juillet (Figure 4.17) car cet indice n'est qu'un simple rapport entre deux bandes. Le LChloC s'étend sur une gamme de valeur assez large allant en été d'environ 3 pour le saule (*Salix alba*) et le robinier (*Robinia pseudoacacia*) à presque 8 pour le frêne (*Fraxinus excelsior*) et le bouleau (*Betula pendula*). Les meilleurs scores obtenus avec cet indice pour identifier le début de saison (BBCH=11) se font avec un seuil fixé à 0,25 de l'amplitude de la saison, et avec un LChloC lissé avec l'algorithme Savitzky-Golay en utilisant une taille de fenêtre de 13 (Figure C.7a et C.7b). Ce qui est encourageant c'est que pour ces deux années, le même paramètre permet d'estimer au mieux le début de saison. Reste à savoir si pour des années futures ces mêmes paramètres seraient à nouveau les plus optimaux.

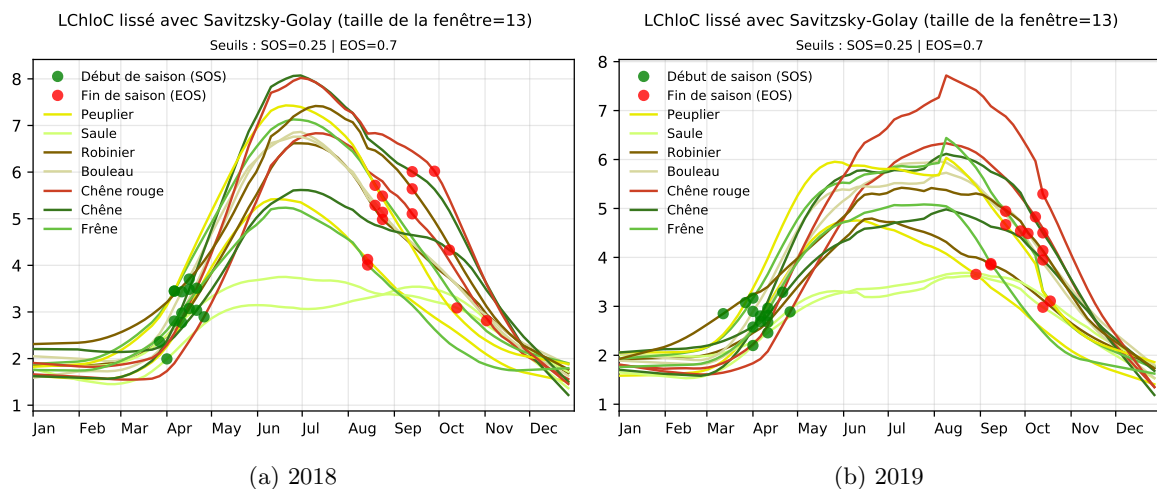


FIGURE 4.17 – Série temporelle du LChloC avec un lissage de Savitzky-Golay utilisant une taille de fenêtre de 13 et un seuil de début de saison fixé à 25% de l'amplitude et 70% pour la fin de saison.

Le NDVI

L'indice NDVI, conformément à la littérature Zeng et al. [2020] montre sa robustesse et son efficacité pour prédire les événements phénologiques. Sa grande force réside dans le fait d'être peu dépendant aux paramètres des seuils et de la taille de la fenêtre du lissage pour détecter le début de saison (Figure C.8a et C.8b). Alors qu'avec un indice comme le LChloC mal paramétré des écarts de plus de 10 jours de précision dans la date du début de saison peuvent apparaître, l'indice NDVI permet une souplesse dans le pourcentage du seuil à fixer. Cela s'explique en grande partie par le fait que le NDVI est un indice très peu étendu dans ses valeurs. En effet cet indice varie de -1 à 1 mais quand il s'agit de végétation active il variera 0,5 et 1 (Figure 4.18), quand le LChloC descend à environ 2 en hiver à plus de 8 en été pour les essences comme le frêne (*Fraxinus excelsior*) ou le bouleau (*Betula pendula*) (Figure 4.17).

Croft et al. [2014] avaient montré que le NDVI saturait lorsque la production de chlorophylle au niveau de l'arbre augmentait, ainsi l'indice n'était plus capable des dizaines de jours à l'avance d'observer la montée. Malheureusement, même si l'indice ACORVI est plus adapté pour des données corrigées atmosphériquement, il n'est pas non plus capable de suivre cette tendance.

En comparant les courbes de l'indice LChloC et de l'indice NDVI pour un pixel de robinier (Annexe C.29m), on remarque bien que le NDVI arrive à son maximum lorsque toutes les feuilles de l'arbre sont sorties (BBCH>19). Or, en observant la courbe du LChloC, le signal continue de s'intensifier pendant plus de 3 mois (de juin à fin août). Ce phénomène s'observe aussi sur la plupart des autres parcelles (Annexe C.29).

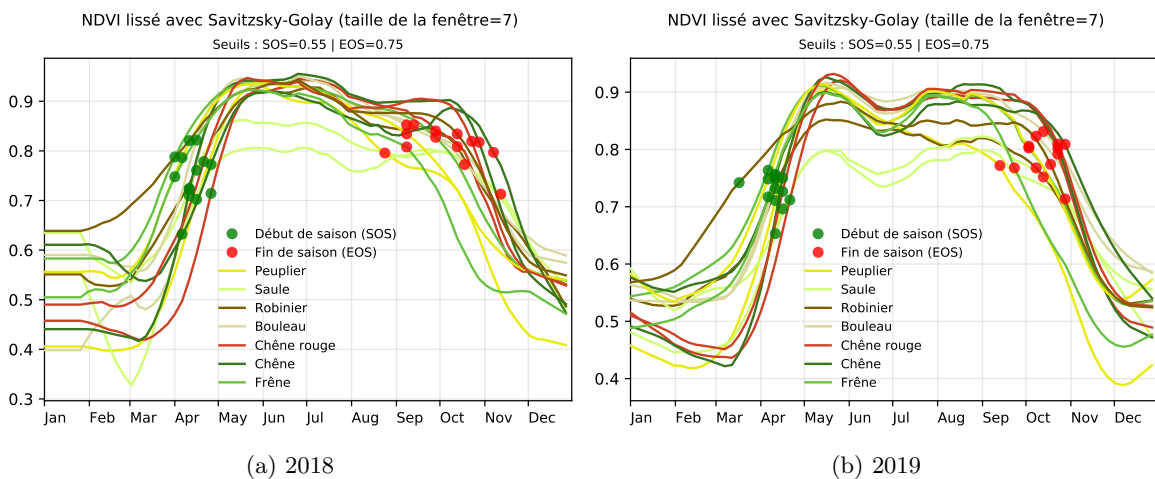


FIGURE 4.18 – Série temporelle du NDVI avec un lissage de Savitzky-Golay utilisant une taille de fenêtre de 7 et un seuil de début de saison fixé à 55% de l'amplitude et de fin de saison à 75%. Ces paramètres permettent de détecter le stade BBCH=19 avec 6,5 jours d'erreur.

À partir d'une fonction de double logistique depuis l'indice *Enhanced Vegetation Index* (EVI) du satellite MODIS Liang et al. [2011] avaient pu observer par télédétection le moment où tous les feuilles d'une forêt composée majoritairement de feuillus sont sorties (FFB pour *Full Bud Burst*, ce qui équivaut au stade BBCH>19). La précision était d'environ deux jours à partir du EVI contre une quinzaine pour le NDVI. Dans notre cas notre précision sur les quatorze parcelles est de 6,5 jours pour les deux années. White et al. [2014] avait montré une erreur d'environ dix jours en utilisant une fonction logistique sur la courbe du EVI à partir du satellite MODIS.

Nos résultats sont donc semblables à la littérature et compte tenu de la résolution temporelle de Sentinel-2, moins de dix jours d'erreurs signifie que l'évènement phénologique a bien été détecté.

4.5.1.2 Fin de saison

Le LCaroC, comme nouvelle référence de fin de saison ?

L'indice LCaroC est le seul des indices qui a pu prédire la fin de saison (BBCH=91) avec moins de dix jours d'erreur pour les années 2018 et 2019 (Tableau 4.5). En analysant les meilleures combinaisons des paramètres de Savitzky-Golay et du seuil concernant l'amplitude de la fin de saison, il s'avère qu'il est possible de garder les mêmes paramètres pour obtenir cette précision. Pour cela la taille de la fenêtre de Savitzky-Golay doit être fixée entre sept et neuf (de préférence à neuf) avec un seuil à 65% de l'amplitude (Figure C.18a et C.18b). Les courbes de l'indice LCaroC par essence à l'aide des meilleurs paramètres ont été tracées afin de faciliter la compréhension (Figure 4.10).

Outre que l'indice LCaroC permet d'obtenir de meilleurs résultats que le NDVI, il est

également beaucoup plus stable avec les mêmes paramètres. En effet, le traditionnel indice NDVI est très dépendant de paramètres précis et malheureusement différents d'une année sur l'autre (Figure C.20a et C.20b). [Lee and Gould \[2002\]](#) montraient que l'anthocyanine était responsable de la couleur rouge des feuilles à l'automne, et l'indice LCaroC semble bien être sensible à cette information.

Nos résultats, surtout pour la fin de saison, sont en accord avec l'étude terrain de [Wong et al. \[2019\]](#) qui ont montré les limites du NDVI pour suivre la phénologie et que les indices de végétation basés sur la teneur en caroténoïde étaient adaptés pour suivre la phénologie du début ou de fin de saison de la végétation.

4.5.2 Influence du sous-étage

À chaque plot et à chaque relevé phénologique, cinq photos de la canopée ont été réalisées avec un fisheye. À partir de ces images hémisphériques, la *Canopy Cover* a été calculé à partir de l'application GLAMA. À l'aide de la photo de la parcelle et du commentaire laissée sur l'état du sous-étage, il est ainsi possible de faire le lien entre l'ouverture de la canopée, l'état d'avancement dans la phénologie de l'essence dominante et du sous-étage. En analysant les résultats, et en comparant les parcelles qui contiennent un sous-étage permanent (comme le fragon), nous n'avons pas observé d'influence particulière de celui-ci. Cependant, comme peut nous le montrer la courbe du LChloC du robinier (Figure C.29a), il semblerait que le sous-étage de graminées très actif de cette parcelle influe sur la signature spectrale. La canopée est à son maximum de fermeture début mai avec environ 50% d'ouverture (indice *Canopy Cover*), mais le LChloC va encore progresser passant de 3,5 début mai à plus de six en août.

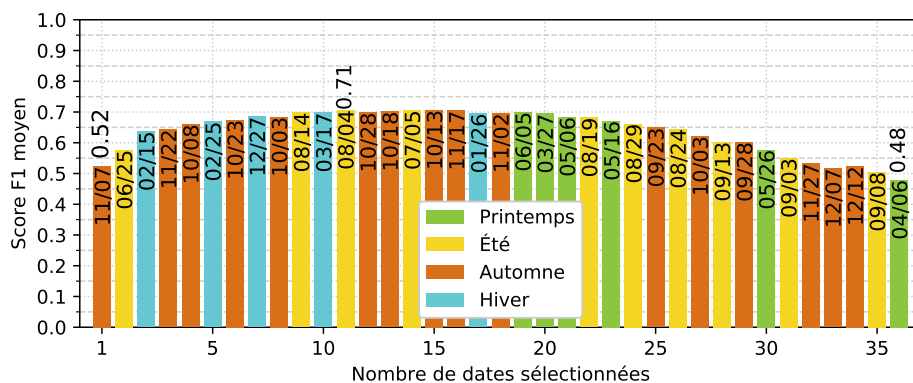
4.5.3 Explication des dates sélectionnées pour l'ensemble des feuillus

L'apport des données multitemporelles avec la sélection de caractéristiques est indéniable pour cartographier les essences forestières. Nos résultats confirment les travaux de [\[Bolin et al., 2018; Grabska et al., 2019; Persson et al., 2018\]](#) et montrent l'importance des dates au printemps avec Sentinel-2. Selon les indices utilisés, nous voyons que pour 2017 les dates du printemps sélectionnées concernent fin mai et début juin, alors que pour 2018 il s'agit plutôt de fin mars et de mi-avril. L'une des explications de la non sélection des dates entre avril et fin mai en 2017 est qu'une seule image très ennuagée était disponible à ces dates. De manière générale, [Hill et al. \[2010\]](#) avaient montré qu'à partir d'un capteur de 2 m de résolution spatiale, la meilleure prédiction était réalisée avec une date mi-mars, une à l'été et une fin octobre, cependant seulement cinq dates étaient disponibles au total. Pour des feuillus de la savane [Madonsela et al. \[2017\]](#) ont montré l'apport des données multitemporelles, la sénescence apparaissant à la transition de la saison humide vers la saison sèche).

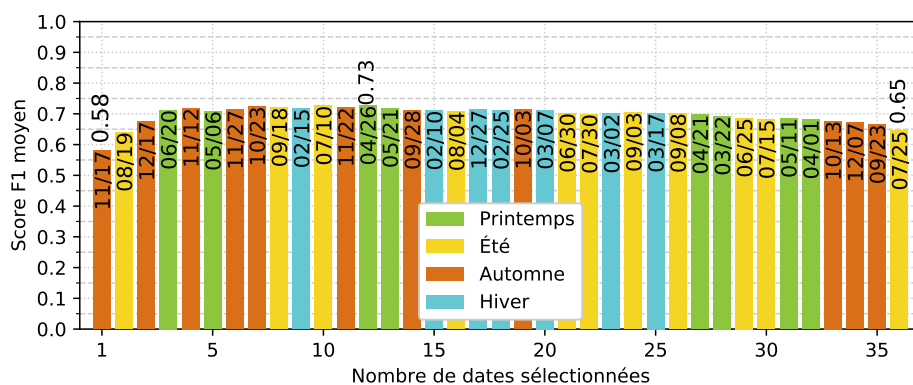
Les mauvaises performances de la série temporelle utilisée dans sa globalité peuvent s'expliquer par du bruit dû à des nuages non détectés et à la dimension de la donnée [\[Hughes, 1968\]](#).

Afin de visualiser l'apport des dates dans la qualité des modèles de chaque année, un graphique résume le résultat des sélections de caractéristiques (Figure 4.19). On observe qu'une date estivale est systématiquement sélectionnée en deuxième position mais que la date la plus discriminante semble être systématiquement durant le mois de novembre. Or, en 2019, il n'y avait aucune image disponible durant ce mois à cause de la couverture nuageuse du territoire étudié. Ainsi en 2019, la date la plus importante n'est donc plus durant le mois de novembre mais durant le mois de février, et le F1 moyen de l'ensemble des essences est de 0,55 en 2019 contre 0,57 en 2017 et 0,52 en 2018. Au final la qualité maximale de l'année est sensiblement la même avec 0,70 en 2019 contre 0,71 et 0,73 en 2017 et 2018. Cela confirme le chapitre 3 qui a montré que si certaines périodes étaient manquantes dans

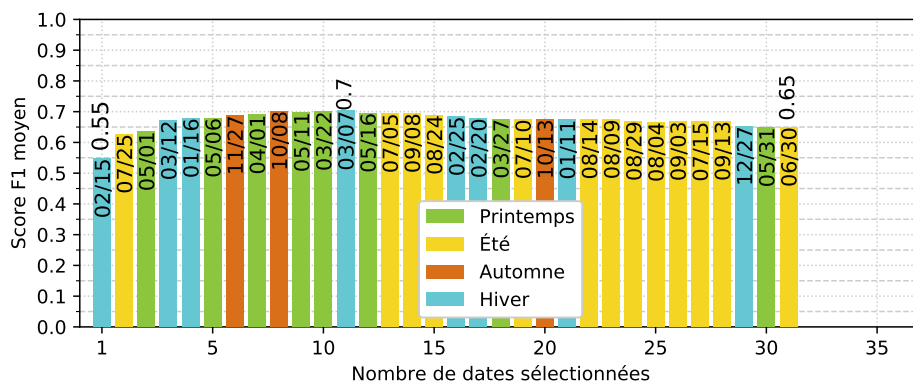
une année, comme l'automne, l'algorithme pouvait atteindre une qualité similaire à d'autres années ayant l'automne mais à l'aide de dates à d'autres saisons.



(a) 2017



(b) 2018



(c) 2019

FIGURE 4.19 – Liste des dates sélectionnées selon le meilleur score F1 moyen pour les années 2017 (a), 2018 (b) et 2019 (c). Chaque saison est représentée par une couleur. Le score de la première meilleure date, du score maximal possible pour l'année et du score de toute la série temporelle sont spécifiés au-dessus des dates choisies.

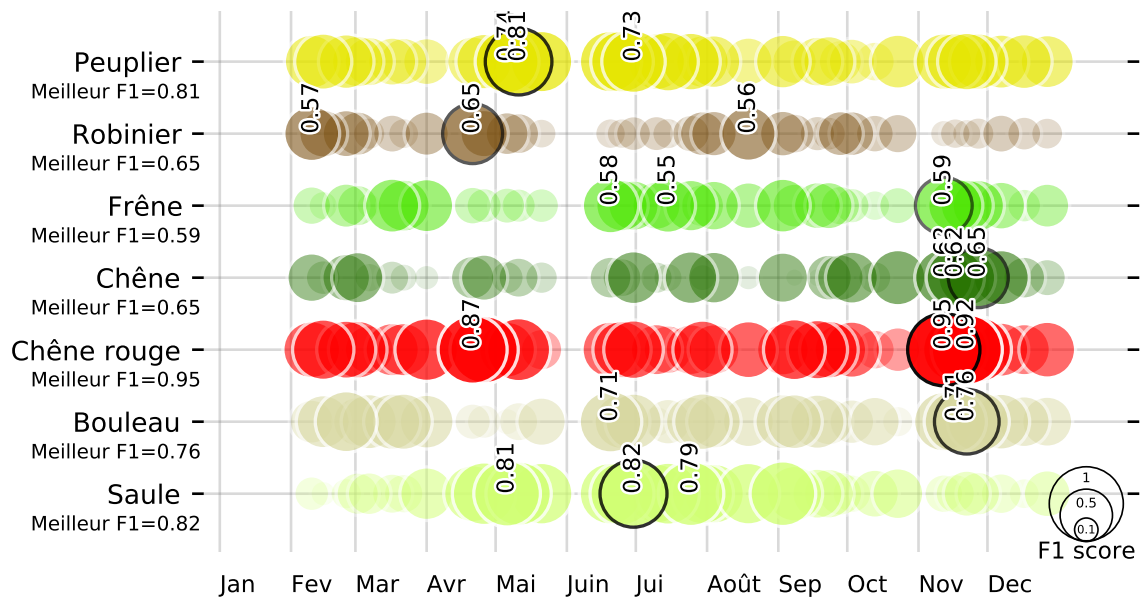
4.5.4 Explication des dates par feuillu

En plus de réduire le nombre de dimensions, la sélection de caractéristiques apporte des informations essentielles à la compréhension de l'algorithme.

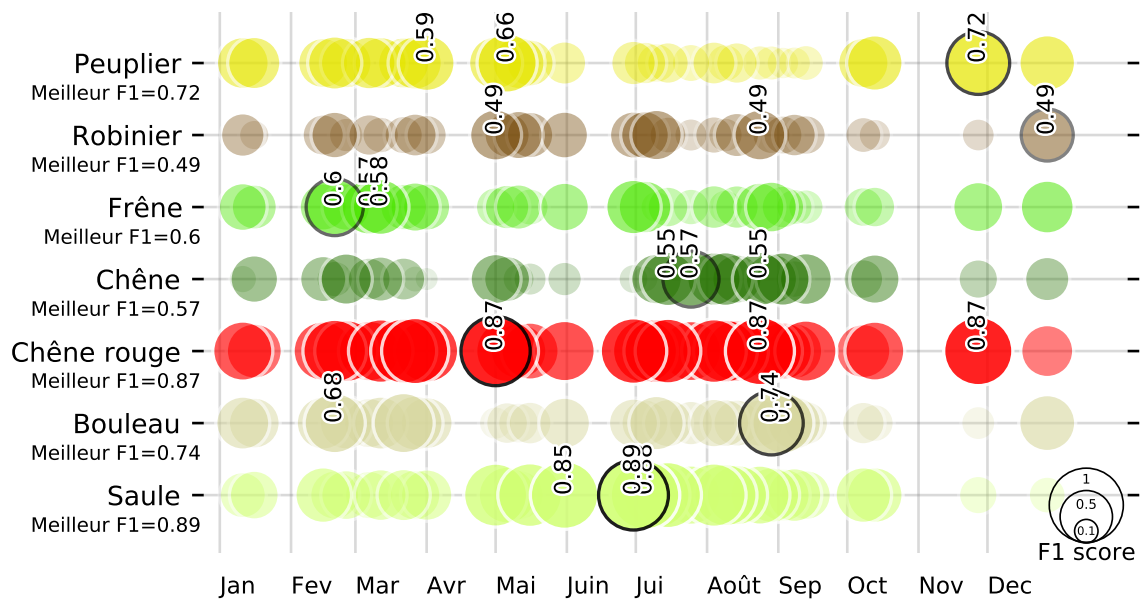
Plusieurs auteurs ont déjà souligné le fait que les dates du début de saison sont importantes pour prédire les essences forestières. [Persson et al. \[2018\]](#) avait détecté la fin du printemps comme période d'importance, alors que [Grabska et al. \[2019\]](#) avaient identifié avril/mai et

mi-octobre. D'autres auteurs signalaient que l'automne était le plus propice pour séparer les essences les unes des autres [Hill et al., 2010; Pasquarella et al., 2018; Schriever and Congalton, 1995]. Cependant les dates obtenus par les précédents auteurs le sont pour classer toutes les essences d'un coup, c'est-à-dire des feuillus et des conifères. Or, il apparaît comme normal que des dates d'automne qui représentent donc la toute fin de saison de végétation ou des dates d'hiver soient les plus utiles pour un algorithme afin de déjà séparer les feuillus des conifères.

En regardant dans le détail des premières dates possibles pour classer chaque essence, on observe que pour le chêne rouge (*Quercus rubra*) les dates de fin avril ou de novembre permettent d'obtenir des scores F1 avoisinant les 0,9 (Figure 4.20).



(a) 2018



(b) 2019

FIGURE 4.20 – Liste de la qualité de prédiction lorsqu'une seule date est utilisée pour prédire une essence (critère de sélection de la date : meilleur score F1 de l'essence en question). Les trois meilleurs résultats par essence sont écrits.

D'après les observations in situ, les dates de début mai et fin novembre du 21 avril 2018 correspondent au début de saison du chêne rouge (*Quercus rubra*), c'est-à-dire qu'au minimum 10% des feuilles étaient sorties.

Pour le robinier, les dates de début mai correspondent à la floraison. Carl et al. [2017] avaient déjà démontré qu'il était possible de détecter la floraison de cette espèce par drone [Carl et al., 2017]. Il est à noter que la floraison avait permis à Boschetti et al. [2007] de bien classer les châtaigniers, non présents dans notre zone d'étude.

Enfin, l'essence pour laquelle il est le plus facile d'interpréter les dates sélectionnées est le saule (*Salix alba*). En effet, en suivant la phénologie in situ, nous avons été dans l'impossibilité pour une parcelle et en difficulté pour une autre de rejoindre le centroïde afin de réaliser les mesures de suivi phénologique. Cela explique certainement en grande partie pourquoi ces parcelles ont le NDVI et le LChloC les plus faibles de toutes les parcelles de début mai jusqu'à fin août car l'eau absorbe grandement l'infrarouge. En regardant les courbes spectrales, on observe bien que les deux parcelles ont une réflectance assez forte en été dans le rouge (B4 de Sentinel-2) et très faible dans l'infrarouge (B8 de Sentinel-2, Annexe C.5).

4.5.4.1 Qualité globale de prédiction

Si les résultats sont encourageants avec des indices de qualité de F1 allant de 0,88 en 2017 et de 0,90 en 2018 lors de la cartographie de l'ensemble des essences, il s'avère que la répartition des feuillus varie parfois de manière importante d'une année sur l'autre (Figure 4.21). En 2017, 19,2% des feuillus prédits étaient des saules contre 12,6% en 2018, alors que cette essence n'a aucune confusion observée lors de l'apprentissage. Les autres essences varient aussi de plusieurs points de pourcentage à l'exception du peuplier. Cela confirme l'hypothèse de la mauvaise représentativité des références terrain et montre que nos échantillons ne sont donc pas suffisamment représentatifs de la diversité des forêts.

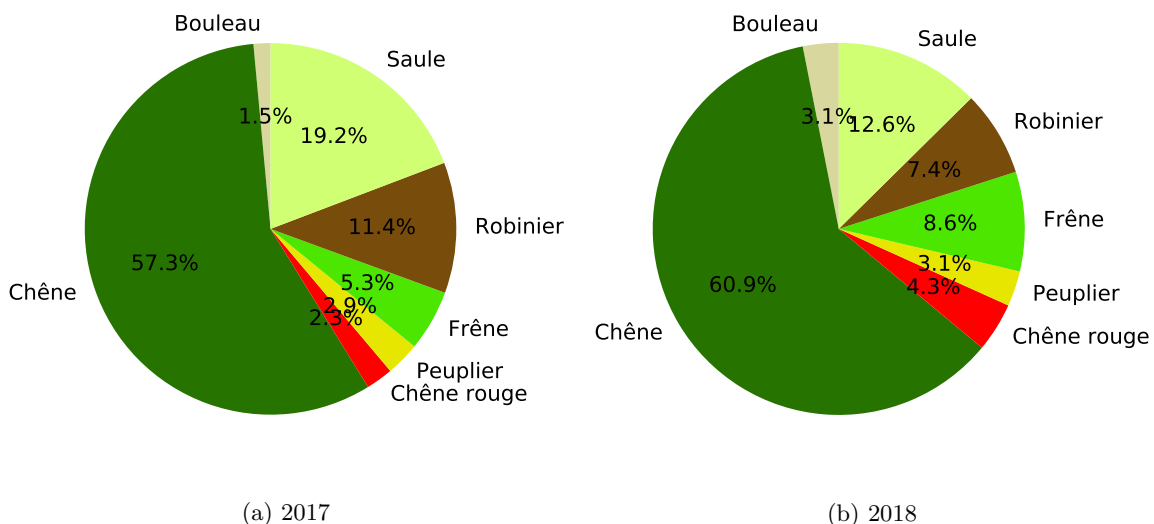


FIGURE 4.21 – Part des feuillus prédite en 2017 et 2018 sur la zone d'étude en classant l'ensemble des essences (incluant les résineux) à partir de la meilleure sélection de caractéristiques.

4.6 Conclusion

La détection de la phénologie par satellite est un domaine assez ancien mais qui se faisait avec un satellite à faible résolution spatiale comme MODIS. L'avantage de ce dernier est sa forte résolution temporelle (quotidienne) alors que Sentinel-2 ne couvre que tous les cinq jours le territoire. Cependant, dans notre zone d'étude qui a tendance à être plutôt ensoleillée, il

apparaît que le duo Sentinel-2 réussit à bien estimer les dates de début et de fin de saison, à condition toutefois de bien choisir ses indices spectraux et les seuils d'amplitude pour les détections. Les résultats ayant été obtenus sur deux années, il conviendrait de continuer ces travaux avec de nouvelles campagnes terrain ou à partir de PhenoCam [Richardson et al., 2011] qui permettraient d'avoir des images quotidiennes sans la charge d'une campagne de terrain. Cela pourrait être un des développements majeurs pour améliorer ces travaux et mieux estimer la précision de détection des métriques phénologiques à partir de données satellites.

En utilisant Sentinel-2 pour détecter le début de saison (BBCH=11, 15 et 19) de plusieurs essences de feuillus, il apparaît que l'indice spectral LChloC (Leaf Chlorophyll Content) est le plus fiable avec une erreur moyenne absolue de cinq jours en 2018 et de 8,5 jours en 2019. Pour la fin de saison (BBCH=91), il s'avère que l'indice LCaroC (Leaf Carotenoid Content) est le plus stable d'une année sur l'autre (8,1 et 8,6 jours d'erreurs pour 2018 et 2019). L'un des points forts de ces deux indices est non seulement leur bons résultats d'une année sur l'autre, mais surtout qu'il est possible d'obtenir ces résultats en gardant les mêmes paramètres (taille de la fenêtre de Savitzky-Golay et seuil fixé selon le pourcentage de l'amplitude). L'indice NDVI, hormis pour la toute fin de saison, est un indice stable entre années et qui, avec des paramètres assez souples du seuil et de la fenêtre de lissage de Savitzky-Golay, permet d'obtenir une bonne précision sous les dix jours (Annexe C.3).



FIGURE 4.22 – Différence de phénologie entre deux cyprès chauves (essence non présente dans la zone d'étude au sud de Toulouse). Photo prise le 12 octobre 2019 à Blagnac dans le parc d'Andromède.

Grâce aux deux satellites Sentinel-2, il a été possible d'obtenir un nombre conséquent d'images dès l'été 2017. La sélection des caractéristiques a été utilisée non seulement pour améliorer la qualité des modèles en réduisant le nombre de dates nécessaires pour atteindre la meilleure prédiction, mais a aussi permis d'analyser les dates privilégiées par l'algorithme. Si de manière générale les dates choisies coïncident plutôt avec la fin de saison, l'interprétation intraspécifique est plus complexe. Pour certaines espèces comme le robinier (*Robinia pseudoacacia*) il a été facile de mettre en relation les dates choisies avec la floraison observée in situ. Pour les autres essences il est plus difficile de comprendre les dates sélectionnées. Sur les trois années, 2017 et 2018 avaient plusieurs images disponibles en novembre, c'est systématiquement ces dates qui ont été sélectionnées en premier et elles correspondent à la toute fin de la saison de végétation selon notre suivi in situ.

Malgré ces limites de compréhension, la qualité globale estimée par le score F1 ou par l'accord global est élevée alors même qu'une validation croisée rigoureuse a évité de sélec-

tionner les échantillons spatialement autocorrélés, ce qui aurait eu pour conséquence un biais optimiste dans la qualité estimée des résultats. Cependant, grâce aux prédictions sur les trois années, on observe que la répartition des essences n'est pas tout à fait la même d'une année sur l'autre. Cela montre que, malgré des résultats qui semblent très satisfaisants, la cartographie est loin d'être optimale et qu'il persiste un biais optimiste.

Il est à noter que les métriques phénologiques (début, fin ou longueur de saison) n'ont pas été utilisées comme base ou supplément d'apprentissage pour prédire les essences forestières. Au vu des deux années de suivi phénologiques et des différences intraspécifiques qu'il peut y avoir, il ne semble pas nécessaire d'y ajouter ces informations. À titre d'illustration de la variabilité intraspécifique, vous pouvez observer sur cette photo (Figure 4.22) deux individus d'une même espèce qui sont pourtant si proches spatialement mais si éloignés sur l'échelle BBCH.

Afin d'améliorer la stabilité des paramètres pour détecter les stades phénologiques, il serait intéressant de mettre en place une fonction de type double logistique sur les indices spectraux afin de vérifier si les seuils fixés manuellement peuvent être les mêmes entre les années. L'inconvénient de l'approche du lissage de Savitzky-Golay est la nécessité de paramétrer la taille de la fenêtre de traitement. Or, selon le bruit présent dans la série temporelle, la taille de la fenêtre doit être changée afin de bien gommer les anomalies.

Enfin, pour palier aux mauvaises performances des séries temporelles sans sélection de caractéristiques, il pourrait être envisagé d'utiliser les synthèses mensuelles de niveau 3A [Hagolle Olivier, 2018] qui, si les résultats sont satisfaisants, faciliterait également la cartographie à l'échelle nationale car elles sont disponibles à l'échelle de la France.

Chapitre 5

Partage des outils et des données : assurer une reproductibilité des travaux

“Happiness only real when shared.”

— Christopher McCandless.

Table des matières

5.1	Introduction	112
5.2	Principes FAIR	112
5.2.1	Findable	113
5.2.2	Accessible	113
5.2.3	Interopérable	113
5.2.4	Re-Usable	114
5.2.5	Limitations du FAIR et du plan national	114
5.3	Mise en œuvre des principes FAIR	114
5.3.1	Les licences choisies	115
5.4	Partage des références terrain	116
5.4.1	Relevés localisés des essences	116
5.4.2	Relevés phénologiques	117
5.5	Partage les outils	118
5.5.1	Pourquoi partager les outils ?	118
5.5.2	Présentation de Museo ToolBox	120
5.5.3	Présentation de Museo Pheno	127
5.5.4	Qualité	130
5.5.5	Un outil facilitant le lien entre la recherche et les étudiants.	135
5.6	Conclusion	136

La reproductibilité des travaux et le partage de la donnée, bien qu'essentiels, sont assez souvent mis de côté. Certaines méthodes scientifiques, bien que déjà écrites en code, ont dû être à nouveau développées dans cette thèse, soit car aucun code ou outil n'était disponible auprès des chercheurs, soit car les codes fournis n'étaient plus à jour.

Dans le contexte de global de la recherche reproductible et du partage des outils et de la donnée, il apparait important de veiller à respecter à ces principes qui, au delà de rendre les travaux plus transparents, permet de les reproduire plus aisément.

Comme vu dans l'ensemble des Chapitres 2 3 4, les outils développés tout au long de cette thèse sont partagées ainsi que leurs données.

Ce chapitre s'intéresse aux méthodes de développement choisies pour respecter les grands principes du partage de la donnée, mais aussi de la fiabilité des bibliothèques écrites en langage

python. Les deux jeux de données créés ou améliorés au long de cette thèse seront présentés brièvement, et les deux bibliothèques python seront expliquées en détail.

- La bibliothèque python Museo ToolBox a fait l’objet d’une publication dans le Journal of Open Source Software [Karasiak, 2020].
- La bibliothèque python Museo Pheno a été publiée uniquement sur la plateforme pip pour python et a son DOI grâce à la plateforme Zenodo (10.5281/zenodo4005981).
- Les références terrain localisées des arbres sont publiées sur Zenodo (10.5281/zenodo.2581399).
- Les suivis phénologiques des feuillus (incluant l’ensemble des photos) sont publiées sur Zenodo (10.5281/zenodo.3994731).

5.1 Introduction

Le plan national pour la science ouverte annoncé le 4 juillet 2018 par Frédérique Vidal, Ministre de l’Enseignement supérieur, de la Recherche et de l’Innovation, rend obligatoire l’accès ouvert pour les publications et pour les données issues de recherches financées sur projets [Ministère de l’Enseignement Supérieur, de la Recherche et de l’Innovation, 2018]. En effet, l’ouverture des données scientifiques est un véritable enjeu pour mettre à bien une recherche reproductible. Cela vise à rendre la science, selon le Ministère de l’Enseignement Supérieur, de la Recherche et de l’Innovation, “*plus cumulative, étayée par les données, transparente, plus rapide et d’accès plus universel*”.

S’il est possible de simplement déposer sur une archive publique un article scientifique et les données qui y sont rattachées, cela ne signifie pas pour autant que le travail sera facilement accessible et reproductible. À titre d’exemple, 64% des tentatives de répliation effectuées sur une sélection de 100 études publiées dans les 3 revues les mieux classées en psychologie n’ont pas retrouvé les résultats originaux [Open Science Collaboration, 2015].

Les objectifs fixés dans ce chapitre sont :

1. présenter la démarche adoptée pour assurer le partage des données et des codes
2. présenter les outils libres élaborés dans le cadre de la thèse pour assurer une reproductibilité des travaux

5.2 Principes FAIR

Si partager ses données est désormais une obligation quand la recherche s’est faite à partir de fonds publics, les moyens et la qualité attendue ne sont pas définis. Dans un contexte où les enjeux liés à la réutilisation des données sont de plus en plus importants, les principes FAIR ont pour objectif de guider ce partage [Wilkinson et al., 2016]. L’acronyme *Findable, Accessible, Interoperable, Reusable* (FAIR), illustré Figure 5.1 correspond à :

1. *Findable* pour Trouvable,
2. *Accessible*, pour Accessible,
3. *Interoperable*, pour Interopérable,
4. *Re-Usable*, pour Réutilisable.

La science ouverte qui applique les principes FAIR constitue “*un levier pour l’intégrité scientifique et favorise la confiance des citoyens dans la science*” [Ministère de l’Enseignement Supérieur, de la Recherche et de l’Innovation, 2018]. Les 4 grands principes du FAIR sont détaillés par la suite.

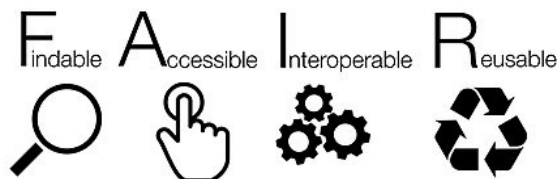


FIGURE 5.1 – Que signifie l’acronyme FAIR ?

5.2.1 Findable

La première étape pour partager les données est de les rendre visibles et trouvables sur internet. Pour que les données soient facilement trouvables par un humain ou par un ordinateur, il faut que les données soient accompagnées de métadonnées. Ces dernières sont indispensables car elles permettent de faire le lien entre la donnée et l'utilisateur qui la recherche, et elle est facilement lisible par un robot d'indexation. C'est une étape clé dans le processus de partage de la donnée.

Le principe **F**indable se résume en 4 grands points [Wilkinson et al., 2016] :

- **F1** : Un identifiant unique et permanent qui permet d'indexer les données et leurs métadonnées.
- **F2** : Les données sont définies avec des métadonnées riches (définies par le **R1**).
- **F3** : Des métadonnées claires et explicites incluent l'identifiant unique de la donnée qu'ils décrivent.
- **F4** : Les métadonnées et leurs données sont indexables par un moteur de recherche.

5.2.2 Accessible

Référencer les données et les outils est la première étape indispensable. S'il est possible de la partager auprès du grand public, il est conseillé de le faire. L'idée est de rendre le plus accessible possible la donnée.

L'accessibilité peut se résumer en 2 grands points :

- **A1**. Les données et métadonnées sont retrouvables à partir d'un identifiant unique en utilisant un protocole standardisé.
 - A1.1** Le protocole doit être ouvert, gratuit, et implémentable.
 - A1.2** Si nécessaire uniquement, l'accès à la donnée peut se faire après une authentification.
- **A2**. Les métadonnées doivent rester accessibles, même si la donnée ne l'est plus.

Des sites comme Hyper Articles en Ligne (HAL) ou Zenodo peuvent être utilisés pour à la fois générer un identifiant unique et partager la donnée. Ces sites vont se charger non seulement d'archiver les données envoyées, mais ils vont également attribuer un Digital Object Identifier (DOI), un identifiant unique numérique qui permet de retrouver la donnée ou la métadonnée.

5.2.3 Interopérable

Une fois que la donnée est indexée et si possible accessible, elle doit être le plus interopérable possible, c'est-à-dire compatible avec le plus d'outils différents (systèmes d'exploitation, chaîne de traitements...). Plus un outil ou une donnée seront utilisables sur de multiples plateformes et compatibles avec des outils existants, plus ils seront considérés comme interopérables.

Ce principe peut se décomposer en 3 grands sous-points :

- **I1**. Les données et métadonnées utilisent un langage formel, accessible, partagé.

- **I2.** Les données et métadonnées utilisent un vocabulaire qui suit les principes FAIR.
- **I3.** Les données et métadonnées incluent, si nécessaire, des références vers d'autres données ou métadonnées.

Lors du partage des données, il est conseillé d'utiliser un format ouvert. Dans cette thèse, les formats utilisés sont :

- pour les tableaux : le format csv.
- pour les données géographiques vectorielles : le format GeoPackage.
- pour les images géoréférencées : le format GeoTiff.

Il convient d'éviter si possible les formats fermés ou propriétaires.

5.2.4 Re-Usable

L'objectif ultime des principes FAIR est de faciliter la réutilisation de la donnée. Pour atteindre cet objectif, les données et métadonnées doivent être bien décrites comme cela elle pourront être répliquées et/ou combinées avec d'autres données.

R1. Les données et métadonnées sont richement décrites avec un nombre d'attributs clair et représentatif.

R1.1. La licence d'utilisation doit être claire et accessible.

R1.2. Les données et les métadonnées détaillent leur provenance.

R1.3. Les données et les métadonnées utilisent les standards de la communauté à laquelle elles appartiennent.

5.2.5 Limitations du FAIR et du plan national

Toutefois on peut dénombrer plusieurs limitations. Lorsque la donnée est à caractère privé (données personnelles, ou des données confidentielles avec du secret défense par exemple), il n'y a en effet plus d'obligation de partage. L'autre soucis réside dans l'absence de sanction en cas de non partage de la donnée. Ce n'est pas parce qu'il s'agit d'une obligation qu'elle sera respectée. Ainsi, il a fallu plus de 3 ans pour que la directive européenne INSPIRE de 2007 qui vise à favoriser le partage des données publiques et géographiques au sein de l'Union Européenne soit transposée dans le droit français¹. Transposée dans la loi française en 2010, cette directive n'est toujours pas appliquée dans tous les domaines, et il est par exemple toujours impossible d'obtenir pour des chercheurs publics les références terrain de l'IGN servant à la création de l'Inventaire Forestier National (IFN).

Pour résumer, le principe sur lequel repose le partage des données et des méthodes repose sur le credo suivant : *“ouvert autant que possible, fermé autant que nécessaire”*[CNRS, 2019].

5.3 Mise en œuvre des principes FAIR

Le travail de cette thèse s'inscrit dans les 4 domaines du principe FAIR :

- **Findable** : Les données et métadonnées sont stockées sur la plateforme Zenodo avec à chaque fois un identifiant unique. Le choix s'est porté sur Zenodo car cette plateforme a été spécifiquement développée pour permettre aux chercheurs de déposer et archiver leurs données et codes (qui peuvent atteindre jusqu'à 50 Go, pour des fichiers plus lourds, il convient de les contacter). La plateforme est développée et hébergée depuis ses débuts en 2013 par le Conseil Européen pour la Recherche Nucléaire (CERN).
- **Accessible** : En plus d'assigner un identifiant unique, la plateforme Zenodo stocke les données et métadonnées et inclue un accès gratuit vers celle-ci. La donnée sauvegardée sur Zenodo est souvent une copie des dépôts GitHub des bibliothèques développées.

1. Ordonnance n° 2010-1232 du 21 octobre 2010 portant diverses dispositions d'adaptation au droit de l'Union européenne en matière d'environnement

- Interoperable : les données sont fournies uniquement avec des formats ouverts. GeoPackage (GPKG) pour les données géographiques vectorielles, GeoTiff pour les données raster, csv pour les données de type tableau.
- Re-Usable : nous incluons des exemples ou un lien vers du code permettant de reproduire des travaux à partir des données fournies.

5.3.1 Les licences choisies

On distingue deux types de licences dans ce chapitre, celles dédiées aux logiciels et celles dédiées aux données. Plusieurs conditions étaient importantes à mes yeux : les données et les codes s'ils sont améliorés ou modifiés doivent rester gratuits et avec la même licence. Et s'il y a une modification ou une amélioration, elle doit profiter à plus grand nombre et donc être partagée.

5.3.1.1 Licence pour les codes

Pour partager les codes python (qui sont valorisés sous forme de bibliothèques), on se réfère alors aux licences pour les logiciels. Les licences d'utilisation pour les logiciels peuvent se classer en deux grandes familles :

- La famille *Berkeley Software Distribution* (BSD) ou la licence MIT : licences ne contenant pas de copyleft.
- La famille qui regroupe les licences *General Public License* (GPL) : licences contenant un copyleft.

Le copyleft est un terme défini en opposition au copyright (droit d'auteur). Le but est de laisser le plus libre possible un programme et ses futures versions. Selon la fondation [GNU \[2015\]](#) : “*Le copyleft est une méthode générale pour rendre libre un programme (ou toute autre œuvre) et obliger toutes les versions modifiées ou étendues de ce programme à être libres également*”.

Au vu des besoins évoqués plus haut, le choix s'est donc naturellement porté vers une licence de type GPL qui s'axe autour de 4 libertés :

- Liberté 0 : La liberté d'exécuter le logiciel, pour n'importe quel usage ;
- Liberté 1 : La liberté d'étudier le fonctionnement d'un programme et de l'adapter à ses besoins, ce qui passe par l'accès aux codes sources ;
- Liberté 2 : La liberté de redistribuer des copies ;
- Liberté 3 : L'obligation de faire bénéficier la communauté des versions modifiées.

L'ensemble des codes libérés dans cette thèse le sont sous la licence GPL v3. Les évolutions des droits d'auteurs et des brevets font que la licence GPL a besoin d'évoluer pour garantir au maximum les libertés d'utilisation. La nouvelle version 3 qui est arrivée en 2007, soit 16 ans après la version 2 (1991) a pour but de garantir une liberté maximale aux développeurs suite aux évolutions des brevets et des nouvelles méthodes de privations des droits. Ainsi la licence V3 a pour but de neutraliser les lois prohibant les logiciels libres lorsqu'ils utilisent par exemple des DRM (gestion numérique des restrictions).

Il est à noter que la proportion de logiciels partagés sous une licence permissive (c'est-à-dire sans copyleft) est de 67% en 2019, alors qu'ils représentaient 41% en 2012 [[Goldstein, A., 2020](#)]. Cela s'explique en partie par le succès de la licence MIT qui représente à elle seule 27% des licences.

5.3.1.2 Licence pour les données

Il existe de nombreuses licences pour libérer les données, nous nous concentrerons sur les principales utilisées à l'internationale, à savoir les Creative Commons (CC) et les Open Data Commons (ODC) [[Labastida and Margoni, 2020](#)].

Les licences de l'ODC, développées par l'Open Knowledge Foundation [Molloy, 2011] permettent, à la condition de partager aux mêmes conditions les nouveaux travaux et de mentionner l'origine de la donnée :

1. Le partage : copie, distribution, utilisation de la base de données
2. La création : production de créations à partir de la base de données
3. L'adaptation : modification, transformation et construction à partir de la base de données

Les licences CC sont disponibles en 6 versions différentes, certaines autorisant un usage à des fins commerciales, d'autres les modifications sans en informer les auteurs initiaux.

Afin de maximiser la compatibilité entre les licences et de permettre une diffusion et l'utilisation maximale des œuvres, nous avons préféré opter pour la licence CC-BY plutôt que la ODB-BY sous les conseils d'un Danny Lämmerhirt [2017], membre de l'ODC. Le BY signifie que lors d'une réutilisation, il convient de créditer la donnée initiale et d'indiquer si des modifications ont été opérées.

5.4 Partage des références terrain

5.4.1 Relevés localisés des essences

Pour partager les références localisées d'arbres situés sur la zone de l'Observatoire Spatial Régional (OSR), l'ensemble des principes du FAIR a été respecté.

Les données sont partagées sur la plateforme Zenodo² sous le format vectoriel GeoPackage qui a été développé par l'Open Geospatial Consortium. Ce format est ouvert, non-propretaire et basé sur le format SQLite 3.

Un fichier README.pdf explique comment la donnée a été récoltée (cf Section 2.2.4) et comment elle a été organisée. En raison de la licence restrictive de la donnée Formosat-2, nous n'avons pas pu partager les séries temporelles. Cependant, et afin de pouvoir reproduire les travaux issus de ce satellite (Chapitre 3), nous avons intégré pour chaque année (de 2006 à 2014) les valeurs spectrales de chaque bande pour l'ensemble des références. S'il n'est pas possible de reproduire la cartographie, il est possible de reproduire les résultats de l'article du chapitre 3 (qualité du modèle, matrice de confusion...).

Le fichier de base localisant les essences contient l'ensemble des informations du Tableau 5.1.

TABLEAU 5.1 – Liste des champs des données de référence localisées des essences. *int* signifie que le champs contient des données de type entier, *str* de type *string* (texte).

champs (type)	définition	exemple
english (<i>str</i>)	Nom de l'espèce (en anglais)	Silver birch
latin (<i>str</i>)	Nom de l'espèce (en latin)	Betula pendula
level1 (<i>int</i>)	type de forêt	1 = feuillu, 2 = conifère
level2 (<i>int</i>)	groupe d'essences	1 = décidu, 2 = feuillu sempervirent, 3 = pin, 4 = autre conifère
level3 (<i>int</i>)	essence	1 = Silver birch, 2 = Oak...

Pour chaque année, un fichier est fourni avec l'ensemble des valeurs spectrales, donc auquel s'ajoute cinq colonnes pour chaque date : une colonne par bande spectrale, et une colonne indiquant si la donnée a été interpolée (donc si un nuage ou une ombre a été détecté). Pour plus de détails se référer au Tableau 5.2.

