



RAPPORT UE 1002 - 703 ATELIER GÉOMATIQUE

ANALYSE DE L'INCERTITUDE DE
CLASSIFICATION DE
PEUPLERAIES PLANTÉES

Réalisé par :

Abir BEN ABDELGHAFFAR

Lucas LIMA

Naly RAKOTOARINDRAZAKA

Table de Matières

1. Introduction	5
2. Matériel et méthodes.....	5
2.1. Zone d' étude	6
2.2. Présentation des données.....	7
2.2.1. Rasters de confidence	7
2.2.2. Raster des métriques LiDAR	8
2.2.3. Données vectorielles sur les parcelles cultivées	8
2.3. Traitements des données	9
2.3.1. Prétraitement des données vectorielles sur les parcelles cultivées	12
2.3.2. Découpe des données rasters	13
2.3.3. Extraction des valeurs et jointure spatial	14
2.3.4. Filtrage des données du département 10 et des données dont la date d' acquisition est différente de celle des données LiDAR	15
2.3.5. Création des boîtes à moustaches avec encoche	16
2.3.6. Création de l' application avec Dash et Plotly	18
3. Résultats.....	20
3.1. Variation du nombre des données suivant les différents filtres.....	20
3.2. Probabilité d' appartenance des pixels à la classe peuplier en fonction de l' âge de la plantation, tous cultivars confondus	22
3.2.1. Probabilité d' appartenance des pixels à la classe peuplier en fonction de l' âge de la plantation, par type de cultivars.....	22
3.2.1. Probabilité d' appartenance par cultivar sans différenciation des dates d'acquisition de données Sentinel	23
3.2.2. Probabilité d' appartenance des pixels à la classe peuplier en fonction de l' âge de la plantation, par cultivar et par années d'acquisition de Sentinel.....	24
3.3. Probabilité d' appartenance des pixels à la classe peuplier, densité de végétation (PAI) et l' ouverture de la canopée (Canopy cover) en fonction de l' âge	26
3.3.1. Visualisation Globale : tous cultivars confondus.....	27
3.3.2. Visualisation détaillée: par type de Cultivar	28
4. Discussion.....	33
4.1. Diminution du nombre de données exploitables	33
4.2. Variation de la probabilité d' appartenance en fonction de l' âge tous cultivars confondus	33
4.3. Variation des valeurs du Canopy Cover (CC) et du Plant Area Index (PAI) en fonction de l' âge tous cultivars confondus	34
4.4. Âge minimum pour une meilleure classification, variable selon le cultivar	35

4.5.	Incohérences de la valeur de probabilité d'appartenance chez les plantations jeunes et âgées.....	37
4.6.	Corrélation entre les variables métriques CC, PAI et la probabilité d' appartenance	40
4.7.	Limites de l' étude	40
5.	Conclusion.....	42
6.	Annexes	43
6.1.	Liens d'accès vers les fichiers PDF contenant tous les graphiques	43
6.2.	Liens d'accès vers les scripts	44
6.3.	Structure finale de l'arborescence du projet	44
7.	Références bibliographiques	45