2. <https://doi.org/10.5281/zenodo.2581399>

TABLEAU 5.2 – Liste des champs complémentaires pour chaque année. Pour 2013 qui est composée de 17 dates, il y a donc 68 colonnes *band_* (17 dates * 4 bandes spectrales) et 17 colonnes *cloud_* allant de 0 à 16. *int* signifie que le champs contient des données de type entier.

champs (type)	définition	exemple
<i>band_ (int)</i>	Valeur spectrale de la bande	<i>band_0</i> est la valeur spectrale de la première date de la bande bleue. <i>band_1</i> est la valeur de la deuxième date de la bande bleue...
<i>cloud_ (int)</i>	Présence (1) ou absence (0) initiale de nuage ou d'ombres	<i>cloud_0</i> est la valeur du masque de nuage à la première date. <i>cloud_1</i> a la deuxième date...

5.4.2 Relevés phénologiques

À l'exception des mesures d'ouverture de la canopée, l'ensemble du suivi phénologique a été saisi dans l'application android Memento. Comme cette application synchronise à notre demande directement le contenu de la base de données du suivi phénologique sur Google Sheet (le tableur en ligne de Google), nous avons pu facilement mettre à jour le tableur avec les mesures d'ouverture de la canopée et récupérer cette information sur la tablette. La donnée terrain des relevés phénologiques a donc toujours été accessible en ligne et à jour pour les utilisateurs ayant eu le lien.

Les données sont partagées sur la plateforme Zenodo³.

Un fichier README.pdf détaille le protocole d'acquisition de la donnée et comment elle est organisée (par parcelle et par date).

L'ensemble des données phénologiques (photos des canopées, des parcelles, et tableur) a été publié. Le tableur ainsi que les images sont fournis dans un format ouvert et non-propriétaire (respectivement csv et jpg).

TABLEAU 5.3 – Liste des champs du fichier *plot.csv* qui contient l'identifiant et le détail de chaque parcelle. *int* signifie que le champs contient des données de type entier, *float* un nombre à virgule et *str* de type *string* (texte).

champs (type)	définition	exemple
<i>plot (int)</i>	identifiant de la parcelle	1 = Peuplier, 2 = Saule...
<i>latin_name (str)</i>	Nom latin de l'essence	<i>Populus spp, Salix alba...</i>
<i>understory (str)</i>	type de sous-étage	nu (<i>bare</i>), herbacée (<i>herbaceous</i>)...
<i>understory_details (str)</i>	Détails du sous-étage	<i>Ruscus aculeatus</i>
<i>comment (str)</i>	Commentaire spécifique	Parcelle dont le suivi s'est arrêté
<i>X (float)</i>	Coordonnées en X	43,49423
<i>Y (float)</i>	Coordonnées en Y	1,22977

Ensuite pour chacune des trois années, un tableau au format csv liste pour chaque parcelle et chaque date l'ensemble des informations disponibles (Tableau 5.4). Toutes les photos ayant permis de calculer les mesures d'ouverture de la canopée sont aussi fournies et sont organisées selon cette nomenclature : *pictures/species_name/plot_number/year/canopy/IMG_date_time.jpg*.

Pour plus de détails sur la façon dont sont calculés les indices de canopée, il est recommandé de consulter la documentation de l'application GLAMA⁴.

3. <https://doi.org/10.5281/zenodo.3994731>

4. <https://www.sci.muni.cz/botany/glama/GLAMA%20manual.pdf>

TABLEAU 5.4 – Liste des champs des relevés phénologiques de chaque année. *int* signifie que le champs contient des données de type entier, *float* un nombre à virgule et *str* de type *string* (texte).

champs (type)	définition	exemple
plot (<i>int</i>)	identifiant de la parcelle	1 = Peuplier, 2 = Saule...
latin_name (<i>str</i>)	Nom latin de l'essence	<i>Populus spp, Salix alba...</i>
date (<i>str</i>)	date	2019/04/23
photo (<i>str</i>)	Chemin de la photo de la parcelle à la date du suivi	pictures/Betula_pendu- la/plot_9/2019/20190423_- field_view.jpg
BBCH (<i>int</i>)	Code BBCH des principaux stades (0x, 1x, 2x, 3x, 9x)	12, 15, 92...
BBCH_flower (<i>int</i>)	Code BBCH des stades de floraison (6x, 7x, 8x)	65, 75...
CACLOSURE (<i>int</i>)	Moyenne des cinq mesures CACLOSURE	65
CACO (<i>int</i>)	Moyenne des cinq mesures de CACO	38
CACOmodified (<i>int</i>)	Moyenne des cinq mesures CACOmodified	35
CHLO (<i>float</i>)	Moyenne des 25 mesures de chlorophylle	28,6
comment (<i>str</i>)	Commentaire spécifique	Merisier à 18 aux alentours

5.5 Partage les outils

5.5.1 Pourquoi partager les outils ?

Le moyen optimal pour reproduire des travaux de recherche est de fournir les codes ayant été utilisés pour réaliser les analyses. La recherche étant dans notre domaine souvent synonyme de création de nouveaux outils et de nouvelles méthodes, il est donc indispensable de partager ces nouvelles créations pour faciliter la reproduction des travaux, mais aussi pour être transparent sur les résultats. Malheureusement rien n'empêche que nos travaux comportent des erreurs dans le développement de la méthode et aujourd'hui les revues scientifiques de notre domaine ne demandent pas à avoir accès à ce matériel.

Pourtant il n'est pas rare de constater des erreurs dans des codes, et parfois dans les bibliothèques ou des logiciels réputés. Ainsi, une erreur dans un logiciel très utilisé a potentiellement invalidé plus de 40 000 études scientifiques, dont plus de 3 000 avec certitude. Pour découvrir cette erreur, [Eklund et al. \[2016\]](#) ont placé un saumon mort dans un IRM et lui ont montré des photos d'êtres humains. Ils ont alors observé que l'animal avait bel et bien une activité cérébrale si l'on se fiait aux données du logiciel. Il s'agissait simplement de la mise en évidence du bug qui affectait les données du logiciel IRM en question. Les faux positifs étaient alors de plus de 70% à cause des mauvais traitements pour débruiter les images. Depuis le bug a été corrigé mais les 3 000 études n'ont pu être reproduites à cause de la non reproductibilité des articles.

L'un des auteurs de l'étude, Tom Nichols, explique sur son blog qu' "*il ne s'agit pas seulement d'un problème de partage des données, c'est une question de transparence. Si les données de ces 3 500 articles scientifiques étaient publiques, il serait beaucoup plus simple de les réviser au fur et à mesure.*"

Malheureusement le cas précédent n'est pas unique. En 2019 [Bhandari Neupane et al. \[2019\]](#) ont trouvé un bug dans un script très utilisé en chimie computationnelle. Selon le système d'exploitation utilisé, le tri n'était pas fait dans le même ordre, et comme le système d'exploitation n'est que rarement spécifié dans les travaux, il est impossible de connaître la

validité des articles et de les reproduire faute d'un code accessible.

Comme les données et les codes qui permettraient de reproduire les travaux ne sont généralement pas fournis avec les publications, il est compliqué de connaître les articles potentiellement invalidés et de les reproduire. Les résultats de ces études seront donc en partie perdus car il n'y a moins de confiance envers ces travaux, alors qu'il aurait été très simple de les reproduire si les revues exigeaient en même temps qu'un article les données et les codes permettant de reproduire les analyses menées.

Si les travaux de cette thèse relèvent moins d'enjeux vitaux sur la santé humaine que les précédents concernant les IRM, il est tout autant important que l'ensemble des domaines scientifiques prenne comme standard le partage des données, des méthodes et des outils permettant de reproduire le plus simplement les travaux. Il ne faut pas oublier non plus que le nombre de bugs détectés est sûrement très nettement inférieur aux nombre de bugs existants et qu'il touche uniquement des bibliothèques ou les logiciels publiés car le code servant uniquement à un article n'est jamais fourni.

Partager les codes pour reproduire toute ou partie de la méthodologie et des résultats d'une étude permettrait :

- D'améliorer significativement la partie méthode et ainsi d'aider les scientifiques à reproduire et/ou améliorer le travail (Figure 5.2).
- De valider la qualité d'un travail et de le faire vérifier auprès des relecteurs.
- De recalculer les résultats en cas de problème détecté plus tard dans une librairie ou dans les données.

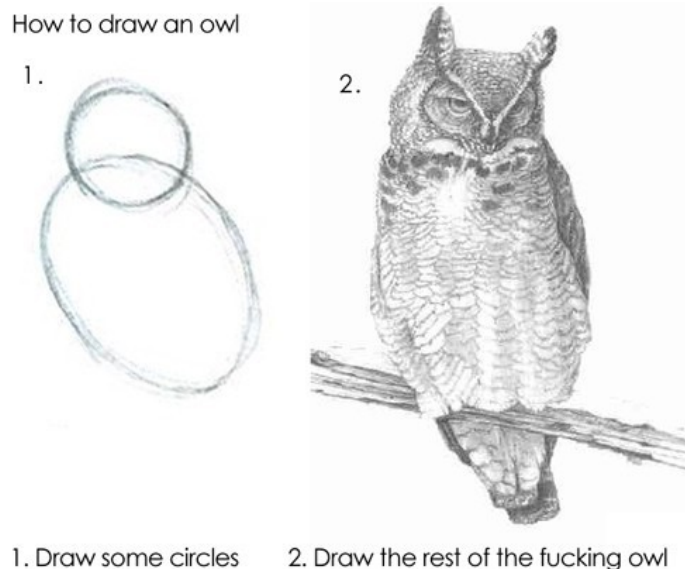


FIGURE 5.2 – Illustration du ressenti que l'on peut avoir à la lecture de la partie consacrée à la méthodologie dans de nombreux travaux. Source inconnue [Brennan, A., 2013].

Pour cela, à chaque article ou chapitre de la thèse, les méthodes permettant de reproduire toute ou partie de la méthodologie et des résultats ont été partagées grâce aux deux bibliothèques python développées : Museo ToolBox et Museo Pheno.

Comme il a été vu plus tôt dans ce chapitre, le partage des codes devraient faire partie intégrante de la publication scientifique. Mais il est possible d'aller encore plus loin et de valoriser un travail en réfléchissant à sa genericité. Après de longues réflexions sur les manques et donc sur les besoins en télédétection, il a été pris la décision de développer des bibliothèques python afin de répondre aux besoins récurrents mais non disponibles dans des outils existants. Le créateur de Linux, Linus Torvalds, disait que "*tout programme n'est bon que s'il est utile à quelque chose*". Partant de ce principe, comme le besoin se faisait ressentir de mon côté, j'ai

donc décidé de me lancer dans l’aventure de la création d’une bibliothèque python généraliste, puis d’une deuxième spécifique aux séries temporelles.

Les bibliothèques créées respectent scrupuleusement les préconisations de la *Python Enhancement Proposal* (PEP) numéro 8⁵. Ainsi les noms donnés aux différentes fonctionnalités suivent la nomenclature suivante :

- `un_module` : un tiret bas sépare chaque mot d’un module (e.g. `cross_validation`)
- `UneClasse` : chaque mot d’une classe commence par une majuscule (e.g. `RasterMath`)
- `une_fonction` : un tiret bas sépare chaque mot d’une fonction (e.g. `add_function`)
- `une_variable` : un tiret bas sépare chaque mot d’une variable (e.g. `x_block_size`)

5.5.2 Présentation de Museo ToolBox

Museo ToolBox est la première et principale bibliothèque python développée dans le cadre de cette thèse. Elle a pour vocation d’être la plus généraliste possible et contient les principales fonctionnalités nécessaires à la reproduction des travaux de thèse. La documentation est disponible en ligne via le site *Read the Docs* (<https://museotoolbox.readthedocs.io>) et le code est hébergé sur GitHub (<https://github.com/nkarasiak/MuseoToolBox>).

Le nom a été choisi en hommage à Orfeo ToolBox (OTB), un logiciel et une bibliothèque python de télédétection initiée par le CNES. Museo serait dans la mythologie grecque le fils d’Orfeo.

La première version de la bibliothèque contenait le module de validation croisée afin de rendre disponible la méthode utilisée dans le chapitre 3 (*Spatial Leave-One-Out*), et le module *processing* qui permet, entre autres, d’extraire la valeur des pixels pour chaque référence terrain. Il se trouve que, dans les bibliothèques existantes, aucune ne permettait de faire facilement le lien entre des données spatiales (vecteur et raster) et des fonctions d’apprentissage qui prennent en entrée des tableaux. Le module *processing* répond donc à cette demande très générique.

On peut regrouper les utilisations de la bibliothèque en plusieurs catégories qui correspondent aux différents modules suivant (Figure 5.3) :

- Traitement et pré-traitement (*processing*)
- Apprentissage (*ai* pour Artificial Intelligence)
- Validation croisée (*cross_validation*)
- Statistiques (*stats*)
- Graphiques (*charts*)
- Données de tests (*datasets*)

Seules les principales fonctionnalités de Museo ToolBox seront développées dans ce manuscrit.

5.5.2.1 Module processing

Classe RasterMath

```
from museotoolbox.processing import RasterMath
```

RasterMath est la clé de voûte de la bibliothèque car cette classe a pour but de gérer les images et de faciliter au maximum la généralité des traitements.

Pour cela, la réflexion part d’un constat simple : qu’est-ce qu’une image géoréférencée si ce n’est un tableau de plusieurs dimensions mais avec en plus une information spatiale ? Partant de ce postulat, tout a été développé pour que l’utilisateur final n’ait à travailler que comme

5. <https://www.python.org/dev/peps/pep-0008/>

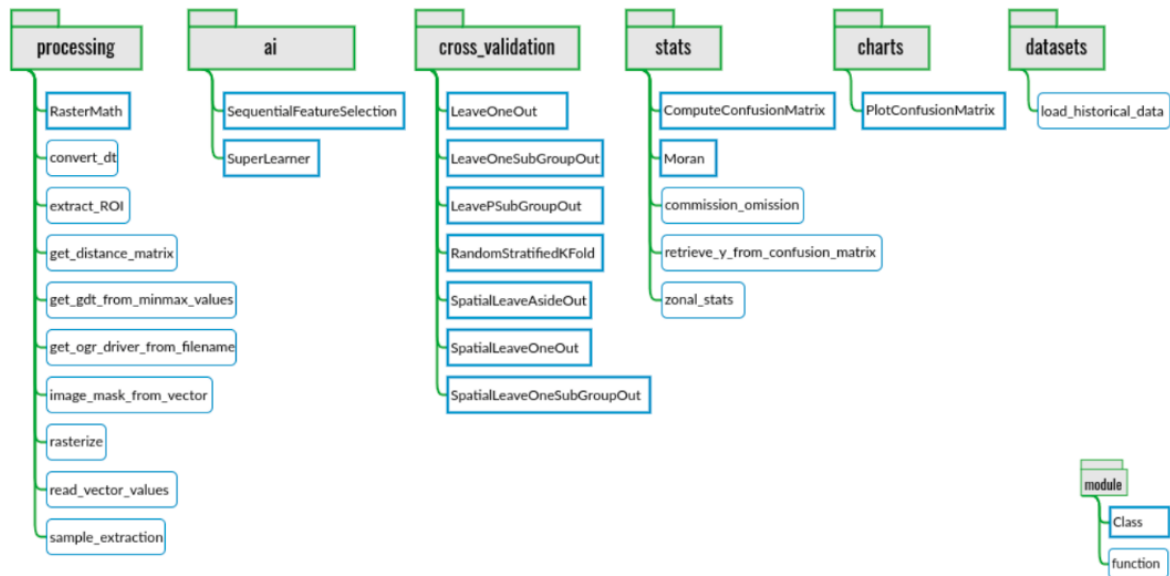


FIGURE 5.3 – Organisation générale de la bibliothèque Museo ToolBox. Plusieurs modules centralisent des traitements par thématique.

s'il ne manipulait que des tableaux, et non des images. La conversion d'une image en de multiples tableaux n'est donc pas à la charge de l'utilisateur, mais à la charge de **RasterMath** (Figure 5.4). L'idée qui en découle est que n'importe quel type de fonction qui s'exécute sur un tableau doit pouvoir fonctionner sur une image en utilisant le moins de lignes possibles.

Une autre limite des outils actuels et qu'ils ne sont généralement pas conçus pour supporter des images volumineuses. Or les séries temporelles d'images Sentinel-2, ne serait-ce que sur une seule tuile, peuvent atteindre plusieurs centaines de gigas. Il est donc impossible de les stocker en mémoire. Il y a donc également un besoin de lire et d'écrire les images morceau par morceau, ce qu'on nomme la lecture et l'écriture par bloc. **RasterMath** va donc fonctionner par défaut avec des blocs de taille carré (256x256 pixels, taille par défaut des blocs selon la bibliothèque GDAL). Ainsi l'image ne sera jamais entièrement chargée en mémoire. Il est toutefois possible de modifier la taille des blocs très facilement pour l'utilisateur.

Une simple fonction prévue pour un tableau (comme une moyenne des bandes par pixel, ou un lissage de série temporelle) sera directement compatible avec **RasterMath** et ce peu importe la dimension de l'image.

L'utilisation de **RasterMath** se résume en trois étapes clés pour l'utilisateur (Code 5.1) :

1. Initialisation de **RasterMath** avec ouverture de l'image
2. Ajout d'une fonction au choix de l'utilisateur
3. Écriture d'une nouvelle image

Selon les besoins, comme lors de l'apprentissage automatique avec des bibliothèques comme *scikit-learn*⁶, la donnée doit être converti en 2D, c'est-à-dire qu'une ligne doit correspondre à un pixel et que le nombre de colonnes doit correspondre au nombre de bandes spectrales de l'image. C'est pour cela que par défaut, et pour optimiser la compatibilité avec des fonctions pour tableaux déjà existantes que **RasterMath** convertira en 2D n'importe quelle image. Il est tout à fait possible de garder la réelle dimension d'un bloc en spécifiant en entrée de **RasterMath** l'argument `return_3d=True`.

Museo ToolBox intègre aussi la gestion des masques. Dans le cas des forêts tempérées, l'illustration est très simple. Seule une petite partie des tuiles Sentinel-2 nous concerne. Il suffira donc de créer un fichier vectoriel ou d'utiliser par exemple la BD Forêt de l'IGN et de spécifier à **RasterMath** de ne travailler qu'à l'intérieur de ces polygones en spécifiant en

6. <https://www.scikit-learn.org>

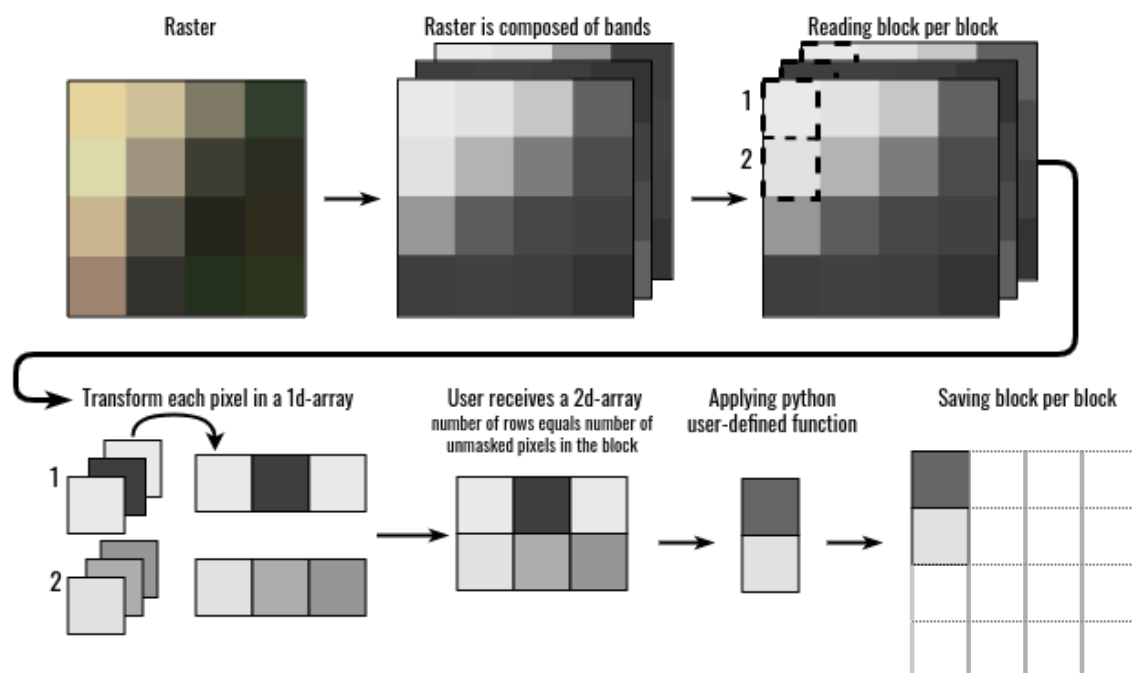


FIGURE 5.4 – Les dessous de la classe `RasterMath`. La lecture et l'écriture d'une image se fait par bloc, chaque bloc étant composé de plusieurs pixels qui sont transformés par défaut 2D afin d'être totalement compatibles avec des fonctions écrites pour des tableaux.

entrée l'argument `in_image_mask='mask.tif'`. Les *no data* (données sans valeur, car elles sont désormais hors masque) auront par défaut la valeur la plus faible du type de la donnée choisie par l'utilisateur. Ainsi si l'utilisateur choisit d'enregistrer son image avec le type entier 16 bits (`int16`, soit de -32768 à +32767), Museo ToolBox assignera la valeur -32768 aux pixels à masquer. Tous ces paramètres sont bien sûr modifiables.

Un exemple simple est le calcul de la moyenne des bandes par chaque pixel (Code 5.1). Pour cela nous utilisons la fonction `mean` de `numpy` qui nous permet de spécifier sur quel axe (qui précise l'ordre d'indexation du tableau) la moyenne doit se calculer. L'axe de dimension 1 correspond aux colonnes, car comme vu précédemment, par défaut Museo ToolBox fonctionne en 2D : cela signifie qu'à chaque pixel correspond une ligne est qu'à chaque colonne correspond une bande spectrale. En spécifiant que la moyenne se calcule sur l'axe 1, cela signifie que chaque ligne n'aura plus qu'un élément qui est la moyenne des colonnes (par réduction d'une dimension du tableau).

```
import numpy as np
import museotoolbox as mtb

raster,vector = mtb.datasets.load_historical_data() # on charge le chemin
              de l'image et du fichier vectoriel de démo

# instantiation de RasterMath
rM = mtb.processing.RasterMath(raster)

# ajout de la fonction mean qui calculera la moyenne selon l'axe 1
rM.add_function(np.mean,out_image='/tmp/mean.tif',axis=1)

# lancement du calcul et écriture de l'image
rM.run()
```

Code 5.1 – Exemple de code utilisant `RasterMath` pour calculer la moyenne des bandes par pixel.

Ce qu'il faut retenir

La classe **RasterMath** permet :

- De lire une image, de calculer une fonction et d'écrire le résultat, en seulement 3 lignes de code et à partir de n'importe quelle fonction prévue pour fonctionner sur un tableau.
- De rendre très accessible la lecture et l'écriture des images pour des domaines scientifiques non spécialistes en télédétection (mathématiciens par exemple).
- Une optimisation du traitement (la lecture et l'écriture se font par bloc, l'image est compressée sans perte automatiquement)
- Le support de plusieurs cœurs pour les calculs les plus chronophages.

Fonction extract_ROI

```
from museotoolbox.processing import extract_ROI
```

Cette fonction permet de faire le lien entre une image et un fichier vectoriel (de type point ou polygone) dans le but d'extraire les valeurs spectrales pour chaque entité. Cette information est essentielle car c'est celle qui sera nécessaire pour réaliser de l'apprentissage automatique. Ainsi la fonction **extract_ROI** va extraire les valeurs spectrales (X) et le label (y) de la ou des classes spécifiées par l'utilisateur (Code 5.2).

```
import museotoolbox as mtb

# on charge le chemin de l'image et du fichier vectoriel de démo
raster, vector = mtb.datasets.load_historical_data()

# on extrait les valeurs spectrales et le label contenu dans la colonne '
  Class'
X, y = mtb.processing.extract_ROI(raster, vector, 'Class')

# on extrait les valeurs spectrales et le label d'autant de colonnes que l'
  on souhaite
X, y, fid = mtb.processing.extract_ROI(raster, vector, 'Class', 'uniquefid')
```

Code 5.2 – Illustration du code permettant d'extraire les labels et les bandes spectrales issus d'un vecteur et d'une image.

5.5.2.2 Module AI : Apprentissage automatique**Classe SuperLearner**

```
from museotoolbox.ai import SuperLearner
```

La classe **SuperLearner** permet de faire le lien entre les données (X) et les labels (y) obtenus par la fonction **extract_ROI** et la bibliothèque d'apprentissage *scikit-learn*. Pour apprendre un modèle à partir de *scikit-learn*, seul les données X et y sont nécessaires. Cependant, il peut être plus compliqué pour prédire ce modèle sur une image entière, surtout si la donnée en entrée (X) a été standardisée.

Elle apporte toutefois une contribution importante en intégrant par défaut des méthodes recommandées selon l'état de l'art comme la standardisation des données spectrales ou les calculs des hyper-paramètres d'un algorithme à partir d'une validation croisée et le calcul des erreurs pour chaque pli de la validation croisée. Une fonctionnalité qui a été beaucoup utilisée dans cette thèse et la possibilité de personnaliser par une fonction comment **SuperLearner** lira une image : il s'agit du même principe que **RasterMath**, ainsi il sera tout à fait possible d'apprendre et de prédire à partir d'une série temporelles de NDVI en ayant uniquement en

entrée la série temporelle spectrale Sentinel-2. Cela permet donc d'économiser beaucoup de temps de calcul, d'écriture et de stockage car la donnée sera générée à la volée.

Ce qu'il faut retenir

SuperLearner permet, entre autre, de gérer :

- L'optimisation des paramètres de l'algorithme à partir d'une validation croisée
- Le calcul des matrices de confusion pour chaque pli de la validation croisée
- La prédiction du modèle sur une image entière
- La personnalisation de la lecture d'une image à travers une fonction 2D (afin de ne pas écrire une nouvelle image)

Classe `SequentialFeatureSelection`

```
from museotoolbox.ai import SequentialFeatureSelection
```

Cette classe se base sur la classe précédente, **SuperLearner**, et permet d'y implémenter une sélection de variables. La plus-value de **SequentialFeatureSelection** ne réside pas dans sa capacité à faire une sélection de caractéristiques variable par variable, car on retrouve cette méthode dans de nombreuses bibliothèques, y compris l'une des plus utilisées qu'est *Mlxnet* [Raschka, 2018]. La plus-value de **SequentialFeatureSelection** réside dans sa capacité à interpréter un ensemble de caractéristiques comme étant une seule et unique variable (agrégats de variables) : il est aisé de spécifier que la sélection de caractéristiques se fait par date (une date peut donc être un ensemble de n bandes spectrales). L'autre avantage est qu'on peut spécifier de sauvegarder les meilleures variables au fur et à mesure, sans attendre que l'ensemble des k meilleures variables sont trouvées.

Ce qu'il faut retenir

SequentialFeatureSelection permet :

- d'améliorer grandement les temps de calculs en diminuant fortement le nombre de caractéristiques uniques.
- d'estimer le fléau de la dimension [Bellman, 1984]
- d'estimer la contribution générale de chaque date

Cependant les limites de cette classe sont :

- La non-gestion d'autres types de sélection de variables (avec remise des variables par exemple)
- La non-gestion de la sélection à l'envers (en partant de l'ensemble des caractéristiques et en les enlevant une à une)

5.5.2.3 Module validation croisée

```
from museotoolbox import cross_validation
```

Les validations croisées présentes dans Museo ToolBox ont toute la particularité d'être stratifiées, c'est-à-dire qu'elles raisonnent par classe. Ainsi, lorsque l'utilisateur demande à séparer son jeu de données avec 50% des effectifs pour la validation et 50% des effectifs pour l'entraînement, la sélection se fera par classe : si la classe **a** est composée de 20 échantillons et une classe **b** de 200 échantillons, la première classe sera divisée en 2 (10 échantillons de chaque) et la seconde aura 100 échantillons pour l'entraînement et 100 pour la validation. Dans le cas

d'une séparation non stratifiée, il aurait été fort probable que la classe très minoritaire, ici la classe **a**, ait été manquante à l'entraînement ou à la validation étant donné son très faible effectif.

Les principales méthodes de validation sont illustrées dans la Figure 5.5. Nous évoquerons ensuite uniquement comment mettre en place une validation croisée spatialisée.

Une validation croisée spatialisée signifie qu'il faut définir une distance minimale entre une référence d'entraînement et une référence de validation. Pour cela, le moyen le plus simple trouvée pour rendre la méthode compatible avec la bibliothèque *scikit-learn* est de définir une instance de la validation croisée à partir d'une matrice de distance pré-calculée. Cela s'explique car pour générer chaque pli de la validation croisée *scikit-learn* va à partir de l'instance choisie (par exemple la classe *Spatial Leave-One-Out*) exécuter la fonction *split* avec comme argument en entrée la variable X, y et groups (si la validation croisée utilise les groupes).

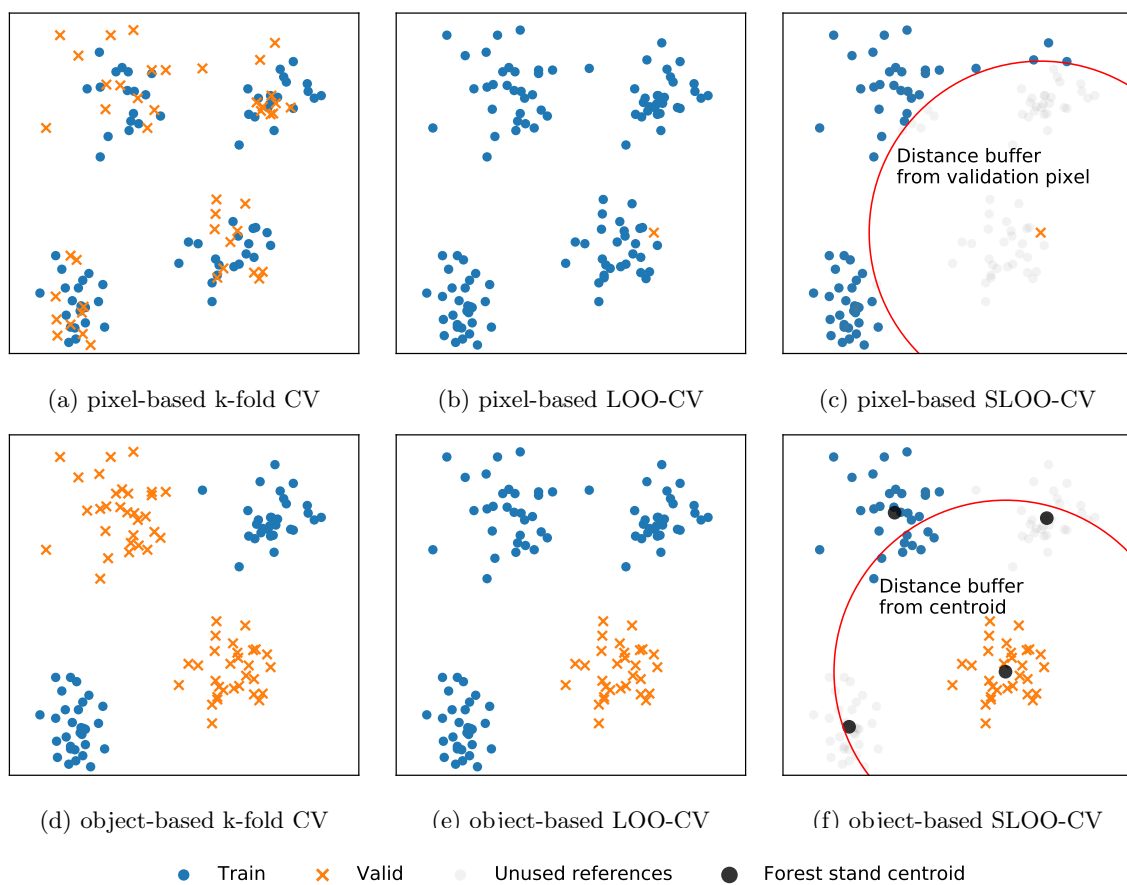


FIGURE 5.5 – Aperçu du fonctionnement des 6 principales méthodes de validation croisée et stratifiée disponibles dans Museo ToolBox.

Pour gérer la distance entre échantillons Museo ToolBox inclut une fonction, **get_distance_matrix**, qui calcule la distance euclidienne entre les points (un point = un pixel) (Code 5.3). Afin de rendre cette fonction la plus générique possible, la distance est calculée en nombre de pixels entre les points, il faut donc donner en entrée de la fonction une image et les références utilisées. Pourquoi ce choix ? Pour deux raisons : la première est que selon les projections géographiques, les distances peuvent se définir en mètres, en inch, ou en minutes, la deuxième est que les données spectrales (X) sont contenues dans un raster et que pour chaque pixel doit être assigné un label (y). Ce choix s'avère donc le plus pratique et rapide à mettre en œuvre.

```
import museotoolbox as mtb

raster,vector = mtb.datasets.load_historical_data()

# calcul une matrice de distance pour chaque référence
distance_matrix = processing.get_distance_matrix(raster,vector)
```

Code 5.3 – Calcul d’une matrice de distance des références terrain.

Ensuite, il suffit de créer une instance de la validation croisée spatialisée, ici nous avons choisi le Spatial Leave-One-Out (Code 5.4).

```
cv = mtb.cross_validation.SpatialLeaveOneOut(distance_threshold=100,
      distance_matrix=distance_matrix)
```

Code 5.4 – Instanciation d’un objet de validation croisée de type Spatial Leave-One-Out.

Une fois la validation croisée initialisée, la solution la plus simple reste de passer cet objet en paramètre d’entrée à *scikit-learn* (notamment via **SuperLearner**).

Il est toutefois possible de générer les plis (*folds*) de la validation croisée (Code 5.5).

```
# X et y sont respectivement les variables et les labels
for train,valid in cv.split(X,y) :
    # affiche la position (indice) des X et y sélectionnés
    print(train,valid)
```

Code 5.5 – Génération de la validation croisée.

Enfin il est aussi possible de générer chaque pli de la validation croisée et d’enregistrer les résultats au format de type vectoriel (comme *Geopackage*). À partir de l’instance initialisé (nommée ici **cv**), il suffit d’exécuter la fonction **save_to_vector** (Code 5.6) :

```
# pour chaque validation/entraînement, un fichier sera enregistré sous le
  format cv_train ou cv_valid suivi de son numéro. e.g. cv_train_1.gpkg.
cv.save_to_vector(vector,'Class',out_vector='cv.gpkg')
```

Code 5.6 – Sauvegarde de chaque pli de la validation croisée vers des fichiers vectoriels.

Ce qu’il faut retenir

- Toutes les validations croisées sont stratifiées (le nombre d’échantillons ou le pourcentage se fait par classe)
- Des validations croisées spatialisées sont disponibles (*Spatial Leave-One-Out*, *Spatial Leave-One-SubGroup-Out*, ou *Spatial Leave-Aside-Out*)
- Possibilité de sauvegarder au format vectoriel chaque validation et entraînement
- Compatibilité totale avec la bibliothèque *scikit-learn*

5.5.2.4 Module graphique

```
from museotoolbox.charts import PlotConfusionMatrix
```

Le module graphique de Museo ToolBox permet de produire des matrices, y compris des matrices de confusion au format pdf, mais en les améliorant avec des éléments supplémentaires comme :

- Le F1 par classe
- L’accord utilisateur et producteur
- La coloration de la diagonale d’une palette de couleurs différente du reste de la matrice

- La moyenne par ligne s'il ne s'agit pas d'une matrice de confusion où X et Y sont de même dimension (e.g. matrices des omissions pour chaque classe et chaque année par exemple)

Un exemple permet de générer graphiquement une matrice de confusion en utilisant un dégradé de vert pour les accords (diagonale) et un dégradé de rouge pour les erreurs (Code 5.7). Il est aussi possible d'ajouter sur le graphique soit l'indice de qualité par classe F1 soit la précision utilisateur et producteur .

La Figure 5.6 montre les deux grands types de matrices de confusion qu'il est possible de réaliser en seulement quelques lignes de codes, et qui permettent de mieux analyser et visualiser les résultats.

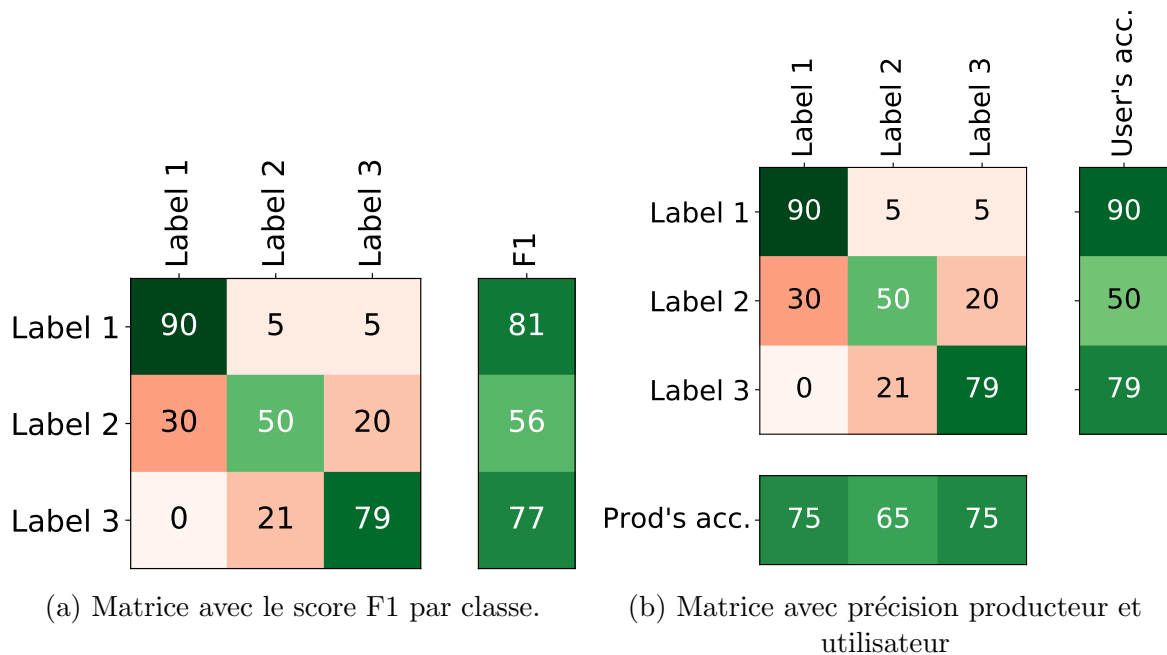


FIGURE 5.6 – Les différents types de matrices de confusion améliorées réalisées avec la classe PlotConfusionMatrix de Museo ToolBox. Ces figures ont été écrites en utilisant le Code 5.7

```
import numpy as np
from museotoolbox import charts

plot = charts.PlotConfusionMatrix(cm=np.array
    ([[90, 5, 5], [30, 50, 20], [0, 21, 79]]), right=1.05)
plot.add_text(font_size = 20)

labels = ['Label 1', 'Label 2', 'Label 3']
plot.add_x_labels(labels)
plot.add_y_labels(labels)

plot.color_diagonal('Greens', 'Reds')
plot.add_f1() # exemple (a)
# plot.add_accuracy() # exemple (b)
```

Code 5.7 – Générer une matrice de confusion améliorée.

5.5.3 Présentation de Museo Pheno

Museo Pheno est venu d'un besoin plus spécifique que Museo ToolBox mais tout autant récurrent, celui des traitements des séries temporelles (comme le calcul d'indices par date, le lissage de la donnée). Comme l'idée de base de Museo ToolBox est d'exploiter au maximum les fonctions pour tableaux et de les rendre très faciles d'accès avec des données raster, Museo

Pheno apporte de nombreuses fonctionnalités dédiées aux tableaux de séries temporelles. Le choix de créer une deuxième bibliothèque s'explique par le fait de laisser la bibliothèque Museo ToolBox la plus généraliste et de ne pas la surcharger. La documentation est disponible en ligne via le site *Read the Docs* (<https://museopheno.readthedocs.io>) et le code est hébergé sur GitHub (<https://github.com/nkarasiak/MuseoPheno>).

On peut lister parmi les besoins récurrents dans l'utilisation des séries temporelles en télédétection :

- Le calcul d'indices spectraux à partir d'une ou plusieurs dates
- Le lissage d'une série temporelle
- Le rééchantillonnage temporel
- Le calcul des métriques phénologiques

Un schéma général de la bibliothèque présente les présentes fonctionnalités et l'organisation de la bibliothèque Museo Pheno (5.7).

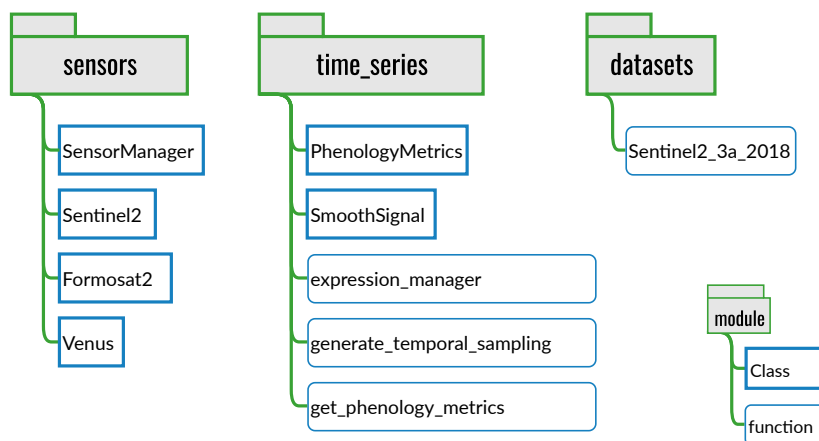


FIGURE 5.7 – La bibliothèque Museo Pheno est organisée autour des deux modules suivants : *sensors* et *time_series*. Le module *datasets* sert uniquement pour fournir un jeu de données test.

5.5.3.1 Module dédié aux capteurs

```
from museopheno import sensors
```

Le module dédié aux capteurs est ce qui permet à Museo Pheno une grande adaptabilité. En utilisant un capteur prédéfini (comme Sentinel-2 ou Formosat-2), l'ordre, le nom des bandes et la liste des indices est déjà prédéfinie (Code 5.8). Si un capteur n'a pas été prédéfini, il est très aisé d'en ajouter un nouveau avec la classe **SensorManager**. Des nouveaux capteurs peuvent être ajoutés côté utilisateur, mais ils peuvent l'être aussi de manière définitive à la bibliothèque en contribuant au code de Museo Pheno et en soumettant une contribution de type *pull request* sur GitHub.

Le module **sensors** définit un capteur par :

- le nom de ses bandes
- l'organisation de ses bandes
- l'organisation des dates par date
- la formule des indices disponibles (NDVI...)

```
import numpy as np
from museopheno import sensors, datasets
```

```

# Création de deux pixels, avec respectivement 1 ou 2 dates
X_singledate = np.array([[5,4,1,9],[4,4,3,9]])
X_twodates = np.array([[5,4,1,9,5,7,1,18],[4,4,3,9,4,7,3,19]])

# Initilisation du capteur avec 4 bandes (bleu, vert, rouge et infrarouge)
sensor = sensors.Sentinel2(n_bands=4)

# Pour connaitre les indices prédéfins, il suffit de taper la commande
sensor.available_indices

ndvi_singledate = s2.generate_index(X_singledate,s2.get_index_expression('
NDVI'))
>>> array([[0.8],[0.5]], dtype=float32)

ndvi_twodates = s2.generate_index(X_twodates,s2.get_index_expression('NDVI'
))
>>> array([[0.8 , 0.8947368 ], [0.5, 0.72727275]],dtype=float32)

# Il est aussi possible de générer un calcul à la volée
s2.generate_index(X_twodates,'B3/B2+1')
>>> array([[1.8, 2.4 ], [2., 2.75]], dtype=float32)

```

Code 5.8 – Générer un indice à partir de Sentinel-2.

Il est à noter que l'ajout des indices pour chaque capteur se fait à l'aide du numéro des bandes (qui peut contenir une lettre) précédé de la lettre B. L'ensemble des opérations nécessaires pour réaliser les calculs sont faits à partir de la fonction **expression_manager** située dans le module **museopheno.time_series**. Afin de ne pas avoir à redefinir à chaque expression l'ordre des bandes ou l'ordre des dates, il est préférable de passer systématiquement par une instance d'un capteur.

5.5.3.2 Module dédié aux séries temporelles

```
from museopheno import time_series
```

Le module **time_series** regroupe l'ensemble des traitement, de l'exécution d'une formule avec **expression_manager**, du rééchantillonnage temporel et du lissage des séries temporelles avec **SmoothSignal**, aux calculs des métriques phénologiques **PhenologyMetrics** et **get_phenology_metrics**.

expression_manager

```
from museopheno.time_series import expression_manager
```

La fonction **expression_manager** s'occupe de l'ensemble des calculs à partir d'un tableau et de l'ordre des bandes. C'est elle qui va analyser l'expression donnée par l'utilisateur soit directement via cette classe soit via les expressions des capteurs et exécuter le calcul.

En se basant sur le code précédent (5.8), il est possible de reproduire le même résultat, mais cela nécessite à chaque calcul de ressaisir les mêmes paramètres (Code 5.9). De plus la module **sensors** se charge de faire le lien entre un capteur prédéfini et l'ensemble des fonctions et classes du module **time_series**. Ainsi, il ne sera demandé qu'une fois les dates s'il y a besoin de rééchantillonner temporellement la donnée ou de calculer les métriques phénologiques par exemple.

```

time_series.expression_manager(X_twodates, bands_order=['2','3','4','8'],
expression='(B8-B4)/(B8+B4)')
>>> array([[0.8 , 0.8947368 ], [0.5 , 0.72727275]], dtype=float32)

```

Code 5.9 – Générer un indice NDVI directement à l'aide expression_manager.

SmoothSignal

```
from museopheno.time_series import SmoothSignal
```

La classe **SmoothSignal** est plus complexe que les classes et fonctions précédentes car elle fait appel à beaucoup plus de fonctions externes et intègre la gestion des dates. Ses principales fonctionnalités sont :

- Le lissage des séries temporelles selon des algorithmes prédéfinis
- Le rééchantillonnage temporel (qui se fait automatiquement selon les dates à obtenir définies par l'utilisateur)

Le rééchantillonnage temporel est effectué directement à partir de la méthode **interpolation** de **SmoothSignal**. Concernant le lissage des séries temporelles, les algorithmes disponibles aujourd'hui sont :

- Savitzky-Golay [Chen et al., 2004; Schafer, 2011]
- Double logistique [Beck et al., 2006; Zhang, 2015]
- Médiane mobile [Reed et al., 1994]

Concernant la double logistique, le code est encore en cours de validation.

PhenologyMetrics

```
from museopheno.time_series import PhenologyMetrics
```

La méthode mise en place dans Museo Pheno pour extraire les métriques phénologiques (début, longueur et fin de saison) se base sur un seuil qui est fonction de l'amplitude entre le minimum saisonnier et le maximum. Afin d'extraire les dates de début et de fin de saison, il a été développé la fonction **get_phenology_metrics**.

5.5.4 Qualité

Créer une bibliothèque python et la certifier, ce n'est pas aussi simple que de sauvegarder son code et de le publier en ligne. Des règles strictes existent afin de valider le plus possible le code et de s'assurer de son bon fonctionnement présent. De plus il convient de s'assurer de la pérennité du fonctionnement d'un code, c'est-à-dire que toutes les nouvelles fonctionnalités d'un code ne devront en aucun cas modifier les précédents comportements de la bibliothèque.

5.5.4.1 Versions et changements

Le versioning

Dans le cadre de cette thèse, nous avons utilisé le logiciel gratuit et open-source git car il permet de versionner n'importe quel type de données, c'est-à-dire que chaque modification apportée à un document sera enregistrée [Loeliger and McCullough, 2012]. Le plus célèbre site qui met à disposition le logiciel git et qui fournit de l'espace disque pour sauvegarder ses projets est GitHub.com qui a été racheté par Microsoft en 2018 pour 7,5 milliards de dollars.

La condition pour pouvoir créer et héberger des dépôts git sur GitHub est de les ouvrir au public car cela participe au partage de la connaissance (les conditions ont évolué récemment et permettent d'avoir quelques dépôts privés). Bien qu'ayant un objectif de rentabilité sûrement plus fort depuis le rachat par Microsoft, la philosophie open source semble toujours dominer aujourd'hui chez GitHub. Le partage en accès libre de tous nos codes fait qu'il n'est pas rare de trouver des réponses à nos problèmes dans des codes déjà existants dans GitHub. En effet, il est fréquent que les problèmes que nous rencontrons ont déjà été rencontrés par d'autres personnes.

En plus d'historiser chaque modification, Git permet de travailler de manière collaborative. Pour cela, la manière la plus simple d'y parvenir est de créer des branches puis de les

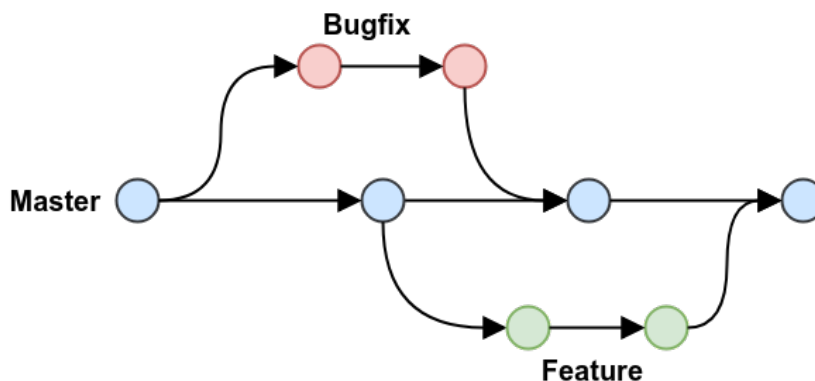


FIGURE 5.8 – Exemple d’usage des branches dans un git. La branche *Master* est la branche principale, et chaque modification peut être effectuée sur une branche distance qui sera par la suite fusionnée sur la branche principale. Cela est particulièrement recommandé lors d’un travail collaboratif. Chaque point représente une modification historisée.

fusionner une fois que l’ensemble des contributeurs auront apporté leur pierre à l’édifice (Figure 5.8). Museo ToolBox et Museo Pheno fonctionnent avec deux branches : la branche *master* contient uniquement les versions publiées, et la branche *develop* contient l’ensemble des modifications qui ont lieu jusqu’à ce qu’une version soit figée.

Une fois que le code a bien évolué et que les fonctionnalités correspondent à un besoin défini, une version sera créée. Ainsi, un utilisateur d’une bibliothèque pourra facilement se repérer dans l’évolution du code et installer une version particulière.

Le nommage des versions

Les versions sont nommées selon la norme semantic versioning numéro 2⁷. Le numéro de la version est composé dans 3 nombres (MAJEUR.MINEUR.PATCH), classés du plus au moins important :

1. la version MAJEURE change si la rétrocompatibilité n’est pas assurée.
2. la version MINEURE ajoute des fonctionnalités tout en assurant la compatibilité avec les précédentes versions.
3. le version du PATCH est utilisée lorsque des corrections de bugs ont été effectuées sans apporter de nouvelles fonctionnalités à l’utilisateur final.

Toutefois, la version majeure numérotée 0 (0.x.y) signifie que la bibliothèque est en cours de développement. Tout peut changer, mais les développeurs doivent tout de même essayer de maintenir la rétrocompatibilité entre les versions. Ainsi, la bibliothèque très utilisée dans le cadre de l’apprentissage automatique, *scikit-learn*, en est après des années de développement à sa version 0.23 au 21 juillet 2020.

En plus de ces trois nombres, les versions de tests (alpha, bêta, ou release candidate) sont spécifiées à la suite des nombres après un tiret. Par exemple 0.12.0-beta.1 signifie que la version majeure 0, mineure 12, et le patch 0 en est à sa première version bêta.

Il est ainsi possible d’installer n’importe quelle version voulue à partir du git, ce qui permet aisément de télécharger à nouveau une version utilisée pour reproduire de manière parfaite les calculs d’une étude publiée en particulier.

L’historique des changements par version

Le *changelog* est un fichier qui résume l’historique des changements en cours ou pour les versions précédentes. Ce fichier est destiné aux humains, il doit donc être très compréhensible et doit permettre à l’utilisateur final de savoir quelles sont les derniers changements du code.

7. <https://semver.org>

Le changelog est écrit selon la norme de *Keep A Changelog* 1.0.0⁸. Il est possible de spécifier pour chaque version ces différents éléments :

- **Added** pour l'ajout de nouvelles fonctionnalités.
- **Changed** pour les changements dans les fonctionnalités existantes.
- **Deprecated** pour les fonctionnalités qui seront supprimées dans un futur proche.
- **Removed** pour les fonctionnalités supprimées.
- **Fixed** pour les bugs corrigés.
- **Security** en cas de vulnérabilité. corrigées.

L'historique des derniers changements de la version de développement est consultable directement sur le GitHub de Museo ToolBox au format *markdown*⁹ ou pour la dernière version publiée directement dans la section Changelog de la documentation¹⁰.

5.5.4.2 Tests et intégration continue

Publier une bibliothèque c'est aussi être responsable des bugs. Comme nous l'avons vu au début de la partie, il arrive que certains bugs affectent potentiellement des dizaines, voire des centaines de publications, et il est rarement possible de valider à nouveau les articles impactés.

Pour éviter au maximum les erreurs, il convient d'établir des protocoles de tests et de les appliquer sur de multiples plateformes (différentes version de python ou de système d'exploitation par exemple) à chaque mise à jour.

La notion de test est cruciale en développement. Il s'agit de s'assurer que les fonctions préalablement écrites fassent bien ce qui est demandé. À chaque nouvelle mise à jour, ce comportement est vérifié sur l'ensemble du code, et si une différence est observée entre ce qui est attendu et ce qui est renvoyé par une fonction, alors l'utilisateur final et le développeur seront avertis par un badge rouge (Figure 5.9). Il est possible de connaître les tests qui ont échoué en cliquant sur ce badge.



(a) Au moins un test a échoué.

(b) Les tests ont réussi avec succès.

FIGURE 5.9 – Un des deux badges apparaît sur les pages de Museo ToolBox afin d'avertir sur la réussite des tests de la bibliothèque. Si le badge est rouge et indique donc *failing*, cela signifie que la version en cours de développement rencontre au moins un problème à corriger.

Les tests sont donc des contrôles définis manuellement et plus une bibliothèque évolue plus le nombre de tests augmente.

À chaque nouvelle modification envoyée sur GitHub, le code est testé par la plateforme d'intégration continue nommée Travis CI¹¹. Il est aussi possible de tester localement la bibliothèque. Travis CI, tout comme GitHub, est gratuit d'utilisation si le projet s'intègre dans une démarche open source. Une procédure a été mise en place pour Museo ToolBox afin de ne soumettre une nouvelle version que si l'ensemble des tests a réussi avec succès. Cette procédure réduit grandement le nombre de bugs possibles, mais est loin d'être infaillible.

Certaines parties du code sans par exemple non testés, et il est possible de savoir lesquels afin de les ajouter dans les tests futurs. Pour cela nous utilisons l'outil *coverage* qui permet de connaître la couverture du code, c'est-à-dire le pourcentage du code testé. Ainsi plus un code aura un haut pourcentage de couverture, moins il sera soumis aux bugs [Brader et al., 2012]. Cependant, la couverture est une notion simple qui vérifie juste si le code a bien été

8. <https://keepachangelog.com>

9. <https://github.com/nkarasiak/MuseoToolBox/blob/develop/CHANGELOG.md>

10. <https://museotoolbox.readthedocs.io/>

11. <https://travis-ci.org/>

exécuté par un protocole de test. Il reste à la charge du développeur de bien définir les tests et de vérifier du mieux possible que toutes les utilisation possibles d'une fonction retourne bien ce qui est attendu. Le pourcentage de couverture du code est calculé systématiquement après chaque test par Travis CI et les résultats sont envoyés sur codecov.io¹² qui permet d'historiser et de consulter les différentes évolution de la couverture du code de la bibliothèque.

Ce qu'il faut retenir

Les bibliothèques python développées pour cette thèse utilisent :

- le protocole git pour enregistrer et historiser l'ensemble des modifications.
- le site GitHub pour héberger les bibliothèques
- des tests pour vérifier à chaque modification le bon fonctionnement du code.
- la norme Semantic Versioning 2.0 pour numéroter les versions.
- la norme de KeepAChangelog pour documenter les changements de chaque version.
- la bibliothèque Sphinx pour générer une documentation automatique à partir d'un code commenté avec la norme NumPy Style.

5.5.4.3 Documentation et exemples

Une documentation est essentielle pour comprendre comment utiliser une bibliothèque. Des exemples aussi bien généralistes que spécifiques permettent aussi de bien cerner l'usage qu'un utilisateur pourra se faire de la bibliothèque.

Tout comme les tests, la documentation est à automatiser afin qu'elle soit générée pour chaque nouvelle version. L'utilisateur d'une version spécifique pourra ainsi consulter la documentation et les exemples dédiés à sa version.

Généré automatiquement à partir d'une configuration et d'exemples écrits en python par le développeur, le site [readthedocs](https://readthedocs.org/)¹³ générera grâce à ces informations une documentation en ligne.

Documentation du code

La *docstring* (de l'anglais *documentation string*) est une chaîne de caractères placée à un endroit spécifique du code (en dessous du nom des fonctions et classes) et qui sert de documentation du code. Il y a plusieurs conventions pour écrire cette documentation. Pour les bibliothèques Museo ToolBox et Museo Pheno nous avons opté pour la *NumPy Python Docstring*¹⁴ en raison de sa facilité d'écriture, de compréhension et sa large utilisation (numpy étant l'une des bibliothèques principale en python).

L'un des grands avantages de respecter une *docstring* est sa compatibilité avec des outils de génération de documentation. Par exemple, dans les *Interface Development Environnement* (IDE) qui sont les outils utilisés par les programmeurs pour écrire leurs codes, il suffit de taper le nom d'une fonction ou d'une classe pour avoir la documentation qui s'affiche avec une mise en page facilitant sa compréhension. Des logiciels comme Sphinx permettent aussi de générer automatiquement un site internet qui contiendra l'ensemble de la docstring qui sera mise en page (Figure 5.10).

En plus de documenter chaque fonction et chaque classe, des multiples exemples ont été écrits en respectant la nomenclature demandée du logiciel Sphinx. Un dossier *examples*, situé à la racine du git de Museo ToolBox, répertorie tous les exemples en utilisant la même organisation que le code original, à savoir qu'à chaque dossier correspond un module python.

12. <https://codecov.io>

13. <https://readthedocs.org/>

14. https://sphinxcontrib-napoleon.readthedocs.io/en/latest/example_numpy.html

```

class RasterMath:
    """
    Read one or multiple rasters per block, and perform one or many functi
    If you want a sample of your data, just call :func:`~museotoolbox.proc

    The default option of rasterMath will return in 2d the dataset :
    - each line is a pixel with in columns its differents values in b

    If you want to have the data in 3d (X,Y,Z), masked data will be given

    Parameters
    -----
    in_image : str.
        Path of a gdal extension supported raster.
    in_image_mask : str or False, optional (default=False).
        If str, path of the raster mask. Value masked are 0, other are cor
        Use ``invert=True`` in :mod:`museotoolbox.processing.image_mask_fr
    return_3d : bool, optional (default=False).
        Default will return a row per pixel (2 dimensions), and axis 2 (b
        If ``return_3d=True``, will return the block without reshape (not
        block_size : list or False, optional (default=[256,256]).
        Define the reading and writing block size. First element is the n
        If False, will use the block size as defined in in_image.
        To define later the block_size, use `custom_block_size`.
    message : str, optional (default='rasterMath...').
        If str, the message will be displayed before the progress bar.
    verbose : bool or int, optional (default=True).
        The higher is the int verbose, the more it will returns informati

    Examples
    -----
    >>> import museotoolbox as mtb
    >>> raster,_= mtb.datasets.load_historical_data()
    >>> rM = mtb.processing.RasterMath(raster)
    Total number of blocks : 15
    >>> rM.add_function(np.mean,out_image='/tmp/test.tif',axis=1,dtype=np.
    Using datatype from numpy table : int16.
    Detected 1 band for function mean.
    >>> rM.run()
    rasterMath... [#####]100%
    Saved /tmp/test.tif using function mean
    """

```

(a) Documentation du code de RasterMath

```

class museotoolbox.processing.RasterMath(in_image, in_image_mask=False, return_3d=False,
block_size=[256, 256], message='rasterMath...', verbose=True)
[source]

```

Read one or multiple rasters per block, and perform one or many functions to one or many geotiff raster outputs. If you want a sample of your data, just call `get_random_block()`.

The default option of rasterMath will return in 2d the dataset :

- each line is a pixel with in columns its differents values in bands so masked data will not be given to this user.

If you want to have the data in 3d (X,Y,Z), masked data will be given too (using numpy.ma).

Parameters

- `in_image` (str.) – Path of a gdal extension supported raster.
- `in_image_mask` (str or False, optional (default=False)) – If str, path of the raster mask. Value masked are 0, other are considered not masked. Use `invert=True` in `museotoolbox.processing.image_mask_from_vector` to mask only what is not in polygons.
- `return_3d` (bool, optional (default=False)) – Default will return a row per pixel (2 dimensions), and axis 2 (bands) are columns. If `return_3d=True`, will return the block without reshape (not suitable to learn with sklearn).
- `block_size` (list or False, optional (default=[256,256])) – Define the reading and writing block size. First element is the number of columns, second element the number of lines per block. If False, will use the block size as defined in `in_image`. To define later the `block_size`, use `custom_block_size`.
- `message` (str, optional (default='rasterMath...')) – If str, the message will be displayed before the progress bar.
- `verbose` (bool or int, optional (default=True)) – The higher is the int verbose, the more it will returns informations.

Examples

```

>>> import museotoolbox as mtb
>>> raster,_= mtb.datasets.load_historical_data()
>>> rM = mtb.processing.RasterMath(raster)
Total number of blocks : 15
>>> rM.add_function(np.mean,out_image='/tmp/test.tif',axis=1,dtype=np.int16)
Using datatype from numpy table : int16.
Detected 1 band for function mean.
>>> rM.run()
rasterMath... [#####]100%
Saved /tmp/test.tif using function mean

```

(b) Page générée d'après la documentation

FIGURE 5.10 – La documentation python de Museo ToolBox est écrite selon le style développé par la bibliothèque Numpy. Cette convention permet de convertir de manière automatique une page internet à partir du code python ou d'afficher l'aide d'une fonction dans l'interface de développement.

Afin que la documentation soit systématiquement à jour, elle sera générée à chaque nouvelle version du code et sera directement publiée en ligne. Si une erreur est détectée dans du calcul de la nouvelle documentation, une erreur apparaîtra alors sur la page de la documentation (Figure 5.11).

docs failing

(a) La documentation a échoué.

docs passing

(b) La documentation a réussi avec succès.

FIGURE 5.11 – Un des deux badges apparaît sur les pages de Museo ToolBox afin d'avertir de l'état de la documentation. Si le badge est indiqué *failing*, cela signifie que la version en cours rencontre un problème à corriger.

Exemples

Une bibliothèque bien documentée c'est aussi une bibliothèque facile à prendre en main. Pour cela, les exemples participent à l'apprentissage de la bibliothèque en fournissant à la fois des éléments de compréhension du fonctionnement du code, mais aussi des idées de son utilisation.

Les exemples permettent aussi de tester le code car ils sont générés à chaque fois qu'une nouvelle version du code est publiée.

Si le logiciel Sphinx est très utilisé pour générer des pages de documentation à partir d'un code c'est qu'il est très performant et très paramétrable. On y trouve de nombreuses

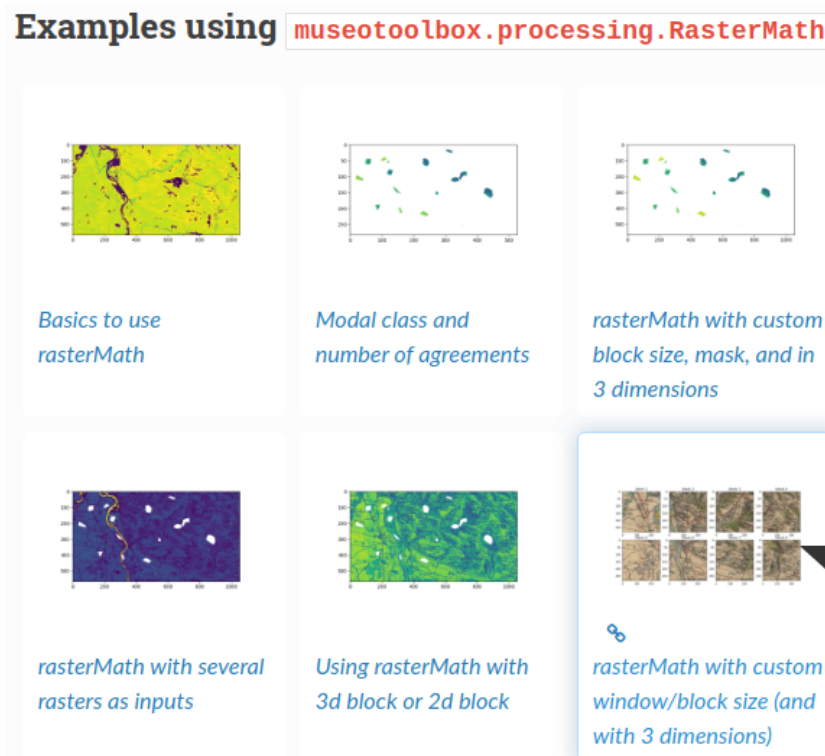


FIGURE 5.12 – Liste des exemples utilisant la classe RasterMath.

extensions qui permettent par exemple d’enrichir le site généré en proposant, pour chaque fonction ou classe, de lister chaque exemple s’y référant (Figure 5.12). Ainsi l’utilisateur qui cherche à utiliser une classe comme **RasterMath** aura non seulement accès à la documentation du code mais aussi à des exemples d’utilisation.

Une fois l’exemple sélectionné, une page explique comment reproduire cet exemple pas à pas. Cette page a été écrite au préalable en python en respectant les normes de Sphinx pour que la conversion puisse se faire automatiquement en HTML. Il est aussi possible de télécharger directement le fichier python de l’exemple.

Ce qu’il faut retenir

Les bibliothèques python développées pour cette thèse utilisent :

- de nombreux exemples, à la fois basiques et spécifiques pour illustrer les possibilités des bibliothèques.
- la norme de numpy pour documenter le code (docstring).
- la bibliothèque sphinx et ses extensions pour générer la documentation (hébergée sur readthedocs.org).

La documentation de Museo ToolBox est accessible sur :

<https://museotoolbox.readthedocs.org>.

La documentation de Museo Pheno est accessible sur :

<https://museopheno.readthedocs.org>.

5.5.5 Un outil facilitant le lien entre la recherche et les étudiants.

5.5.5.1 Un outil pédagogique...

La création de Museo ToolBox a permis d’aborder beaucoup plus aisément la transition entre les cours d’introduction à la télédétection (via des logiciels comme QGIS et OTB)

et les cours d'approfondissement qui nécessitent un usage de la programmation en python. En effet, grâce à la fonction `extract_ROI` et à la classe `RasterMath`, nous avons pu nous focaliser sur l'usage de la bibliothèque `scikit-learn`. D'un point de vue pédagogique, Museo ToolBox est donc un outil qui permet de se concentrer sur les parties les plus essentielles (comment choisir un algorithme, ses paramètres, comment choisir ses références terrain...), car l'interfaçage entre un algorithme et des données de télédétection est désormais plus aisée. Cela n'a pas empêché de voir avec les étudiants les bases de l'ouverture, de la lecture et de l'écriture d'une image à partir de la bibliothèque GDAL (sur laquelle repose Museo ToolBox).

Un autre point important de Museo ToolBox est de faire le lien entre la recherche et l'enseignement. Comme la bibliothèque Museo ToolBox a été développée pour être générique et avec une documentation conséquente pour assurer la reproductibilité des travaux de recherche, elle a été vite adoptée par les étudiants. La mise à disposition dans Museo ToolBox des méthodes de validations croisées spatiales développées lors de cette thèse, leur compatibilité avec des bibliothèques comme `scikit-learn` qui sont elles aussi développées par des chercheurs permet de montrer que la recherche peut être très opérationnelle.

5.5.5.2 ...amélioré par les étudiants

Museo ToolBox est un outil en développement qui répond aux principaux besoins de la télédétection que j'avais identifié dans ma thèse. Cependant, mes besoin de thèses ne sont pas représentatifs des besoins globaux de ce domaine. Si les fonctions de base sont génériques et utiles à la plupart des utilisateurs, mais certains travaux ne pourront pas se faire à cause de certaines limitations dans le code. Afin de pallier à ces blocages et d'améliorer significativement la généricité de Museo ToolBox il a été proposé un atelier aux étudiants du Master 2 SIGMA dans lequel j'ai enseigné la télédétection durant ma thèse. Cet atelier, dédié à la classe `RasterMath`, s'est divisé en deux parties : la première consacrée à la parallélisation des calculs, et la seconde à la gestion des filtres spatiaux (e.g. une moyenne tenant en compte de la valeur des voisins) [Antonetti et al., 2020].

Reprendre un code est souvent compliqué car il a été pensé par une personne, et il y a un besoin de comprendre le fonctionnement, ce qui peut prendre du temps de mise en route. Malgré les nombreux commentaires et la documentation complète de Museo ToolBox, certaines fonctions de `RasterMath` n'étaient pas suffisamment spécifiques pour un usage précis et elles regroupaient de nombreuses opérations, ce qui a rendu difficile la mise en route de l'atelier. Les étudiants ont dû réécrire avec ma collaboration une partie du code afin de séparer au maximum toutes les opérations de `RasterMath`, c'est-à-dire la lecture, l'écriture, la gestion des masques, le calcul par bloc, etc... Une fois cette réécriture effectuée, les étudiants ont pu apporter la parallélisation du code assez facilement. Pour les calculs utilisant des filtres spatiaux, la mise en production du code n'a pu aboutir car l'atelier se divisait en deux parties et le temps consacré à Museo ToolBox était déjà dépassé. Un code fonctionnel gérant les filtres spatiaux est cependant opérationnel dans la branche `spatial_filter`¹⁵ du GitHub de Museo ToolBox mais elle n'inclut pas les grands changements effectués par l'autre groupe (meilleure séparation du code par fonction).

5.6 Conclusion

La création des outils puis leur partage dans le cadre de cette thèse m'apparaît comme un besoin indispensable pour plusieurs raisons. Au delà de simplement respecter les principes FAIR, le partage des données et des codes apporte inconditionnellement un sentiment de qualité et de transparence.

La figure 5.13 résume l'ensemble des moyens mis en œuvre pour respecter les principes FAIR du mieux que possible. Les données sont indexées et trouvables via la plateforme Ze-

15. https://github.com/nkarasiak/MuseoToolBox/tree/spatial_filter

nodo, qui non seulement fournit un DOI mais aussi un espace d'archivage pour chaque nouvelle version des codes et des données. Les licences de partage sont assez permissives pour favoriser la réutilisation mais incluent un copyleft (GPL V3 pour les codes et CC-BY pour les données). Afin de reproduire les travaux, les bibliothèques python Museo ToolBox et Museo Pheno ont été développées et sont fournies avec un grand nombre d'exemple pour se les approprier le plus rapidement et facilement possible. Enfin, l'ensemble des données est partagé sous des formats non-propriétaires et ouverts (csv, jpg, gpkg, geotiff).

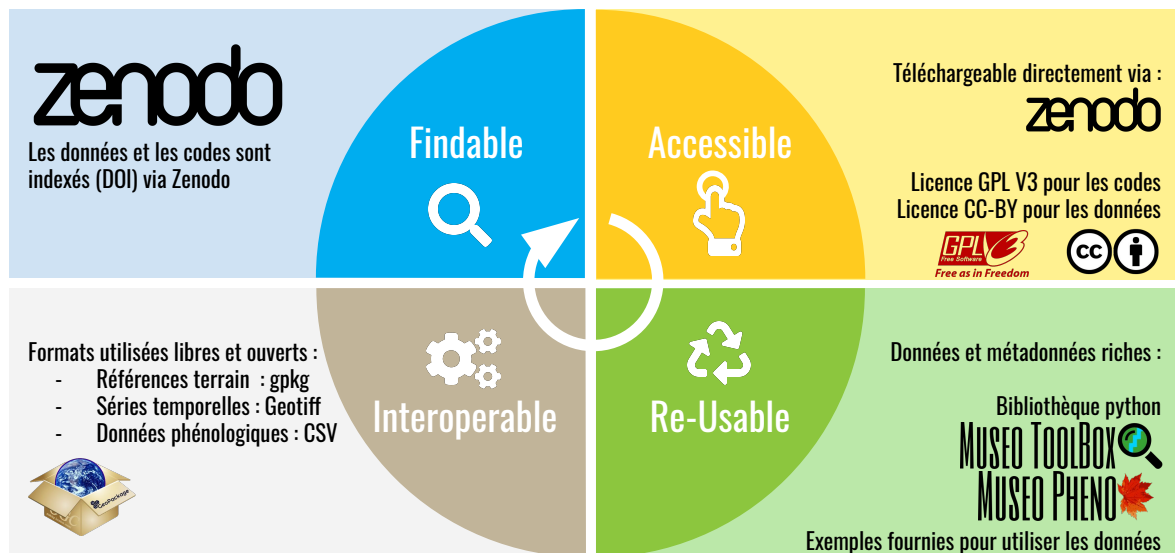


FIGURE 5.13 – Schéma résumant les principes FAIR auxquels cette thèse satisfait.

Le partage des méthodes, comme c'est le cas pour les validations croisées de type *Spatial Leave-One-Out* ou le *Spatial Leave-One-Sub-Group-Out*, permet à d'autres chercheurs de profiter de nos avancées. J'ai d'ailleurs eu des retours assez positifs de chercheuses et chercheurs, dont Patrick Schratz [Schratz et al., 2019] qui m'a dit qu'il allait très certainement considéré Museo ToolBox pour son prochain article scientifique¹⁶.

Concernant les données, il manque aujourd'hui une grande base de références terrain dédiée à la télédétection. Cette base pourrait recenser, et si possible fournir, des données terrain contenant des attributs comme la date d'acquisition, la précision géographique et la thématique de la donnée. Il serait par exemple possible de demander sur une zone données les références de forêt, et si la le type de forêt (feuillu/conifère) voire remonter même à l'essence. Cela pourrait se faire selon une nomenclature universelle. Cette base pourrait se construire à partir d'autres bases de données ouvertes mais ajouterait une valeur ajoutée en uniformisant les jeux de données entre eux. Aujourd'hui il existe quelques jeux de données pour la télédétection mais ils sont plutôt dédiés aux réseaux de neurones [Sumbul et al., 2019] ou plutôt généralistes et recensent plusieurs bases de données mais sans les agglomérer et apporter une valeur ajoutée sur la donnée en elle-même [Alemohammad, 2019].

16. <https://twitter.com/nkarasiak/status/1261558248538288130>

Chapitre 6

Conclusion générale

“The only advice, indeed, that one person can give another about reading is to take no advice, to follow your own instincts, to use your own reason, to come to your own conclusions.”

— Virginia Woolf. *The Second Common Reader*

Table des matières

6.1 Conclusion	140
6.2 Perspectives	141

6.1 Conclusion

L'objectif général de la thèse vise à identifier les blocages dans la production des cartes des essences forestières pour la cartographie à large échelle et à partir d'une méthode d'apprentissage supervisée. Cette thèse s'est intéressée spécifiquement à l'apport des séries temporelles d'images satellitaires (SITS) à partir de capteurs à haute résolution spatiale et spectrale comme celles du capteur Sentinel-2. Deux principaux défis ont été identifiés dans le travail de cette thèse, le premier consiste à prendre en compte l'autocorrélation spatiale entre les données d'entraînement et de validation des modèles et à l'évaluer. Le second défi est d'estimer l'apport de la phénologie, c'est-à-dire du cycle annuel de la végétation, dans la cartographie des essences forestières.

Pour notre problématique, le choix d'une méthode de validation doit permettre d'estimer au mieux la capacité prédictive d'un algorithme, c'est-à-dire estimer également sa capacité à extrapoler. L'état de l'art réalisé au Chapitre 1 et Chapitre 2 a mis en avant le manque de recherche sur la stabilité spatiale et annuel des cartographies. Les références terrain des essences forestières sont souvent peu nombreuses, de l'ordre de quelques dizaines à quelques centaines par essence, et très agglomérées spatialement, ce qui induit une grande autocorrélation spatiale des données. Or, les données utilisées par un algorithme de classification correspondent aux valeurs spectrales issues de données satellitaires. Donc une forte agrégation des références signifie une forte ressemblance des données.

L'arrivée de satellites à haute résolution spatiale et temporelle sans coût d'acquisition pour l'utilisateur final est une véritable opportunité pour notre domaine. En effet les différences entre essences peuvent être importantes à certaines dates. L'exemple le plus parlant est entre une essence de type feuille et une de type conifère. Si l'on regarde les deux essences en février, il est très facile de les distinguer à distance car l'une est avec aiguilles, alors que l'autre n'a plus aucune feuille.

Pour étudier la stabilité annuelle des classifications, neuf années de séries temporelles du capteur Formosat-2 ont été exploitées sur la zone d'étude de l'Observatoire Spatial Régional (OSR). Le Chapitre 2 a ainsi pu mettre en évidence un biais important de surestimation de la qualité des prédictions dû en grande partie à la présence d'une forte autocorrélation dans le jeu de données. Une méthode de validation tenant compte de la ressemblance spectrale des échantillons, dans notre cas à partir d'une mesure de distance, est donc fortement recommandée afin de diminuer le plus possible ce biais. Ce chapitre a aussi mis en évidence les limites de l'algorithme SVM sur des séries temporelles à forte densité. Celles-ci sont plus susceptibles de contenir du bruit (nuages non détectés par l'algorithme par exemple) ce qui perturbe grandement la qualité de prédiction, et ce malgré une très grande majorité de dates de bonne qualité. On évalue la surestimation de la qualité des travaux similaires aux nôtres qui n'auraient pas pris en compte l'effet de l'autocorrélation spatiale d'environ 0,40 points d'accord global.

Pour estimer en détail l'influence de l'autocorrélation spatiale sur de plus larges échelles et en fonction d'un plus grand nombre de références, le Chapitre 3 a été dédié à cette problématique. Comme il n'est pas possible d'avoir des dizaines de milliers de références terrain par essence, cette partie s'est faite en appliquant un modèle à deux classes : feuillus et conifères. Six méthodes de validation différentes ont été mises en place, allant de la moins restrictive à la plus restrictive en terme d'autocorrélation spatiale. Le nombre de données d'entraînement a été augmenté de manière non linéaire, allant de 2 à près de 2000 peuplements pour chaque type de forêt.

Après avoir étudié en détail l'impact de l'autocorrélation spatiale et la stabilité spatiale et annuelle des cartographies des forêts, la méthodologie était donc prête pour évaluer le potentiel du nouveau capteur Sentinel-2. Lancé en 2015 et 2017, ces deux capteurs observent aujourd'hui au minimum tous les 5 jours n'importe quel endroit de la terre. Depuis 2017, la répétitivité temporelle est donc à son maximum, et c'est pour cela que les années 2017, 2018

et 2019 ont été utilisées indépendamment pour estimer l'apport des dates et de la phénologie dans la cartographie des essences.

Le Chapitre 4 contient deux thématiques. La première est dédiée à la cartographie des feuillus et une méthode spécifique feuillu par feuillu a été développée afin d'estimer quelles sont les dates les plus propices pour détecter chaque essence. Seuls les feuillus ont été choisis car leur phénologie est très marquée et qu'un protocole de suivi phénologique a été mise en place dès l'automne 2017 dans le cadre de cette thèse pour suivre sept essences de feuillus. Ainsi, tous les dix jours en automne et en printemps, des relevés ont été effectués dans les forêts pour noter l'évolution de la phénologie des arbres, l'ouverture de la canopée et l'état du sous-étage. Ces mesures ont pour objectif de mieux comprendre ce qui influence le signal à cette échelle spatiale.

La deuxième partie du Chapitre 4 étudie le potentiel d'extraction des stades phénologiques. Avec une résolution temporelle de cinq jours, les observations comme la pousse des feuilles, leur coloration ou leur chute devraient être observables. Plusieurs tailles de fenêtre de Savitzky-Golay ont été comparées afin de détecter pour chaque indice (*Leaf Anthocyanid Content* (LAnthoC), *Leaf Carotenoid Content* (LCaroC), *Leaf Chlorophyll Content* (LChloC), NDVI) si un seuil optimal permettait d'observer ex situ les stades phénologiques. Conformément aux travaux de ce domaine, le début de saison se détecte avec un peu moins de dix jours d'erreur que ce soit avec le NDVI ou d'autres indices comme le LChloC qui a été pensé pour être sensible à la teneur en chlorophylle des feuilles. Cependant pour la fin de saison, il apparaît que le LCaroC semble particulièrement sensible à la sénescence. De manière global le NDVI permet toutefois d'obtenir de bons résultats et est assez stable d'une année à l'autre.

Enfin, le Chapitre 5 détaille comment cette thèse s'assure que l'ensemble des données récoltées et des méthodes développées soit accessible au plus grand nombre afin d'assurer une reproductibilité maximale des travaux. Pour cela les données récoltées (références localisées des essences et suivi phénologique) sont partagées sous la licence CC-BY sur la plateforme Zenodo avec un Digital Object Identifier (DOI). Les méthodes (e.g. traitement générique des images géoréférencées, validation croisée ou encore sélection de caractéristiques par date) ont fait l'objet de deux bibliothèques python. La principale, Museo ToolBox, a été validée par les pairs et publiée dans la revue *Journal of Open Source Software*. La seconde, Museo Pheno, est dédiée aux séries temporelles et permet de reproduire la deuxième partie de la thèse.

Les principaux résultats de cette thèse ont montré, malgré un potentiel de cartographie pour certaines essences comme le chêne rouge (*Quercus rubra*) ou le saule (*Salix alba*), des difficultés pour cartographier les essences forestières, et ce malgré une littérature plutôt optimiste dans ce domaine scientifique. Plusieurs raisons ont été recensées et évaluées pour expliquer ce phénomène. Tout d'abord, le nombre de références terrain est souvent insuffisant en plus d'être réparti très localement sur un territoire. Ensuite, pour avoir un échantillonnage temporel régulier sur l'ensemble du territoire tout en ayant beaucoup de dates pour bénéficier de l'apport de la phénologie, l'usage des séries temporelles à haute densité est donc un pré-requis. Il a été noté que dans cette thèse la mauvaise détection de quelques rares nuages pouvait impacter fortement la qualité d'un modèle, parfois de l'ordre de 40% l'accord global, soit une qualité presque divisée par deux entre les meilleures dates sélectionnées et l'ensemble de la série temporelle.

6.2 Perspectives

Plusieurs pistes de travail sont à considérer pour diminuer au maximum ces limites qui empêchent aujourd'hui la réalisation de carte des essences forestières à large échelle. Notons toutefois un grand espoir pour certaines essences présentes dans notre zone d'étude comme le chêne rouge (*Quercus rubra*) ou le saule (*Salix alba*) qui ont des signatures spectrales assez particulières comparées aux autres essences.

La première piste consiste à augmenter significativement le nombre de références localisées

d'essences forestières. Pour cela, il conviendrait de renforcer les collaborations avec des acteurs majeurs de la forêt en France. Des premiers contacts et échanges très satisfaisants ont été réalisés avec le comité des forêts. Ces premiers échanges ont été très constructifs, notamment du fait de leur sensibilité aux problématiques de la télédétection.

La deuxième piste, sur du plus long terme et à plus large échelle, consisterait à mettre en place une base de données de références localisées pour la télédétection qui inclurait :

- La qualification de l'objet selon une nomenclature déjà reconnue
- La date à laquelle la référence a été prise ou validée pour la dernière fois
- La distance approximative de validation (exemple : validité dans un rayon de 10 m par exemple)

Cette base ne serait pas spécialement dédiée aux essences forestières, mais pourrait être utilisée pour tout type d'information spatiale.

Plus spécifiquement sur la méthodologie, l'approche par les séries temporelles d'images optiques a été envisagée dans cette thèse à partir d'un capteur Sentinel-2 avec 10 m de résolution spatiale. La société Planet Labs opère depuis 2017 le satellite SkySat avec une résolution quotidienne et quatre bandes spectrales, dans le bleu (450–515 nm), vert (515–595 nm), rouge (605–695 nm), et infrarouge (740–900 nm) à deux mètres de résolution spatiale. Profiter de l'avancée dans les réseaux de neurones qui sont capables temporelles [Pelletier et al., 2019] mais aussi texture Zhang and Xie [2012].

Il pourrait être aussi envisagé de coupler les séries temporelles optiques avec des images issues de capteurs actifs comme celles du radar Sentinel-1. Cela permettrait de profiter d'informations relatives à la biomasse. Certains travaux prometteurs sur les cultures ou l'occupation du sol ont mis en évidence l'apport de cette méthode [Ienco et al., 2019; Van Tricht et al., 2018].

Pour une cartographie à l'échelle nationale, cela peut se faire dans un premier temps en étudiant la faisabilité de la cartographie pour les deux essences qui sont très bien détectées dans notre zone d'étude comme le chêne rouge (*Quercus rubra*) ou le saule (*Salix spp.*). Afin d'avoir le même échantillonnage temporel sur l'ensemble du territoire, condition sine qua non en apprentissage automatique, il pourrait être envisagé d'utiliser les produits Sentinel-2 de niveau 3A du pôle THEIA. Il s'agit de synthèses mensuelles [Hagolle Olivier, 2018] qui sont calculées à partir la moyenne des réflectances de surface observées par ciel clair, pendant une période de 45 jours [Hagolle et al., 2018] pour chaque pixel et chaque bande spectrale. La donnée à la date centrale (si la synthèse mensuelle est celle du mois de juillet il s'agit de la date du 15 juillet) aura plus de poids que les dates les plus éloignées.

Produire une première carte à l'échelle nationale de certaines essences forestières pourrait démontrer le potentiel de la télédétection. Cela favoriserait la communication autour de ces produits, ce qui pourrait déboucher sur des échanges et des contacts dans le domaine forestier afin d'enrichir la base de données des références d'essences localisées mais aussi d'avoir un retour sur la qualité des produits.

Bibliographie

- Ahl, D. E., Gower, S. T., Burrows, S. N., Shabanov, N. V., Myneni, R. B., and Knyazikhin, Y. (2006). Monitoring spring canopy phenology of a deciduous broadleaf forest using MODIS. *Remote Sensing of Environment*, 104(1) :88–95. Cité page 80, 82
- Alemohammad, H. (2019). Radiant mlhub : A repository for machine learning ready geospatial training data. *AGUFM*, 2019 :IN11A–05. Cité page 137
- Alexandre, S. (2017). Rapport de mission de la déléguée interministérielle à la forêt et au bois. Technical report, Ministère de l’Environnement, de l’Énergie et de la Mer. Cité page 5
- Anderegg, W. R. L., Konings, A. G., Trugman, A. T., Yu, K., Bowling, D. R., Gabbitas, R., Karp, D. S., Pacala, S., Sperry, J. S., Sulman, B. N., and Zenes, N. (2018). Hydraulic diversity of forests regulates ecosystem resilience during drought. *Nature*, 561(7724) :538–541. Cité page 7
- Anselin, L. (1988). *Spatial Econometrics : Methods and Models*. Kluwer Academic Publishers, Dordrecht. Cité page 64
- Anselin, L. (1989). What is special about spatial data : Alternative perspectives on spatial data analysis. Technical Paper 89-4, National Center for Geographic Information and Analysis, Santa Barbara, CA : NCGIA. Cité page 64
- Antonetti, S., Duflos, A., Ternisien de Boiville, H., and Tronc Laures, A.-S. (2020). Museo toolbox : Atelier programmation et réseaux de neurones. Technical report, Master SIGMA, Université Toulouse Jean Jaures & ENSAT. http://sigma.univ-toulouse.fr/_resources/th%C3%A9matiques/pdf/Projets2020/Atelier_MuseoToolbox-et-Reseaux-de-neurones_Rapport.pdf. Cité page 136
- Aono, Y. and Kazui, K. (2008). Phenological data series of cherry tree flowering in Kyoto, Japan, and its application to reconstruction of springtime temperatures since the 9th century. *International Journal of Climatology*, 28(7) :905–914. Cité page 8, 9
- Aragones, D., Rodriguez-Galiano, V., Caparros-Santiago, J., and Navarro-Cerrillo, R. (2019). Could land surface phenology be used to discriminate mediterranean pine species? *International Journal of Applied Earth Observation and Geoinformation*, 78 :281–294. Cité page 32, 41, 81
- Atkinson, P. and Lewis, P. (2000). Geostatistical classification for remote sensing : an introduction. *Computers & Geosciences*, 26(4) :361–371. Cité page 64, 66
- Atkinson, P. M. (2004). Spatially weighted supervised classification for remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, 5(4) :277–291. Cité page 64, 66

-
- Atkinson, P. M., Jeganathan, C., Dash, J., and Atzberger, C. (2012). Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology. *Remote Sensing of Environment*, 123 :400–417. Cité page 88
- Aval, J., Fabre, S., Zenou, E., Sheeren, D., Fauvel, M., and Briottet, X. (2019). Object-based fusion for urban tree species classification from hyperspectral, panchromatic and ndsm data. *International Journal of Remote Sensing*, 40(14) :5339—5365. Cité page 41
- Baccini, A., Goetz, S. J., Walker, W. S., Laporte, N. T., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P. S. A., Dubayah, R., Friedl, M. A., Samanta, S., and Houghton, R. A. (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change*, 2(3) :182–185. Cité page 2
- Badeau, V., Bonhomme, M., Bonne, F., Carré, J., Cecchini, S., Chuine, I., Ducatillion, C., Jean, F., Lebourgeois, F., and Seguin, B. (2017). *Les plantes au rythme des saisons : guide d'observation phénologique*. Biotope. Cité page 26, 84, 85, 87
- Baetens, L., Desjardins, C., and Hagolle, O. (2019). Validation of copernicus sentinel-2 cloud masks obtained from maja, sen2cor, and fmask processors using reference cloud masks generated with a supervised active learning procedure. *Remote Sensing*, 11(4). Cité page 70
- Bai, H., Cao, F., Atkinson, M. P., Chen, Q., Wang, J., and Ge, Y. (2020). Incorporating spatial association into statistical classifiers : local pattern-based prior tuning. *International Journal of Geographical Information Science*, pages 1–38. Cité page 84
- Baillarin, S., Meygret, A., Dechoz, C., Petrucci, B., Lacherade, S., Trémas, T., Isola, C., Martimort, P., and Spoto, F. (2012). Sentinel-2 level 1 products and image processing performances. In *2012 IEEE international geoscience and remote sensing symposium*, pages 7003–7006. IEEE. Cité page 16
- Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., Samain, O., Roujean, J. L., and Leroy, M. (2007). LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION. *Remote Sensing of Environment*, 110(3) :275–286. Cité page 32
- Barré, R., Lavoux, T., Piveteau, V., Theys, J., and Cairn (Service en ligne) (2015). *Un demi-siècle d'environnement entre science, politique et prospective : en l'honneur de Jacques Theys*. Quae. Cité page 32
- Bastin, J.-F., Berrahmouni, N., Grainger, A., Maniatis, D., Mollicone, D., Moore, R., Patriarca, C., Picard, N., Sparrow, B., Abraham, E. M., Aloui, K., Atesoglu, A., Attore, F., Bassullu, C., Bay, A., Garzuglia, M., García-Montero, L. G., Groot, N., Guerin, G., Laestadius, L., Lowe, A. J., Mamane, B., Marchi, G., Patterson, P., Rezende, M., Ricci, S., Salcedo, I., Diaz, A. S.-P., Stolle, F., Surappaeva, V., and Castro, R. (2017). The extent of forest in dryland biomes. *Science*, 356(6338) :635–638. Cité page 2
- Beale, C., Lennon, J., Yearsley, J., Brewer, M., and Elston, D. (2010). Regression analysis of spatial data. *Ecology Letters*, 13(2) :246–264. Cité page 64
- Beaudoin, A., Bernier, P., Guindon, L., Villemaire, P., Guo, X., Stinson, G., Bergeron, T., Magnussen, S., and Hall, R. (2014). Mapping attributes of Canada's forests at moderate resolution through kNN and MODIS imagery. *Canadian Journal of Forest Research*, 44(5) :521–532. Cité page 32

-
- Beck, P. S., Atzberger, C., Høgda, K. A., Johansen, B., and Skidmore, A. K. (2006). Improved monitoring of vegetation dynamics at very high latitudes : A new method using MODIS NDVI. *Remote Sensing of Environment*, 100(3) :321–334. Cité page 88, 102, 130
- Bellman, R. (1984). *Dynamic programming*. Princeton Univ. Pr, Princeton, NJ. Cité page 83, 124
- Berberoglu, S., Curran, P., Lloyd, C., and Atkinson, P. (2007). Texture classification of Mediterranean land cover. *International Journal of Applied Earth Observation and Geoinformation*, 9(3) :322–334. Cité page 64
- Bergès, L. and Dupouey, J.-L. (2017). Écologie historique et ancienneté de l'état boisé : concepts, avancées et perspectives de la recherche. *Revue Forestière Française*, (4) :297. Cité page 10
- Bhandari Neupane, J., Neupane, R. P., Luo, Y., Yoshida, W. Y., Sun, R., and Williams, P. G. (2019). Characterization of Leptazolines A–D, Polar Oxazolines from the Cyanobacterium *Leptolyngbya* sp., Reveals a Glitch with the “Willoughby–Hoye” Scripts for Calculating NMR Chemical Shifts. *Organic Letters*, 21(20) :8449–8453. Cité page 118
- Bojinski, S., Verstraete, M., Peterson, T. C., Richter, C., Simmons, A., and Zemp, M. (2014). The Concept of Essential Climate Variables in Support of Climate Research, Applications, and Policy. *Bulletin of the American Meteorological Society*, 95(9) :1431–1443. Cité page 8
- Bolyn, C., Michez, A., Gaucher, P., Lejeune, P., and Bonnet, S. (2018). Forest mapping and species composition using supervised per pixel classification of Sentinel-2 imagery. *BASE*, 22(3) :172–187. Cité page 32, 33, 41, 42, 56, 58, 105
- Bonnor, G. M. (1967). Estimation of Ground Canopy Density from Ground Measurements. *Journal of Forestry*, 65(8) :544–547. Cité page 30
- Borchert, R., Robertson, K., Schwartz, M. D., and Williams-Linera, G. (2005). Phenology of temperate trees in tropical climates. *International Journal of Biometeorology*, 50(1) :57–65. Cité page 8
- Boschetti, M., Boschetti, L., Oliveri, S., Casati, L., and Canova, I. (2007). Tree species mapping with Airborne hyper-spectral MIVIS data : the Ticino Park study case. *International Journal of Remote Sensing*, 28(6) :1251–1261. Cité page 108
- Boureau, J.-G. (2008). *Manuel d'interprétation des photographies aériennes infrarouges : application aux milieux forestiers et naturels*. Inventaire forestier national, Nogent-sur-Vernisson, la documentation française edition. Cité page 11, 20
- Boyd, D.S.Danson, F. (2005). Satellite remote sensing of forest resources : three decades of research development. *Progress in Physical Geography*, 29(1) :1–26. Cité page 41
- Boyd, W. (2010). Ways of Seeing in Environmental Law : How Deforestation Became an Object of Climate Governance. SSRN Scholarly Paper ID 1694612, Social Science Research Network, Rochester, NY. Cité page 3
- Brader, L., Hilliker, H. F., and Wills, A. C. (2012). *Testing for continuous delivery with Visual Studio 2012*. Microsoft patterns andamp ; practices. Cité page 132