Liste de Figures

Figure 1 : Carte avec les emprises des parties de quatre départements abordées.....	7
Figure 2 : Diagramme de la chaîne de traitement pour la génération du résultat final	11
Figure 3 : Capture d'écran montrant l'incohérence entre la date de plantation renseignée dans la table attributaire et ce qui peut être constaté par photo-interprétation sur la BD Ortho de 2021 16	
Figure 4 : Capture d'écran de l'application créée pour explorer les graphiques de manière interactive	19
Figure 5 : Capture d'écran de l'application créée pour comparer deux graphiques de manière interactive	19
Figure 6 : Probabilité d'appartenance en fonction de l'âge de plantation tous cultivars confondus.....	22
Figure 7 : Probabilité d'appartenance en fonction de l'âge de plantation pour le cultivar Koster.....	23
Figure 8 : Probabilité d'appartenance en fonction de l'âge pour le cultivar Polargo	24
Figure 9 : Probabilité d'appartenance en fonction de l'âge pour le cultivar I214	24
Figure 10 : Probabilité d'appartenance en fonction de l'âge entre les années de 2017 à 2020 pour le cultivar Koster	25
Figure 11 : Probabilité d'appartenance en fonction de l'âge entre les années de 2017 à 2020 pour le cultivar Polargo	26
Figure 12 : Probabilité d'appartenance en fonction de l'âge entre les années de 2017 à 2020 pour le cultivar I214	26
Figure 13 : Canopy cover en fonction de l'âge tous cultivars confondus	27
Figure 14 : Densité de végétation (PAI) en fonction de l'âge tous cultivars confondus.....	28
Figure 15 : Canopy cover et probabilité d'appartenance en fonction de l'âge de plantation pour le cultivar Koster	29
Figure 16 : Densité de végétation (PAI) et probabilité d'appartenance en fonction de l'âge de plantation pour le cultivar Koster	30
Figure 17 : Canopy cover et probabilité d'appartenance en fonction de l'âge de plantation pour le cultivar I45/51	31
Figure 18 : Densité de végétation (PAI) et probabilité d'appartenance en fonction de l'âge de plantation pour le cultivar I45/51.....	32
Figure 19 : Canopy cover et Plant Area Index (PAI) en fonction de l'âge pour le cultivar I214	35
Figure 20 : Pixels à faibles valeurs de probabilité de classification à l'âge de 8 ans pour le cultivar I214.....	38
Figure 21 : Capture d'écran de l'affichage sur l'application de pixels à faibles valeurs de probabilité de classification pour le cultivar I214 à l'âge de 8 ans	38
Figure 22 : Capture d'écran de la parcelle dep47-111 avec des faibles valeurs de probabilité d'appartenance à l'âge de 7ans, Ortho Photo 2017	39

Liste de Tableaux

Tableau 1 : Variation de nombre de parcelles par département tout au long de chaîne de traitement	20
Tableau 2 : Quantité de données en pixel tout au long de chaîne de traitement	20
Tableau 3 : Nombre de cultivars et nombre de pixels restants après l'application de la fonction top_cultivars pour le traçage du graphique de probabilité d'appartenance en fonction de l'âge de la plantation	21
Tableau 4 : Nombre de cultivars et nombre de pixels restants après l'application de la fonction top_cultivars pour le traçage du graphique des métriques LiDAR	21

1. Introduction

Les peupliers jouent un rôle central en raison de leur importance économique, étant des arbres à croissance rapide répondant à la demande mondiale croissante de produits en bois. Le laboratoire DYNAFOR, dans le cadre de la thèse de Y. Hamrouni, a développé un indice spectral nommé Poplar Detection Index (PI), décliné en quatre variantes dédiées à la détection des peupleraies. Cet indice a permis la production de cartes des peupleraies françaises pour les années 2017 à 2022.

Ces cartes, élaborées pour répondre aux besoins de la filière peuplier en France, sont principalement destinées aux partenaires forestiers du CNP et du CNPF. Cependant, la faible densité d'arbres dans les jeunes plantations, les effets du sol et du sous-étage herbacé perturbent les résultats de la classification.

Ainsi, ce projet a pour objectif principal d'étudier l'évolution de la probabilité d'appartenance des pixels à la classe peuplier, notamment dans les jeunes plantations. Pour cela, les objectifs spécifiques suivants ont été mis en place

- Étudier l'évolution de la probabilité d'appartenance des pixels à la classe peuplier en fonction de l'âge de la plantation, tous cultivars de peupliers confondus
- Étudier l'évolution de la probabilité d'appartenance des pixels à la classe peuplier en fonction de l'âge de la plantation par type de cultivars.
- Étudier l'évolution de la probabilité d'appartenance des pixels à la classe peuplier en fonction de la densité de végétation (Plant Area Index) et l'ouverture de la canopée (Canopy cover) estimées à partir de données LiDAR HD de l'IGN.

2. Matériel et méthodes

Le développement de ce projet s'est appuyé sur trois ensembles de jeux de données :

- Des rasters de certitude contenant la probabilité d'appartenance issue de la classification réalisée avec l'indice PI pour six années consécutives, de 2017 à 2022
- Des rasters de cinq métriques dérivées d'un relevé LiDAR réalisé entre 2021 et 2023, selon la région

- Des données vectorielles des parcelles des zones d'étude, contenant des informations sur les cultivars de peupliers plantés, les dates de plantation, les dates des relevés LiDAR, etc

Ces données ont été prétraitées et combinées pour produire différents diagrammes en boîte à moustaches (box plots), qui constituent les résultats finaux attendus pour les analyses statistiques.