-
- Brasier, C. M. and Scott, J. K. (1994). European oak declines and global warming : a theoretical assessment with special reference to the activity of *Phytophthora cinnamomi*. *EPPO Bulletin*, 24(1) :221–232. Cité page 7
- Breda, N. (1999). L'Indice foliaire des couverts forestiers : mesure, variabilité et rôle fonctionnel. *Revue Forestière Française*, page 135. Cité page 29
- Breda, N. J. J. (2003). Ground-based measurements of leaf area index : a review of methods, instruments and current controversies. *Journal of Experimental Botany*, 54(392) :2403–2417. Cité page 29
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1) :5–32. Cité page 19, 68
- Brennan, A. (2013). How To Draw an Owl.
<https://knowyourmeme.com/memes/how-to-draw-an-owl>. Cité page 119
- Brenning, A. (2012). Spatial cross-validation and bootstrap for the assessment of prediction rules in remote sensing : The R package *sperrorest*. In *2012 IEEE International Geoscience and Remote Sensing Symposium*, pages 5372–5375. Cité page 55, 65, 66
- Brockhoff, E. G., Jactel, H., Parrotta, J. A., Quine, C. P., and Sayer, J. (2008). Plantation forests and biodiversity : oxymoron or opportunity? *Biodiversity and Conservation*, 17(5) :925–951. Cité page 2
- Brown, R. and Gillespie, T. (1995). Microclimatic landscape design : Creating thermal comfort and energy efficiency. *New York*. Cité page 23
- Bunker, D. E., Declerck, F., Bradford, J. C., Colwell, R. K., Perfecto, I., Phillips, O. L., Sankaran, M., and Naeem, S. (2005). Species loss and aboveground carbon storage in a tropical forest. *Science*, 310(5750) :1029–1031. Cité page 41
- Burgat, F. (2020). *Qu'est-ce qu'une plante ? essai sur la vie végétale*. La Couleur des idées. Éditions du Seuil, Paris. Cité page 3
- Bórnez, K., Descals, A., Verger, A., and Peñuelas, J. (2020). Land surface phenology from VEGETATION and PROBA-V data. Assessment over deciduous forests. *International Journal of Applied Earth Observation and Geoinformation*, 84 :101974. Cité page 81
- Cano, E., Denux, J.-P., Bisquert, M., Hubert-Moy, L., and Chéret, V. (2017). Improved forest-cover mapping based on MODIS time series and landscape stratification. *International Journal of Remote Sensing*, 38(7) :1865–1888. Cité page 32, 41
- Cao, R., Chen, Y., Shen, M., Chen, J., Zhou, J., Wang, C., and Yang, W. (2018). A simple method to improve the quality of NDVI time-series data by integrating spatiotemporal information with the Savitzky-Golay filter. *Remote Sensing of Environment*, 217 :244–257. Cité page 88
- Carl, C., Landgraf, D., van der Maaten-Theunissen, M., Biber, P., and Pretzsch, H. (2017). *Robinia pseudoacacia* L. Flower Analyzed by Using An Unmanned Aerial Vehicle (UAV). *Remote Sensing*, 9(11) :1091. Cité page 108
- Carleer, A. and Wolff, E. (2004). Exploitation of very high resolution satellite data for tree species identification. *Photogrammetric Engineering & Remote Sensing*, 70(1) :135–140. Cité page 41
- Carter, G. A. (1991). Primary and Secondary Effects of Water Content on the Spectral Reflectance of Leaves. *American Journal of Botany*, 78(7) :916. Cité page 25

-
- Catry, T., Li, Z., Roux, E., Herbreteau, V., Gurgel, H., Mangeas, M., Seyler, F., and Dessay, N. (2018). Wetlands and Malaria in the Amazon : Guidelines for the Use of Synthetic Aperture Radar Remote-Sensing. *International Journal of Environmental Research and Public Health*, 15(3) :468. Cité page 13
- Caudullo, G., Welk, E., and San-Miguel-Ayanz, J. (2017). Chorological maps for the main European woody species. *Data in Brief*, 12 :662–666. Cité page 7
- Chazdon, R. L., Brancalion, P. H. S., Laestadius, L., Bennett-Curry, A., Buckingham, K., Kumar, C., Moll-Rocek, J., Vieira, I. C. G., and Wilson, S. J. (2016). When is a forest a forest ? Forest concepts and definitions in the era of forest and landscape restoration. *Ambio*, 45(5) :538–550. Cité page 2
- Chen, D. and Wei, H. (2009). The effect of spatial autocorrelation and class proportion on the accuracy measures from different sampling designs. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(2) :140–150. Cité page 42, 55, 62, 65, 68
- Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B., and Eklundh, L. (2004). A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. *Remote Sensing of Environment*, 91(3-4) :332–344. Cité page 88, 130
- Chouaib, H. (2011). Sélection de caractéristiques : méthodes et applications. *Paris Descartes University : Paris, France*. Cité page 92
- Chuine, I., Yiou, P., Viovy, N., Seguin, B., Daux, V., and Ladurie, E. L. R. (2004). Grape ripening as a past climate indicator. *Nature*, 432(7015) :289–290. Cité page 8
- Cinotti, B. (1996). Evolution des surfaces boisées en France : proposition de reconstitution depuis le début du XIX^e siècle. *Revue forestière française*. Cité page 3
- Clark, M. L., Buck-Diaz, J., and Evens, J. (2018). Mapping of forest alliances with simulated multi-seasonal hyperspectral satellite imagery. *Remote Sensing of Environment*, 210 :490–507. Cité page 33
- Cleland, E. E., Chuine, I., Menzel, A., Mooney, H. A., and Schwartz, M. D. (2007). Shifting plant phenology in response to global change. *Trends in Ecology & Evolution*, 22(7) :357–365. Cité page 8
- Clerici, N., Weissteiner, C. J., and Gerard, F. (2012). Exploring the Use of MODIS NDVI-Based Phenology Indicators for Classifying Forest General Habitat Categories. *Remote Sensing*, 4(6) :1781–1803. Cité page 81
- CNRS (2019). Feuille de route du CNRS pour la science ouverte. Cité page 114
- Comber, A., Fisher, P., Brunsdon, C., and Khmag, A. (2012). Spatial analysis of remote sensing image classification accuracy. *Remote Sensing of Environment*, 127 :237–246. Cité page 62
- Comber, A., Fisher, P., and Wadsworth, R. (2005). What is Land Cover ? *Environment and Planning B : Planning and Design*, 32(2) :199–209. Cité page 3
- Comber, A. J., Harris, P., and Tsutsumida, N. (2016). Improving land cover classification using input variables derived from a geographically weighted principal components analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 119 :347–360. Cité page 64, 66

-
- Congalton, R. (1998). A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data. *Photogrammetric Engineering & Remote Sensing*, 54(5) :593–600. Cité page 62, 63
- Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37(1) :35–46. Cité page 55, 62
- Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3) :273–297. Cité page 19
- Croft, H., Chen, J., and Zhang, Y. (2014). Temporal disparity in leaf chlorophyll content and leaf area index across a growing season in a temperate deciduous forest. *International Journal of Applied Earth Observation and Geoinformation*, 33 :312–320. Cité page 104
- Cánovas-García, F., Alonso-Sarría, F., Gomariz-Castillo, F., and Oñate-Valdivieso, F. (2017). Modification of the random forest algorithm to avoid statistical dependence problems when classifying remote sensing imagery. *Computers & Geosciences*, 103 :1–11. Cité page 55, 65, 66, 77
- Dai, A. (2011). Drought under global warming : a review. *WIREs Climate Change*, 2(1) :45–65. Cité page 7
- Dainville, F. d. (1955). La carte de Cassini et son intérêt géographique. *Bulletin de l'Association de Géographes Français*, 32(251) :138–147. Cité page 9
- Dale, M. R. and Fortin, M.-J. (2014). *Spatial Analysis : A Guide For Ecologists*. Cambridge University Press, Cambridge, 2 edition. Cité page 48, 91
- Danny Lämmerhirt (2017). Avoiding data use silos : How governments can simplify the open licensing landscape. <https://research.okfn.org/avoiding-data-use-silos/>. Cité page 116
- Daubemire, R. (1959). Canopy coverage method of vegetation analysis. *Northwest Sci*, 33 :39–64. Cité page 30
- De Beurs, K. M. and Henebry, G. M. (2005). Land surface phenology and temperature variation in the International Geosphere-Biosphere Program high-latitude transects. *Global Change Biology*, 11(5) :779–790. Cité page 80
- Dechesne, C., Mallet, C., Le Bris, A., and Gouet-Brunet, V. (2017). Semantic segmentation of forest stands of pure species combining airborne lidar data and very high resolution multispectral imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 126 :129–145. Cité page 11
- Dedieu, G., Karnieli, A., Hagolle, O., Jeanjean, H., Cabot, F., Ferrier, P., and Yaniv, Y. (2006). Venus : A joint Israel-French Earth Observation scientific mission with High spatial and temporal resolution capabilities. In *Recent Advances in Quantitative Remote Sensing*, pages 517–521, Valencia, Spain. J.A.Sobrino. Cité page 43
- Delbart, N., Kergoat, L., Le Toan, T., Lhermitte, J., and Picard, G. (2005). Determination of phenological dates in boreal regions using normalized difference water index. *Remote Sensing of Environment*, 97(1) :26–38. Cité page 81, 82, 89
- Delbart, N., Le Toan, T., Kergoat, L., and Fedotova, V. (2006). Remote sensing of spring phenology in boreal regions : A free of snow-effect method using NOAA-AVHRR and SPOT-VGT data (1982–2004). *Remote Sensing of Environment*, 101(1) :52–62. Cité page 81, 89, 102

-
- Delgado, R. and Tibau, X.-A. (2019). Why Cohen’s Kappa should be avoided as performance measure in classification. *PLOS ONE*, 14(9) :e0222916. Cité page 34
- Deville, J. (2009). *L’incendie meurtrier dans la forêt des Landes en août 1949*. Les Éditions des Pompiers de France, Paris. Cité page 4
- Di Mauro, N., Vergari, A., Basile, T., Ventola, F., and Esposito, F. (2017). End-to-end learning of deep spatio-temporal representations for satellite image time series classification. In *2017 ECML/PKDD Discovery Challenges*. Cité page 63
- Diao, C. and Wang, L. (2016). Incorporating plant phenological trajectory in exotic saltcedar detection with monthly time series of Landsat imagery. *Remote Sensing of Environment*, 182(Supplement C) :60–71. Cité page 41
- Dormann, C. (2007). Effects of incorporating spatial autocorrelation into the analysis of species distribution data. *Global Ecology and Biogeography*, 16(2) :129–138. Cité page 64
- Dray, S., Legendre, P., and Peres-Neto, P. (2006). Spatial modelling : a comprehensive framework for principal coordinate analysis of neighbour matrices (pcnm). *Ecological Modelling*, 196(3) :483–493. Cité page 64
- Du puy, S., Derrière, N., and Wurpillot, S. (2017). La forêt plantée en France : état des lieux. L’IF 40, IGN. Cité page 4
- Duputié, A., Rutschmann, A., Ronce, O., and Chuine, I. (2015). Phenological plasticity will not help all species adapt to climate change. *Global Change Biology*, 21(8) :3062–3073. Cité page 7
- Eilers, P. H. C. (2003). A Perfect Smoother. *Analytical Chemistry*, 75(14) :3631–3636. Cité page 57, 88
- Eklund, A., Nichols, T. E., and Knutsson, H. (2016). Cluster failure : Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, 113(28) :7900–7905. Cité page 118
- Elatawneh, A., Wallner, A., Rehush, N., Schneider, T., and Knoke, T. (2013). *Forest tree species identification using phenological stages and RapidEye data a case study in the forest of Freising*, page 21–38. GITO Verlag. Cité page 20, 33
- ESA (2015). Sentinel-2 User Handbook. https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook. Cité page 13
- ESA (2020). Copernicus Sentinel Data Access 2019, Annual report. Cité page 17
- F. Dormann, C., M. McPherson, J., B. Araújo, M., Bivand, R., Bolliger, J., Carl, G., G. Davies, R., Hirzel, A., Jetz, W., Daniel Kissling, W., Kühn, I., Ohlemüller, R., R. Peres-Neto, P., Reineking, B., Schröder, B., M. Schurr, F., and Wilson, R. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data : a review. *Ecography*, 30(5) :609–628. Cité page 64
- FAO (2001). *Global Forest Resources Assessment 2000 : Main Report*. Food and Agriculture Organization of the United Nations. Cité page 2
- Fassnacht, F. E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L. T., Straub, C., and Ghosh, A. (2016). Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment*, 186 :64–87. Cité page 19, 32, 33, 34, 41, 42, 55, 82

-
- Fauvel, M., Tarabalka, Y., Benediktsson, J. A., Chanussot, J., and Tilton, J. C. (2013). Advances in spectral-spatial classification of hyperspectral images. *Proceedings of the IEEE*, 101(3) :652–675. Cité page 64, 66
- Féret, J.-B. (2009). *Apport de la modélisation pour l'estimation de la teneur en pigments foliaires par télédétection*. PhD thesis, Paris 6. Cité page 88
- Féret, J. B. and Asner, G. P. (2012). Tree species discrimination in tropical forests using airborne imaging spectroscopy. *IEEE Transactions on Geoscience and Remote Sensing*, 51(1) :73–84. Cité page 41
- Ferreira, M. P., Zortea, M., Zanotta, D. C., Shimabukuro, Y. E., and de Souza Filho, C. R. (2016). Mapping tree species in tropical seasonal semi-deciduous forests with hyperspectral and multispectral data. *Remote Sensing of Environment*, 179(Supplement C) :66–78. Cité page 19
- Fichtner, A., Schnabel, F., Bruelheide, H., Kunz, M., Mausolf, K., Schuldt, A., Härdtle, W., and Oheimb, G. (2020). Neighbourhood diversity mitigates drought impacts on tree growth. *Journal of Ecology*, 108(3) :865–875. Cité page 7
- Fisher, J. I. and Mustard, J. F. (2007). Cross-scalar satellite phenology from ground, Landsat, and MODIS data. *Remote Sensing of Environment*, 109(3) :261–273. Cité page 81
- Foody, G. (2005). Local characterization of thematic classification accuracy through spatially constrained confusion matrices. *International Journal of Remote Sensing*, 26(6) :1217–1228. Cité page 62
- Foody, G. (2008). Harshness in image classification accuracy assessment. *International Journal of Remote Sensing*, 29(11) :3137–3158. Cité page 62
- Foody, G. (2009a). Sample size determination for image classification accuracy assessment and comparison. *International Journal of Remote Sensing*, 30(20) :5273–5291. Cité page 62
- Foody, G. and Hill, R. (1996). Classification of tropical forest classes from Landsat TM data. *International Journal of Remote Sensing*, 17(12) :2353–2367. Cité page 41
- Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80(1) :185–201. Cité page 62
- Foody, G. M. (2009b). Sample size determination for image classification accuracy assessment and comparison. *International Journal of Remote Sensing*, 30(20) :5273–5291. Cité page 55
- Foody, G. M. (2020a). Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification. *Remote Sensing of Environment*, 239 :111630. Cité page 34
- Foody, G. M. (2020b). Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification. *Remote Sensing of Environment*, 239 :111630. Cité page 62, 63, 69
- Fox, W. (1990). *Toward a Transpersonal Ecology : Developing New Foundations for Environmentalism*. State University of New York Press. Cité page 3
- Franklin, S., Hall, R., Moskal, L., Maudie, A., and Lavigne, M. (2000). Incorporating texture into classification of forest species composition from airborne multispectral images. *International Journal of Remote Sensing*, 21(1) :61–79. Cité page 64, 66

-
- Freeman (2008). An Overview of Plant Defenses against Pathogens and Herbivores. *The Plant Health Instructor*. Cité page 23
- Freschet, G., Roumet, C., Violle, C., and Garnier, E. (2018). *Interactions entre le sol et la végétation : structure des communautés de plantes et fonctionnement du sol.*, pages 83–99. ISTE. Cité page 31
- Gamfeldt, L., Snäll, T., Bagchi, R., Jonsson, M., Gustafsson, L., Kjellander, P., Ruiz-Jaen, M. C., Fröberg, M., Stendahl, J., Philipson, C. D., Mikusiński, G., Andersson, E., Westerlund, B., Andrén, H., Moberg, F., Moen, J., and Bengtsson, J. (2013). Higher levels of multiple ecosystem services are found in forests with more tree species. *Nature Communications*, 4(1) :1340. Cité page 6, 41
- Gamon, J. A., Huemmrich, K. F., Wong, C. Y. S., Ensminger, I., Garrity, S., Hollinger, D. Y., Noormets, A., and Peñuelas, J. (2016). A remotely sensed pigment index reveals photosynthetic phenology in evergreen conifers. *Proceedings of the National Academy of Sciences*, 113(46) :13087–13092. Cité page 28
- Garnier, M. (2019). La BD FORÊT® V2 : Une cartographie forestière nationale pour les territoires. L’IF 46, IGN. Cité page 11
- Garrity, S. R., Bohrer, G., Maurer, K. D., Mueller, K. L., Vogel, C. S., and Curtis, P. S. (2011). A comparison of multiple phenology data sources for estimating seasonal transitions in deciduous forest carbon exchange. *Agricultural and Forest Meteorology*, 151(12) :1741–1752. Cité page 80
- Gautreau, P. and Viard-Crétat, A. (2020). Forêt - Enjeux environnementaux. In *Dictionnaire critique de l’anthropocène*, pages 409–412. CNRS. Cité page 3
- Geiß, C., Aravena Pelizari, P., Schrade, H., Brenning, A., and Taubenböck, H. (2017). On the effect of spatially non-disjoint training and test samples on estimated model generalization capabilities in supervised classification with spatial features. *IEEE Geoscience and Remote Sensing Letters*, 14(11) :2008–2012. Cité page 65, 66
- Ghamisi, P., Maggiori, E., Li, S., Souza, R., Tarablaka, Y., Moser, G., De Giorgi, A., Fang, L., Chen, Y., Chi, M., Serpico, S. B., and Benediktsson, J. A. (2018). New frontiers in spectral-spatial hyperspectral image classification : The latest advances based on mathematical morphology, markov random fields, segmentation, sparse representation, and deep learning. *IEEE Geoscience and Remote Sensing Magazine*, 6(3) :10–43. Cité page 64, 65, 66
- Ghimire, B., Rogan, J., and Miller, J. (2010). Contextual land-cover classification : incorporating spatial dependence in land-cover classification models using random forests and the getis statistic. *Remote Sensing Letters*, 1(1) :45–54. Cité page 64, 66
- Ghiyammat, A. and Shafri, H. (2010). A review on hyperspectral remote sensing for homogeneous and heterogeneous forest biodiversity assessment. *International Journal of Remote Sensing*, 31(7) :1837–1856. Cité page 41
- Giam, X. (2017). Global biodiversity loss from tropical deforestation. *Proceedings of the National Academy of Sciences*, 114(23) :5775–5777. Cité page 2
- Gill, S. J., Biging, G. S., and Murphy, E. C. (2000). Modeling conifer tree crown radius and estimating canopy cover. *Forest Ecology and Management*, 126(3) :405–416. Cité page 30
- GNU (2015). Qu’est-ce que le copyleft ? <https://www.gnu.org/licenses/copyleft.fr.html>. Cité page 115

-
- Goldstein, A. (2020). Open Source Licenses in 2020 : Trends and Predictions. <https://resources.whitesourcesoftware.com/blog-whitesource/top-open-source-licenses-trends-and-predictions>. Cité page 115
- Gonsamo, A., D'odorico, P., and Pellikka, P. (2013). Measuring fractional forest canopy element cover and openness - definitions and methodologies revisited. *Oikos*, 122(9) :1283–1291. Cité page 86
- Goëau, H., Boujemaa, N., Molino, J.-F., Duché, G., Péronnet, A., Bonnet, P., Joly, A., Bakić, V., Barbe, J., Yahiaoui, I., Selmi, S., Carré, J., and Barthélémy, D. (2013). Pl@ntNet mobile app. In *Proceedings of the 21st ACM international conference on Multimedia - MM '13*, pages 423–424, Barcelona, Spain. ACM Press. Cité page 22
- Grabska, E., Hostert, P., Pflugmacher, D., and Ostapowicz, K. (2019). Forest Stand Species Mapping Using the Sentinel-2 Time Series. *Remote Sensing*, 11(10) :1197. Cité page 20, 32, 33, 82, 105, 106
- Granier, A., Cochard, H., and Lemoine, D. (2002). Fonctionnement hydrique et hydraulique des arbres forestiers. *La Houille Blanche*, pages 18–23. Cité page 27
- Graves, S., Asner, G., Martin, R., Anderson, C., Colgan, M., Kalantari, L., and Bohlman, S. (2016). Tree Species Abundance Predictions in a Tropical Agricultural Landscape with a Supervised Classification Model and Imbalanced Data. *Remote Sensing*, 8(2) :161. Cité page 47
- Green, A. A., Berman, M., Switzer, P., and Craig, M. D. (1988). A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing*, 26(1) :65–74. Cité page 57
- Griffith, D. A. and Chun, Y. (2016). Spatial autocorrelation and uncertainty associated with remotely-sensed data. *Remote Sensing*, 8(7). Cité page 55
- Grizonnet, M., Michel, J., Poughon, V., Inglada, J., Savinaud, M., and Cresson, R. (2017). Orfeo ToolBox : open source processing of remote sensing images. *Open Geospatial Data, Software and Standards*, 2(1) :15. Cité page 45
- Guanter, L., Kaufmann, H., Segl, K., Foerster, S., Rogass, C., Chabrillat, S., Kuester, T., Hollstein, A., Rossner, G., Chlebek, C., Straif, C., Fischer, S., Schrader, S., Storch, T., Heiden, U., Mueller, A., Bachmann, M., Mühle, H., Müller, R., Habermeyer, M., Ohndorf, A., Hill, J., Buddenbaum, H., Hostert, P., van der Linden, S., Leitão, P., Rabe, A., Doerffer, R., Krasemann, H., Xi, H., Mauser, W., Hank, T., Locherer, M., Rast, M., Staenz, K., and Sang, B. (2015). The EnMAP Spaceborne Imaging Spectroscopy Mission for Earth Observation. *Remote Sensing*, 7(7) :8830–8857. Cité page 14
- Guigues, L., Trias-Sanz, R., Chehata, N., Taillandier, F., and Deveau, M. (2006). B. 4 segmentation multi-échelles d'images : théorie et applications. *Bulletin d'Information Scientifique & Technique de l'IGN n*, 75(1) :41. Cité page 11
- Guillot, G. (2014). *Guide des fleurs des forêts*. Belin, Paris. Cité page 5
- Guo, Q., Fei, S., Potter, K. M., Liebhold, A. M., and Wen, J. (2019). Tree diversity regulates forest pest invasion. *Proceedings of the National Academy of Sciences*, 116(15) :7382–7386. Cité page 7
- Guyon, D. and Breda, N. (2016). Imagerie optique multispectrale satellitaire pour des applications forestières. In *Observation des surfaces continentales par télédétection*, volume 1 of *Collection Système Terre Environnement - Série Télédétection pour l'observation des surfaces continentales*, pages 247–319. ISTE Editions. Cité page 14

-
- Hagolle, O. (2019). Using NDVI with atmospherically corrected data. <https://labo.obs-mip.fr/multitemp/using-ndvi-with-atmospherically-corrected-data/>.
Cité page 88
- Hagolle, O., Dedieu, G., Mougenot, B., Debaecker, V., Duchemin, B., and Meygret, A. (2008). Correction of aerosol effects on multi-temporal images acquired with constant viewing angles : Application to formosat-2 images. *Remote Sensing of Environment*, 112(4) :1689–1701.
Cité page 45
- Hagolle, O., Huc, M., Pascual, D., and Dedieu, G. (2010). A multi-temporal method for cloud detection, applied to formosat-2, venus, landsat and sentinel-2 images. *Remote Sensing of Environment*, 114(8) :1747–1755.
Cité page 45
- Hagolle, O., Huc, M., Pascual, D., and Dedieu, G. (2015). A multi-temporal and multi-spectral method to estimate aerosol optical thickness over land, for the atmospheric correction of formosat-2, landsat, vens and sentinel-2 images. *Remote Sensing*, 7(3) :2668–2691.
Cité page 45, 70, 87
- Hagolle, O., Morin, D., and Kadiri, M. (2018). Detailed Processing Model for the Weighted Average Synthesis Processor (WASP) for Sentinel-2. <https://doi.org/10.5281/zenodo.1401360>.
Cité page 142
- Hagolle Olivier (2018). Les synthèses mensuelles de niveau 3A de Theia pour Sentinel-2. <https://labo.obs-mip.fr/multitemp/les-syntheses-mensuelles-de-niveau-3a-de-theia/>.
Cité page 110, 142
- Hammond, T. O. and Verbyla, D. L. (1996). Optimistic bias in classification accuracy assessment. *International Journal of Remote Sensing*, 17(6) :1261–1266. Cité page 42, 47, 55, 65
- Hancock, S., Armston, J., Hofton, M., Sun, X., Tang, H., Duncanson, L. I., Kellner, J. R., and Dubayah, R. (2019). The GEDI simulator : A large-footprint waveform lidar simulator for calibration and validation of spaceborne missions. *Earth and Space Science*.
Cité page 12
- Harris, J. (2009). Soil microbial communities and restoration ecology : facilitators or followers? *Science*, 325(5940) :573–574.
Cité page 41
- Harris, N. L., Brown, S., Hagen, S. C., Saatchi, S. S., Petrova, S., Salas, W., Hansen, M. C., Potapov, P. V., and Lotsch, A. (2012). Baseline Map of Carbon Emissions from Deforestation in Tropical Regions. *Science*, 336(6088) :1573–1576. Cité page 2
- Hartigan, J. A. and Wong, M. A. (1979). Algorithm AS 136 : A K-Means Clustering Algorithm. *Applied Statistics*, 28(1) :100.
Cité page 92
- Hermance, J. F., Jacob, R. W., Bradley, B. A., and Mustard, J. F. (2007). Extracting phenological signals from multiyear AVHRR NDVI time series : Framework for applying high-order annual splines with roughness damping. *IEEE Transactions on Geoscience and Remote Sensing*, 45(10) :3264–3276.
Cité page 57
- Herrault, P.-A., Sheeren, D., Fauvel, M., and Paegelow, M. (2015). Vectorisation automatique des forêts dans les minutes de la carte d'état-major du 19^e siècle. *Revue Internationale de Géomatique*, 25(1) :35–51.
Cité page 10
- Hill, R. A., Wilson, A., George, M., and Hinsley, S. (2010). Mapping tree species in temperate deciduous woodland using time-series multi-spectral data. *Applied Vegetation Science*, 13(1) :86–99.
Cité page 41, 42, 58, 82, 105, 107

-
- Hościło, A. and Lewandowska, A. (2019). Mapping Forest Type and Tree Species on a Regional Scale Using Multi-Temporal Sentinel-2 Data. *Remote Sensing*, 11(8) :929. Cité page 32, 33, 58
- Hubbell, S. P., He, F., Condit, R., Borda-de Agua, L., Kellner, J., and ter Steege, H. (2008). How many tree species are there in the Amazon and how many of them will go extinct? *Proceedings of the National Academy of Sciences*, 105(Supplement 1) :11498–11504. Cité page 7
- Hufkens, K., Friedl, M., Sonnentag, O., Braswell, B. H., Milliman, T., and Richardson, A. D. (2012). Linking near-surface and satellite remote sensing measurements of deciduous broadleaf forest phenology. *Remote Sensing of Environment*, 117 :307–321. Cité page 82
- Hughes, G. (1968). On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory*, 14(1) :55–63. Cité page 58, 88, 105
- Ienco, D., Interdonato, R., Gaetano, R., and Ho Tong Minh, D. (2019). Combining Sentinel-1 and Sentinel-2 Satellite Image Time Series for land cover mapping via a multi-source deep learning architecture. *ISPRS Journal of Photogrammetry and Remote Sensing*, 158 :11–22. Cité page 142
- IGN (2019). Mémento de l’inventaire forestier. https://inventaire-forestier.ign.fr/IMG/pdf/memento_2019_web-2.pdf. Cité page 5
- Immitzer, M., Atzberger, C., and Koukal, T. (2012). Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. *Remote Sensing*, 4(9) :2661–2693. Cité page 20, 32, 41, 58
- Immitzer, M., Neuwirth, M., Böck, S., Brenner, H., Vuolo, F., and Atzberger, C. (2019). Optimal Input Features for Tree Species Classification in Central Europe Based on Multi-Temporal Sentinel-2 Data. *Remote Sensing*, 11(22) :2599. Cité page 33, 82
- Immitzer, M., Neuwirth, M., Böck, S., Vuolo, F., Brenner, H., and Atzberger, C. (2018). The added value of multi-temporal Sentinel-2 data for tree species classification in the Wienerwald Biosphere Reserve . Cité page 33
- Immitzer, M., Vuolo, F., and Atzberger, C. (2016). First experience with sentinel-2 data for crop and tree species classifications in central europe. *Remote Sensing*, 8(3) :166. Cité page 41, 42
- Inglada, J. (2018). Machine learning for land cover map production - Follow-up on the TiSeLaC challenge. <http://www.cesbio.ups-tlse.fr/multitemp/?p=13332>. Cité page 63
- 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. *Remote Sensing*, 9(1) :95. Cité page 12, 19, 33, 65, 66, 82
- IPBES (2019). Summary for policymakers of the global assessment report on biodiversity and ecosystem services. Cité page 6
- Isaacson, B. N., Serbin, S. P., and Townsend, P. A. (2012). Detection of relative differences in phenology of forest species using Landsat and MODIS. *Landscape Ecology*, 27(4) :529–543. Cité page 81

-
- Jacquemoud, S. and Ustin, S. (2019). *Leaf Optical Properties*. Cambridge University Press, 1 edition. Cité page 23, 26
- Jactel, H., Bauhus, J., Boberg, J., Bonal, D., Castagneyrol, B., Gardiner, B., Gonzalez-Olabarria, J. R., Koricheva, J., Meurisse, N., and Brockerhoff, E. G. (2017). Tree Diversity Drives Forest Stand Resistance to Natural Disturbances. *Current Forestry Reports*, 3(3) :223–243. Cité page 7
- Jain, A. K., Murty, M. N., and Flynn, P. J. (1999). Data clustering : a review. *ACM Computing Surveys*, 31(3) :264–323. Cité page 91
- Jennings, S. (1999). Assessing forest canopies and understorey illumination : canopy closure, canopy cover and other measures. *Forestry*, 72(1) :59–74. Cité page 30
- Jensen, J. R. (2000). *Remote sensing of the environment : an earth resource perspective*. Prentice Hall series in geographic information science. Prentice Hall, Upper Saddle River, N.J. Cité page 24
- Johnson, B., Tateishi, R., and Xie, Z. (2012). Using geographically weighted variables for image classification. *Remote Sensing Letters*, 3(6) :491–499. Cité page 65, 66
- Jönsson, P. and Eklundh, L. (2004). Timesat—a program for analyzing time-series of satellite sensor data. *Computers & geosciences*, 30(8) :833–845. Cité page 81
- Julien, Y. and Sobrino, J. A. (2009). Global land surface phenology trends from GIMMS database. *International Journal of Remote Sensing*, 30(13) :3495–3513. Cité page 88
- Jönsson, A., Eklundh, L., Hellström, M., Barring, L., and Jönsson, P. (2010). Annual changes in MODIS vegetation indices of Swedish coniferous forests in relation to snow dynamics and tree phenology. *Remote Sensing of Environment*, 114(11) :2719–2730. Cité page 80, 82
- Kalvane, G., Romanovskaja, D., Briede, A., and Baksiene, E. (2009). Influence of climate change on phenological phases in Latvia and Lithuania. *Climate Research*, 39 :209–219. Cité page 94, 95
- Kandasamy, S., Baret, F., Verger, A., Neveux, P., and Weiss, M. (2013). A comparison of methods for smoothing and gap filling time series of remote sensing observations – application to modis lai products. *Biogeosciences*, 10(6) :4055–4071. Cité page 45
- Karasiak, N. (2020). Museo ToolBox : A Python library for remote sensing including a new way to handle rasters. *Journal of Open Source Software*, 5(48) :1978. Cité page 69, 112
- Karasiak, N., Dejoux, J.-F., Fauvel, M., Willm, J., Monteil, C., and Sheeren, D. (2019). Statistical Stability and Spatial Instability in Mapping Forest Tree Species by Comparing 9 Years of Satellite Image Time Series. *Remote Sensing*, 11(21) :2512. Cité page 65
- Karasiak, N., Lomellini, A., Herrault, P.-A., Fauvel, M., Ladet, S., and Sheeren, D. (2016). Historical Map : un plug-in expérimental QGIS pour la vectorisation de l’occupation des sols dans les minutes de la carte d’État-Major datant du 19ème siècle. Cité page 10
- Kavzoglu, T. and Colkesen, I. (2009). A kernel functions analysis for support vector machines for land cover classification. *International Journal of Applied Earth Observation and Geoinformation*, 11(5) :352–359. Cité page 19, 47, 90

-
- Key, T., Warner, T., McGraw, J., and Fajvan, M. (2001). A comparison of multispectral and multitemporal information in high spatial resolution imagery for classification of individual tree species in a temperate hardwood forest. *Remote Sensing of Environment*, 75(1) :100–112. Cité page 41
- Khaliq, A., Peroni, L., and Chiaberge, M. (2018). Land cover and crop classification using multitemporal sentinel-2 images based on crops phenological cycle. In *2018 IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS)*, pages 1–5, Salerno. IEEE. Cité page 33
- Khanduri, V. P., Sharma, C. M., and Singh, S. P. (2008). The effects of climate change on plant phenology. *The Environmentalist*, 28(2) :143–147. Cité page 8
- Khatami, R., Mountrakis, G., and Stehman, S. (2017). Mapping per-pixel predicted accuracy of classified remote sensing images. *Remote Sensing of Environment*, 191 :156–167. Cité page 63
- Kokaly, R. F., Despain, D. G., Clark, R. N., and Livo, K. E. (2003). Mapping vegetation in Yellowstone National Park using spectral feature analysis of AVIRIS data. *Remote Sensing of Environment*, 84(3) :437–456. Cité page 26
- Kolb, T. E., Fettig, C. J., Ayres, M. P., Bentz, B. J., Hicke, J. A., Mathiasen, R., Stewart, J. E., and Weed, A. S. (2016). Observed and anticipated impacts of drought on forest insects and diseases in the United States. *Forest Ecology and Management*, 380 :321–334. Cité page 7
- Korhonen, L., Korhonen, K., Rautiainen, M., and Stenberg, P. (2006). Estimation of forest canopy cover : a comparison of field measurement techniques. *Silva Fennica*, 40(4). Cité page 30
- Koutsias, N. (2003). An autologistic regression model for increasing the accuracy of burned surface mapping using landsat thematic mapper data. *International Journal of Remote Sensing*, 24(10) :2199–2204. Cité page 64, 66
- Kurtto, A., Sennikov, A., and Lampinen, R. (2013). *Atlas Florae Europaeae : Distribution of Vascular Plants in Europe*. The Committee for Mapping the Flora of Europe & Societas Biologica Fennica Vanamo. Cité page 7
- Kymlicka, W. and Donaldson, S. (2014). Animals and the Frontiers of Citizenship. *Oxford Journal of Legal Studies*, 34(2) :201–219. Cité page 3
- Körner, C. and Basler, D. (2010). Phenology Under Global Warming. *Science*, 327(5972) :1461–1462. Cité page 8
- Kühn, I. (2007). Incorporating spatial autocorrelation may invert observed patterns. *Diversity and Distributions*, 13(1) :66–69. Cité page 64
- Kühn, I. and Dormann, C. (2012). Less than eight (and a half) misconceptions of spatial analysis. *Journal of Biogeography*, 39(5) :995–998. Cité page 64
- Labastida, I. and Margoni, T. (2020). Licensing FAIR Data for Reuse. *Data Intelligence*, 2(1-2) :199–207. Cité page 115
- Lallemant, T., Touzet, T., and Gervaise, A. (2017). Une méthodologie nationale pour le géoréférencement et la vectorisation des cartes d'état-major, minutes au 1/40 000. *Revue Forestière Française*, page 341. Cité page 10

-
- Larrieu, J.-F. (2019). Fertilisation raisonnée en arboriculture fruitière. Guide, Chambre d'Agriculture de Tarn-et-Garonne. Cité page 31
- Larrieu, L. and Gonin, P. (2008). L'indice de biodiversité potentielle (ibp) : une méthode simple et rapide pour évaluer la biodiversité potentielle des peuplements forestiers. *Revue Forestière Française*. Cité page 6
- Le Rest, K., Pinaud, D., Monestiez, P., Chadoeuf, J., and Bretagnolle, V. (2014). Spatial leave-one-out cross-validation for variable selection in the presence of spatial autocorrelation. *Global Ecology and Biogeography*, 23(7) :811–820. Cité page 47, 55, 65, 66, 77, 91
- Lebourgeois, F. (2018). La phénologie.
<https://www6.nancy.inrae.fr/silva/Equipes-de-recherche/EcoSILVA/Personnels/Pages-personnelles/LEBOURGEOIS-Francois/La-Phenologie/>. Cité page 26
- Lee, D. W. and Gould, K. S. (2002). Why leaves turn red : pigments called anthocyanins probably protect leaves from light damage by direct shielding and by scavenging free radicals. *American Scientist*, 90(6) :524–531. Cité page 88, 105
- Lemaître, G., Nogueira, F., and Aridas, C. K. (2017). Imbalanced-learn : A python toolbox to tackle the curse of imbalanced datasets in machine learning. *Journal of Machine Learning Research*, 18(17) :1–5. Cité page 92
- Lepeltier, T. (2020). L'écologie va-t-elle tuer l'antisécisme ?
<https://lamorce.co/lecologie-va-t-elle-tuer-lantispecisme/>. Cité page 3
- Liang, L., Schwartz, M. D., and Fei, S. (2011). Validating satellite phenology through intensive ground observation and landscape scaling in a mixed seasonal forest. *Remote Sensing of Environment*, 115(1) :143–157. Cité page 104
- Lieth, H. (1974). Purposes of a phenology book. In *Phenology and seasonality modeling*, pages 3–19. Springer. Cité page 80
- Lin, C., Popescu, S., Thomson, G., Tsogt, K., and Chang, C. (2015). Classification of tree species in overstorey canopy of subtropical forest using QuickBird images. *PLoS ONE*, 10(5). Cité page 41
- Lin, W.-C., Tsai, C.-F., Hu, Y.-H., and Jhang, J.-S. (2017). Clustering-based undersampling in class-imbalanced data. *Information Sciences*, 409-410 :17–26. Cité page 92
- Lisein, J., Michez, A., Claessens, H., and Lejeune, P. (2015). Discrimination of deciduous tree species from time series of Unmanned Aerial System imagery. *PLoS ONE*, 10(11) :1–20. Cité page 41, 58, 59
- Liu, C., Frazier, P., and Kumar, L. (2007). Comparative assessment of the measures of thematic classification accuracy. *Remote Sensing of Environment*, 107(4) :606–616. Cité page 62
- Liu, Y., Gong, W., Hu, X., and Gong, J. (2018). Forest type identification with random forest using sentinel-1a, sentinel-2a, multi-temporal landsat-8 and dem data. *Remote Sensing*, 10(6). Cité page 33, 41, 42
- Loeliger, J. and McCullough, M. (2012). Version control with Git. Cité page 130
- Lund, G. (2018). Definitions of Forest State, Stage, Origin, and Management.
<http://doi.org/10.13140/RG.2.2.14093.03042>. Cité page 3

-
- Lyons, M. B., Keith, D. A., Phinn, S. R., Mason, T. J., and Elith, J. (2018). A comparison of resampling methods for remote sensing classification and accuracy assessment. *Remote Sensing of Environment*, 208 :145–153. Cité page 55, 63, 65, 66
- Madonsela, S., Cho, M. A., Mathieu, R., Mutanga, O., Ramoelo, A., Kaszta, Z., Kerchove, R. V. D., and Wolff, E. (2017). Multi-phenology WorldView-2 imagery improves remote sensing of savannah tree species. *International Journal of Applied Earth Observation and Geoinformation*, 58 :65–73. Cité page 105
- Magnussen, S., Boudewyn, P., and Wulder, M. (2004). Contextual classification of landsat tm images to forest inventory cover types. *International Journal of Remote Sensing*, 25(12) :2421–2440. Cité page 64, 66
- Mallinis, G. and Koutsias, N. (2008). Spectral and spatial-based classification for broad-scale land cover mapping based on logistic regression. *Sensors*, 8(12) :8067–8085. Cité page 64, 66
- Mathews, F. (2012). The Anguish of Wildlife Ethics. *New Formations*, 76(76) :114–131. Cité page 3
- Matile, P. (2000). Biochemistry of Indian summer : physiology of autumnal leaf coloration. *Experimental Gerontology*, 35(2) :145–158. Cité page 27
- McConnell, B. R. and Smith, J. G. (1970). Response of Understory Vegetation to Ponderosa Pine Thinning in Eastern Washington. *Journal of Range Management*, 23(3) :208. Cité page 30
- McIver, D. K. and Friedl, M. A. (2001). Estimating pixel-scale land cover classification confidence using nonparametric machine learning methods. *IEEE Transactions on Geoscience and Remote Sensing*, 39(9) :1959–1968. Cité page 63
- Meier, U. (1985). Die merkblattserie 27 „entwicklungsstadien von pflanzen“. biologische bundesanstalt für land-und forstwirtschaft. *Nachrichtenbl. Deut. Pflanzenschutzd*, 37 :76–77. Cité page 84
- Meier, U. (1997). *Growth stages of mono-and dicotyledonous plants*. Blackwell Wissenschafts-Verlag. Cité page 94, 95
- Meier, U., Bleiholder, H., Buhr, L., Feller, C., Hack, H., Heß, M., Lancashire, P., Schnock, U., Stauß, R., Boom, T., Weber, E., and Zwerger, P. (2009). The bbch system to coding the phenological growth stages of plants-history and publications. *Journal für Kulturpflanzen*, 61 :41–52. Cité page 84
- Melgani, F. and Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 42(8) :1778–1790. Cité page 19, 58
- Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., and Nauss, T. (2018). Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. *Environmental Modelling & Software*, 101 :1–9. Cité page 42, 47
- Meyer, H., Reudenbach, C., Wöllauer, S., and Nauss, T. (2019). Importance of spatial predictor variable selection in machine learning applications – moving from data reproduction to spatial prediction. *Ecological Modelling*, 411 :108815. Cité page 63, 65, 66, 77

-
- Meyera, P., Staenzb, K., and Ittena, K. I. (1996). Semi-automated procedures for tree species identification in high spatial resolution data from digitized colour infrared-aerial photography. *ISPRS Journal of Photogrammetry and Remote Sensing*, 51(1) :5–16. Cité page 41
- Mickelson, J. G., Civco, D. L., and Silander, J. (1998). Delineating forest canopy species in the northeastern united states using multi-temporal tm imagery. *Photogrammetric Engineering and Remote Sensing*, 64 :891–904. Cité page 33
- Millard, K. and Richardson, M. (2015). On the importance of training data sample selection in random forest image classification : A case study in peatland ecosystem mapping. *Remote Sensing*, 7(7) :8489–8515. Cité page 55, 65
- Miller, J., Franklin, J., and Aspinall, R. (2007). Incorporating spatial dependence in predictive vegetation models. *Ecological Modelling*, 202(3-4) :225–242. Cité page 64
- Ministère de l’Enseignement Supérieur, de la Recherche et de l’Innovation (2018). Plan national pour la science ouverte. Cité page 112
- Molloy, J. C. (2011). The open knowledge foundation : open data means better science. *PLoS Biol*, 9(12) :e1001195. Cité page 116
- Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37(1) :17–23. Cité page 47, 91
- Morin, X., Lechowicz, M. J., Augspurger, C., O’Keefe, J., Viner, D., and Chuine, I. (2009). Leaf phenology in 22 North American tree species during the 21st century. *Global Change Biology*, 15(4) :961–975. Cité page 8
- Mountrakis, G., Im, J., and Ogole, C. (2011). Support vector machines in remote sensing : A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3) :247–259. Cité page 19, 47, 90
- Moutinho, P., Santilli, M., Schwartzman, S., and Rodrigues, L. (2005). Pourquoi ignorer la déforestation tropicale? une proposition visant à inclure la conservation des forêts dans le protocole de kyoto. *Unasylva (FAO)*. Cité page 2
- Mu, X., Hu, M., Song, W., Ruan, G., Ge, Y., Wang, J., Huang, S., and Yan, G. (2015). Evaluation of sampling methods for validation of remotely sensed fractional vegetation cover. *Remote Sensing*, 7(12) :16164–16182. Cité page 55
- Myint, S. W., Wentz, E. A., and Purkis, S. J. (2007). Employing spatial metrics in urban land-use/land-cover mapping. *Photogrammetric Engineering & Remote Sensing*, 73(12) :1403–1415. Cité page 64, 66
- Myneni, R. B., Keeling, C. D., Tucker, C. J., Asrar, G., and Nemani, R. R. (1997). Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature*, 386(6626) :698–702. Cité page 82, 89
- Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., and Wulder, M. A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, 148 :42–57. Cité page 33, 62
- Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251). Cité page 112

-
- Oshiro, T. M., Perez, P. S., and Baranauskas, J. A. (2012). How Many Trees in a Random Forest? In Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J. M., Mattern, F., Mitchell, J. C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen, B., Sudan, M., Terzopoulos, D., Tygar, D., Vardi, M. Y., Weikum, G., and Perner, P., editors, *Machine Learning and Data Mining in Pattern Recognition*, volume 7376, pages 154–168. Springer Berlin Heidelberg, Berlin, Heidelberg. Cité page 20
- Pal, M. and Foody, G. M. (2010). Feature selection for classification of hyperspectral data by svm. *IEEE Transactions on Geoscience and Remote Sensing*, 48(5) :2297–2307. Cité page 58
- Paletto, A. and Tosi, V. (2009). Forest canopy cover and canopy closure : comparison of assessment techniques. *European Journal of Forest Research*, 128(3) :265–272. Cité page 30
- Palmer, K. F. and Williams, D. (1974). Optical properties of water in the near infrared. *Journal of the Optical Society of America*, 64(8) :1107. Cité page 25
- Pasquarella, V. J., Holden, C. E., and Woodcock, C. E. (2018). Improved mapping of forest type using spectral-temporal Landsat features. *Remote Sensing of Environment*, 210 :193–207. Cité page 33, 41, 58, 81, 107
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn : Machine learning in python. *Journal of Machine Learning Research*, 12(Oct) :2825–2830. Cité page 47, 69, 90
- Pelletier, C., Valero, S., Inglada, J., Champion, N., and Dedieu, G. (2016). Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas. *Remote Sensing of Environment*, 187 :156–168. Cité page 20, 68
- Pelletier, C., Valero, S., Inglada, J., Champion, N., Marais Sicre, C., and Dedieu, G. (2017). Effect of Training Class Label Noise on Classification Performances for Land Cover Mapping with Satellite Image Time Series. *Remote Sensing*, 9(2) :173. Cité page 20
- Pelletier, C., Webb, G., and Petitjean, F. (2019). Temporal Convolutional Neural Network for the Classification of Satellite Image Time Series. *Remote Sensing*, 11(5) :523. Cité page 142
- Pelletier, M. (1990). *La Carte de Cassini : l'extraordinaire aventure de la carte de France*. Presses Ponts et Chaussées. Cité page 9
- Persson, M., Lindberg, E., and Reese, H. (2018). Tree species classification with multi-temporal sentinel-2 data. *Remote Sensing*, 10(11). Cité page 20, 33, 41, 42, 56, 58, 105, 106
- Plaza, J., Plaza, A., and Barra, C. (2009). Multi-Channel Morphological Profiles for Classification of Hyperspectral Images Using Support Vector Machines. *Sensors*, 9(1) :196–218. Cité page 21
- Pohjankukka, J., Pahikkala, T., Nevalainen, P., and Heikkonen, J. (2017). Estimating the prediction performance of spatial models via spatial k-fold cross validation. *International Journal of Geographical Information Science*, 0(0) :1–19. Cité page 47, 55, 65, 66, 91
- Pontius, R. G. and Millones, M. (2011). Death to Kappa : birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing*, 32(15) :4407–4429. Cité page 34, 62