2.1. Zone d'étude

Les données concernent cinq tuiles du Sentinel-2 couvrant des parties de quatre départements : Lot-et-Garonne (47), Tarn-et-Garonne (82), Savoie (73) et Aube (10).

Les parcelles situées dans le sud-ouest de la France (47 et 82) sont exposées à un stress hydrique plus important par rapport à celles du nord (10) et du sud-est (73). Cela implique un besoin accru de nettoyage au pied des jeunes peupliers pour limiter la compétition inter-spécifique, ce qui amplifie l'effet du sol lors des observations par télédétection, notamment en augmentant la réflectance du sol et en réduisant la distinction entre les classes.

En revanche, dans les plantations situées plus au nord (notamment dans le département 10), ce nettoyage n'est pas pratiqué avec la même intensité, entraînant la présence d'un sous-étage herbacé important. Ce sous-étage interfère principalement avec les indices spectraux utilisés pour la détection des peupliers et peut entraîner des confusions dans l'identification des classes, surtout dans les années suivant la plantation.

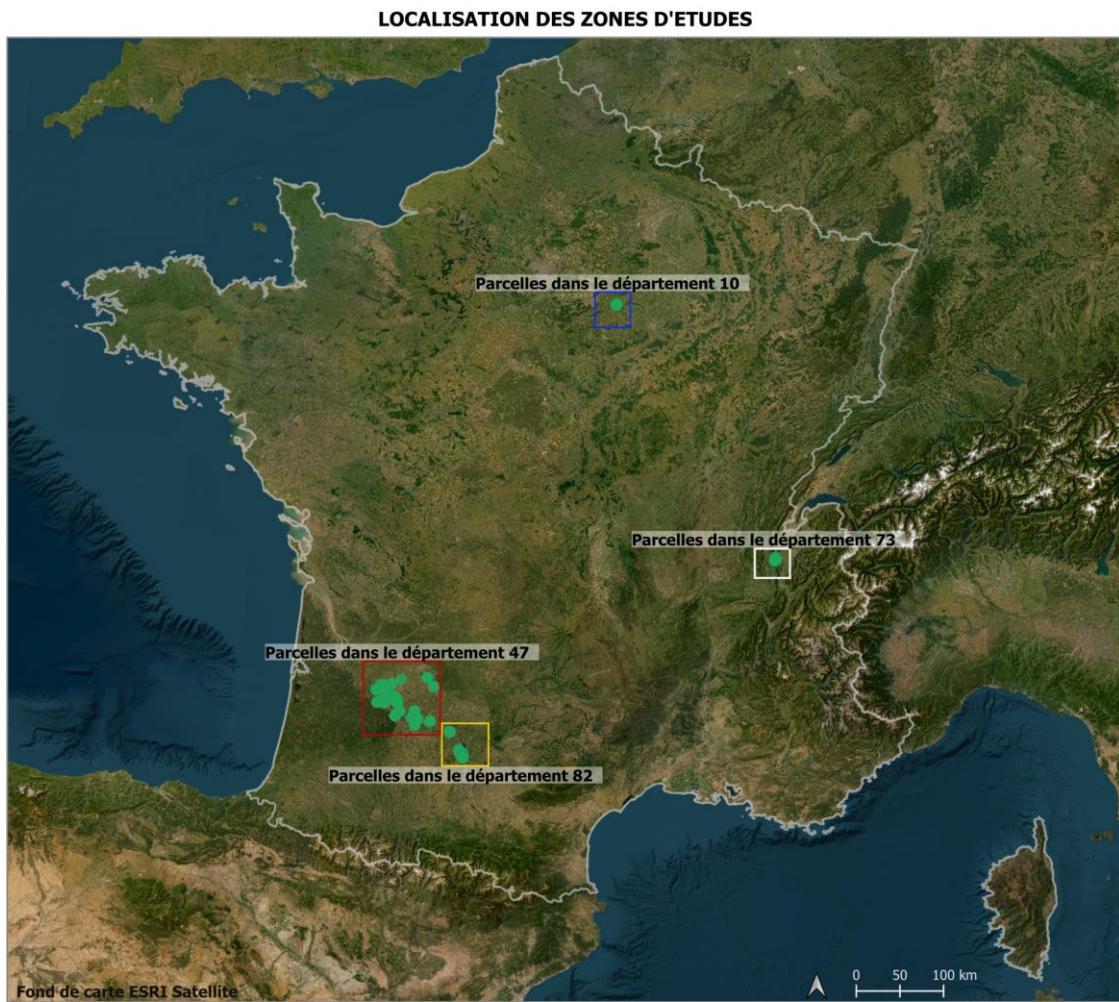


Figure 1 : Carte avec les emprises des parties de quatre départements abordées

2.2. Présentation des données

2.2.1. Rasters de confiance

Les rasters de certitude ont été générés à partir des cinq tuiles des données Sentinel-2. Les pixels ont des valeurs allant de 0 à 100 et représentent la probabilité d'appartenance à la classe « peuplier », calculée à partir de l'indice PI. Les cinq tuiles ont été générées séparément pour chaque année de l'étude (2017 à 2022), totalisant 30 rasters. Elles sont toutes projetées dans le système de coordonnées Lambert-93 (EPSG : 2154) avec une résolution spatiale de 10 mètres.

Pour les peupliers matures, l'indice PI présente un niveau de certitude assez élevé. C'est pourquoi notre étude s'est concentrée sur les 12 premières années de plantation, une période clé marquée par la phase d'installation des peupleraies et leur croissance active.

La distinction des jeunes peupliers par les algorithmes de classification est particulièrement difficile en raison de leur petite taille. Les jeunes arbres peuvent partager des caractéristiques spectrales similaires à celles d'autres types de végétation, et leur faible couvert végétal est davantage influencé par le sous-étage herbacé et le sol, ce qui augmente le risque d'erreurs de classification.

2.2.2. Raster des métriques LiDAR

Les rasters des métriques LiDAR ont également une résolution de 10 mètres, mais sont projetés en WGS 84 UTM Zone 31 (EPSG : 32631). Ils couvrent l'ensemble des départements étudiés, à l'exception du département 10 où les données LiDAR sont indisponibles. Cependant, toutes les parcelles ne disposent pas des métriques calculées. La date d'acquisition de ces données peut varier de 2021 à 2023.