-
- Probst, P. and Boulesteix, A.-L. (2017). To tune or not to tune the number of trees in random forest. *J. Mach. Learn. Res.*, 18(1) :6673–6690. Cité page 20
- Provost, D. (2019). Vers une gestion antispéciste et utilitariste de nos forêts. <https://lamorce.co/vers-une-gestion-antispeciste-et-utilitariste-de-nos-forets/>. Cité page 2, 3
- Puissant, A., Hirsch, J., and Weber, C. (2005). The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery. *International Journal of Remote Sensing*, 26(4) :733–745. Cité page 64, 66
- Putz, F. E. and Redford, K. H. (2010). The Importance of Defining ‘Forest’ : Tropical Forest Degradation, Deforestation, Long-term Phase Shifts, and Further Transitions : Importance of Defining ‘Forest’. *Biotropica*, 42(1) :10–20. Cité page 2
- Ramezan, A., Warner, A., and Maxwell, A. (2019). Evaluation of Sampling and Cross-Validation Tuning Strategies for Regional-Scale Machine Learning Classification. *Remote Sensing*, 11(2). Cité page 55, 62, 65
- Raschka, S. (2018). Mlxtend : Providing machine learning and data science utilities and extensions to python’s scientific computing stack. *Journal of Open Source Software*, 3(24) :638. Cité page 124
- Reed, B. C., Brown, J. F., VanderZee, D., Loveland, T. R., Merchant, J. W., and Ohlen, D. O. (1994). Measuring phenological variability from satellite imagery. *Journal of Vegetation Science*, 5(5) :703–714. Cité page 88, 130
- Reed, B. C., Schwartz, M. D., and Xiao, X. (2009). Remote Sensing Phenology. In Noormets, A., editor, *Phenology of Ecosystem Processes*, pages 231–246. Springer New York, New York, NY. Cité page 82
- Richardson, A. D., Friedl, M. A., Frohling, S., Pless, R., and PhenoCam Collaborators (2011). Phenocam : A continental-scale observatory for monitoring the phenology of terrestrial vegetation. In *AGU Fall Meeting Abstracts*, volume 2011, pages B11D–0517. Cité page 82, 109
- Rivas, H., Delbart, N., Otlé, C., Maignan, F., and Vaudour, E. (2019). Monitoring phenology of crops at the parcel scale : combining high and medium spatial resolution data. In *Geophysical Research Abstracts*, volume 21. Cité page 81
- Rivers, M. (2019). *European Red List of Trees*. IUCN, International Union for Conservation of Nature. Cité page 5
- Rivers, M., Beech, E., Bazos, I., Bogunić, F., Buirra, A., Caković, D., Carapeto, A., Carta, A., Cornier, B., Fenu, G., et al. (2019). European red list of trees. In *European Red list assessment review workshop : trees and shrubs*. IUCN Global Species Programme Red List Unit. Cité page 7
- Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillerá-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., and Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8) :913–929. Cité page 55, 63, 64, 65, 66, 77
- Rocha, A. D., Groen, T. A., and Skidmore, A. K. (2019). Spatially-explicit modelling with support of hyperspectral data can improve prediction of plant traits. *Remote Sensing of Environment*, 231 :111200. Cité page 65

-
- Rodriguez-Galiano, V., Ghimire, B., Rogan, J., Chica-Olmo, M., and Rigol-Sanchez, J. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67 :93–104. Cité page 68
- Rudel, T. K., Coomes, O. T., Moran, E., Achard, F., Angelsen, A., Xu, J., and Lambin, E. (2005). Forest transitions : towards a global understanding of land use change. *Global Environmental Change*, 15(1) :23–31. Cité page 3
- Sacchetti, A., Cisbani, A., Babini, G., and Galeazzi, C. (2010). The italian precursor of an operational hyperspectral imaging mission. In *Small Satellite Missions for Earth Observation*, pages 73–81. Springer. Cité page 14
- Safou, O. and Saint-Martin, M. (1989). Le trichome foliaire de quelques quercus périméditerranéens. *Bulletin de la Société Botanique de France. Lettres Botaniques*, 136(4-5) :291–304. Cité page 24
- San-Miguel-Ayanz, J., Rigo, D. d., Caudullo, G., Durrant, T. H., and Mauri, A. (2016). *European atlas of forest tree species*. Publication Office of the European Union, Luxembourg, 2016 edition. Cité page 7
- Sandmann, G., Römer, S., and Fraser, P. D. (2006). Understanding carotenoid metabolism as a necessity for genetic engineering of crop plants. *Metabolic Engineering*, 8(4) :291–302. Cité page 23
- Schafer, R. (2011). What Is a Savitzky-Golay Filter ? [Lecture Notes]. *IEEE Signal Processing Magazine*, 28(4) :111–117. Cité page 130
- Schmid, M., Bosser, J., Ségalen, P., Dabin, B., Maignien, R., Combeau, A., Bachelier, G., Schmid, M., Bosser, J., Guinard, M., and Verdier, P. (1970). La végétation et les conditions édaphiques. In *Pédologie et développement*, number 10 in Techniques Rurales en Afrique, pages 149–164. ORSTOM, Paris. Cité page 31
- Schratz, P., Muenchow, J., Iturritxa, E., Richter, J., and Brenning, A. (2019). Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. *Ecological Modelling*, 406 :109–120. Cité page 42, 47, 55, 63, 65, 66, 77, 137
- Schriever, J. R. and Congalton, R. G. (1995). Evaluating seasonal variability as an aid to cover-type mapping from landsat thematic mapper data in the northeast. *Photogrammetric Engineering and Remote Sensing*, 61(3) :321–327. Cité page 33, 107
- Secretariat, G. (2009). Implementation plan for the global observing system for climate in support of the unfccc (2010 update). In *Proceedings of the Conference of the Parties (COP), Copenhagen, Denmark*, pages 7–19. Citeseer. Cité page 8
- Seidl, R., Thom, D., Kautz, M., Martin-Benito, D., Peltoniemi, M., Vacchiano, G., Wild, J., Ascoli, D., Petr, M., Honkaniemi, J., Lexer, M., Trotsiuk, V., Mairota, Paola Svoboda, M., Fabrika, M., and Nagel, Thomas A. Reyer, C. P. O. (2017). Forest disturbances under climate change. *Nature Climate Change*, 7 :395–402. Cité page 41
- Serbin, S. P. and Townsend, P. A. (2020). Scaling Functional Traits from Leaves to Canopies. In Cavender-Bares, J., Gamon, J. A., and Townsend, P. A., editors, *Remote Sensing of Plant Biodiversity*, pages 43–82. Springer International Publishing, Cham. Cité page 28
- Sergey, R. (2017). Temporal and spatial approaches for land cover classification. In *2017 ECML/PKDD Discovery Challenges*. Cité page 63

-
- Shahidan, M. F. and Salleh, E. (2007). Effects of tree canopies on solar radiation filtration in a tropical microclimatic environment. Cité page 23
- Shao, Y., Lunetta, R. S., Wheeler, B., Iames, J. S., and Campbell, J. B. (2016). An evaluation of time-series smoothing algorithms for land-cover classifications using MODIS-NDVI multi-temporal data. *Remote Sensing of Environment*, 174 :258–265. Cité page 88
- Sheeren, D., Bastin, N., Ouin, A., Ladet, S., Balent, G., and Lacombe, J.-P. (2009). Discriminating small wooded elements in rural landscape from aerial photography : a hybrid pixel/object-based analysis approach. *International Journal of Remote Sensing*, 30(19) :4979–4990. Cité page 64
- Sheeren, D., Fauvel, M., Josipović, V., Lopes, M., Planque, C., Willm, J., and Dejoux, J.-F. (2016). Tree Species Classification in Temperate Forests Using Formosat-2 Satellite Image Time Series. *Remote Sensing*, 8(9) :734. Cité page 19, 20, 33, 35, 42, 55, 56, 58, 90
- Shekhar, S., Schrater, P. R., Vatsavai, R. R., Weili Wu, and Chawla, S. (2002). Spatial contextual classification and prediction models for mining geospatial data. *IEEE Transactions on Multimedia*, 4(2) :174–188. Cité page 64, 66
- Shunlin Liang and Jindi Wang (2012). A Systematic View of Remote Sensing. In *Advanced Remote Sensing*, pages 1–31. Elsevier. Cité page 13
- Solberg, A. H. S., Taxt, T., and Jain, A. K. (1996). A markov random field model for classification of multisource satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 34(1) :100–113. Cité page 64, 66
- Soltner, D. (2017). *Les bases de la production végétale : phytotechnie générale : le sol, le climat, la plante*. Science et techniques agricoles. Cité page 23
- Sousa-Silva, R., Verheyen, K., Ponette, Q., Bay, E., Sioen, G., Titeux, H., Van de Peer, T., Van Meerbeek, K., and Muys, B. (2018). Tree diversity mitigates defoliation after a drought-induced tipping point. *Global Change Biology*, 24(9) :4304–4315. Cité page 7
- Spracklen, B. D. and Spracklen, D. V. (2019). Identifying european old-growth forests using remote sensing : A study in the ukrainian carpathians. *Forests*, 10(2). Cité page 33, 41, 42
- Stehman, S. (2009). Sampling designs for accuracy assessment of land cover. *International Journal of Remote Sensing*, 30(20) :5243–5272. Cité page 55, 62
- Stehman, S. and Wickham, J. (2011). Pixels, blocks of pixels, and polygons : Choosing a spatial unit for thematic accuracy assessment. *Remote Sensing of Environment*, 115(12) :3044–3055. Cité page 62
- Stehman, S. V. and Foody, G. M. (2019). Key issues in rigorous accuracy assessment of land cover products. *Remote Sensing of Environment*, 231 :111199. Cité page 62, 63
- Strauss, M. (2017). Planet Earth to get a daily selfie. *Science*, 355(6327) :782–783. Cité page 15
- Sugden, A. M. (2018). Tree diversity improves forest productivity. *Science*, 362(6410) :41.14–43. Cité page 7
- Sugden, A. M. (2020). Tree diversity relieves drought impacts. *Science*, 367(6483) :1208.1–1209. Cité page 7

-
- Sumbul, G., Charfuelan, M., Demir, B., and Markl, V. (2019). Bigearthnet : A large-scale benchmark archive for remote sensing image understanding. In *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, pages 5901–5904. IEEE. Cité page 137
- Sun, Y., Wong, A. K., and Kamel, M. S. (2009). Classification of imbalanced data : a review. *International Journal of Pattern Recognition and Artificial Intelligence*, 23 :687–719. Cité page 56
- Taiz, L., Alkon, D., Draguhn, A., Murphy, A., Blatt, M., Hawes, C., Thiel, G., and Robinson, D. G. (2019). Plants Neither Possess nor Require Consciousness. *Trends in Plant Science*, 24(8) :677–687. Cité page 3
- Thompson, I., Mackey, B., McNulty, S., Mosseler, A., et al. (2009). Forest resilience, biodiversity, and climate change. In *A synthesis of the biodiversity/resilience/stability relationship in forest ecosystems. Secretariat of the Convention on Biological Diversity, Montreal. Technical Series*, volume 43, page 67. Cité page 41
- Thompson, I. D., Okabe, K., Tylianakis, J. M., Kumar, P., Brockerhoff, E. G., Schellhorn, N. A., Parrotta, J. A., and Nasi, R. (2011). Forest Biodiversity and the Delivery of Ecosystem Goods and Services : Translating Science into Policy. *BioScience*, 61(12) :972–981. Cité page 6, 41
- Tichý, L. (2016). Field test of canopy cover estimation by hemispherical photographs taken with a smartphone. *Journal of Vegetation Science*, 27(2) :427–435. Cité page 86, 87
- Tigges, J., Lakes, T., and Hostert, P. (2013). Urban vegetation classification : Benefits of multitemporal rapideye satellite data. *Remote Sensing of Environment*, 136 :66–75. Cité page 41, 42
- Touzet, T. and Lecordix, F. (2010). La carte forestière sans papier. *Journal du comité français de cartographie*, 206 :53–62. Cité page 11
- Trenberth, K. E., Dai, A., van der Schrier, G., Jones, P. D., Barichivich, J., Briffa, K. R., and Sheffield, J. (2014). Global warming and changes in drought. *Nature Climate Change*, 4(1) :17–22. Cité page 7
- Trichon, V. and Julien, M.-P. (2006). Tree species identification on large-scale aerial photographs in a tropical rain forest, french guiana—application for management and conservation. *Forest Ecology and Management*, 225(1) :51–61. Cité page 41
- Ulrich, E. (1995). Le réseau renecofor : objectifs et réalisation. *Revue forestière française*. Cité page 85
- Ustin, S., Gitelson, A., Jacquemoud, S., Schaepman, M., Asner, G., Gamon, J., and Zarco-Tejada, P. (2009). Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. *Remote Sensing of Environment*, 113. Cité page 41
- Valavi, R., Elith, J., Lahoz-Monfort, J. J., and Guillera-Arroita, G. (2019). blockcv : An r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. *Methods in Ecology and Evolution*, 10(2) :225–232. Cité page 65, 66
- Vallauri, D., editor (2003). *Livre blanc sur la protection des forêts naturelles en France : forêts métropolitaines*. Tec & doc, Paris. Cité page 6

-
- Vallauri, D., Grel, A., Granier, E., and Dupouey, J.-L. (2012). Les forêts de Cassini. Analyse quantitative et comparaison avec les forêts actuelles. Research Report, WWF. Cité page 10
- Van Tricht, K., Gobin, A., Gilliams, S., and Piccard, I. (2018). Synergistic Use of Radar Sentinel-1 and Optical Sentinel-2 Imagery for Crop Mapping : A Case Study for Belgium. *Remote Sensing*, 10(10) :1642. Cité page 142
- Vapnik, V. N. (1998). Adaptive and learning systems for signal processing communications, and control. *Statistical learning theory*. Cité page 19, 47, 90
- Vidalou, J.-B. (2017). *Être forêts : habiter des territoires en lutte*. La Découverte. Cité page 3
- Vitasse, Y., Porté, A. J., Kremer, A., Michalet, R., and Delzon, S. (2009). Responses of canopy duration to temperature changes in four temperate tree species : relative contributions of spring and autumn leaf phenology. *Oecologia*, 161(1) :187–198. Cité page 8
- Vrieling, A., Meroni, M., Darvishzadeh, R., Skidmore, A. K., Wang, T., Zurita-Milla, R., Oosterbeek, K., O'Connor, B., and Paganini, M. (2018). Vegetation phenology from Sentinel-2 and field cameras for a Dutch barrier island. *Remote Sensing of Environment*, 215 :517–529. Cité page 33
- Walsh, S. J. (1980). Coniferous tree species mapping using landsat data. *Remote Sensing of Environment*, 9(1) :11–26. Cité page 41
- Wang, J., Xu, R., and Yang, S. (2009). Estimation of plant water content by spectral absorption features centered at 1,450 nm and 1,940 nm regions. *Environmental Monitoring and Assessment*, 157(1-4) :459–469. Cité page 25
- Wang, L., Shi, C., Diao, C., Ji, W., and Yin, D. (2016). A survey of methods incorporating spatial information in image classification and spectral unmixing. *International Journal of Remote Sensing*, 37(16) :3870–3910. Cité page 64
- Wareing, P. F. (1956). Photoperiodism in Woody Plants. *Annual Review of Plant Physiology*, 7(1) :191–214. Cité page 80
- Waser, L., Ginzler, C., Kuechler, M., Baltsavias, E., and Hurni, L. (2011). Semi-automatic classification of tree species in different forest ecosystems by spectral and geometric variables derived from airborne digital sensor (ADS40) and RC30 data. *Remote Sensing of Environment*, 115 :76–85. Cité page 41
- Way, D. A. (2011). Tree phenology responses to warming : spring forward, fall back? *Tree Physiology*, 31(5) :469–471. Cité page 80
- Wessel, M., Brandmeier, M., and Tiede, D. (2018). Evaluation of Different Machine Learning Algorithms for Scalable Classification of Tree Types and Tree Species Based on Sentinel-2 Data. *Remote Sensing*, 10(9) :1419. Cité page 33
- White, K., Pontius, J., and Schaberg, P. (2014). Remote sensing of spring phenology in northeastern forests : A comparison of methods, field metrics and sources of uncertainty. *Remote Sensing of Environment*, 148 :97–107. Cité page 104
- Whitney, A. (1971). A direct method of nonparametric measurement selection. *IEEE Transactions on Computers*, C-20(9) :1100–1103. Cité page 83, 92

-
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J., Groth, P., Goble, C., Grethe, J. S., Heringa, J., 't Hoen, P. A., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., and Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1) :160018. <http://www.nature.com/articles/sdata201618>. Cité page 112, 113
- Wolter, P., Mladenoff, D., Host, G., and Crow, T. (1995). Improved forest classification in the northern lake states using multi-temporal Landsat imagery. *Photogrammetric Engineering & Remote Sensing*, 61(9) :1129–1143. Cité page 33, 41
- Wong, C. Y., D'Odorico, P., Bhathena, Y., Arain, M. A., and Ensminger, I. (2019). Carotenoid based vegetation indices for accurate monitoring of the phenology of photosynthesis at the leaf-scale in deciduous and evergreen trees. *Remote Sensing of Environment*, 233 :111407. Cité page 105
- Wulder, M. and Boots, B. (1998). Local spatial autocorrelation characteristics of remotely sensed imagery assessed with the getis statistic. *International Journal of Remote Sensing*, 19(11) :2223–2231. Cité page 63
- Wulf, H. and Stuhler, S. (2015). Sentinel-2 : Land Cover, Preliminary User Feedback on Sentinel-2A Data In Proceedings of the Sentinel-2A Expert Users Technical Meeting, Frascati, Italy, 29–30 September 2015. Cité page 88
- Xiao, X., Hagen, S., Zhang, Q., Keller, M., and Moore, B. (2006). Detecting leaf phenology of seasonally moist tropical forests in South America with multi-temporal MODIS images. *Remote Sensing of Environment*, 103(4) :465–473. Cité page 80, 82
- Yan, E., Wang, G., Lin, H., Xia, C., and Sun, H. (2015). Phenology-based classification of vegetation cover types in Northeast China using MODIS NDVI and EVI time series. *International Journal of Remote Sensing*, 36(2) :489–512. Cité page 81
- Ye, S., Pontius, R., and Rakshit, R. (2018). A review of accuracy assessment for object-based image analysis : From per-pixel to per-polygon approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*, 141 :137–147. Cité page 82
- Zeng, L., Wardlow, B. D., Xiang, D., Hu, S., and Li, D. (2020). A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. *Remote Sensing of Environment*, 237 :111511. Cité page 81, 82, 89, 103
- Zhang, C. and Xie, Z. (2012). Combining object-based texture measures with a neural network for vegetation mapping in the Everglades from hyperspectral imagery. *Remote Sensing of Environment*, 124 :310–320. Cité page 142
- Zhang, X. (2015). Reconstruction of a complete global time series of daily vegetation index trajectory from long-term AVHRR data. *Remote Sensing of Environment*, 156 :457–472. Cité page 130
- Zhang, X., Friedl, M. A., Schaaf, C. B., Strahler, A. H., Hodges, J. C., Gao, F., Reed, B. C., and Huete, A. (2003). Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, 84(3) :471–475. Cité page 32, 80

-
- Zhang, X., Wang, J., Gao, F., Liu, Y., Schaaf, C., Friedl, M., Yu, Y., Jayavelu, S., Gray, J., Liu, L., Yan, D., and Henebry, G. M. (2017). Exploration of scaling effects on coarse resolution land surface phenology. *Remote Sensing of Environment*, 190 :318–330. Cité page 88
- Zhao, W. and Du, S. (2016). Spectral-spatial feature extraction for hyperspectral image classification : A dimension reduction and deep learning approach. *IEEE Transactions on Geoscience and Remote Sensing*, 54(8) :4544–4554. Cité page 65, 66
- Zhen, Z., Quackenbush, L. J., Stehman, S. V., and Zhang, L. (2013). Impact of training and validation sample selection on classification accuracy and accuracy assessment when using reference polygons in object-based classification. *International Journal of Remote Sensing*, 34(19) :6914–6930. Cité page 42
- Zhu, X. and Liu, D. (2014). Accurate mapping of forest types using dense seasonal landsat time-series. *ISPRS Journal of Photogrammetry and Remote Sensing*, 96 :1–11. Cité page 41



Acronymes

CEOS *Committee on Earth Observation Satellites*. 24

ECV *Essential Climate Variables* ou Variables Climatiques Essentielles en français. 11

EOSDIS Earth Observing System Data and Information System. 24

ESA European Space Agency. 26

EUFORGEN *European Forest Genetic Resources Programme*. 10

FAO *Food and Agriculture Organization of the United Nations* ou, en français, l'Organisation des Nations Unies pour l'Alimentation et l'Agriculture. 6

GBIF *Global Biodiversity Information Facility*. 10

GCOS *Global Climate Observing System*. 11

IBP Indice de Biodiversité Potentielle. 9

IFN Inventaire Forestier National. 6, 8

IGN Institut National de l'information Géographique et forestière. xi, 6, 7, 13

IPBES Plateforme Intergouvernementale scientifique et politique sur la Biodiversité et les Services Écosystémiques. 9

KFCV *K-Fold Cross-Validation*. 56

LAI *Leaf Area Index*, appelé aussi en français indice de surface foliaire. 20

MCCV *Monte-Carlo Cross-Validation*. 56

MNT Modèle Numérique de Terrain. 25

NASA National Aeronautics and Space Administration. 24

ONG Organisation Non Gouvernementale. 6, 8

PNACC Plan National d'Adaptation au Changement Climatique. 8

RF *Random Forest* ou en français, Forêts Aléatoires. 27

SIG Système d'Information Géographique. 12, 28

SITS *Satellite Image Time Series* ou en français, Série Temporelle d'Images Satellites. 28

SVM *Support Vector Machines* ou en français, Machines à Vecteur de Support. 26, I

TOA *Top Of Atmosphere*. 25

TOC *Top Of Canopy*. 26

UICN Union Internationale pour la Conservation de la Nature. xi, 8–10

WRM *World Rainforest Movement*. 6



Annexes

Annexe A

Compléments du chapitre 2

A.1 Tree species map

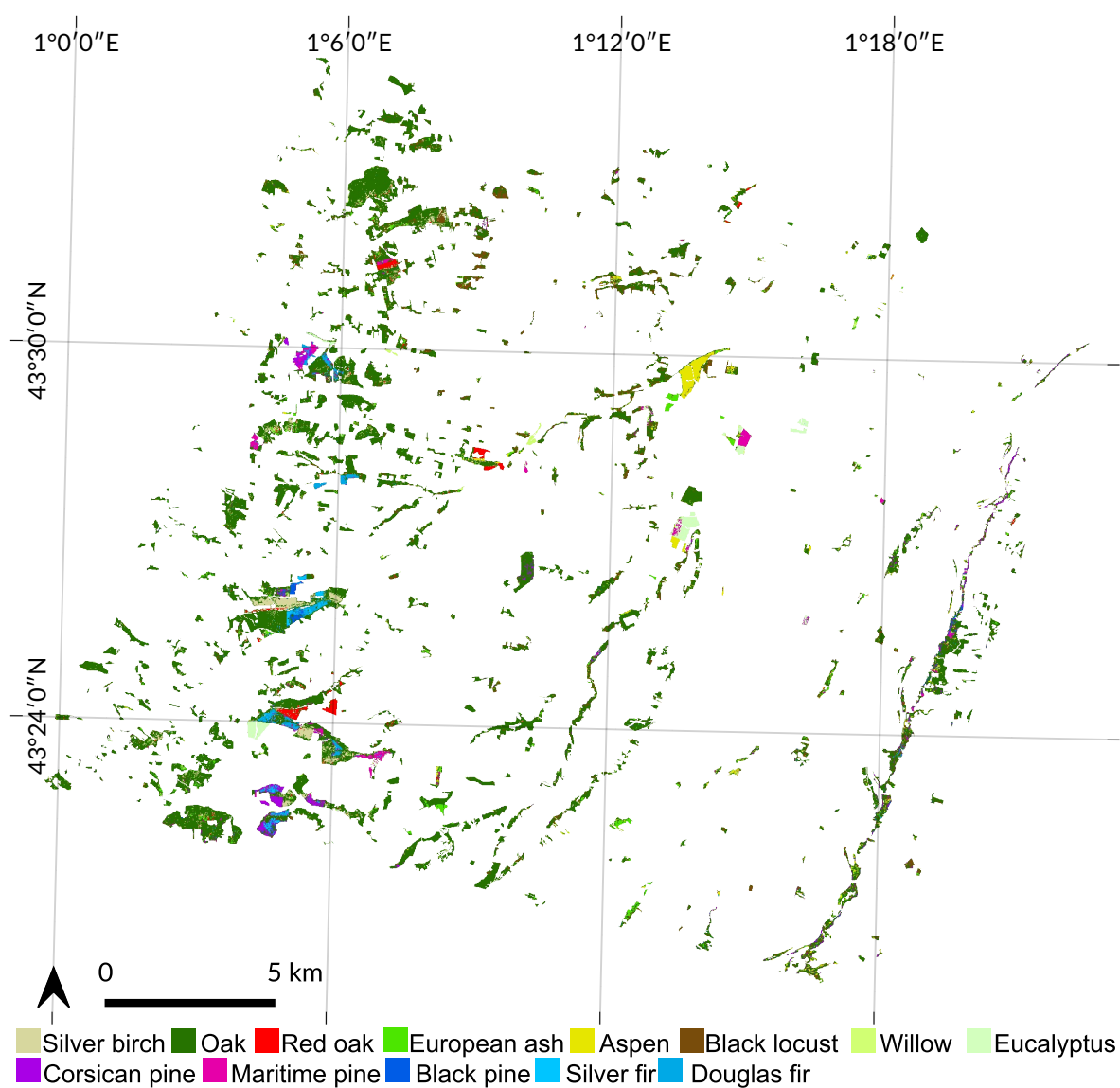


FIGURE A.1 – Map of the most predicted species using the 10 SITS with the Spatial Leave-One-Out cross-validation method (SLOO-CV).

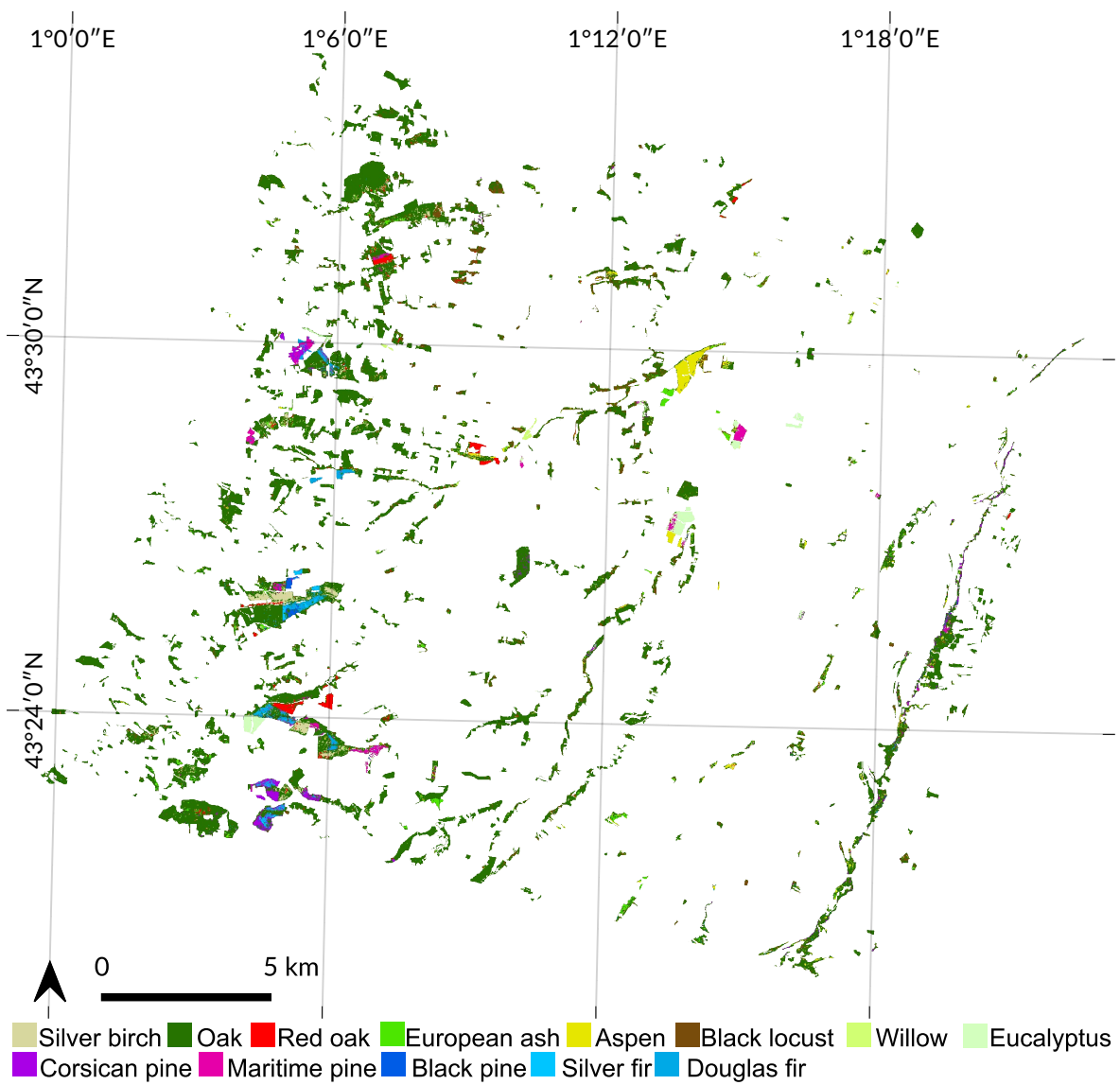


FIGURE A.2 – Map of the most predicted species using the 10 SITS with the Leave-One-Out cross-validation method (LOO-CV).

A.2. SIGNIFICANCE TABLES FOR PREDICTION BETWEEN YEARS

A.2 Significance tables for prediction between years

TABLEAU A.1 – Wilcoxon signed-rank test significance table for SLOO-CV computed from the overall accuracy of the 50 predictions for each single-year SITS. In **bold**, where $p < 0.05$.

	2006	2006bis	2007	2008	2009	2010	2011	2012	2013	2014
2006	<i>nan</i>	244	242	255	302	354	313	194	193	309
2006bis	244	<i>nan</i>	198	551	335	407	307	556	384	459
2007	242	198	<i>nan</i>	143	112	90	208	120	57	144
2008	255	551	143	<i>nan</i>	349	345	365	433	275	257
2009	302	335	112	349	<i>nan</i>	288	299	348	173	229
2010	354	407	90	345	288	<i>nan</i>	382	284	215	384
2011	313	307	208	365	299	382	<i>nan</i>	315	166	250
2012	194	556	120	433	348	284	315	<i>nan</i>	356	451
2013	193	384	57	275	173	215	166	356	<i>nan</i>	360
2014	309	459	144	257	229	384	250	451	360	<i>nan</i>

TABLEAU A.2 – Wilcoxon signed-rank test significance table for LOO-CV computed from the overall accuracy of the 50 predictions for each single-year SITS. In **bold**, where $p < 0.05$.

	2006	2006bis	2007	2008	2009	2010	2011	2012	2013	2014
2006	<i>nan</i>	0	0	0	0	0	0	0	0	6
2006bis	0	<i>nan</i>	6	20	18	30	29	30	10	7
2007	0	6	<i>nan</i>	4	10	0	6	5	0	6
2008	0	20	4	<i>nan</i>	22	27	25	42	4	5
2009	0	18	10	22	<i>nan</i>	30	11	32	16	11
2010	0	30	0	27	30	<i>nan</i>	16	30	0	12
2011	0	29	6	25	11	16	<i>nan</i>	16	6	7
2012	0	30	5	42	32	30	16	<i>nan</i>	4	13
2013	0	10	0	4	16	0	6	4	<i>nan</i>	7
2014	6	7	6	5	11	12	7	13	7	<i>nan</i>

A.3 Effect of clouds and cloud shadows

TABLEAU A.3 – Percent of reference pixels affected by clouds or shadows detected by the MACCS algorithm for each species and each year.

Species	2006	2006bis	2007	2008	2009	2010	2011	2012	2013	2014
Silver birch	26	0	17	1	9	0	0	0	4	0
Oak	24	0	20	11	6	5	1	0	2	0
Red Oak	21	0	11	0	1	6	3	0	0	0
Aspen	29	0	8	0	0	1	4	3	0	0
European Ash	22	0	11	6	5	1	0	0	3	1
Black locust	25	0	11	1	0	0	5	0	2	0
Willow	27	0	6	2	2	0	6	2	0	0
Eucalyptus	23	0	7	0	0	5	4	0	2	0
Corsican Pine	16	0	16	11	2	5	5	0	1	4
Maritime Pine	25	0	9	9	3	3	2	0	0	0
Black Pine	30	0	17	5	10	1	0	0	2	0
Silver Fir	28	0	17	5	8	5	0	0	3	0
Douglas	26	0	10	9	5	4	4	0	0	0

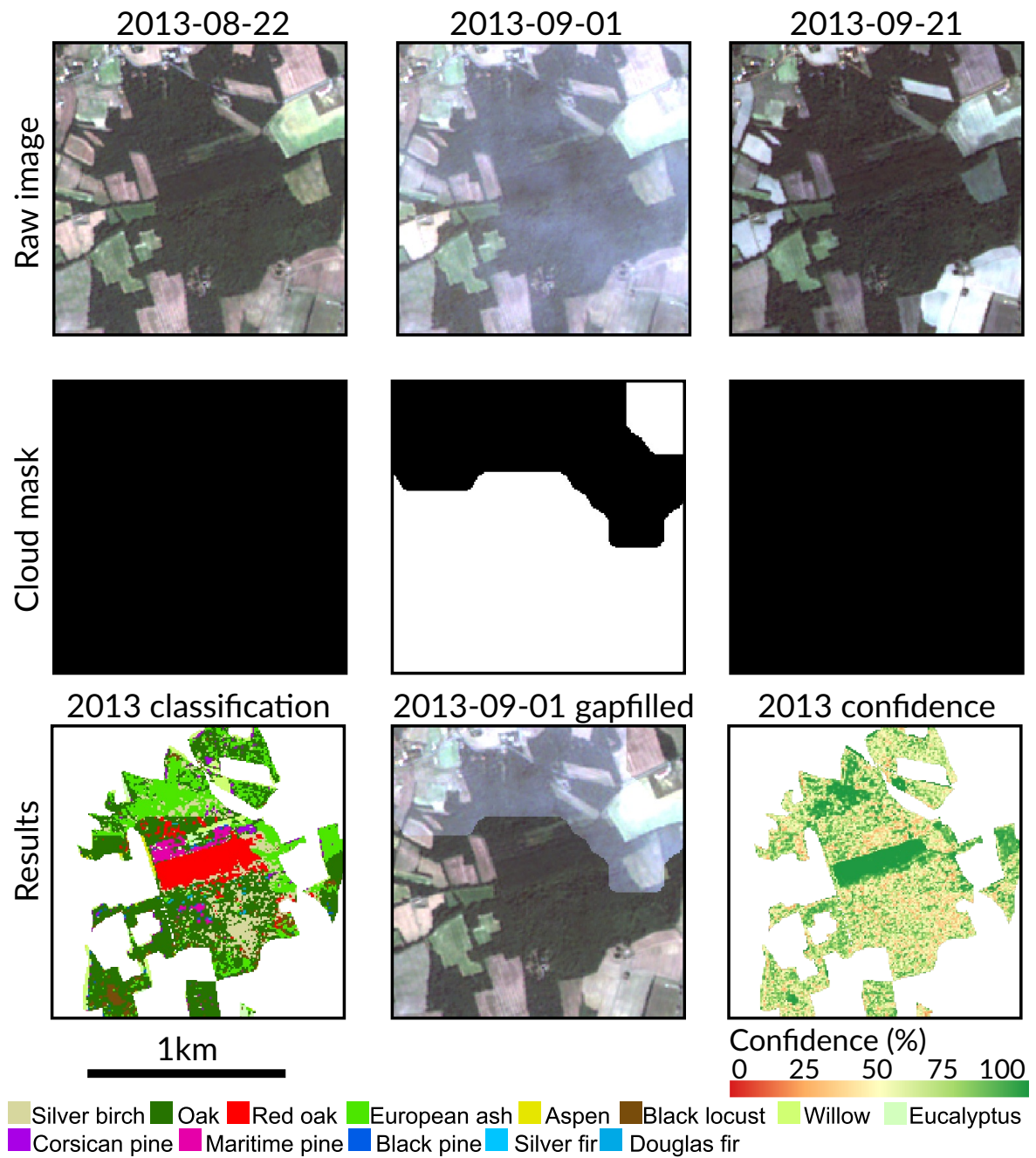


FIGURE A.3 – Example of misclassified forest stand (for European ash, in the north) due to under-detected clouds in 2013-09-01. In the cloud mask, pixels in black are cloud and shadow free. Pixels in white are cloudy or shady.

A.4 Training size per species

TABLEAU A.4 – Number of training pixels for each cross-validation method. For both methods, value per species is the mean from the 50-folds. The split between test and train set is exactly the same for each year.

Species	SLOO-CV	LOO-CV
Broadleaf		
Silver birch	35	35
Oak	97	97
Red oak	118	118
Aspen	142	142
European ash	50	50
Black locust	50	50
Willow	21	21
Eucalyptus	85	85
Conifer		
Corsican pine	33	33
Maritime pine	79	79
Black pine	26	26
Silver fir	54	54
Douglas fir	46	46
Total	836	836

Annexe B

Compléments du chapitre 3

B.1 Moran's I correlogram on Sentinel-2 data

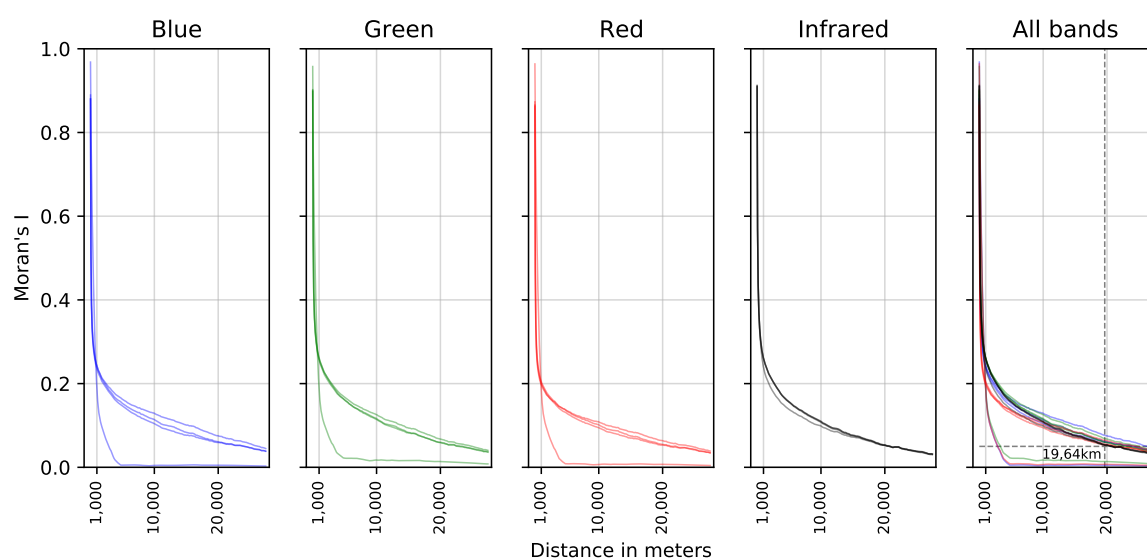


FIGURE B.1 – Moran's I correlograms of each Sentinel-2 spectral band, for pixels representing forests. Each curve represents one spectral band at one date in the dataset. For a Moran's I threshold value of 0.05, spatial independence between nearby pixels was assumed. This threshold was reached at 1,964 pixels (i.e. 19.64 km) on average. The grey dashed line represents the mean distance value (in x) where Moran's I = 0.05 (in y).

B.2 Relative frequency distribution of NDVI values per class

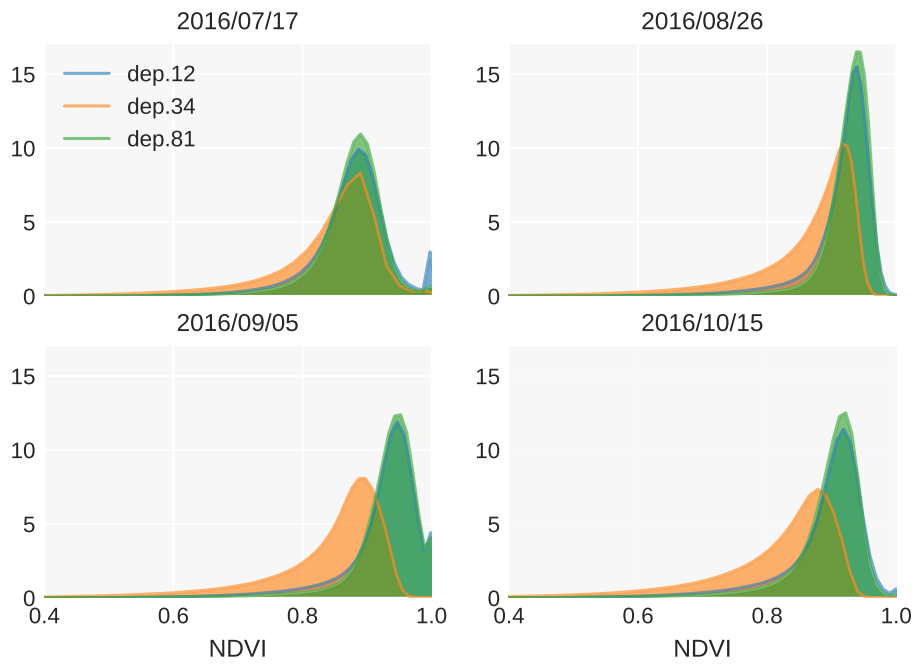


FIGURE B.2 – Relative frequency distribution (in %) of NDVI values for broadleaf class.

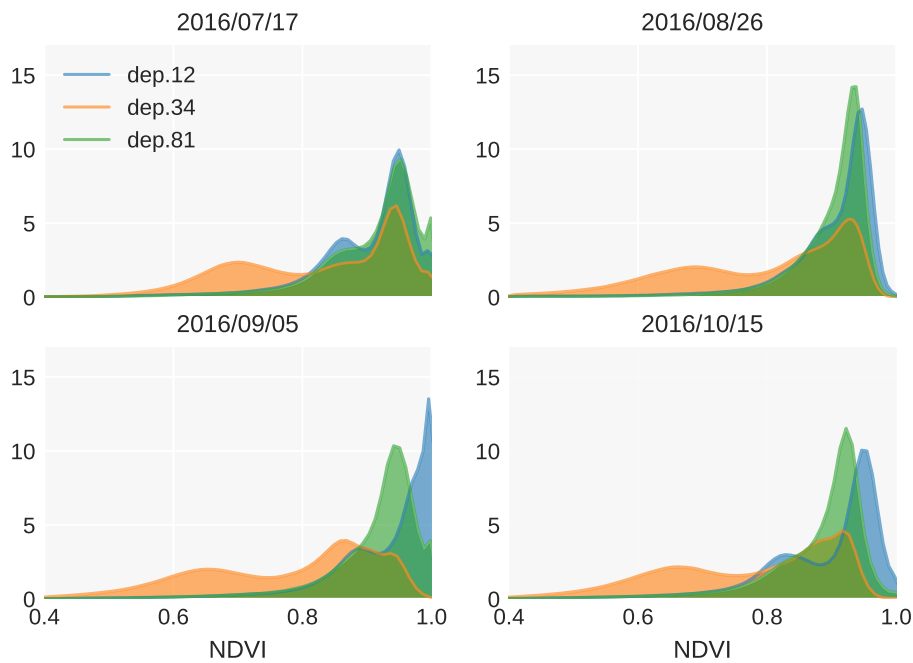
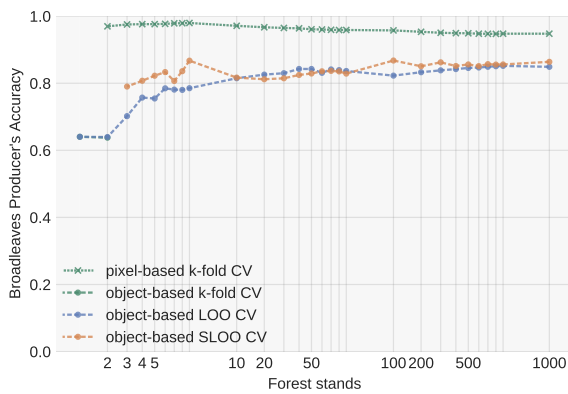
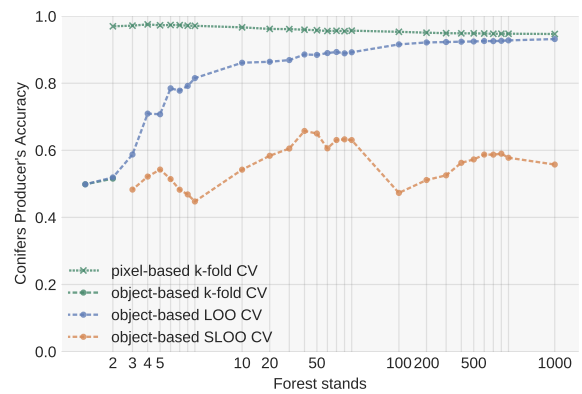


FIGURE B.3 – Relative frequency distribution (in %) of NDVI values for coniferous class.

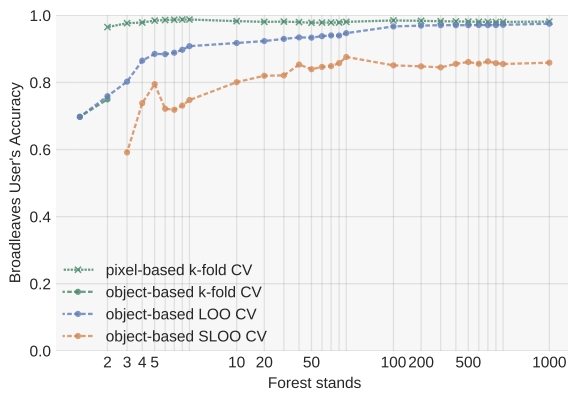
B.3 Class-specific performances and class prevalence



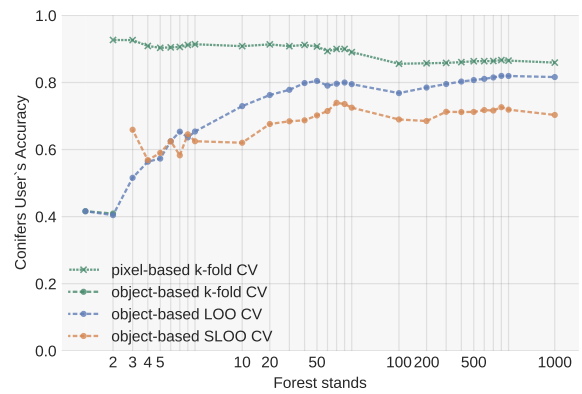
(a) Producer's accuracy for broadleaf class



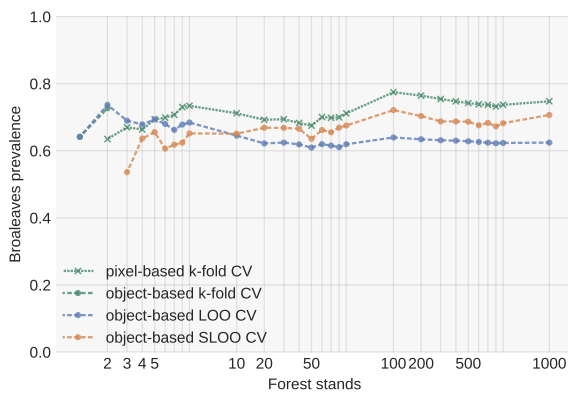
(b) Producer's accuracy for coniferous class



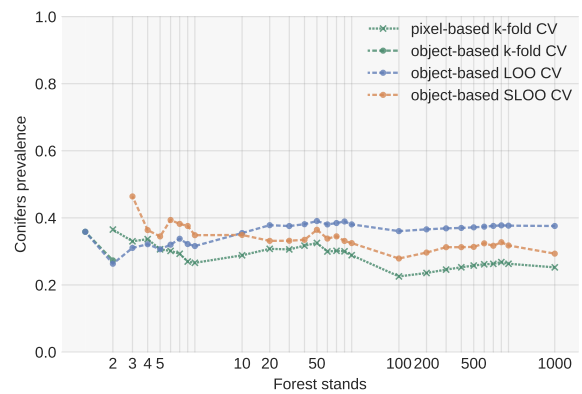
(c) User's accuracy for broadleaf class



(d) User's accuracy for coniferous class



(e) Prevalence for broadleaf class in test set



(f) Prevalence for coniferous class in test set

FIGURE B.4 – Average accuracy metrics and prevalence in the test set for broadleaf and coniferous classes (in %).