Ces données ont été utilisées pour évaluer l'évolution des caractéristiques LiDAR en fonction de l'âge des peupliers. Les indicateurs sélectionnés sont pertinents pour décrire la dynamique structurelle des peupliers et leur croissance au fil des années. Les métriques abordées sont les suivantes :

Canopy cover : Cette métrique reflète l'expansion et l'ouverture du canopée, directement liée à la croissance des plantes, et donc à leur âge. On s'attend à une augmentation progressive de cette métrique jusqu'à atteindre un certain seuil, reflétant la maturité du peuplier. La couverture de canopée est un déterminant majeur de la réflectance des forêts mesurée par télédétection optique ([Dawson et al. 1999](#)).

PAI (Plant Area Index) : Cet indicateur mesure la surface foliaire totale et les éléments végétatifs par unité de surface. Il est étroitement corrélé au développement structurel des peupliers. Le PAI fournit également des informations importantes sur la reprise de la végétation et la détection du changement dans les zones perturbées, ce qui est directement lié à l'évaluation de la dynamique de parcelle en fonction du temps ([Siriruk Pimmasarn et al. 2020](#)).

Ces métriques spécifiques, nous cherchons à établir des corrélations entre les caractéristiques structurelles des peupliers et les performances de l'algorithme de classification, en tenant compte de leur évolution avec l'âge.

2.2.3. Données vectorielles sur les parcelles cultivées

Les données des parcelles sont projetées dans le système de coordonnées Lambert-93 (EPSG : 2154) et sont disponibles pour chacun des quatre départements étudiés. Cependant,

les parcelles situées dans le département 82 chevauchent deux tuiles et ont été divisées en deux zones distinctes : Saint-Pierre et Borde Basse.

Les données vectorielles incluent une table attributaire contenant les informations suivantes :

- Identifiant de la parcelle
- Variété(s) de cultivar(s) présente(s) dans chaque parcelle
- Date de plantation
- Métriques LiDAR calculées pour l'ensemble de la parcelle, comprenant des statistiques telles que le minimum, le maximum, la moyenne et l'écart-type
- Date d'acquisition des données LiDAR

Les parcelles du département 10 ne disposent pas de métriques LiDAR. Par conséquent, les informations provenant de ce département ont uniquement contribué à l'analyse statistique de l'incertitude de classification en fonction de l'âge de la plantation.