B.4 Confusion matrices

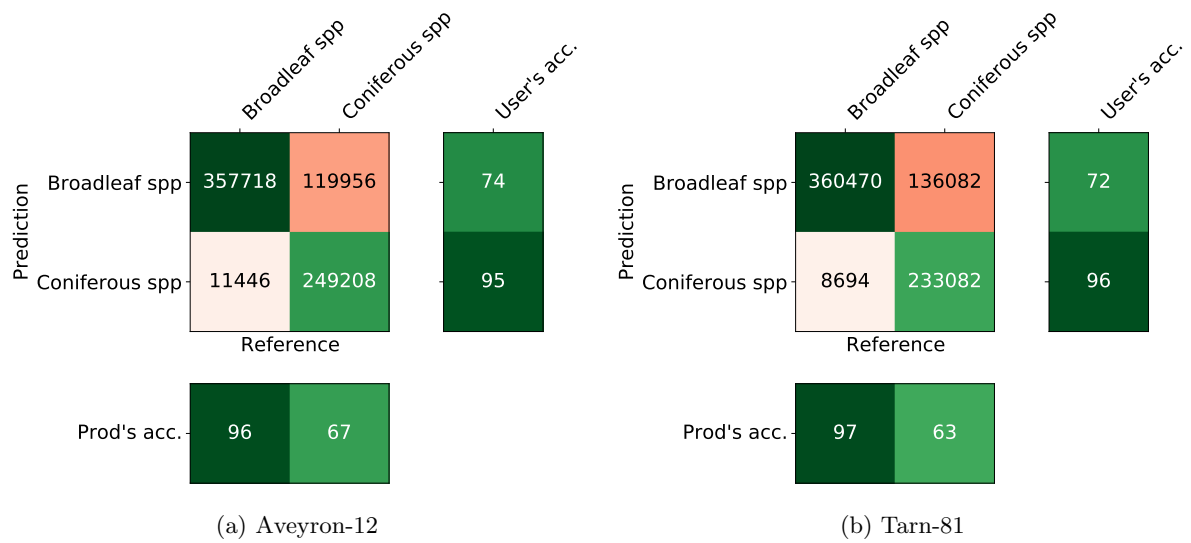


FIGURE B.5 – Confusion matrix of the test set in Aveyron-12 (left) and in Tarn-81 (right) from the Herault-34 full model fitted with the maximum training set size of 1,972 forest stands. The average accuracy metrics have been estimated from 369,164 reference pixels for each class after undersampling repeated 30 times. The values in the cells are in pixels. The units of Producer and User's accuracy are in %.

Annexe C

Compléments du chapitre 4

C.1 Protocole de suivi phénologique



Protocole pour le suivi phénologique des essences

Cadre : Thèse sur la télédétection hypertemporelle des essences forestières

Initié par : Nicolas Karasiak

Date de début de la campagne : Jeudi 28 septembre 2017

Afin de mieux comprendre les images satellite utilisées pour la classification des essences, une campagne de terrain va être réalisée afin de suivre l'évolution de plusieurs essences forestières présentes sur la zone de l'Observatoire Spatial Régionale sur l'année 2017/2018/2019. La fréquence des visites sera d'au minimum 5 jours et au maximum 10 jours.

Les essences à suivre

La phénologie des 7 essences forestières de feuillus seront analysées :
Chêne, Chêne rouge, Frêne, Bouleau, Peuplier, Robinier, Saule.



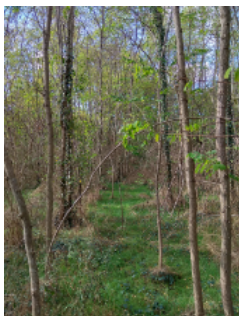
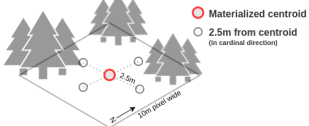



L'équipement à prévoir

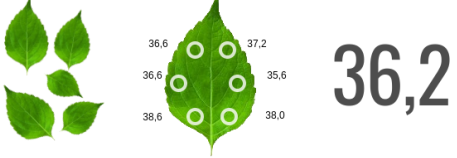
Le matériel à avoir impérativement à chaque sortie est :

- ✓ Tablette Samsung Galaxy Tab S3
- ✓ Fisheye à clipser sur la tablette
- ✓ Si photo des feuilles (pas en automne 2018) : Support maison pour prendre les feuilles en photo, incluant :
 - ✓ Plexiglas
 - ✓ Fond blanc avec carré rouge (2x2cm)
- ✓ Jumelles afin de bien observer l'étape phénologique
- ✓ Échenilloir avec perche télescopique pour prendre les feuilles
- ✓ SPAD 502 Chlorophyll meter
- ✓ Livre détaillant le code BBCH : *Les plantes au rythme des saisons - Guide d'observation phénologique* (Badeau et al, 2017).

Les données à saisir pour chaque point d'intérêt

Pour chaque peuplement (et chaque essence), devra être rempli à chaque sortie le tableau suivant :

Élément à relever	Exemple
Code BBCH	62 (début de floraison, 20 % de fleurs ouvertes)
<p>Photo du lieu (en mode portrait)</p>  <p>Flèche P notée sur le piquet symbolisant le centroid. Possibilité de voir les précédentes photos en regardant les précédents relevés de la parcelle sur Memento Database.</p> 	
<p>% ouverture de la canopée (Indice CACO) (5 photos de la canopée à 2m du sol, horizontal, au centroid et aux 4 points cardinaux, à 2,5m du centroid')</p> <p>Take a canopy cover photography from :</p>  <p>Avec le logiciel Level Camera, prendre la photo à 2m du sol, à l'horizontal. Sauvegarder le résultat sous la forme "Numéro de la parcelle + Première lettre du point cardinal" dans l'application GLAMA, ex: 2C (Saufe, centroid) 5E (Frêne, est)...</p>	
<p>Photo d'une feuille représentative</p>  <p>idéalement feuille du 1/3 supérieur (plus jeune feuille adulte) Photo sur fond blanc a4/a3 avec carré rouge (2x2cm), le tout aplati par une vitre plexiglas. Si plusieurs feuilles avec phénologie différente, prendre une photo de l'ensemble des feuilles.</p>	

<p>Mesure de la chlorophylle avec le SPAD Prendre 5 feuilles du 1/4 supérieur (plus jeune feuille adulte) et 6 mesures par feuille. Si moins de 6 mesures par feuille possibles, prendre autant de feuilles possibles pour arriver à 25 mesures.</p>	<p>For each leaf 6 mesures per leaf Mean for 30 measures</p> 
Code BBCH sous-étage	91 (Sénescence : 10% des feuilles sont colorées ou tombées)
Sous-étage (type, % grille de Prodon*)	Laisser un commentaire

Les relevés terrain : comment procéder ?

1. Se diriger vers les peuplements suivis :
 - a. Itinéraire à partir de l'application Memento Database
 - b. OU à partir des planches A3 localisant la placette
 Si difficulté à localiser le fichier, dans Memento cliquer sur l'adresse de la parcelle puis sur le logo Google Maps. Ensuite choisissez l'application Locus Map quand on vous le demande. Dans Locus Map, appuyer sur la flèche (à droit des coordonnées et du i, et choisissez Boussole).
2. Localiser le piquet de la placette suivie et prendre une photo en mode portrait, dans *Memento Database*, en suivant l'orientation indiquée par la lettre P sur le piquet (exemple de photos disponibles dans les relevés précédents).
3. Prendre une photo de la canopée avec le logiciel *Photo Level*
4. Au piquet, avancez-vous de 2,5m à chaque point cardinal puis :
 - a. Prendre en photo la canopée, à 2m du sol (tablette à l'horizontal, capteur photo équipé du fisheye et orienté vers le ciel).
5. Ouvrir le logiciel *GLAMA*, et calculer les indices d'ouverture de canopée en sauvegardant les résultats sous le nom : "Numéro de la parcelle" suivi de la "première lettre du point cardinal" (ou C pour le centroid).
Exemple : 1C pour Peuplier 1 Centroid, 1E pour Peuplier 1 Est...
6. Si un arbre est marqué (à la bombe de peinture), prendre quelques feuilles (6 de préférence) au nord le plus haut possible afin de relever 25 mesures de chlorophylle

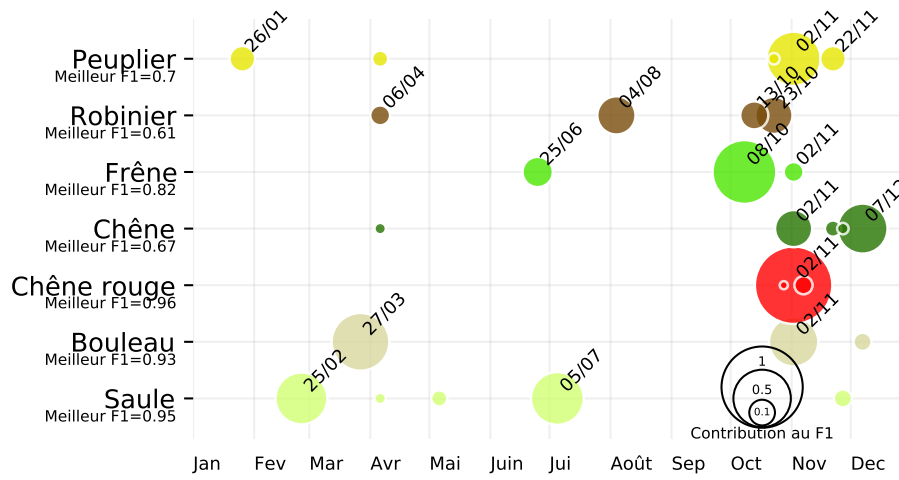
¹ Selon Tichý (2016) : *“At three regularly spaced positions forming an equilateral triangle with sides of 5 m within each plot, three hemispheric photographs were taken in the vertical direction, with the camera placed 1.5 m above the ground.”*

Bibliographie

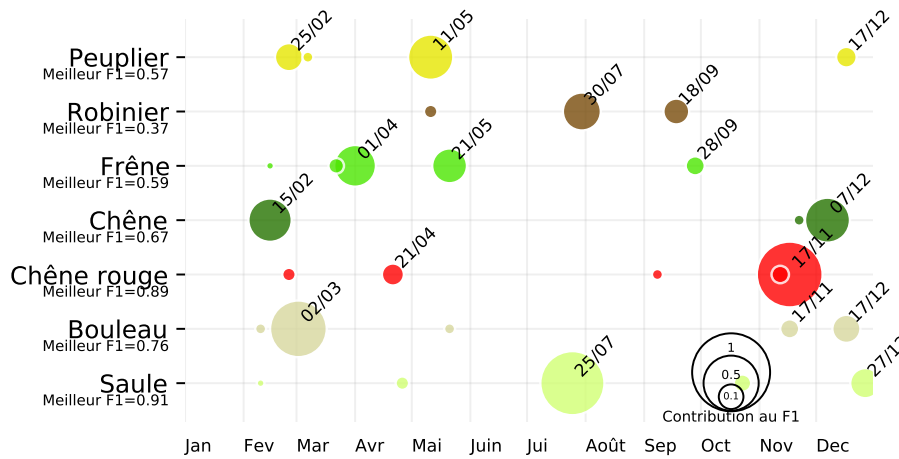
- Anger, C., Jean, F., 2014. Présentations des méthodes d'observation de la filière forêt.
- Easlon, H.M., Bloom, A.J., 2014. Easy Leaf Area: Automated Digital Image Analysis for Rapid and Accurate Measurement of Leaf Area. *Applications in Plant Sciences* 2, 1400033. doi:10.3732/apps.1400033
- Landert, K.A., 2016. Comparing Photographic and GIS-based Applications for Estimating Canopy Cover in Southern Appalachian Bogs. 2016 NCUR 0.
- Tichý, L., 2016. Field test of canopy cover estimation by hemispherical photographs taken with a smartphone. *J Veg Sci* 27, 427–435. doi:10.1111/jvs.12350
- Vincent Badeau, Marc Bonhomme, Fabrice Bonne, Jennifer Carré, Sébastien Cecchini, Isabelle Chuine, Catherine Ducatillon, Frédéric Jean, François Lebourgeois, 2017. *Les plantes au rythme des saisons*, Biotope Editions. ed.
- Zadoks, J.C., Chang, T.T., Konzak, C.F., 1974. A decimal code for the growth stages of cereals. *Weed Research* 14, 415–421. doi:10.1111/j.1365-3180.1974.tb01084.x

C.2 Les meilleures dates pour les indices spectraux

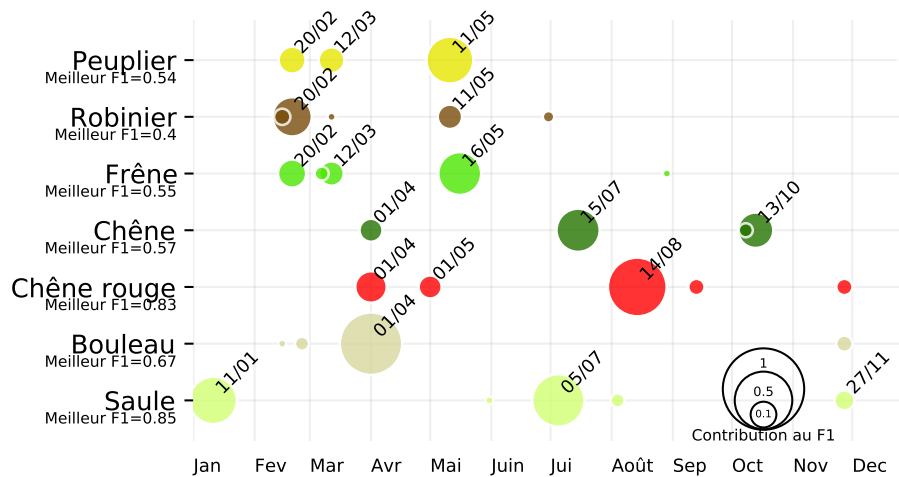
C.2.1 LAnthoC



(a) 2017.



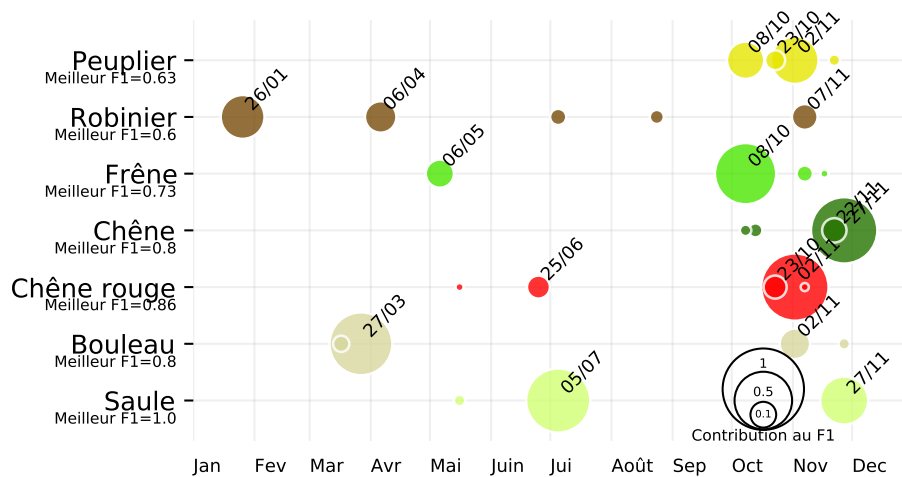
(b) 2018.



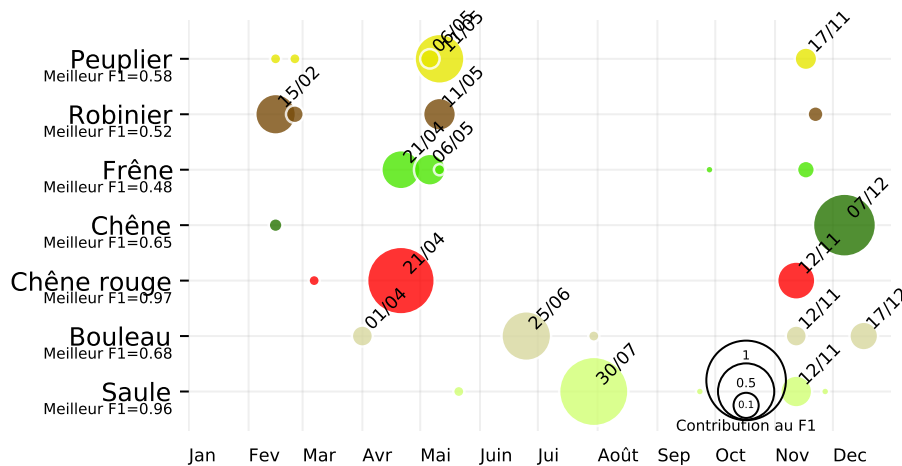
(c) 2019.

FIGURE C.1 – Liste des cinq meilleures dates pour classer chaque essence à partir de l'indice LAnthoC de l'année 2017 (a), 2018 (b) et 2019 (c).

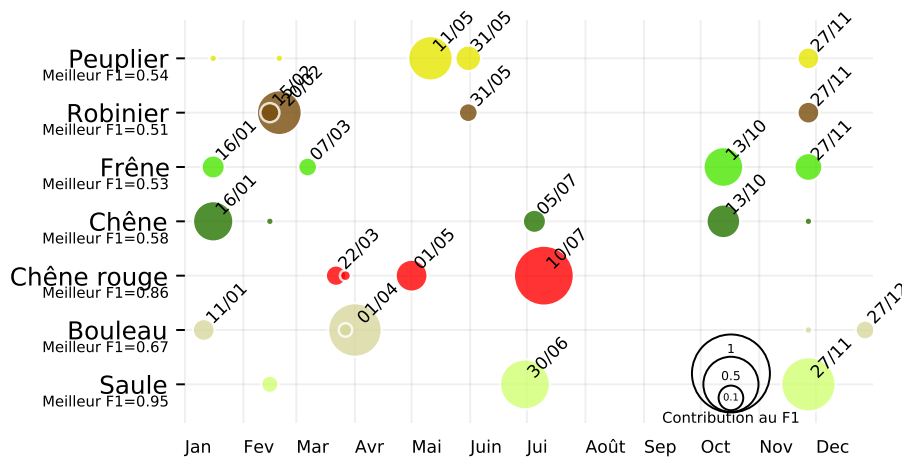
C.2.2 LCaroC



(a) 2017.



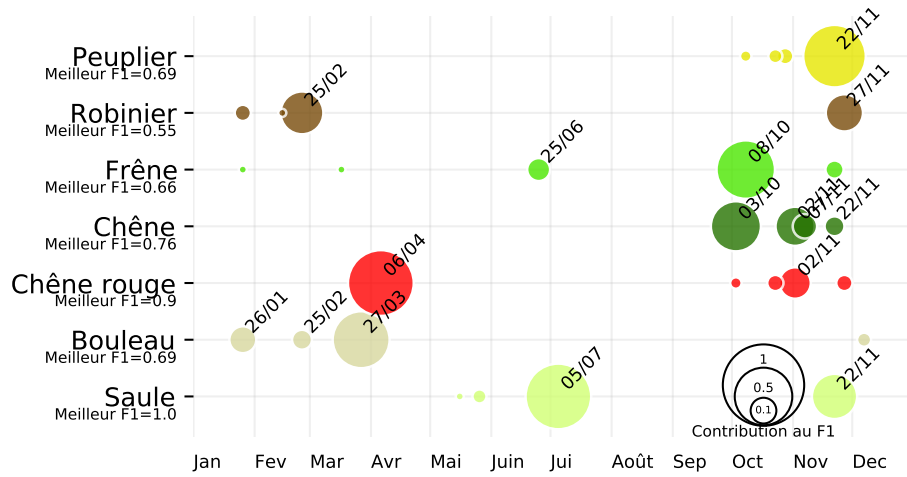
(b) 2018.



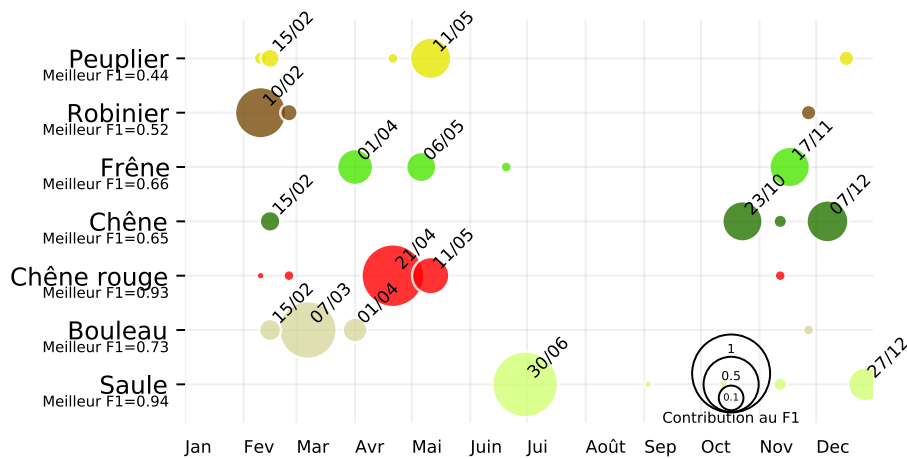
(c) 2019.

FIGURE C.2 – Liste des cinq meilleures dates pour classer chaque essence à partir de l'indice LCaroC de l'année 2017 (a), 2018 (b) et 2019 (c).

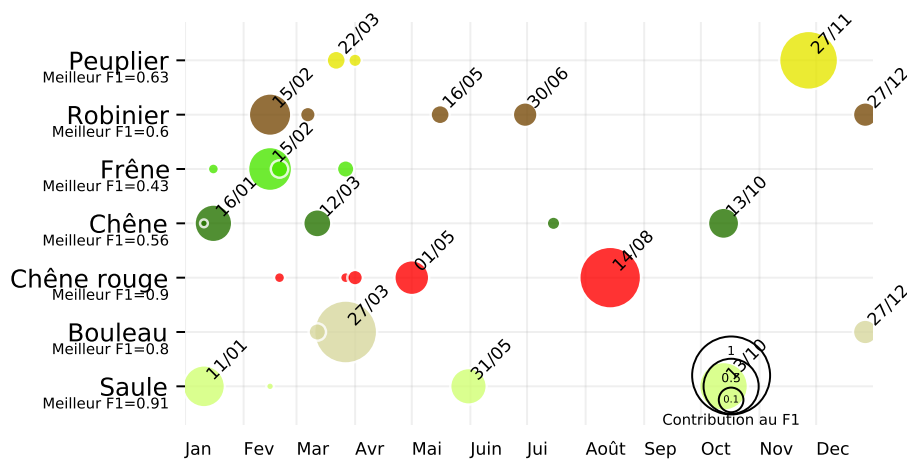
C.2.3 LChloC



(a) 2017.



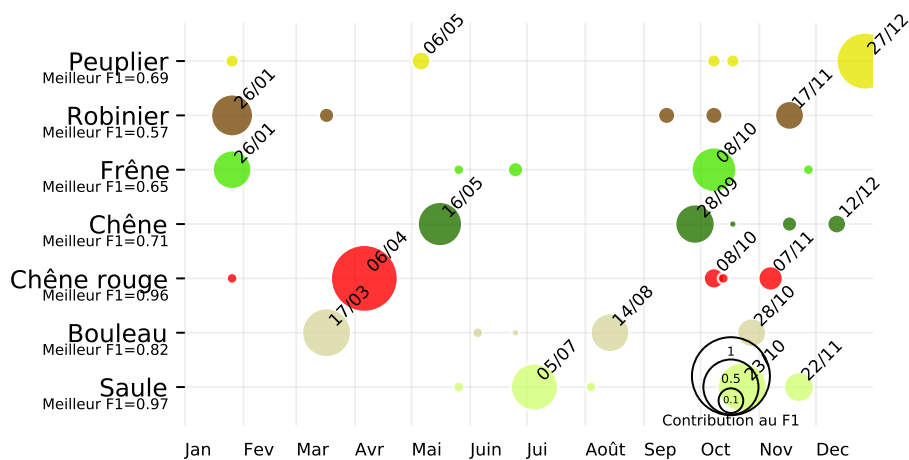
(b) 2018.



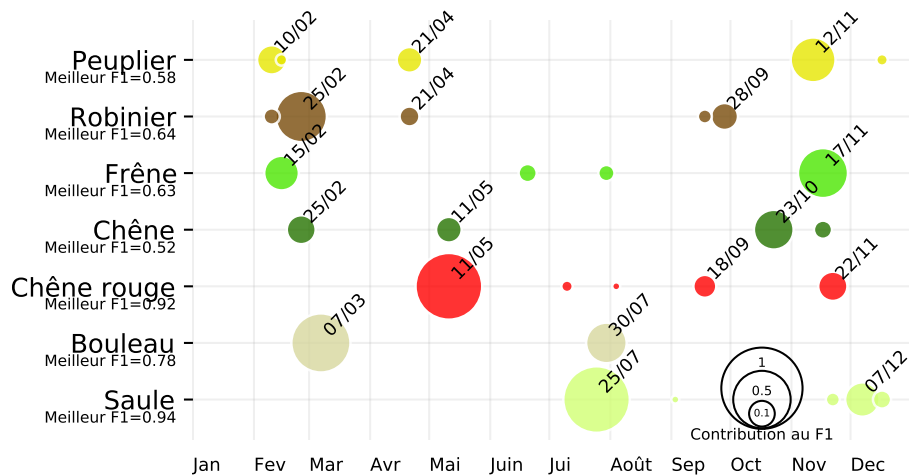
(c) 2019.

FIGURE C.3 – Liste des cinq meilleures dates pour classer chaque essence à partir de l'indice LChloC de l'année 2017 (a), 2018 (b) et 2019 (c).

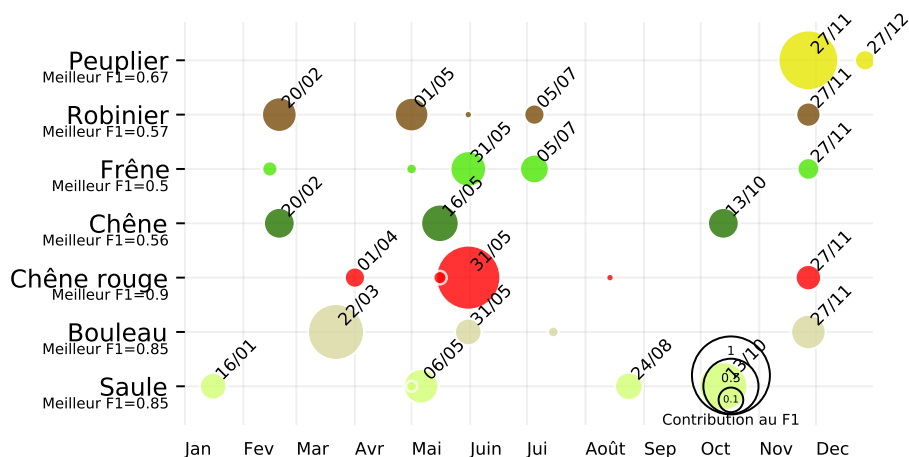
C.2.4 NDVI



(a) 2017.



(b) 2018.



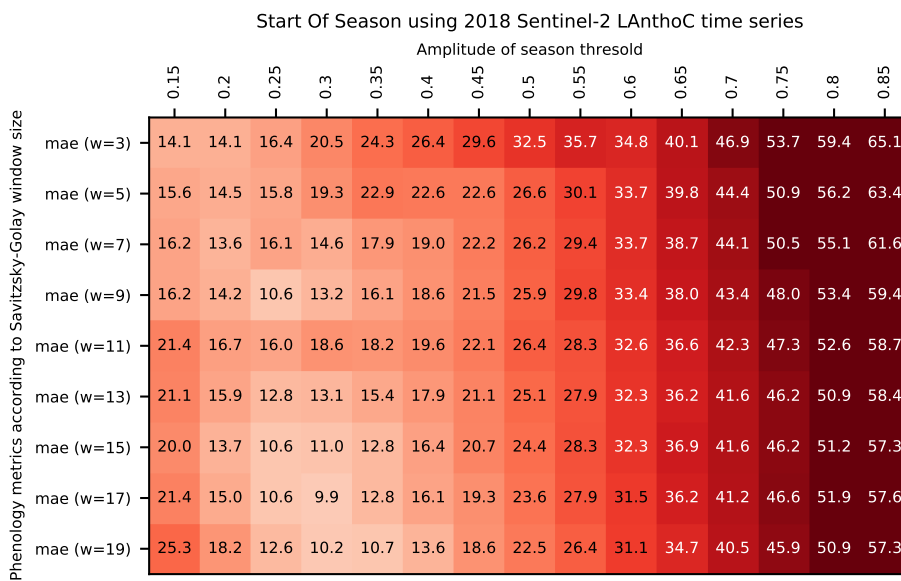
(c) 2019.

FIGURE C.4 – Liste des cinq meilleures dates pour classer chaque essence à partir de l'indice NDVI de l'année 2017 (a), 2018 (b) et 2019 (c).

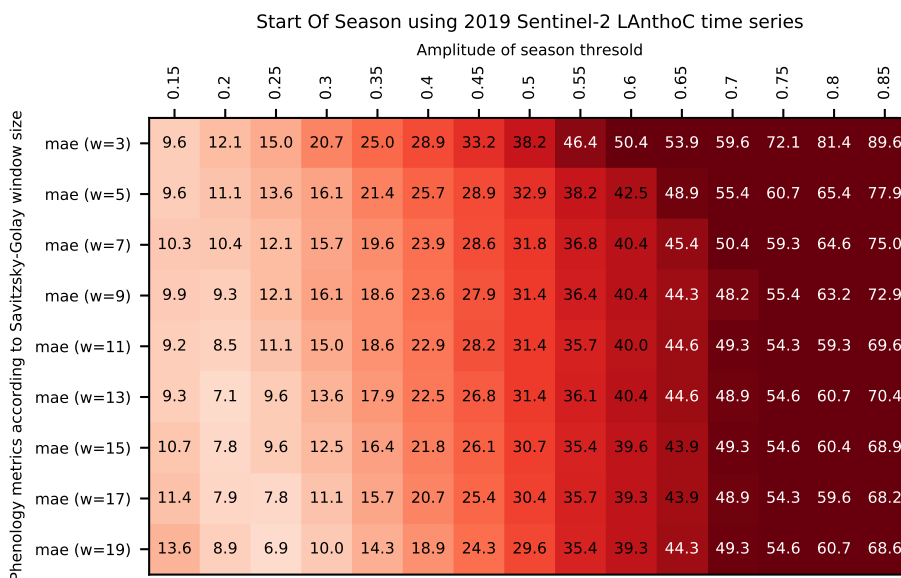
C.3 Les paramètres des indices et de Savitzsky-Golay

C.3.1 Pour le début de saison (BBCH=11)

C.3.1.1 LAnthoC



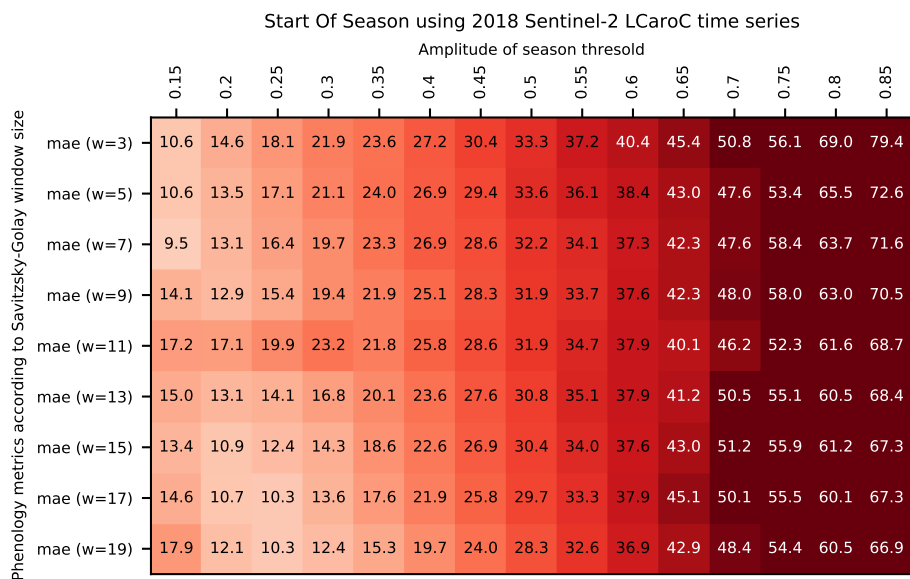
(a) 2018



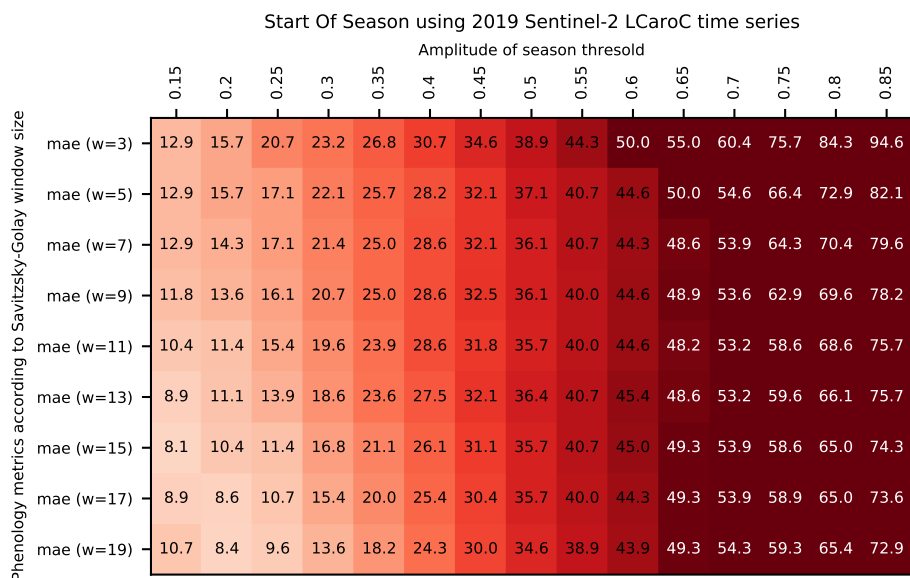
(b) 2019

FIGURE C.5 – Erreur moyenne absolue (MAE) pour le début de saison (BBCH=11) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LAnthoC.

C.3.1.2 LCaroC



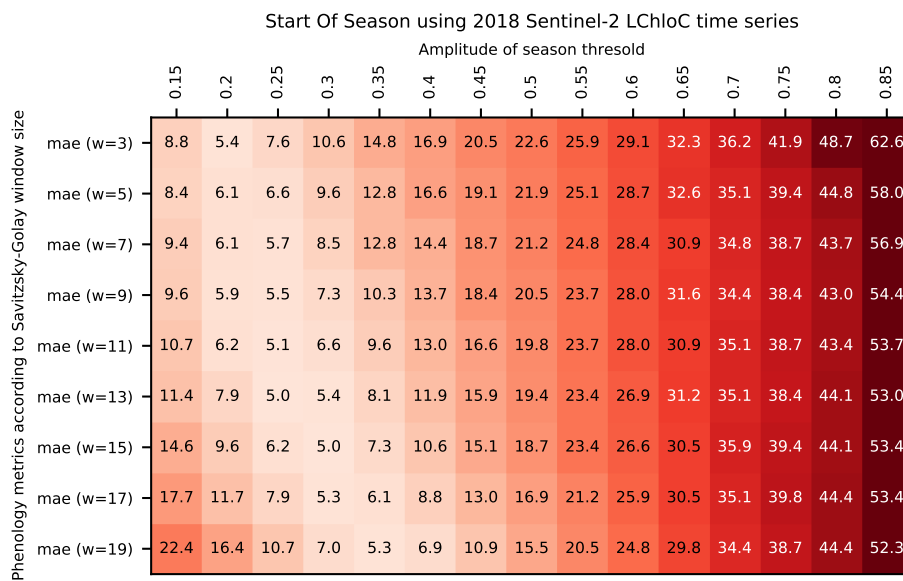
(a) 2018



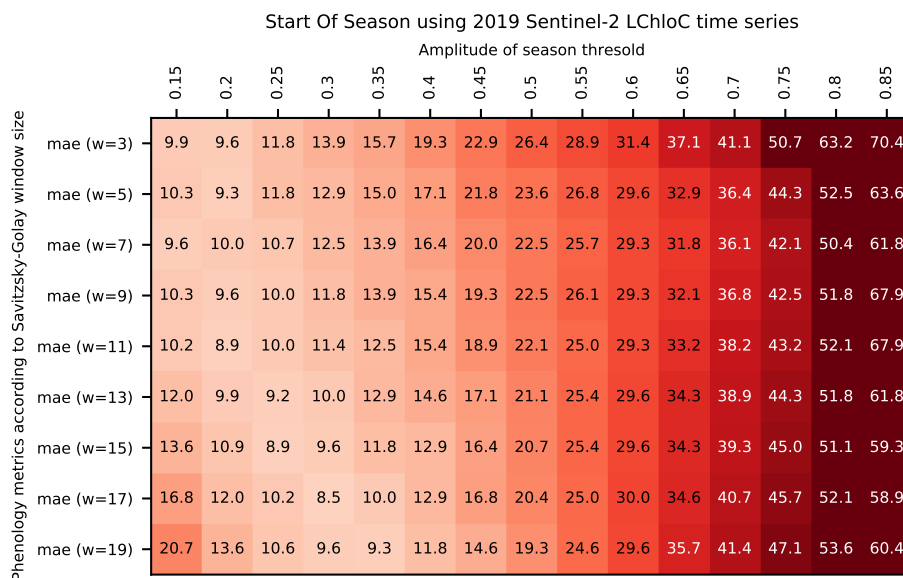
(b) 2019

FIGURE C.6 – Erreur moyenne absolue (MAE) pour le début de saison (BBCH=11) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l’amplitude du LCaroC.

C.3.1.3 LChloC



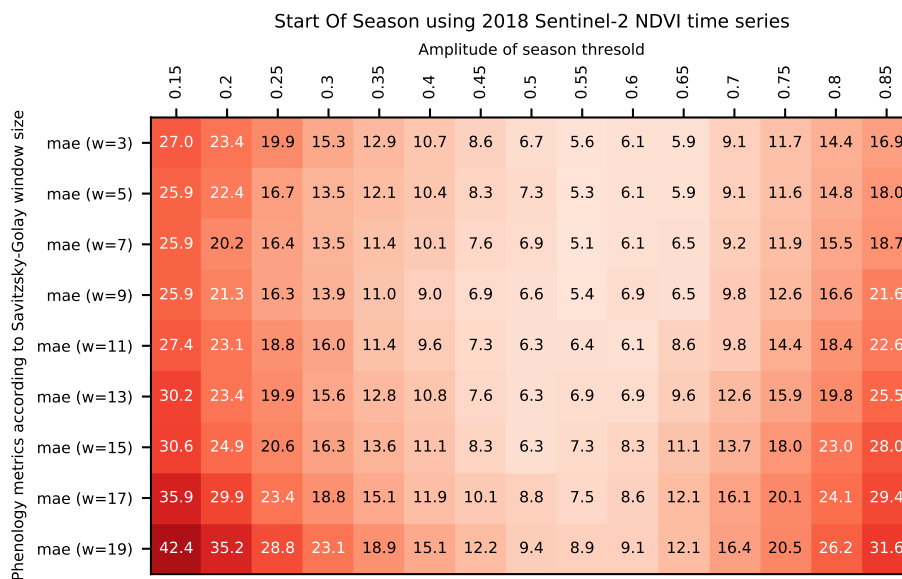
(a) 2018



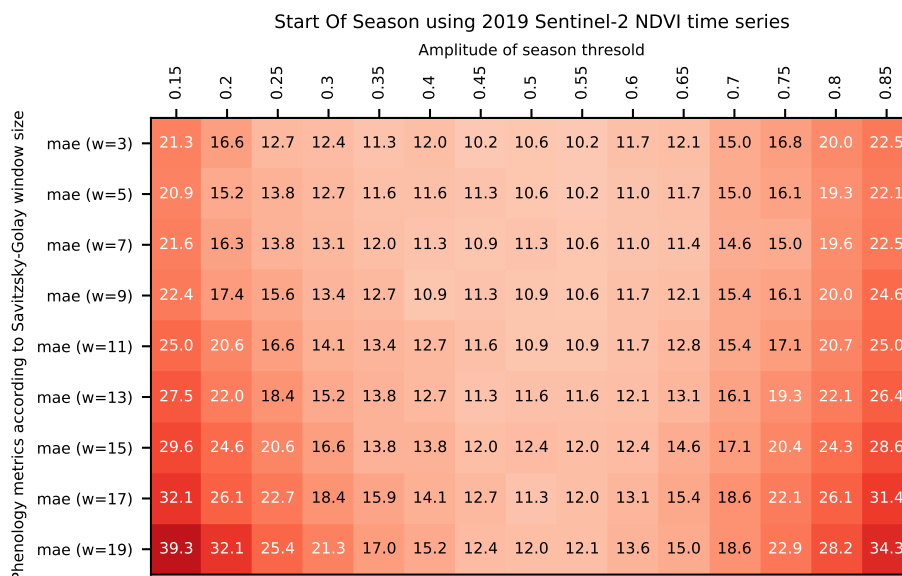
(b) 2019

FIGURE C.7 – Erreur moyenne absolue (MAE) pour le début de saison (BBCH=11) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LChloC.

C.3.1.4 NDVI



(a) 2018

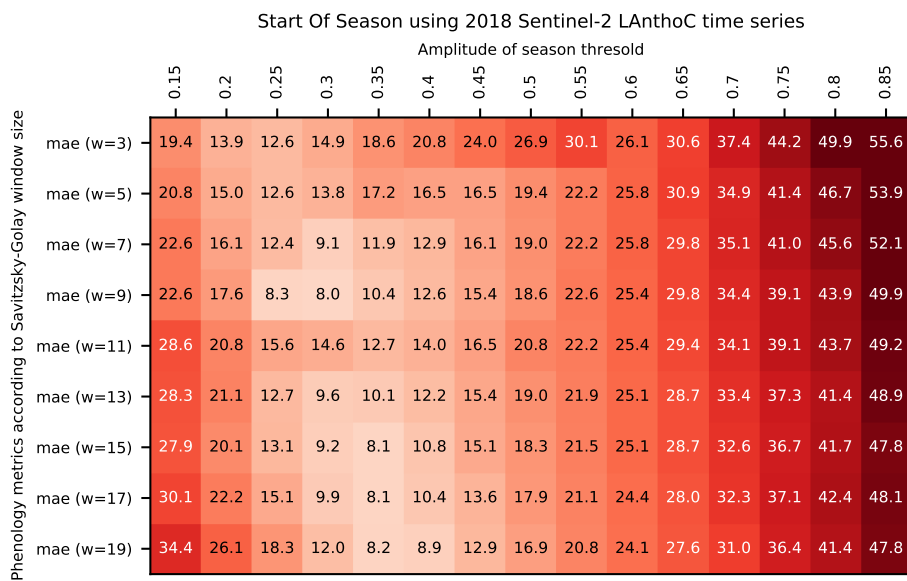


(b) 2019

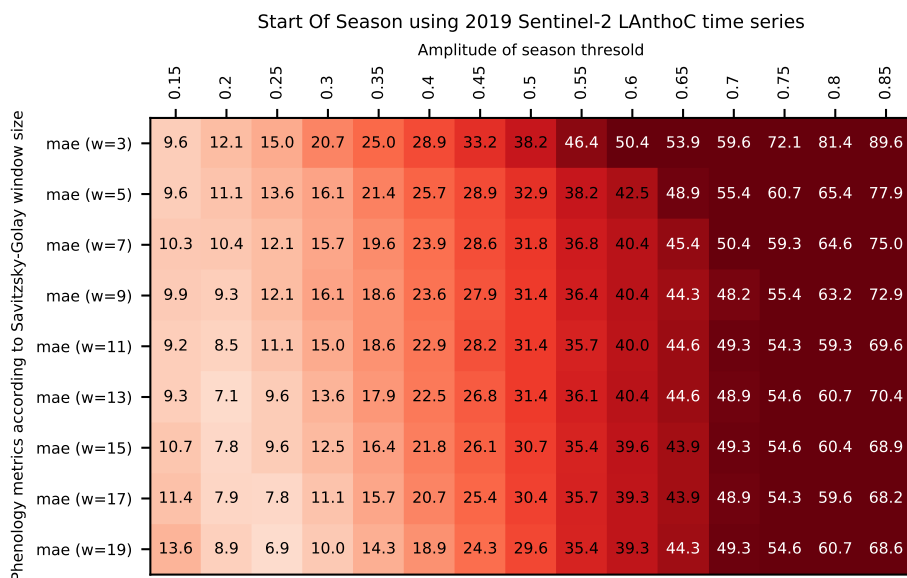
FIGURE C.8 – Erreur moyenne absolue (MAE) pour le début de saison (BBCH=11) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l’amplitude du NDVI pour l’année.

C.3.2 Pour le début de saison (BBCH=15)

C.3.2.1 LAnthoC



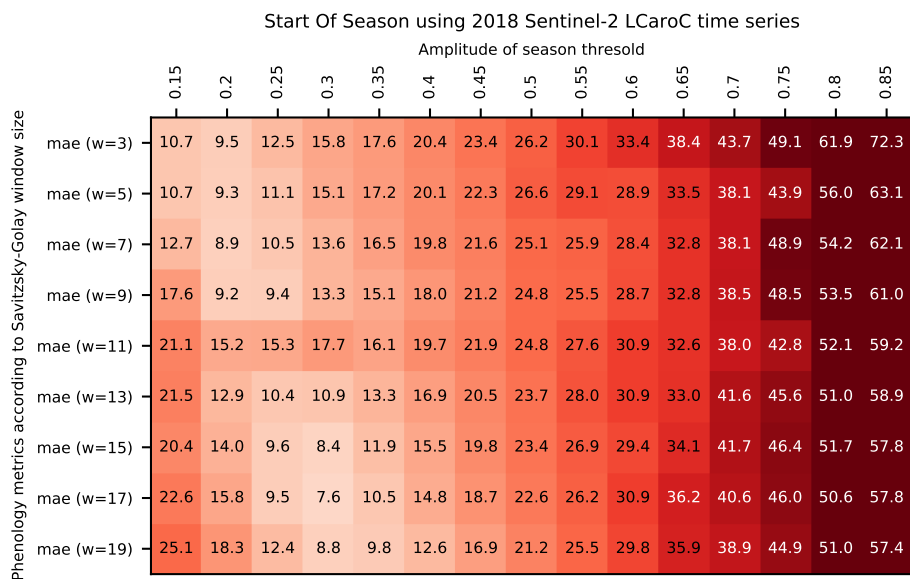
(a) 2018



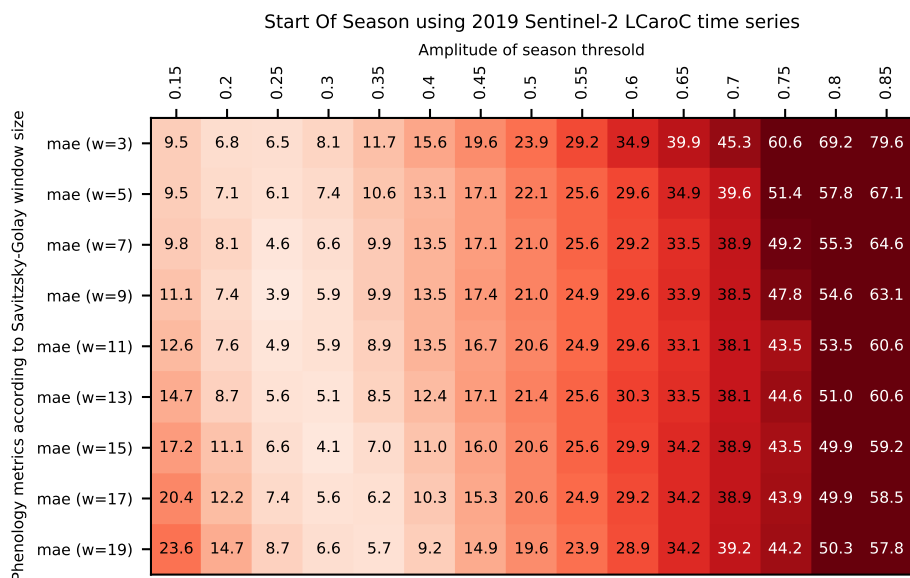
(b) 2019

FIGURE C.9 – Erreur moyenne absolue (MAE) pour le début de saison (BBCH=11) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LAnthoC.