2.3. Traitements des données

Comme mentionné dans l'introduction, l'objectif principal du projet est d'étudier l'évolution de la probabilité d'appartenance des pixels à la classe peuplier de l'indice PI dans 4 régions françaises. Pour atteindre cet objectif, il a été nécessaire de traiter différentes données et ensuite les combiner. Le résultat de cette combinaison a été un tableau au format .csv, qui a permis de générer les diagrammes en boîte à moustaches.

Les traitements ont été réalisés en Python (version 3.10) à l'aide de l'éditeur de code VSCode (version 1.96). Le Jupyter Notebook (.ipynb) a été utilisé pour explorer et traiter les données grâce à des fonctions spécifiques créées pour répondre aux objectifs de l'étude, exécutées depuis des fichiers Python (.py). Dans ce rapport, nous ne détaillerons pas chaque fonction, car elles contiennent déjà des **docstrings** explicites qui les expliquent. Ainsi, nous nous concentrerons sur les objectifs principaux de chaque traitement et mettrons l'accent sur les choix méthodologiques lorsque cela sera nécessaire.

Le logiciel QGIS (version LTR 3.34) a été utilisé pour faciliter le filtrage, la visualisation ainsi que la vérification des données.

La chaîne de traitement a été divisée en 5 étapes principales et est représentée dans le diagramme ([Figure 2](#)) ci-dessous :

- i. Nettoyage et mise en forme des données vectorielles ;
- ii. Reprojection, rééchantillonnage et découpe des rasters ;

- iii. Extraction des valeurs des rasters et jointure spatiale avec les données vectorielles ;
- iv. Filtrage des valeurs dans le fichier .csv ;
- v. Visualisation des résultats.

Pour que la chaîne de traitement fonctionne, il est essentiel que les scripts soient exécutés dans l'ordre suivant :

- 1_Nettoyage_gpkg.ipynb
- 2_Masquage_rasters.ipynb
- 3_Extraction_valeurs.ipynb
- 4_Filtrage_csv.ipynb

Après l'exécution du script de filtrage, il est possible d'exécuter n'importe quel script de traçage (Boxnotch*) pour générer les résultats finaux ou tout script lançant les applications créées (App*), qui permettent d'explorer les données finales.