C.3.2.2 LCaroC



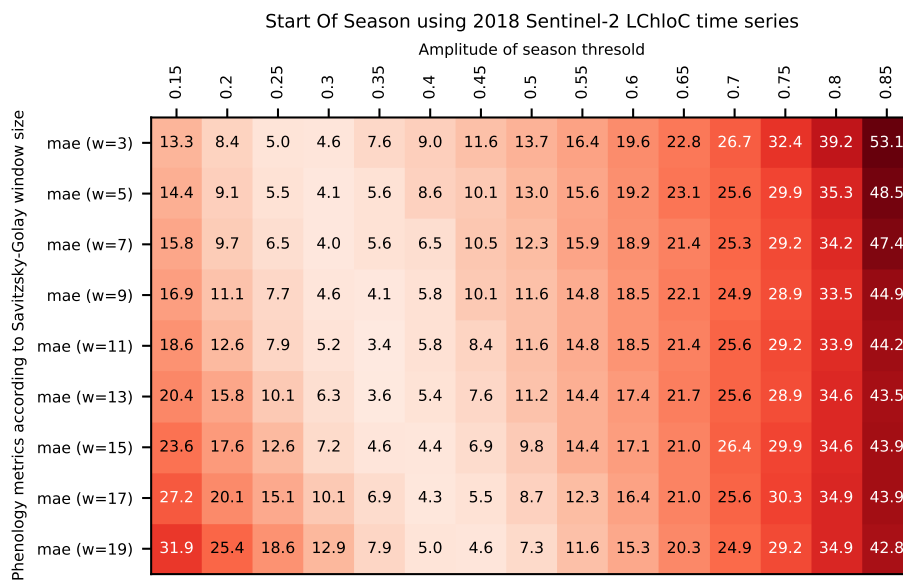
(a) 2018



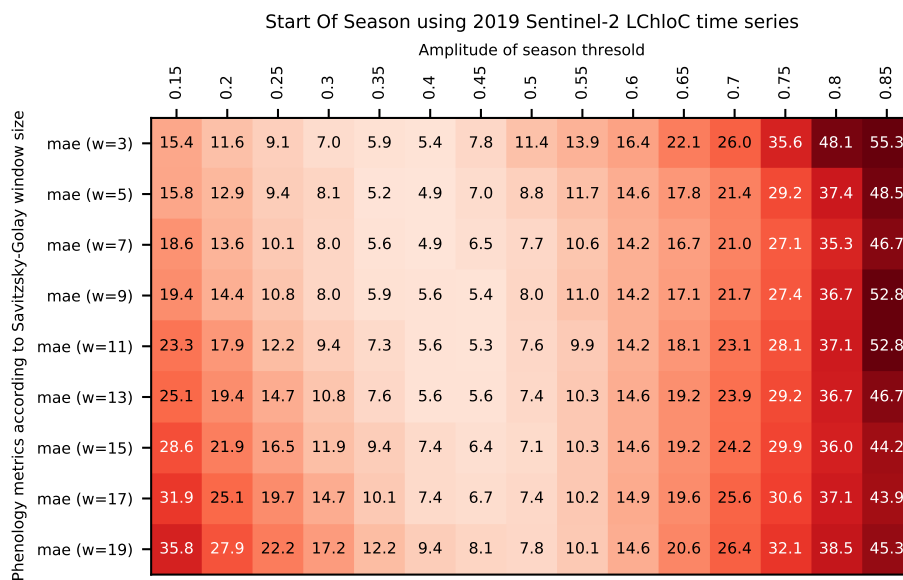
(b) 2019

FIGURE C.10 – Erreur moyenne absolue (MAE) pour le début de saison (BBCH=15) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l’amplitude du LCaroC.

C.3.2.3 LChloC



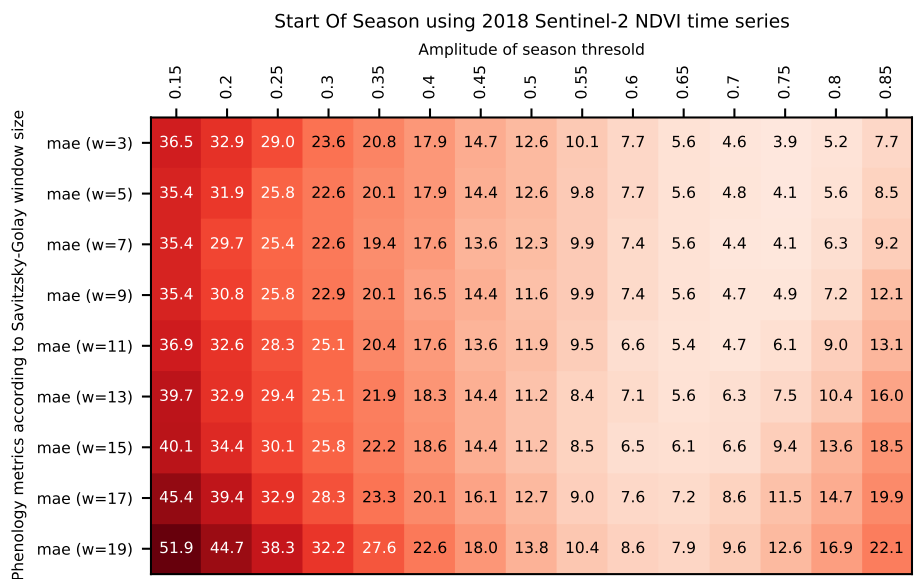
(a) 2018



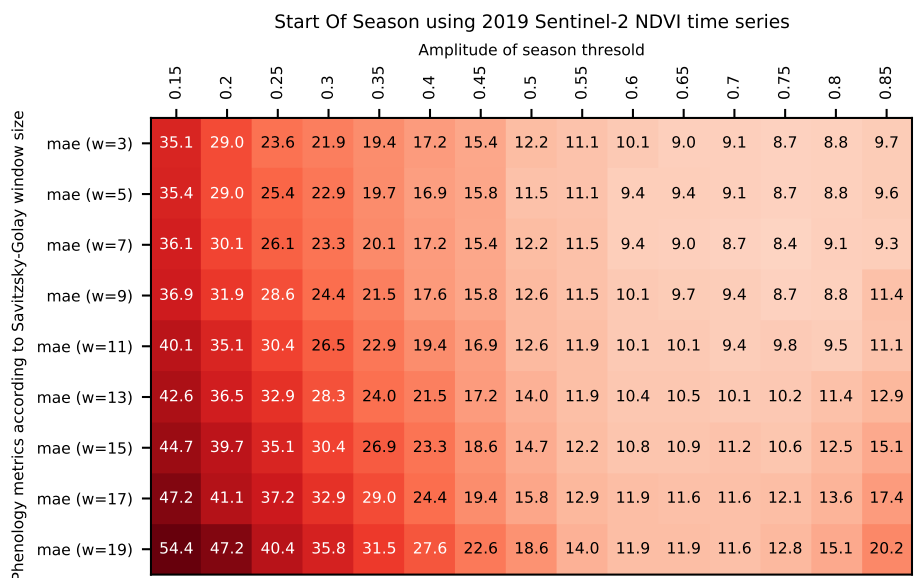
(b) 2019

FIGURE C.11 – Erreur moyenne absolue (MAE) pour le début de saison (BBCH=15) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LChloC.

C.3.2.4 NDVI



(a) 2018

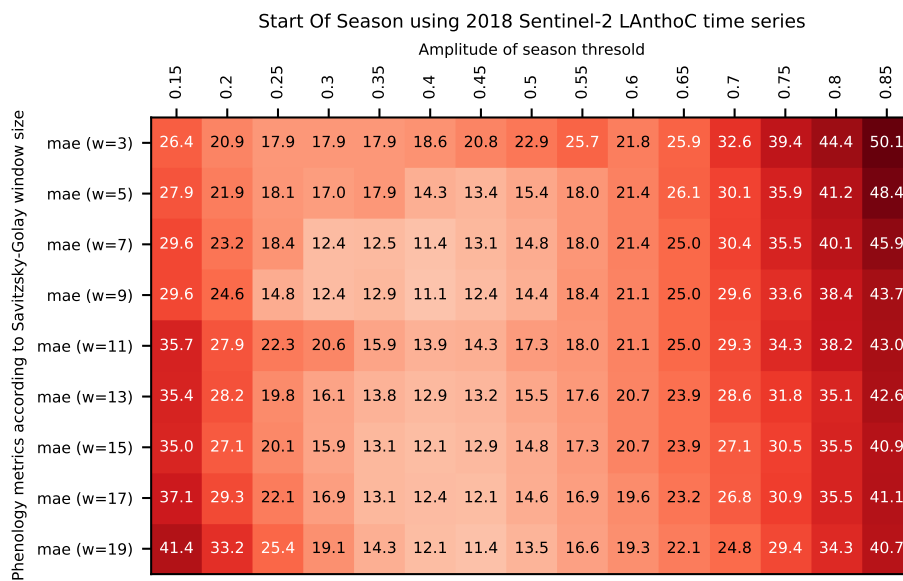


(b) 2019

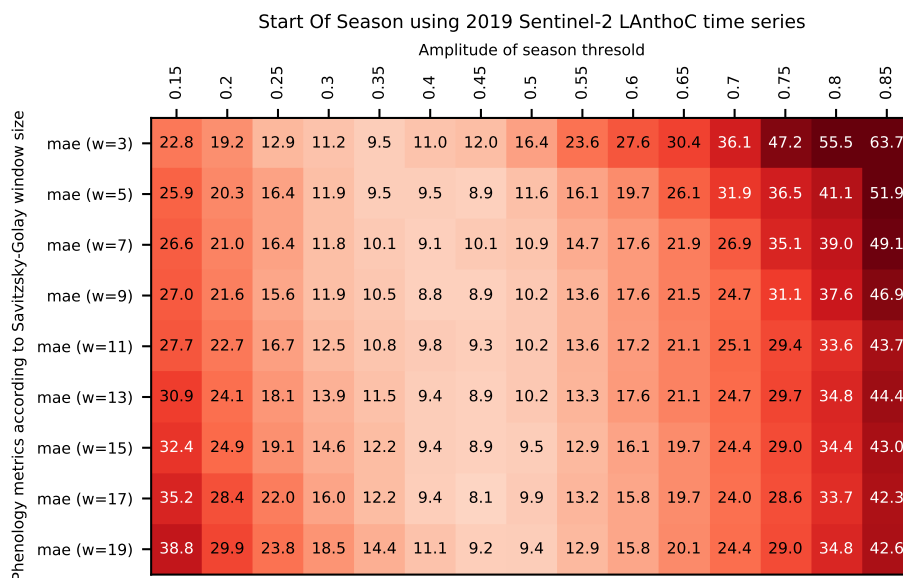
FIGURE C.12 – Erreur moyenne absolue (MAE) pour le début de saison (BBCH=15) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du NDVI.

C.3.3 Pour le début de saison (BBCH=19)

C.3.3.1 LAnthoC



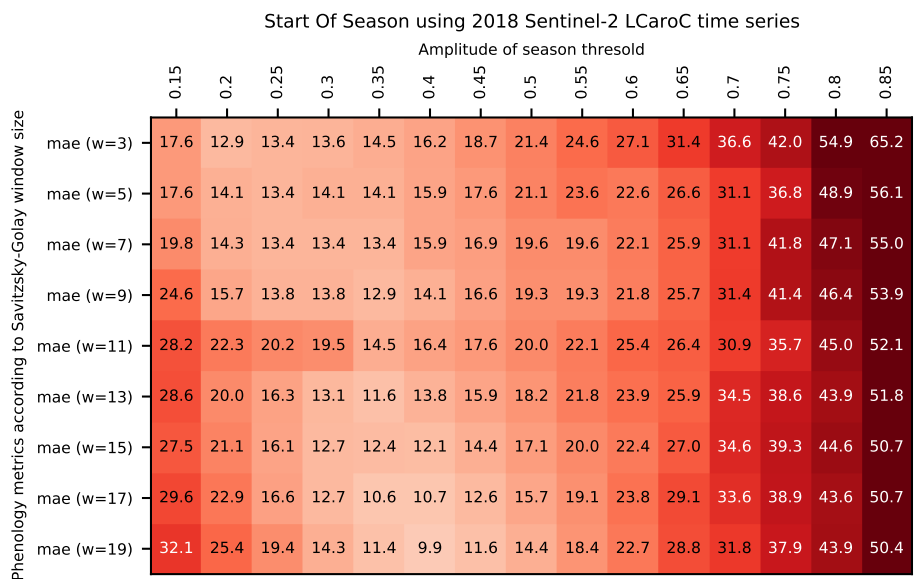
(a) 2018



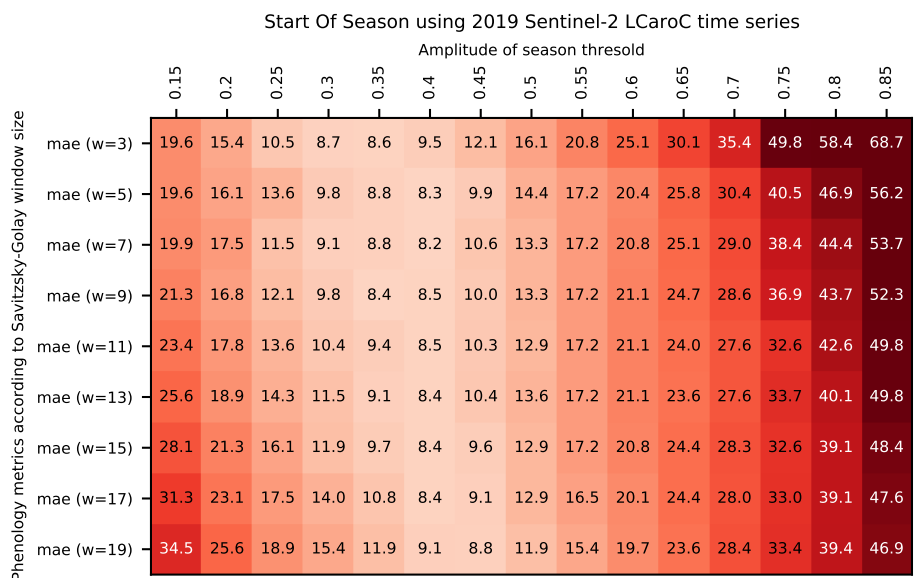
(b) 2019

FIGURE C.13 – Erreur moyenne absolue (MAE) pour le début de saison (BBCH=19) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LAnthoC.

C.3.3.2 LCaroC



(a) 2018



(b) 2019

FIGURE C.14 – Erreur moyenne absolue (MAE) pour le début de saison (BBCH=19) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LCaroC.

C.3.3.3 LChloC

Start Of Season using 2018 Sentinel-2 LChloC time series

Amplitude of season threshold

	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85
mae (w=3)	20.4	15.4	10.4	7.6	6.8	8.2	8.8	10.2	11.9	13.6	16.6	19.8	25.4	32.1	46.1
mae (w=5)	21.4	16.1	11.6	8.0	7.5	7.9	8.9	9.5	10.7	13.3	16.9	18.7	22.9	28.2	41.4
mae (w=7)	22.9	16.8	13.0	8.9	8.2	7.1	8.6	8.8	10.9	12.9	15.1	18.4	22.1	27.1	40.4
mae (w=9)	23.9	18.2	14.6	10.7	8.0	7.1	8.2	9.5	9.9	12.6	15.9	18.0	21.8	26.4	37.9
mae (w=11)	25.7	19.6	15.0	11.3	7.9	7.9	7.9	8.8	9.9	12.3	14.4	18.6	22.1	26.8	37.1
mae (w=13)	27.5	22.9	17.1	13.2	9.6	8.2	7.9	8.4	9.5	11.6	14.8	18.6	21.8	27.5	36.4
mae (w=15)	30.7	24.6	19.6	14.3	10.7	8.4	7.1	7.9	9.5	10.7	14.1	19.3	22.9	27.5	36.8
mae (w=17)	34.3	27.1	22.1	17.1	13.4	9.1	7.9	7.5	8.8	10.0	13.9	18.6	23.2	27.9	36.8
mae (w=19)	38.9	32.5	25.7	20.0	15.0	11.1	8.0	7.5	8.8	9.6	13.4	17.9	22.1	27.9	35.7

Phenology metrics according to Savitzky-Golay window size

(a) 2018

Start Of Season using 2019 Sentinel-2 LChloC time series

Amplitude of season threshold

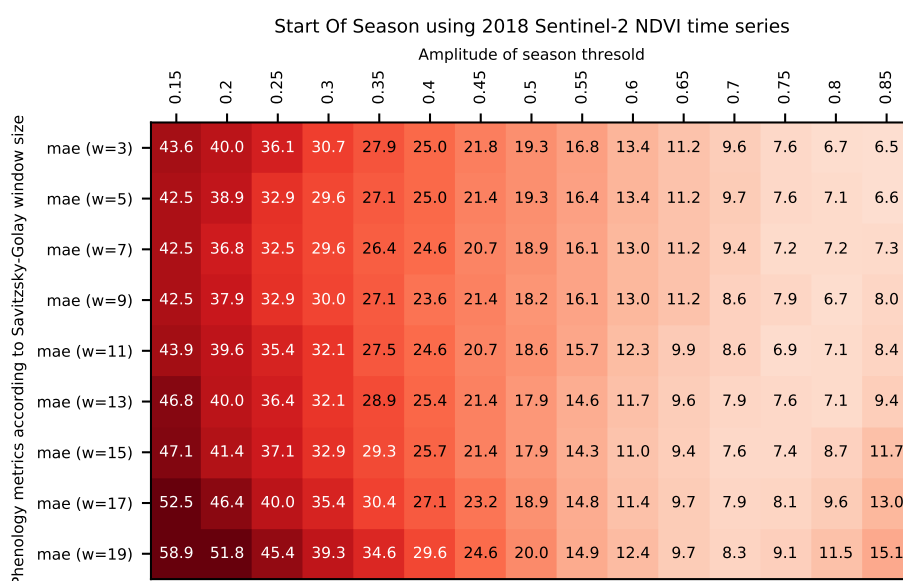
	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85
mae (w=3)	26.3	22.4	19.3	16.4	13.9	9.4	8.1	7.1	7.3	8.8	13.6	16.9	25.8	38.3	45.4
mae (w=5)	26.6	23.8	19.6	17.5	13.9	11.5	8.0	8.1	7.4	8.0	9.9	12.2	20.1	28.3	37.9
mae (w=7)	29.5	24.5	20.2	17.4	14.3	11.5	9.4	8.0	7.4	7.9	9.1	12.6	18.6	26.1	36.1
mae (w=9)	30.2	25.2	21.6	18.1	14.6	12.6	10.1	8.7	7.7	8.2	10.2	13.3	18.3	26.9	41.9
mae (w=11)	34.1	28.8	23.1	19.5	16.0	12.6	10.4	9.1	8.8	8.2	10.3	14.0	18.3	26.5	41.9
mae (w=13)	35.9	30.2	25.6	20.9	16.4	13.6	12.2	10.1	8.9	8.6	11.1	14.7	19.4	26.1	35.9
mae (w=15)	39.5	32.7	27.4	22.0	18.9	16.1	12.9	10.8	8.9	8.1	11.4	15.1	20.1	25.1	33.4
mae (w=17)	42.7	35.9	30.6	25.6	20.2	16.1	13.3	11.1	9.9	8.5	10.7	15.8	20.1	26.2	33.0
mae (w=19)	46.6	38.8	33.1	28.1	22.4	18.1	15.4	12.2	10.3	9.6	11.8	16.5	21.5	27.6	34.4

Phenology metrics according to Savitzky-Golay window size

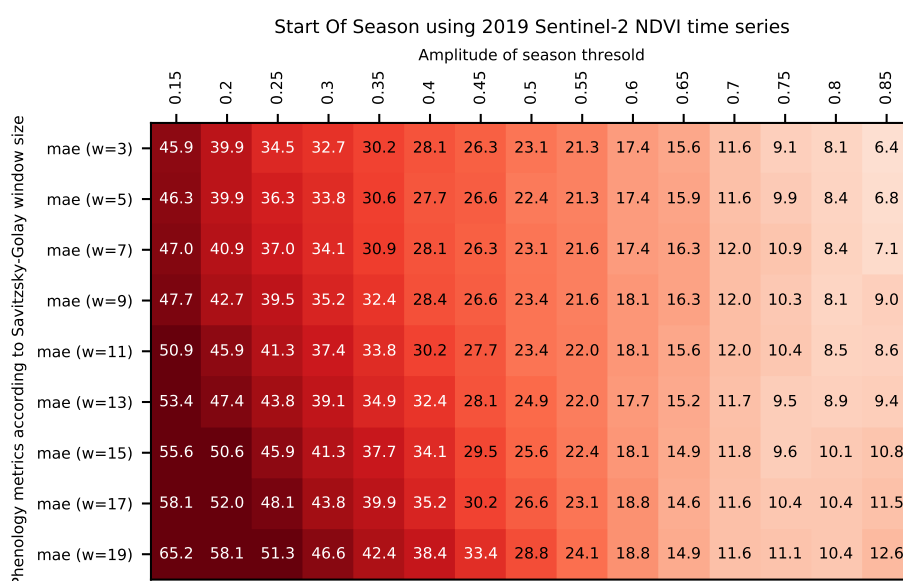
(b) 2019

FIGURE C.15 – Erreur moyenne absolue (MAE) pour le début de saison (BBCH=19) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LChloC.

C.3.3.4 NDVI



(a) 2018

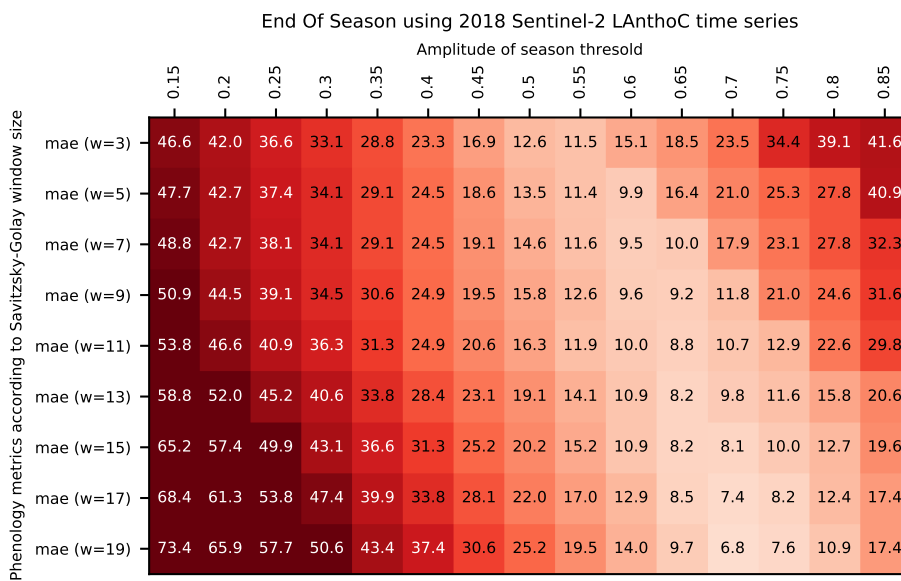


(b) 2019

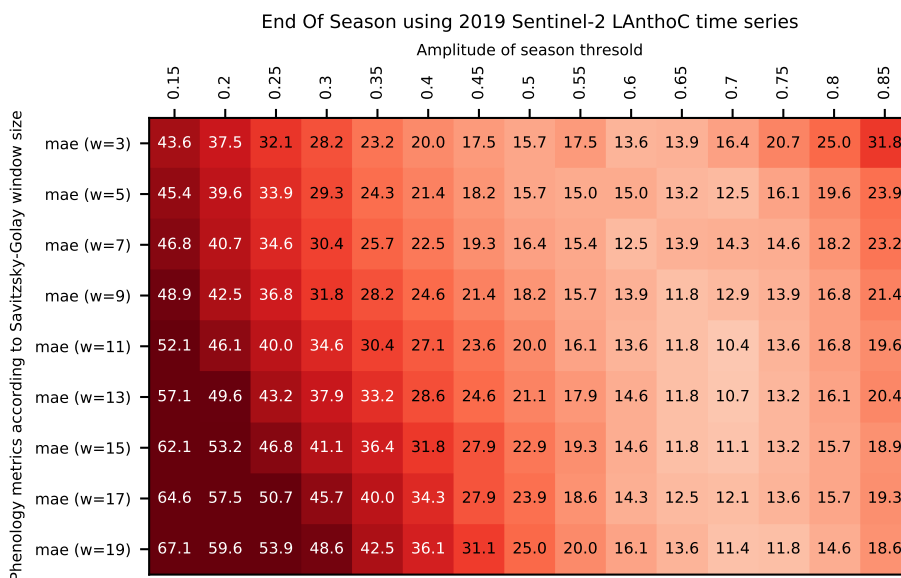
FIGURE C.16 – Erreur moyenne absolue (MAE) pour le début de saison (BBCH=19) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l’amplitude du NDVI pour l’année.

C.3.4 Pour la fin de saison (BBCH=91)

C.3.4.1 LAnthoC



(a) 2018



(b) 2019

FIGURE C.17 – Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=91) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LAnthoC.

C.3.4.2 LCaroC

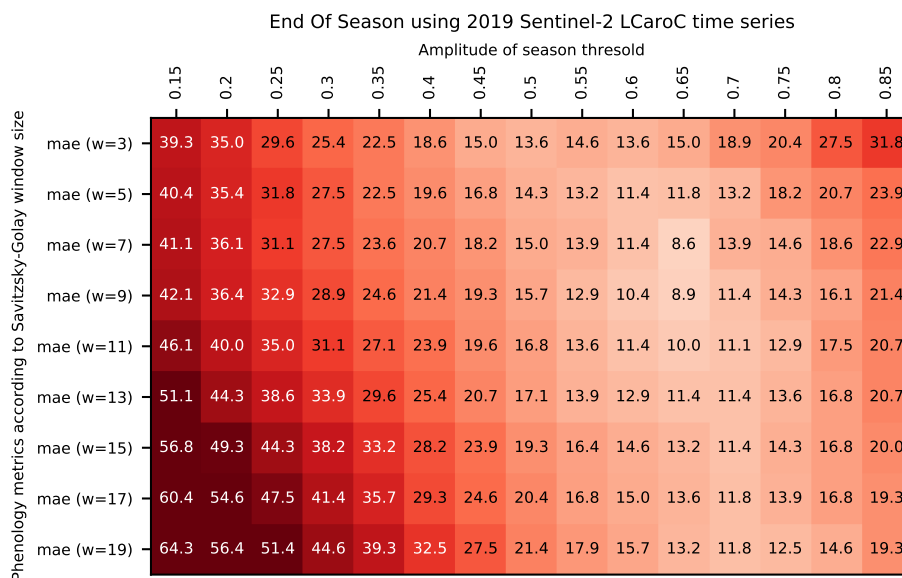
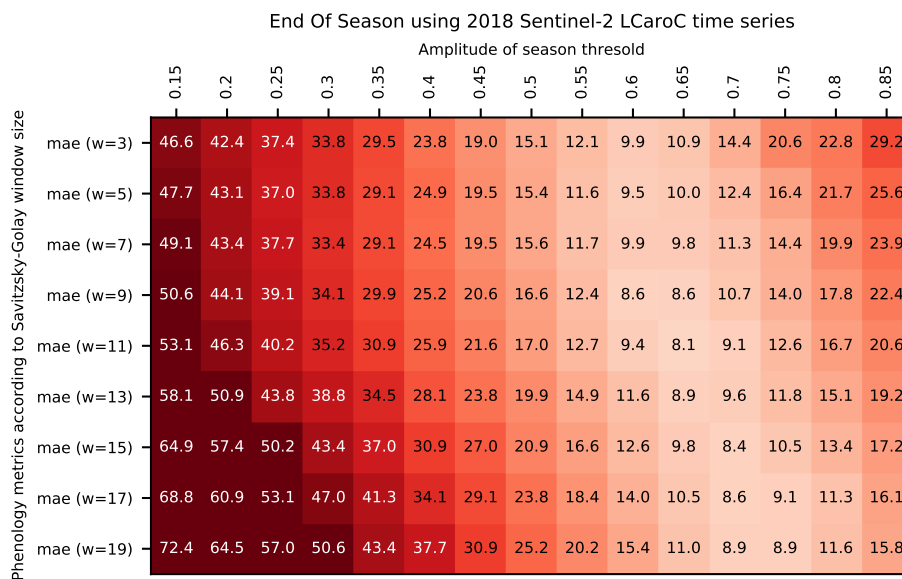
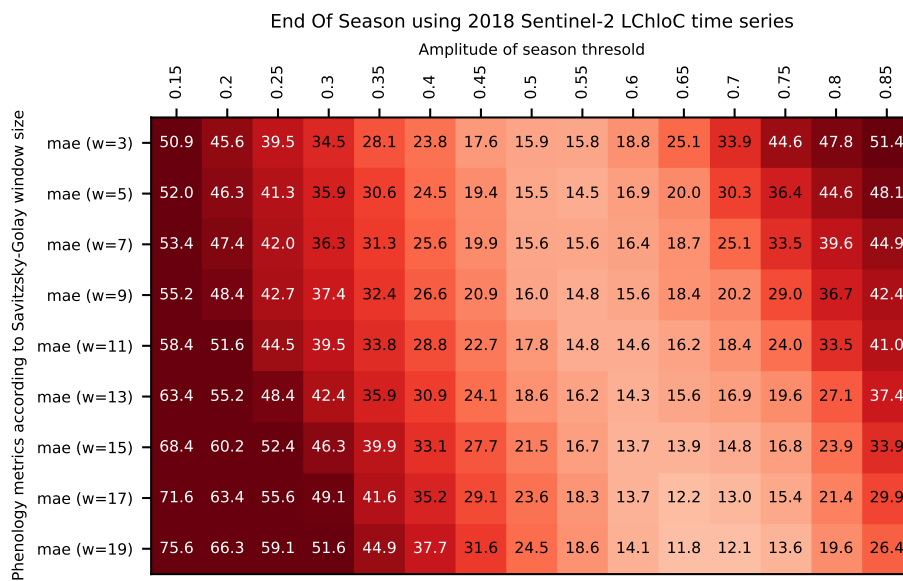
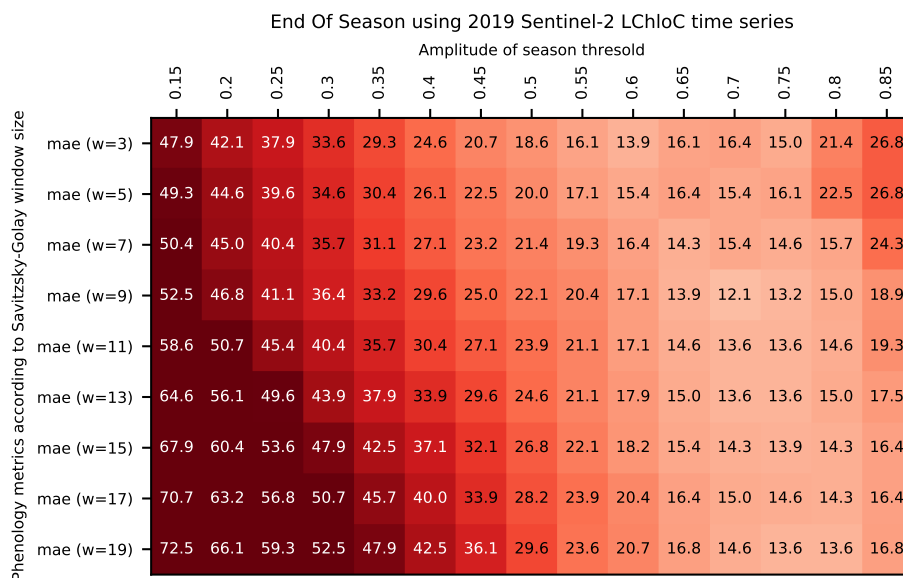


FIGURE C.18 – Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=91) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l’amplitude du LCaroC.

C.3.4.3 LChloC



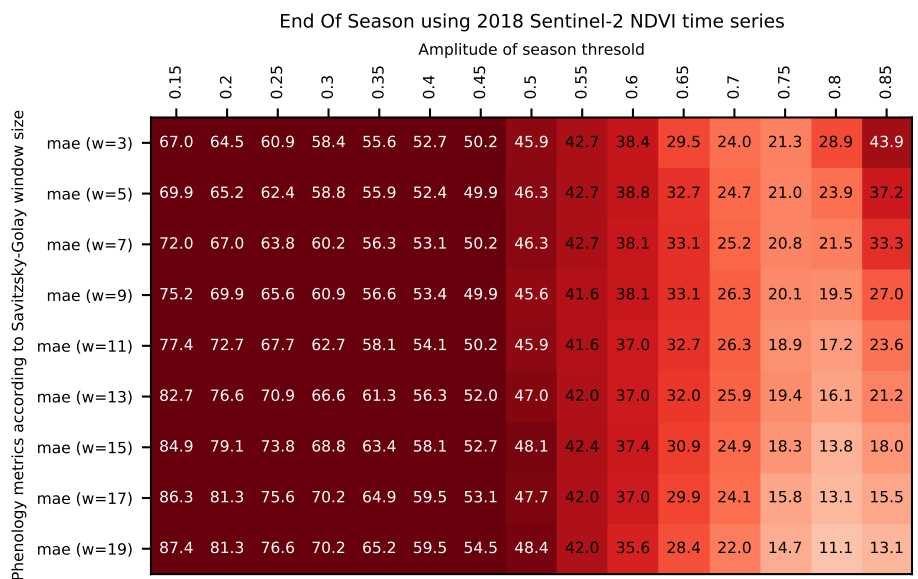
(a) 2018



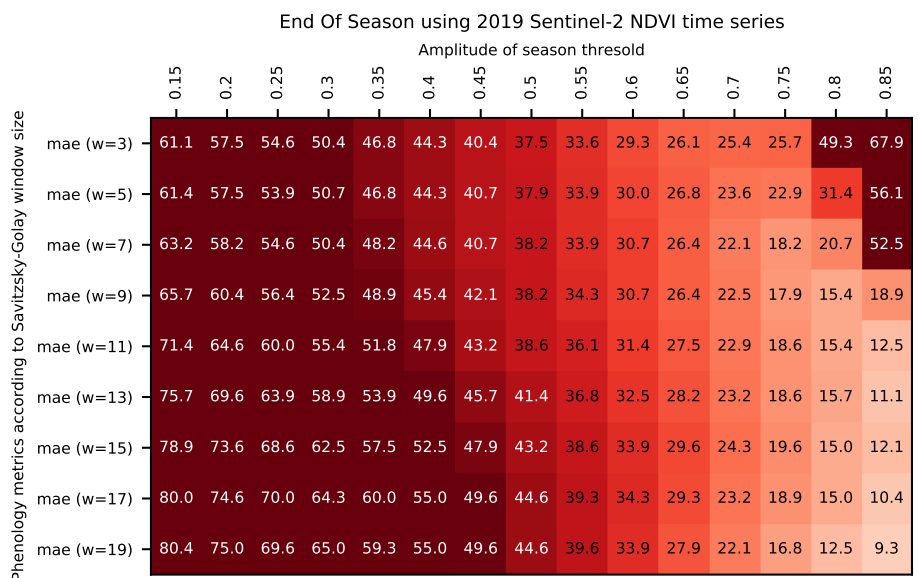
(b) 2019

FIGURE C.19 – Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=91) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LChloC.

C.3.4.4 NDVI



(a) 2018

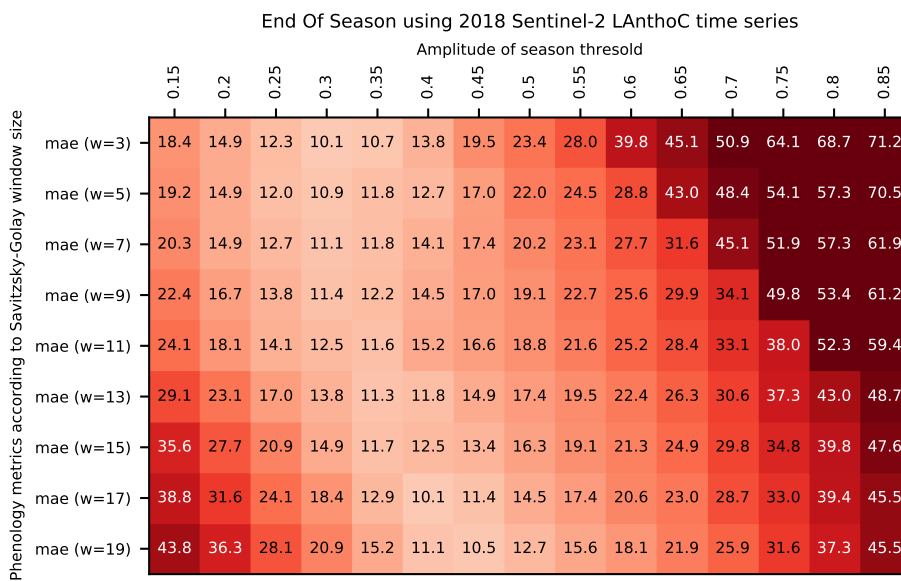


(b) 2019

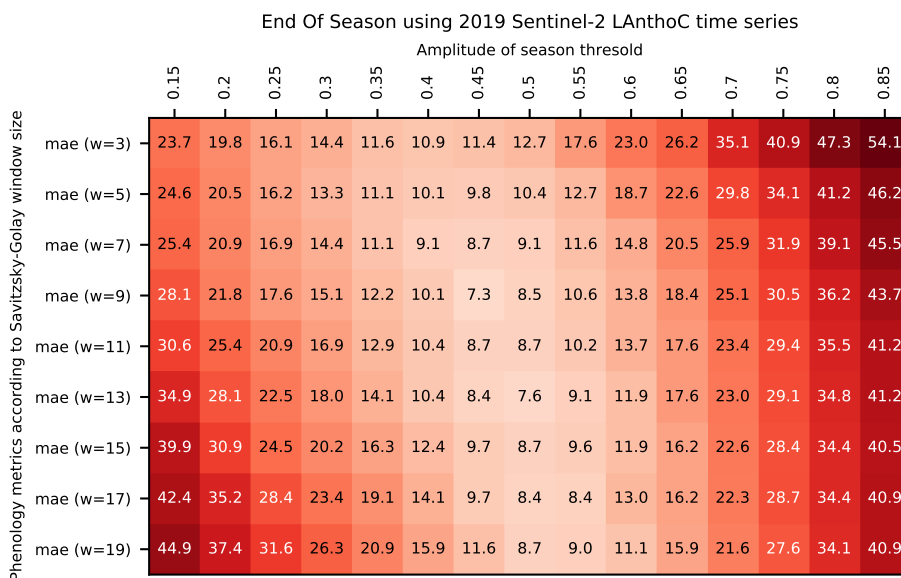
FIGURE C.20 – Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=91) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du NDVI.

C.3.5 Pour la fin de saison (BBCH=95)

C.3.5.1 LAnthoC



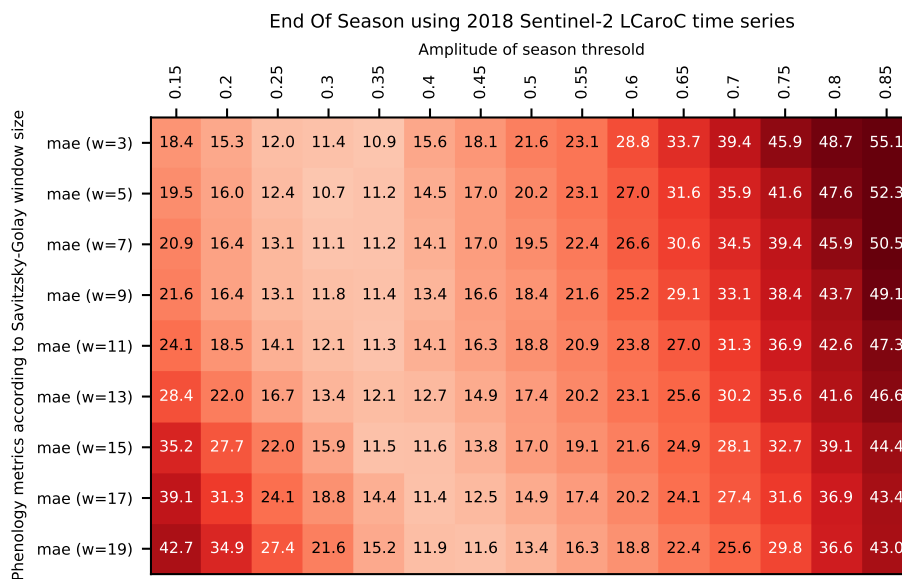
(a) 2018



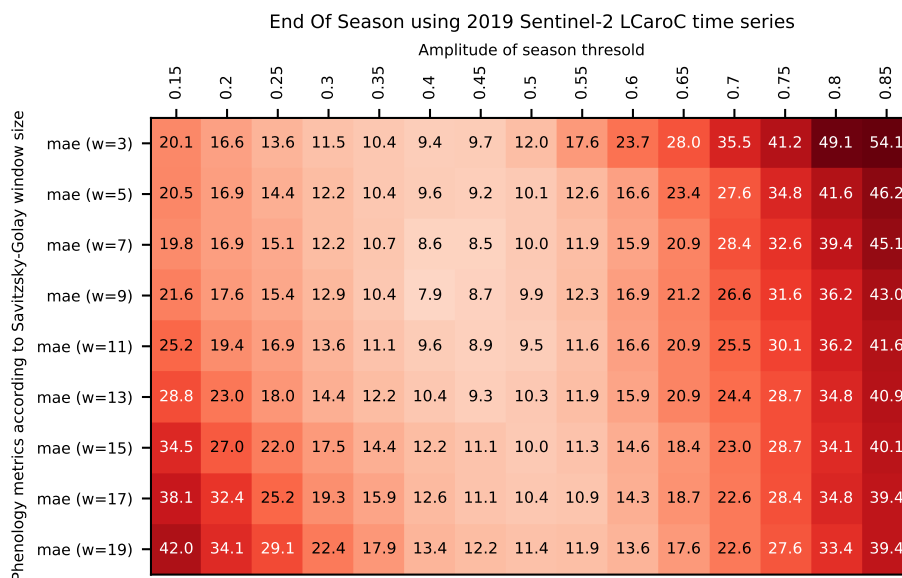
(b) 2019

FIGURE C.21 – Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=95) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LAnthoC.

C.3.5.2 LCaroC



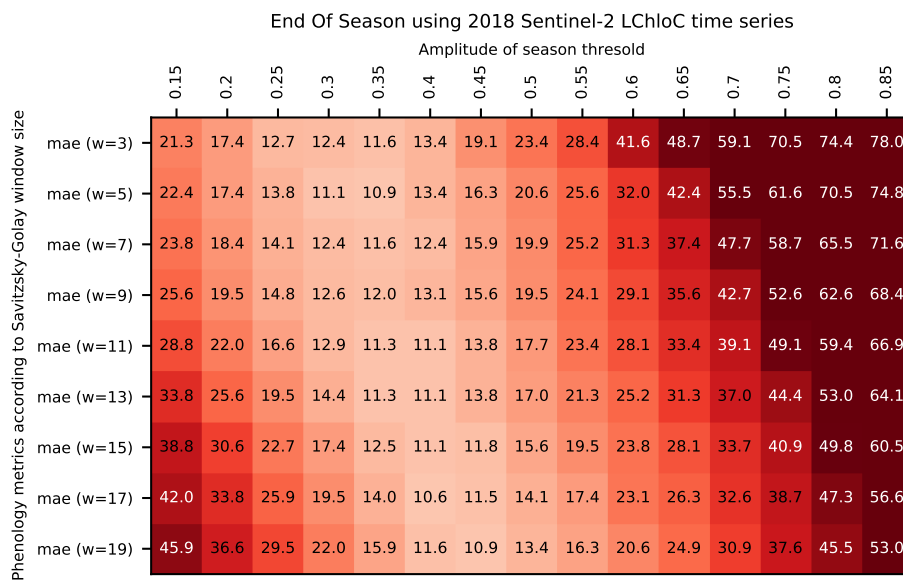
(a) 2018



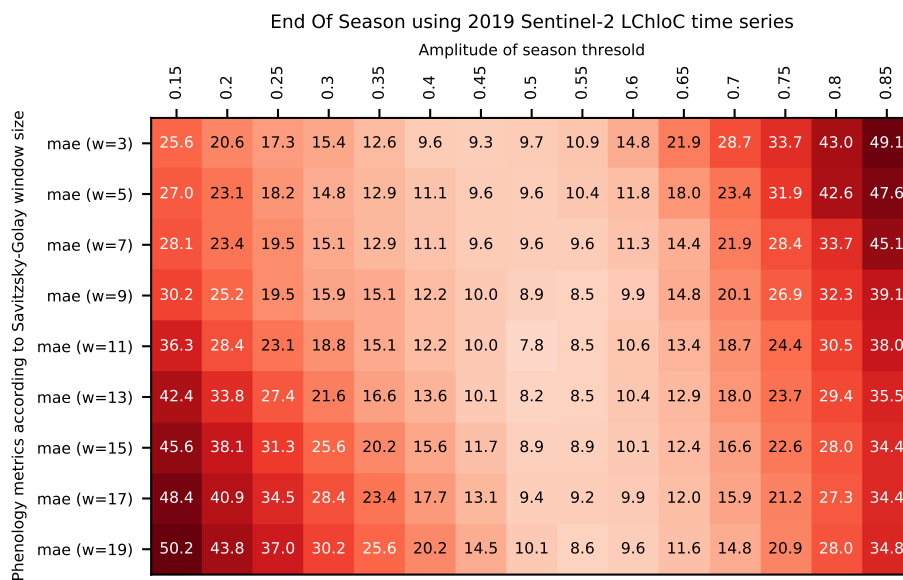
(b) 2019

FIGURE C.22 – Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=95) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LCaroC.

C.3.5.3 LChloC



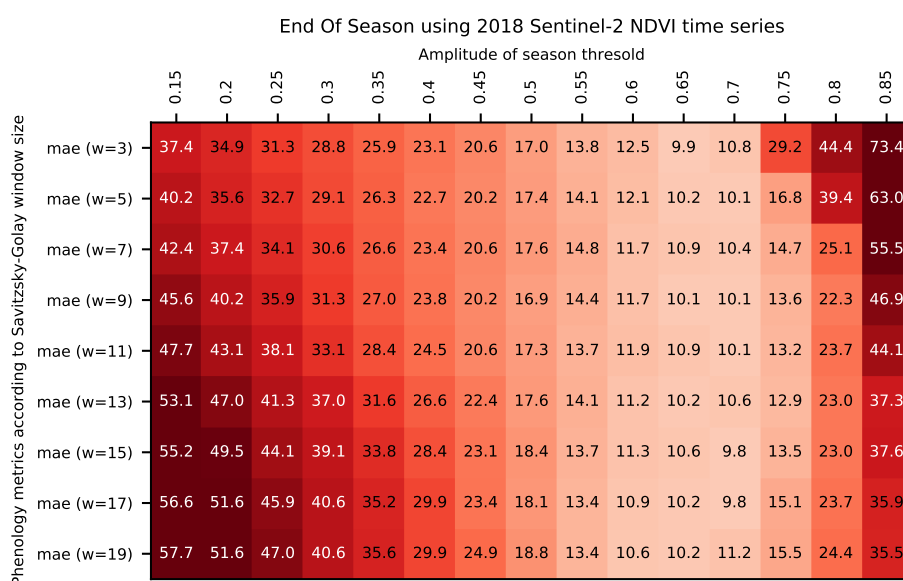
(a) 2018



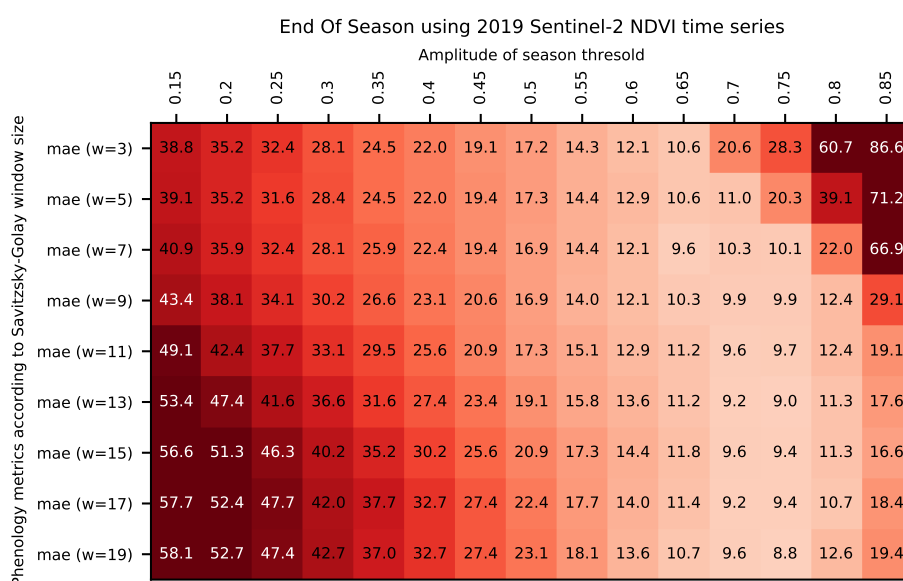
(b) 2019

FIGURE C.23 – Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=95) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LChloC.