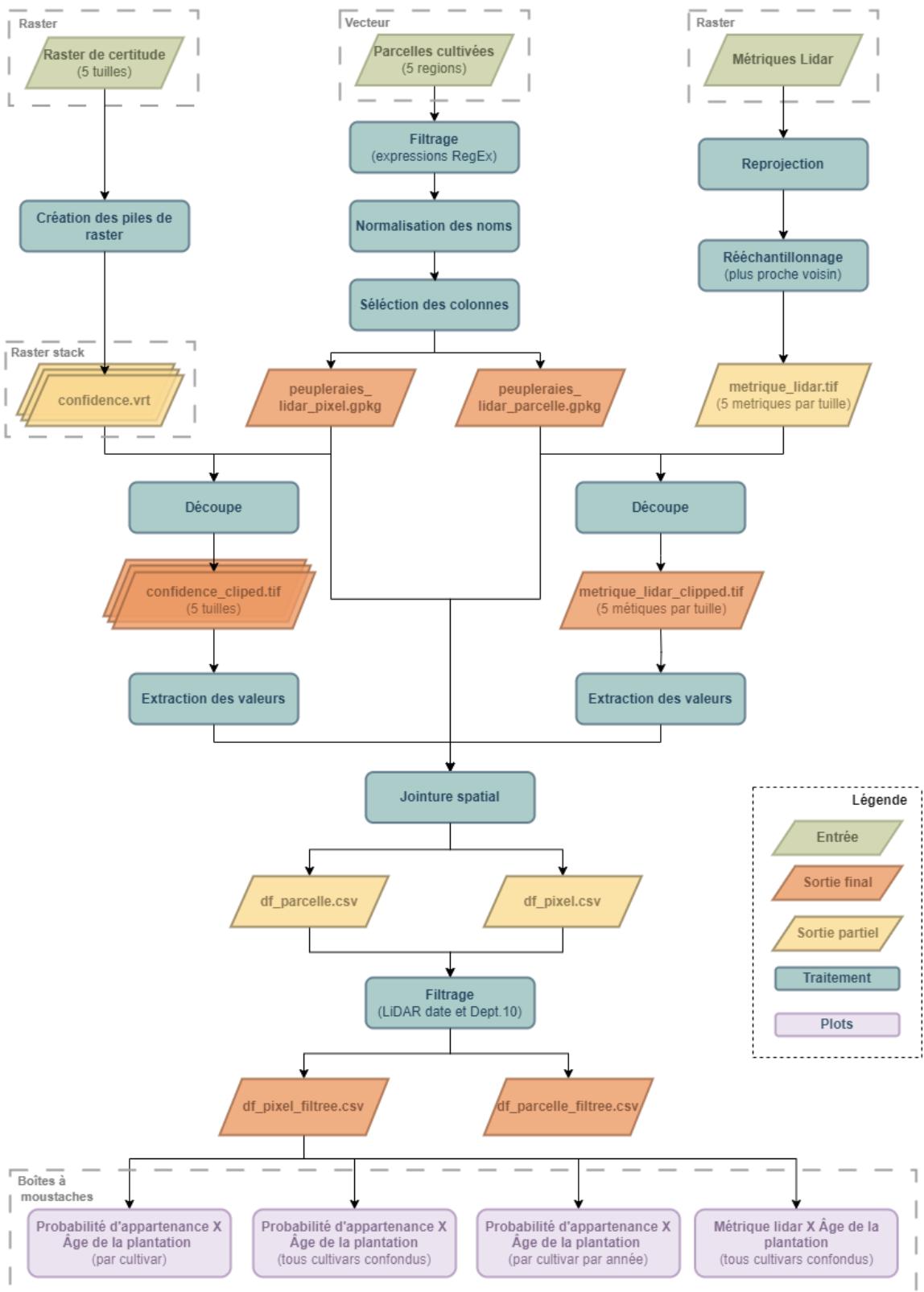


Figure 2 : Diagramme de la chaîne de traitement pour la génération du résultat final

2.3.1. Prétraitement des données vectorielles sur les parcelles cultivées

Le traitement des données vectorielles a été réalisé à l'aide du script 1_Nettoyage_gpkg.ipynb, qui utilise les fonctions définies dans le fichier functions_nettoyage.py. Cette étape avait deux objectifs principaux :

- i. Nettoyer et filtrer les données pour ne conserver que les parcelles contenant un seul cultivar, en excluant celles ayant plusieurs cultivars plantés ou sans cultivar.
- ii. Préparer les colonnes de la table attributaire afin qu'elles soient exploitables lors de la jointure spatiale avec les informations issues des rasters.

Une première phase exploratoire, conduite à l'aide du script 0_Exploration_vecteur_gpkg.ipynb, a permis d'identifier les incohérences et erreurs de frappe devant être corrigées.

La partie de nettoyage et de filtrage a été réalisée à l'aide d'expressions régulières (RegEx) grâce à la fonction seul_cultivar, qui a supprimé les parcelles sans informations sur les cultivars ou celles contenant plusieurs cultivars. Avec la fonction corriger_cultivar, une normalisation des noms des cultivars a été effectuée.

Un identifiant unique a été créé pour chaque parcelle, et une colonne source a été ajoutée pour indiquer à quelle zone la parcelle appartenait. Enfin, un filtrage et une normalisation des noms de colonnes ont été réalisés pour permettre la fusion de toutes les données vectorielles en une seule couche vectorielle à l'aide de la fonction concat de la bibliothèque Geopandas. L'objectif de fusionner en une seule couche vectorielle était de simplifier le processus de découpe qui serait effectué sur les rasters.

Le traitement des données vectorielles a permis de générer deux couches vectorielles distinctes :

- **peupleraies_merged_parcelle** : Regroupe les informations de toutes les parcelles et inclut les métriques LiDAR moyennes calculées pour chaque parcelle.
- **peupleraies_merged_pixel (sans données LiDAR calculées)** : Contient exactement les mêmes parcelles avec les mêmes informations que la première couche, à l'exception des données LiDAR calculées. Cette couche vectorielle a été créée pour permettre l'extraction des valeurs LiDAR pour chaque pixel à partir des rasters.

Avec les couches vectorielles prêtes, il était possible d’effectuer la découpe des rasters pour ne retenir que les pixels se trouvant spatialement dans les parcelles contenant des informations exploitables.

2.3.2. Découpe des données rasters

La découpe des rasters a dû être réalisée de deux manières en fonction des données, car les rasters de certitude et ceux des métriques LiDAR ont des systèmes de coordonnées et des emprises différents.