C.3.5.4 NDVI



(a) 2018

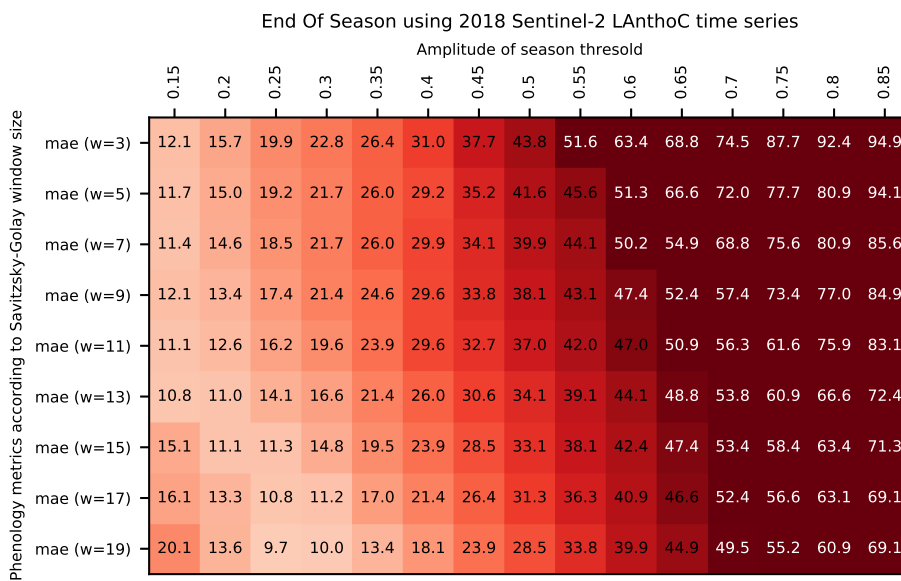


(b) 2019

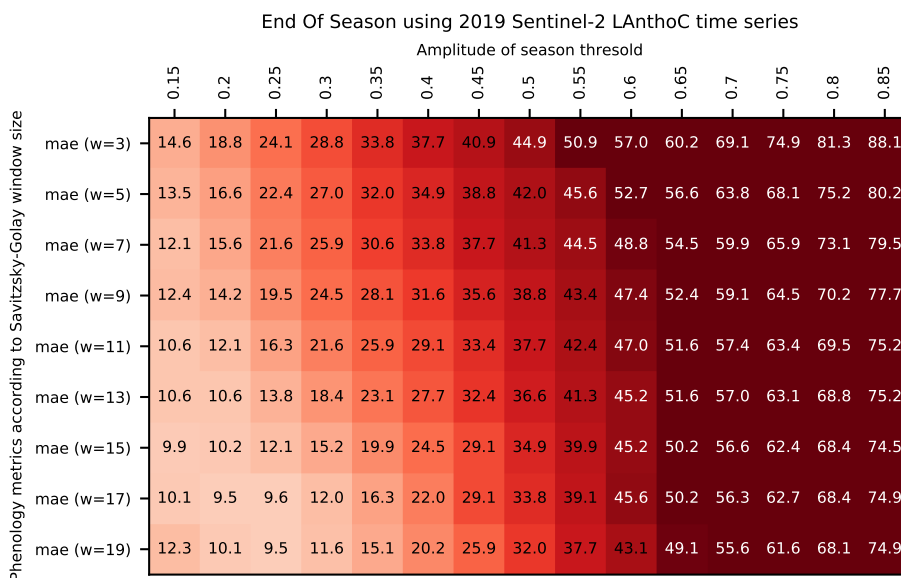
FIGURE C.24 – Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=95) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l’amplitude du NDVI.

C.3.6 Pour la fin de saison (BBCH=99)

C.3.6.1 LAnthoC



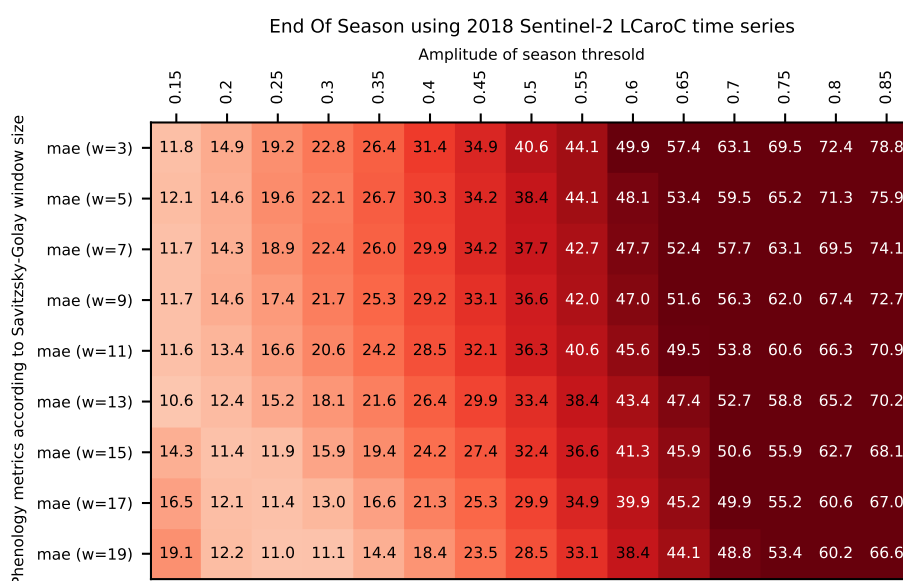
(a) 2018



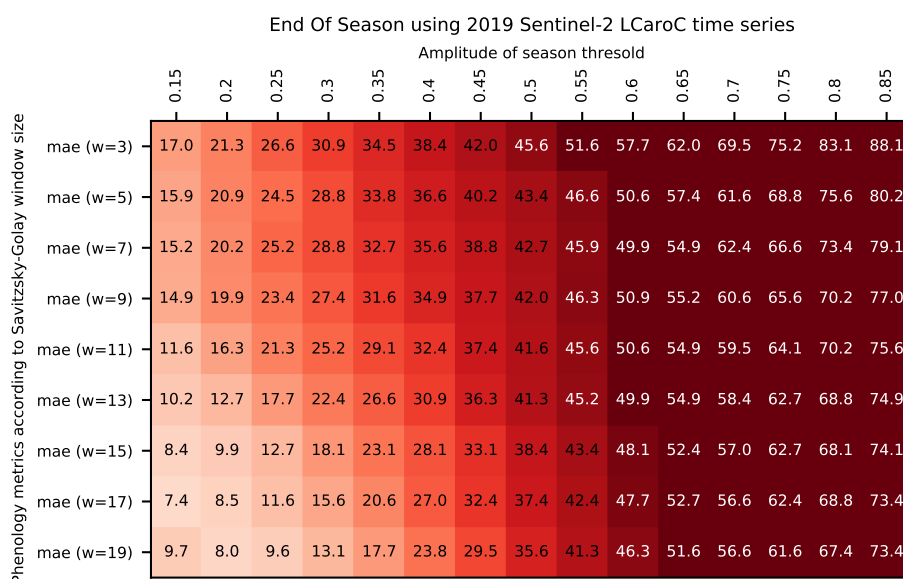
(b) 2019

FIGURE C.25 – Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=99) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LAnthoC.

C.3.6.2 LCaroC



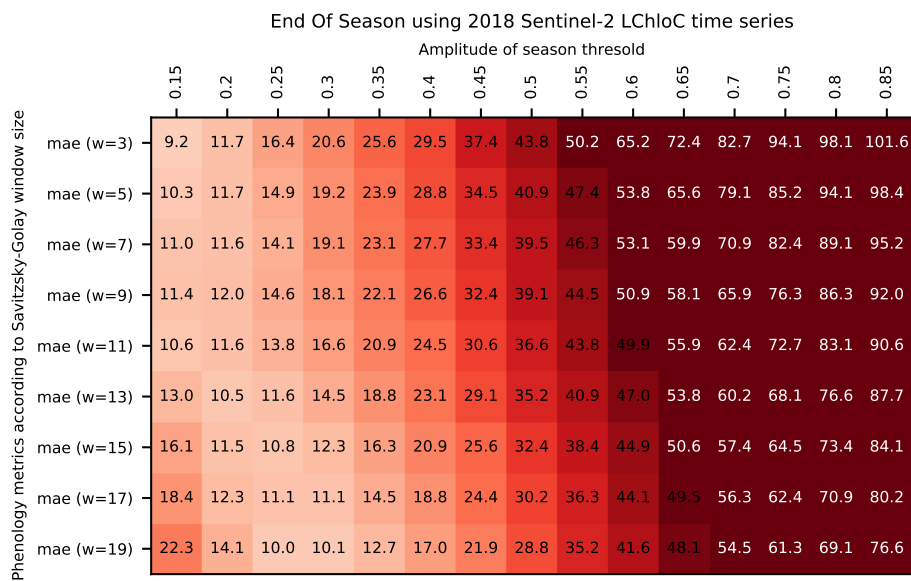
(a) 2018



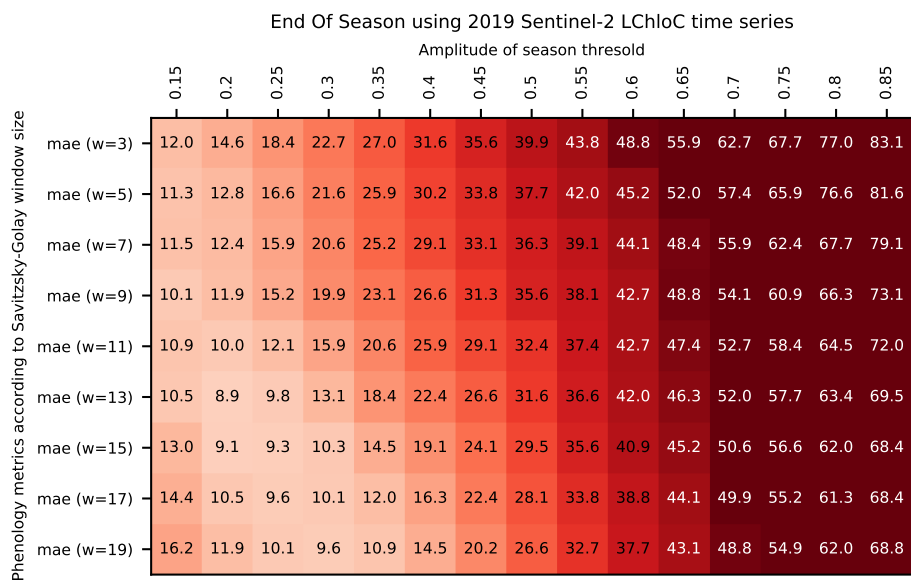
(b) 2019

FIGURE C.26 – Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=99) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l’amplitude du LCaroC.

C.3.6.3 LChloC



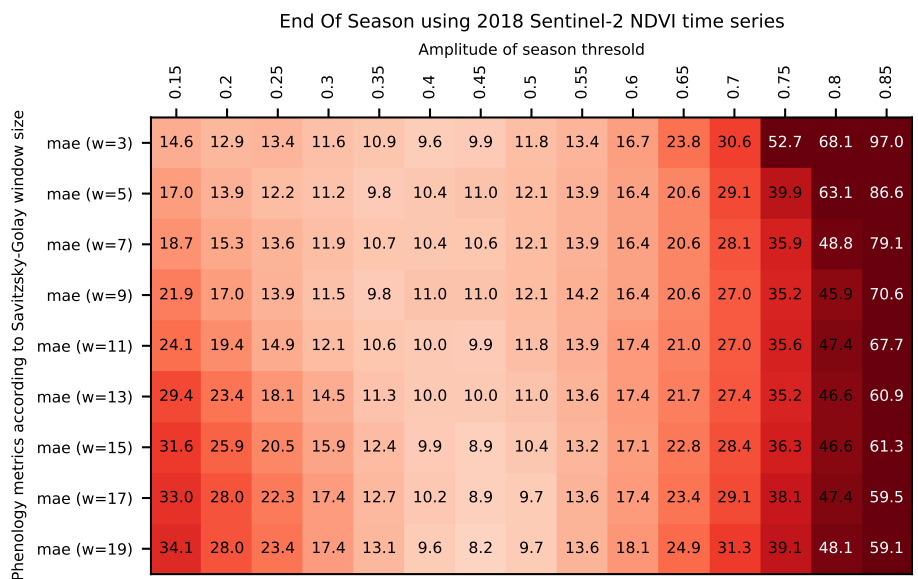
(a) 2018



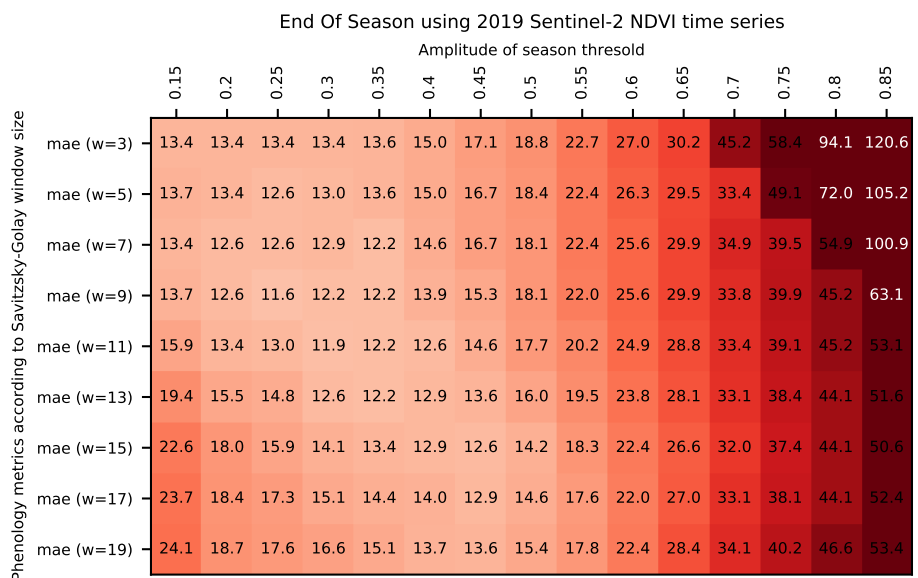
(b) 2019

FIGURE C.27 – Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=99) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du LChloC.

C.3.6.4 NDVI



(a) 2018



(b) 2019

FIGURE C.28 – Erreur moyenne absolue (MAE) pour la fin de saison (BBCH=99) en fonction de la taille du filtre de Savitzky-Golay et du pourcentage de l'amplitude du NDVI.

C.4 Indice spectral avec *Canopy Cover* et BBCH

C.4.1 LChloC vs NDVI

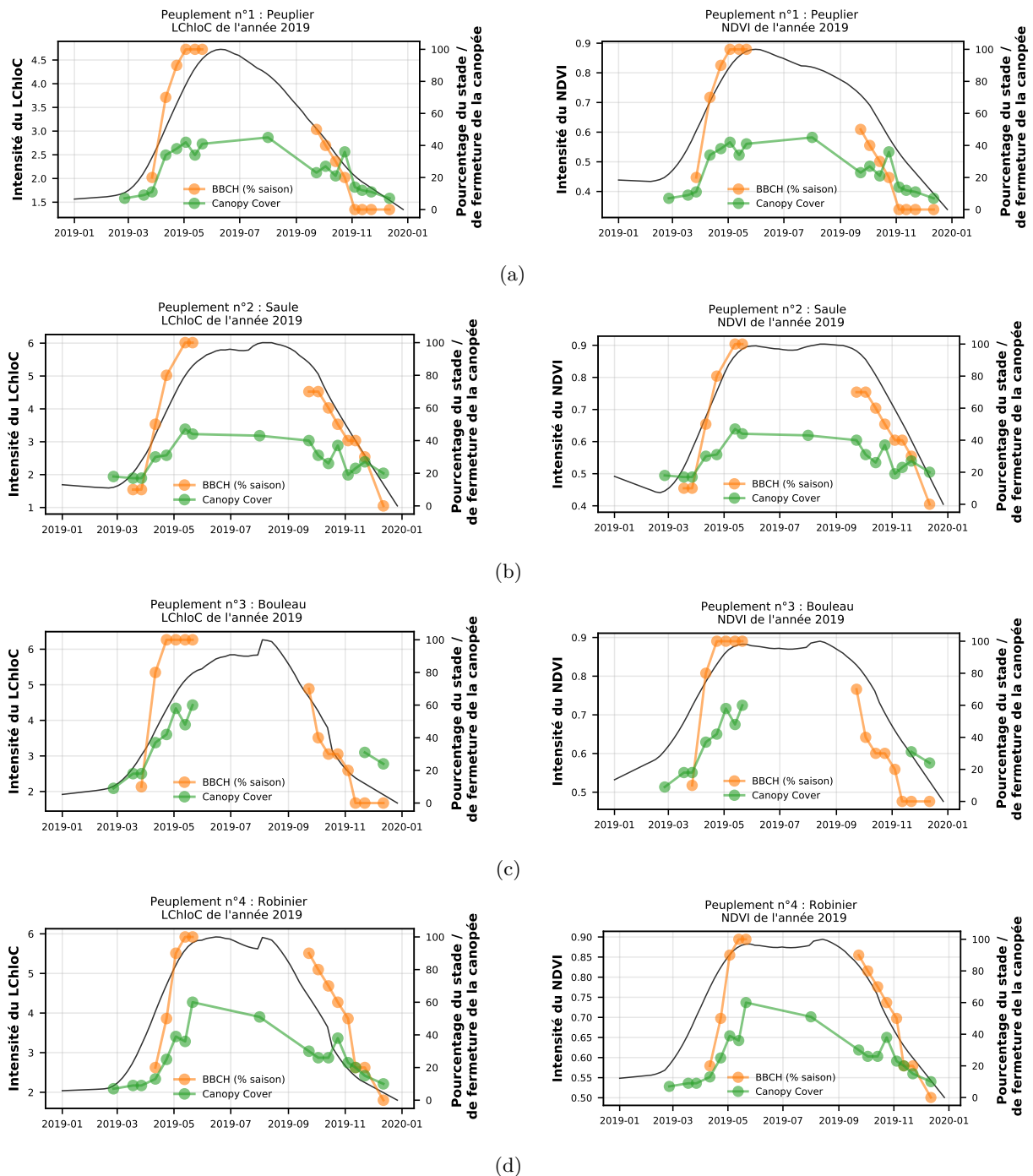


FIGURE C.29 – Différence de signal entre le LChloC et le NDVI en comparaison avec le BBCH et le *Canopy Cover* (1/3).

C.4. INDICE SPECTRAL AVEC CANOPY COVER ET BBCH

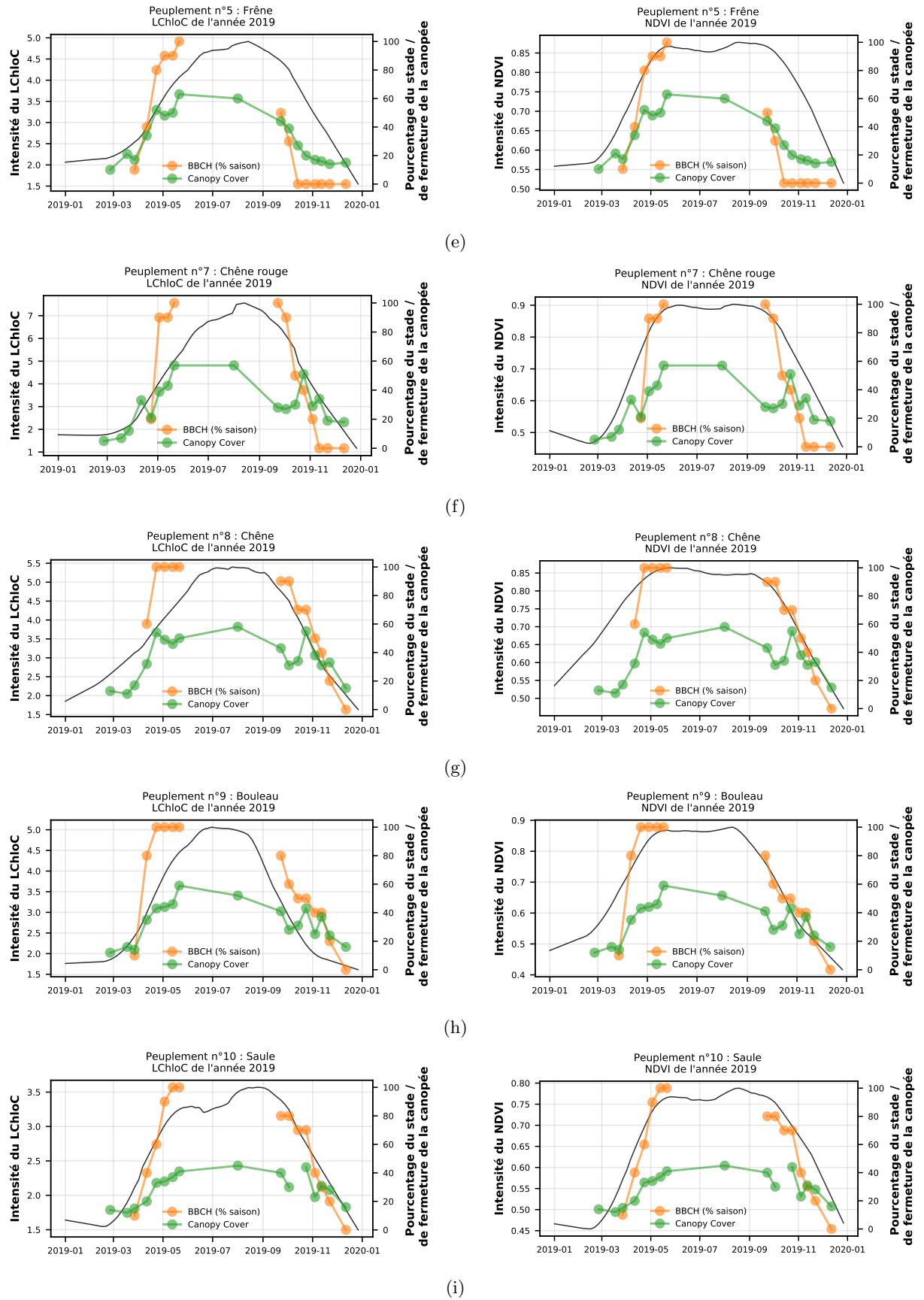
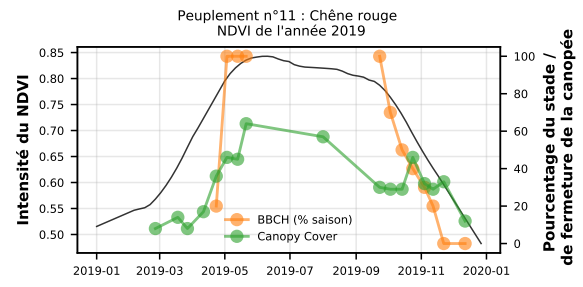
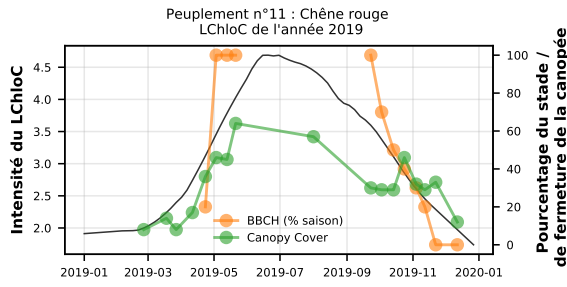
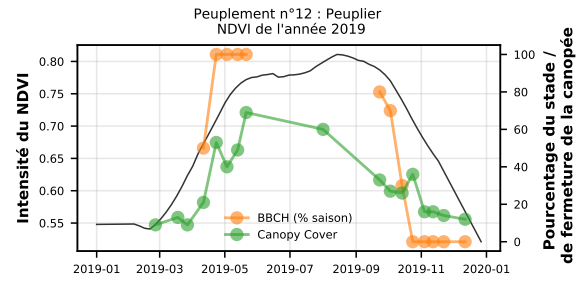
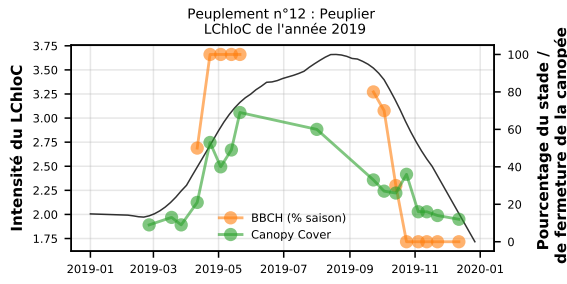


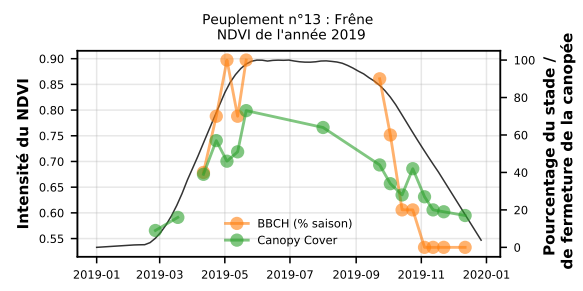
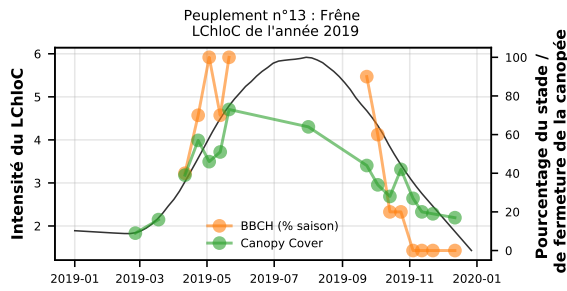
FIGURE C.29 – Différence de signal entre le LChloC et le NDVI en comparaison avec le BBCH et le Canopy Cover (2/3).



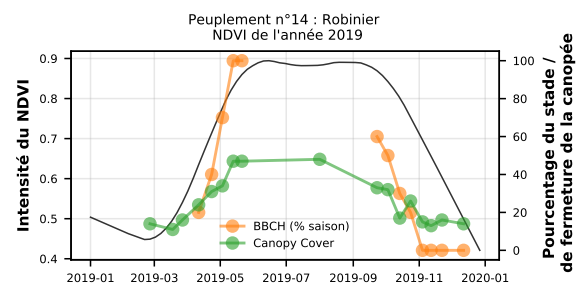
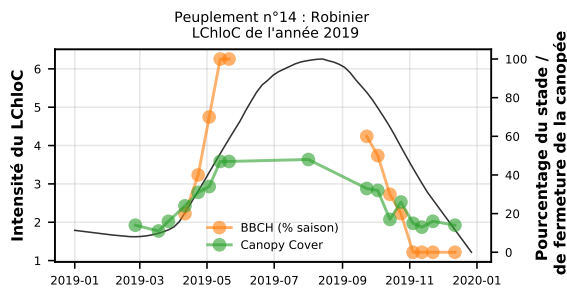
(j)



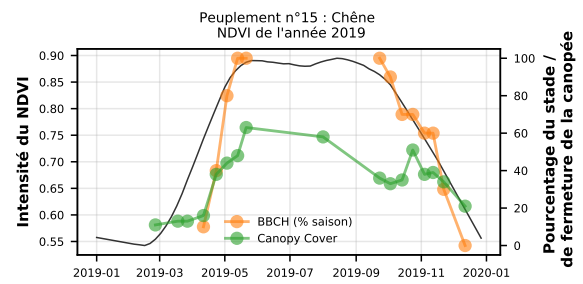
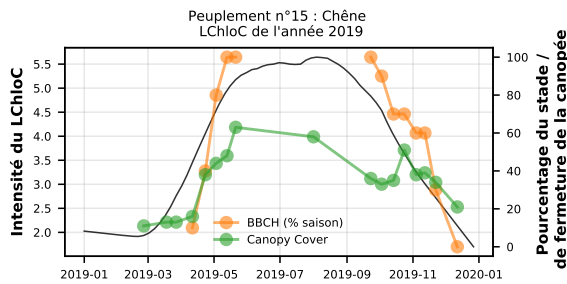
(k)



(l)



(m)



(n)

FIGURE C.29 – Différence de signal entre le LChloC et le NDVI en comparaison avec le BBCH et le *Canopy Cover* (3/3).

C.5 Évolution des spectre des feuillus

C.5.1 En 2017

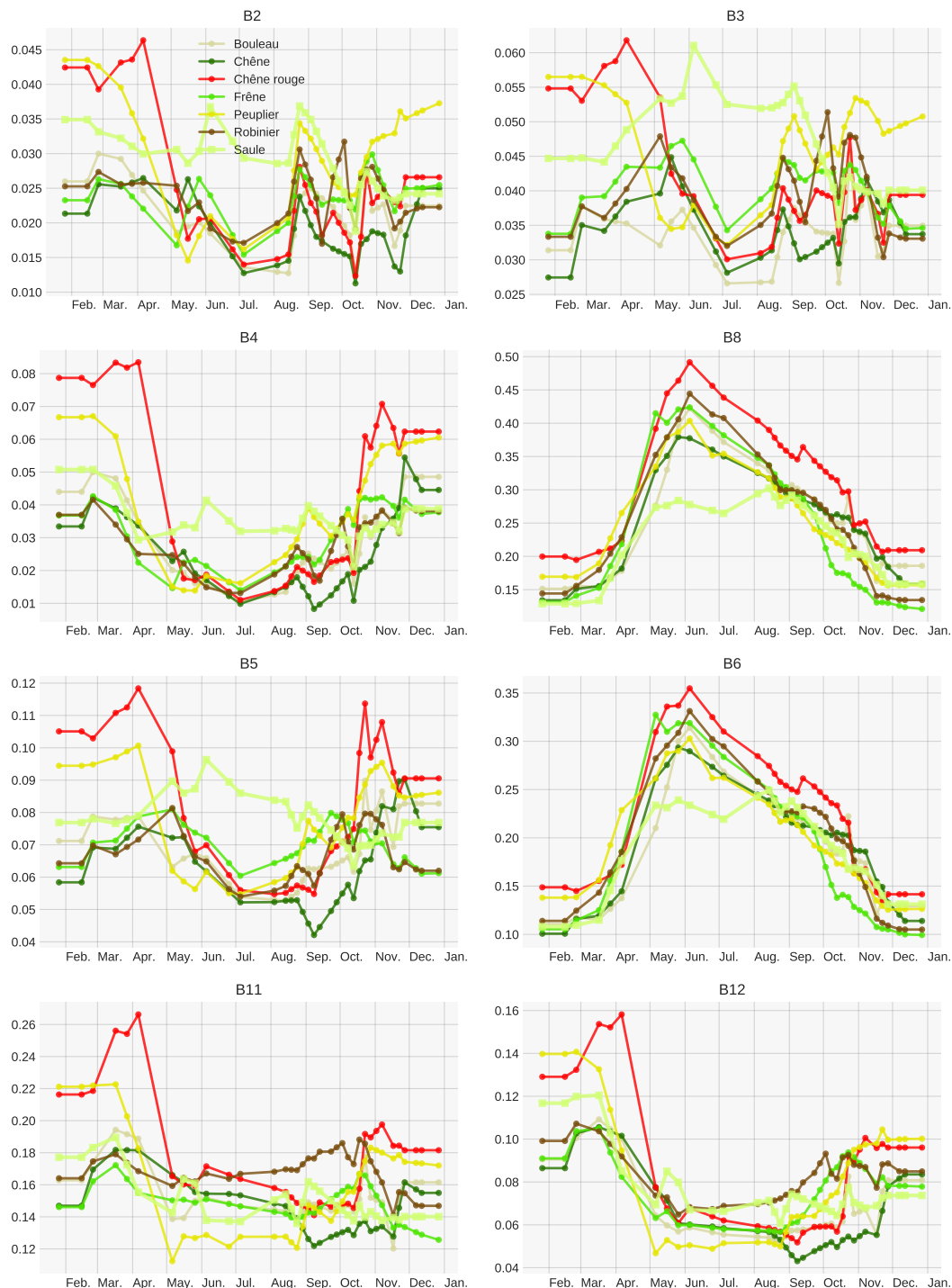


FIGURE C.30 – B2, B3, B4 et B8 sont respectivement la bande bleue, verte, rouge, et proche infra-rouge à 10 m de résolution spatiale. B5 (red-edge, 705nm), B6 (red-edge, 740nm), B11 (1610 nm) et B12 (2190 nm) sont des bandes à 20 m, rééchantillonnées à 10 m pour les besoins de cette étude.

C.5.2 En 2018

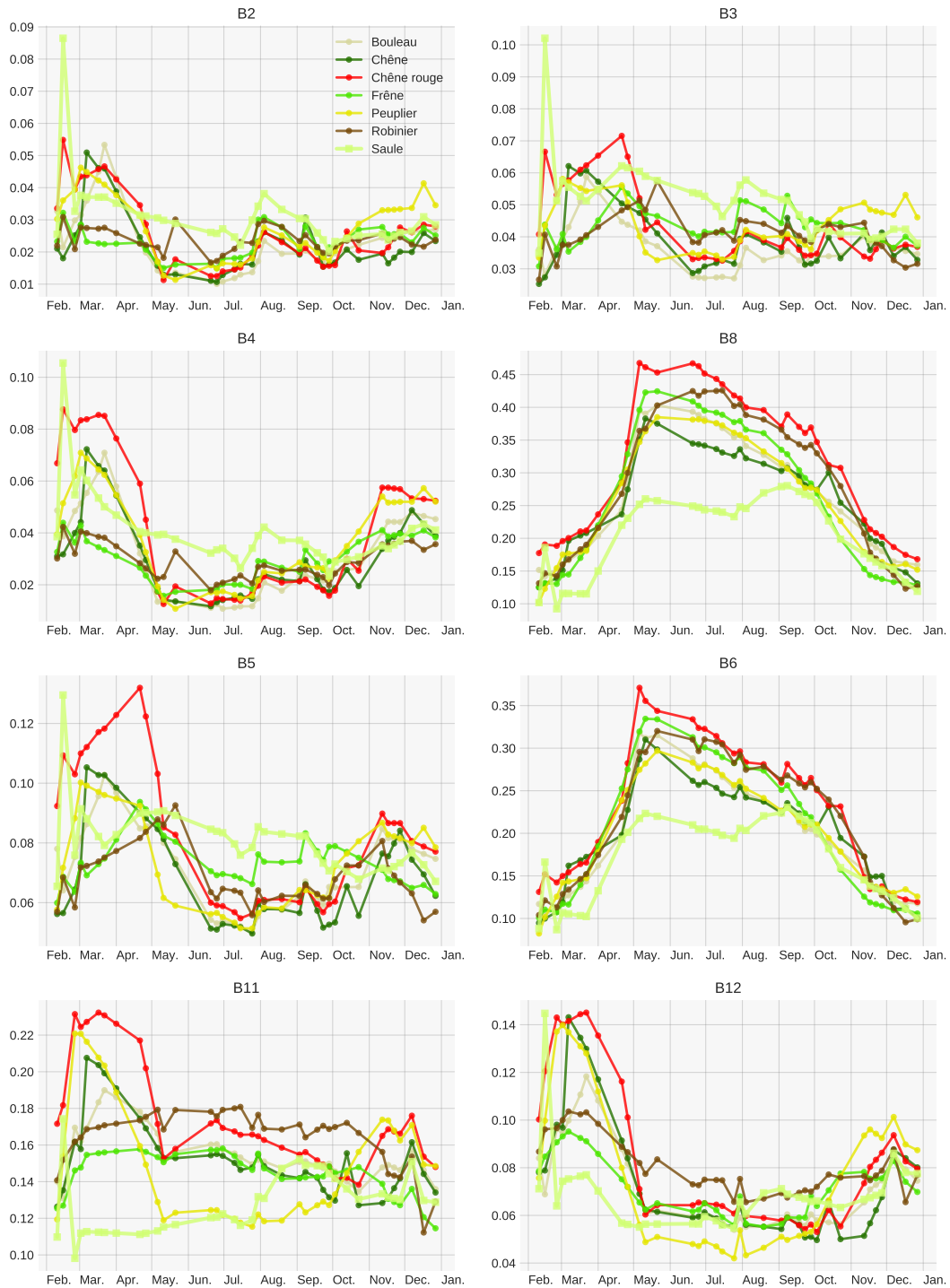


FIGURE C.31 – B2, B3, B4 et B8 sont respectivement la bande bleue, verte, rouge, et proche infra-rouge à 10 m de résolution spatiale. B5 (red-edge, 705nm), B6 (red-edge, 740nm), B11 (1610 nm) et B12 (2190 nm) sont des bandes à 20 m, rééchantillonnées à 10 m pour les besoins de cette étude.

C.5.3 En 2019

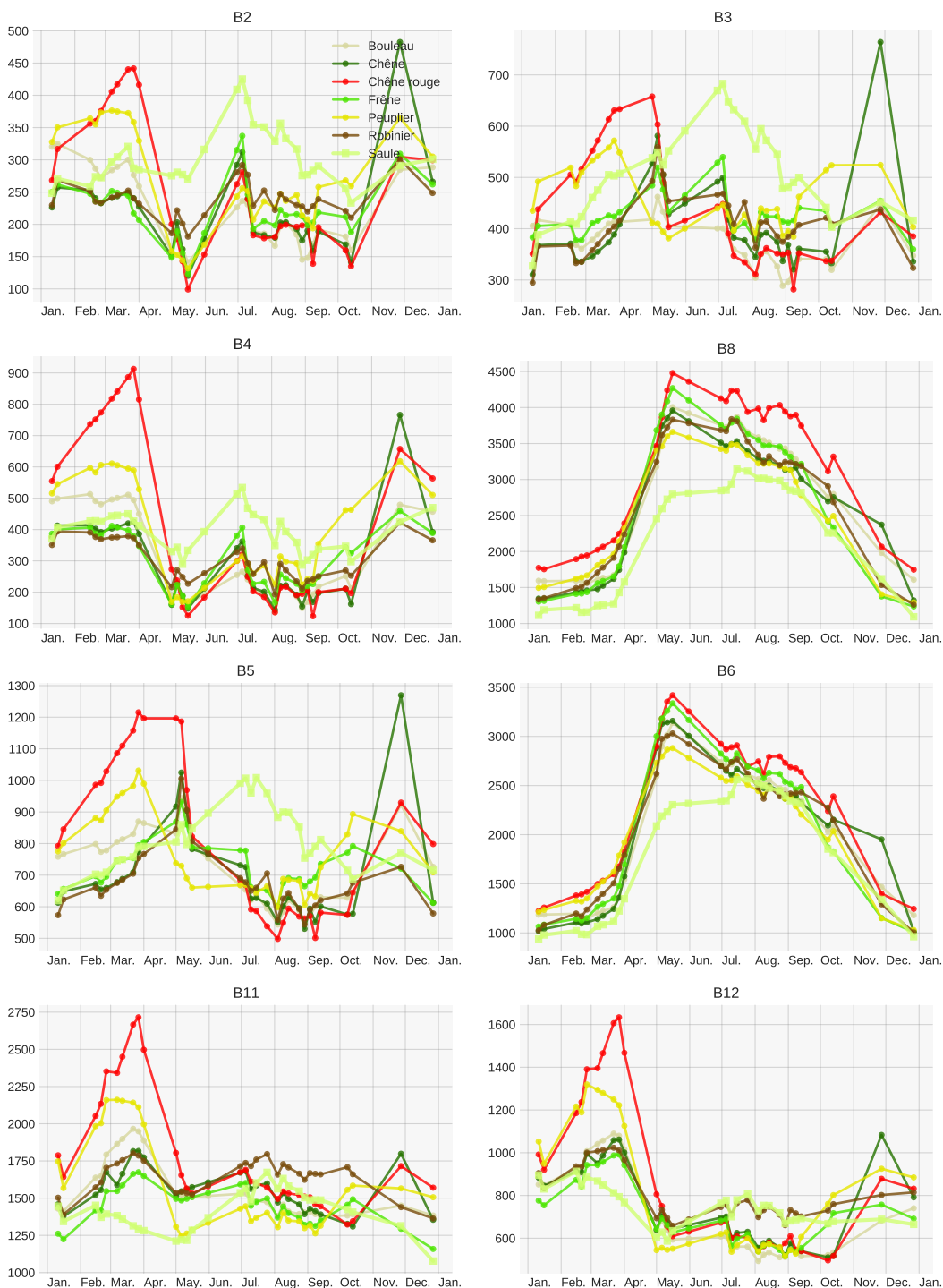


FIGURE C.32 – B2, B3, B4 et B8 sont respectivement la bande bleue, verte, rouge, et proche infra-rouge à 10 m de résolution spatiale. B5 (red-edge, 705nm), B6 (red-edge, 740nm), B11 (1610 nm) et B12 (2190 nm) sont des bandes à 20 m, rééchantillonnées à 10 m pour les besoins de cette étude.

C.6 Observations phénologiques

C.6.1 Lien entre BBCH et ouverture de la canopée

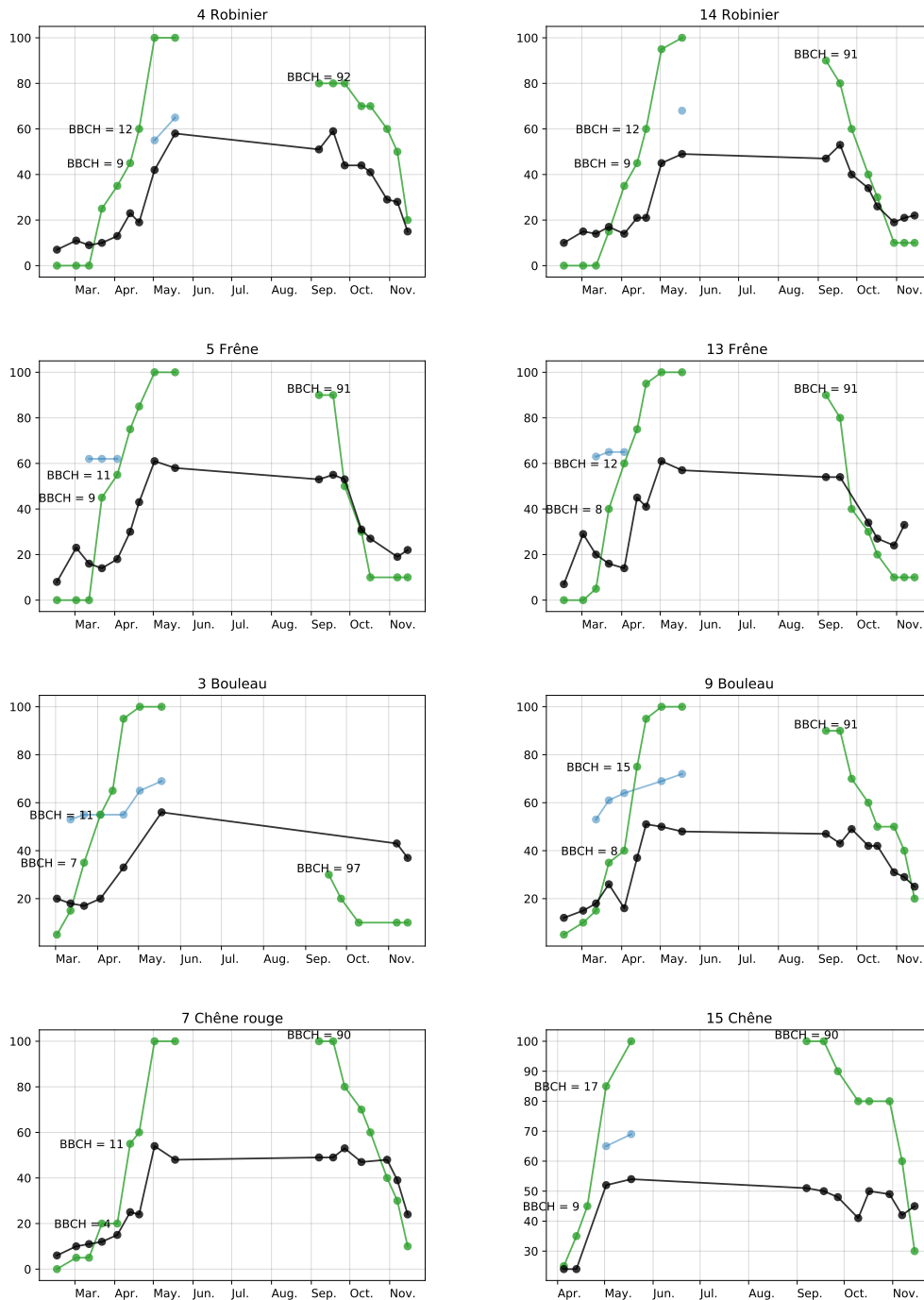


FIGURE C.33 – Évolution de l'ouverture de la canopée (indice Canopy Cover en noir) et du code BBCH (en vert). En bleu, période où a été observée la floraison.

Annexe D

Liste des publications

D.1 Journal international à comité de lecture

Karasiak N., Dejoux, J.-F., Monteil, C., and Sheeren, D. (Accepté en révision mineure au 13 août 2020). Spatial dependence between training and test sets : another pitfall of classification accuracy assessment in remote sensing. *Machine Learning*

Karasiak, N. (2020). Museo ToolBox : A Python library for remote sensing including a new way to handle rasters. *Journal of Open Source Software*, 5(48) :1978

Karasiak, N., Dejoux, J.-F., Fauvel, M., Willm, J., Monteil, C., and Sheeren, D. (2019). Statistical Stability and Spatial Instability in Mapping Forest Tree Species by Comparing 9 Years of Satellite Image Time Series. *Remote Sensing*, 11(21)

D.2 Conférence internationale à comité de lecture

Karasiak, N., Fauvel, M., Dejoux, J.-F., Monteil, C., and Sheeren, D. (2020). Optimal Dates For Deciduous Tree Species Mapping Using Full Years Sentinel-2 Time Series In South West France. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, V-3-2020 :469–476, Nice, France

Karasiak, N., Sheeren, D., Dejoux, J.-F., Feret, J.-B., Willm, J., and Monteil, C. (2018). Is ground-based phenology of deciduous tree species consistent with the temporal pattern observed from sentinel-2 time series? In *ForestSAT 2018*, page 192 p., Maryland, United States

Karasiak, N., Sheeren, D., Fauvel, M., Willm, J., Dejoux, J.-F., and Monteil, C. (2017). Mapping tree species of forests in southwest France using Sentinel-2 image time series. In *2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp)*, pages 1–4, Brugge, Belgium

D.3 Conférence nationale à comité de lecture

Karasiak, N., Sheeren, D., Dejoux, J.-F., and Monteil, C. (2019). Estimer l'apport de la phénologie dans la classification des essences forestières. In *SAGEO*, Clermont-Ferrand, France

Karasiak, N., Sheeren, D., Dejoux, J.-F., and Monteil, C. (2019). Impact de l'autocorrélation spatiale sur la qualité des modèles d'apprentissage automatique : étude dans le cadre de la classification d'essences forestières à partir de données satellitaires. In *ThéoQuant*, Besançon, France

Karasiak, N., Sheeren, D., Dejoux, J.-F., and Monteil, C. (2018). Monitoring and understanding the green-leaf phenology of tree species with Sentinel-2. In *THEIA Workshop for Sentinel-2 L2A MAJA products*, page 46 p., Toulouse, France

D.4 Chapitre d'ouvrage

Karasiak, N. and Perbet, P. (2018). Remote Sensing of Distinctive Vegetation in Guiana Amazonian Park. In Baghdadi, N., Mallet, C., and Zribi, M., editors, *QGIS and Applications in Agriculture and Forest*, pages 215–245. John Wiley & Sons, Inc

D.5 Revue de vulgarisation

Karasiak, N. (2019). La phénologie des arbres : une donnée incontournable pour cartographier nos forêts? *Forêt entreprise*, (247) :20–24

Karasiak N. (2019). Mapping deforestation with Sentinel Hub. <https://medium.com/sentinel-hub/mapping-deforestation-from-sentinel-hub-de6aae67f817>

D.6 Divers

Trois prix gagnés sur les trois participations au concours de télédétection organisée par Snergise (Sentinel-Hub, <https://www.sentinel-hub.com/contest/>) dans les catégories suivantes :

- Neige et glaciers (1er prix)
- Agriculture et forêt (2ème prix)
- Autres (2ème prix)