Pour les rasters de certitude, un « stack » de toute la série temporelle a été effectué, en empilant les rasters selon la date de production (2017 à 2022). Un Virtual Raster (.vt) a été créé à cet effet, puis découpé à l’aide de la couche vectorielle, ce qui a été simplifié, car les deux utilisaient le même système de projection (EPSG : 2154).

Pour la découpe, un tampon négatif de 10 mètres a été utilisé afin d’exclure les pixels situés en bordure des parcelles, susceptibles de biaiser les analyses, car ils présentent une probabilité de confiance plus faible étant situés dans les franges des parcelles. Lors de la découpe, la valeur -999 a été attribuée pour NoData, et le raster a été enregistré au format *int16* avec une compression au format *LZW*.

Pour les métriques LiDAR, il a été nécessaire de créer une boucle de traitement en utilisant le “stack” des rasters de certitude comme raster de référence. Dans la première étape, une reprojection vers l’EPSG : 2154 a été effectuée. Dans la deuxième étape, un rééchantillonnage utilisant la méthode du plus proche voisin a été réalisé. Cette méthode sélectionne la valeur du pixel le plus proche du nouveau centre sans effectuer d’interpolation, ce qui est idéal pour les données catégoriques, car elle conserve les valeurs originales sans introduire de variations numériques.

Dans la troisième étape, une découpe a été effectuée pour conserver les extensions correspondant à chaque tuile, après alignement parfait entre le raster de référence et le raster des métriques LiDAR.

Enfin, une dernière découpe a été réalisée pour chaque métrique à l’aide des données vectorielles, avec le même tampon négatif de 10 mètres, assurant un alignement parfait avec les rasters de certitude. Le résultat final est un raster par tuile (à noter que le département 10 ne dispose pas de données LiDAR) avec une valeur NoData fixée à -999 et un type de données défini en *float32*.

2.3.3. Extraction des valeurs et jointure spatial

Une fois que tous les rasters ont été alignés et découpés de manière identique, il a été temps d'extraire les valeurs de chacun des rasters et de les combiner avec les attributs des données vectorielles afin de générer un fichier .csv permettant la production des boîtes à moustaches. Deux boucles distinctes d'extraction et de jointures spatiales ont été réalisées :

- La première boucle a traité uniquement les rasters de confiance, qui ont ensuite été combinés avec les données vectorielles contenant les métriques LiDAR calculées à l'échelle des parcelles.
- La deuxième boucle a extrait non seulement les valeurs des rasters de certitude, mais également celles des métriques LiDAR, afin de les combiner avec les données vectorielles qui ne disposaient pas de métriques LiDAR calculées, mais uniquement des informations sur les parcelles cultivées. Dans ces données, chaque pixel possède une valeur distincte pour chaque métrique LiDAR.

L'extraction a été réalisée à l'aide de deux fonctions. La fonction extract_confidence_values traite les rasters de certitude et récupère les informations suivantes:

- Coordonnées X et Y des pixels
- Valeur de confiance
- Date du raster (récupérée à partir de l'ordre des bandes dans le raster)
- Tuile (extraite à partir du nom du fichier raster)

La deuxième fonction, extract_lidar_values traite les rasters des métriques LiDAR et récupère :

- Coordonnées X et Y des pixels
- Valeur de la métrique

Les informations extraites des rasters étaient associées aux parcelles à l'aide d'une jointure spatiale (sjoin) avec le prédictat spatial intersect. Lors de cette jointure, l'âge de la plantation a également été calculé en soustrayant la date du raster à la date de plantation renseignée dans la table attributaire des données vectorielles. D'ailleurs les âges de plantation négatifs, qui indiquent une coupe de la plantation durant la série temporelle, ont été supprimés.

Cette chaîne de traitement a abouti à deux fichiers .csv où chaque ligne représente un pixel :

- Dans le premier, les valeurs des métriques LiDAR provenaient de la table attributaire des données vectorielles, et chaque pixel partageait la même valeur pour une parcelle donnée (colonne unique_id).
- Dans le second, les valeurs des métriques étaient directement extraites des rasters, chaque pixel ayant une valeur distincte pour chaque métrique.

2.3.4. Filtrage des données du département 10 et des données dont la date d'acquisition est différente de celle des données LiDAR

Avec le fichier .csv généré, nous avons effectué deux filtrages supplémentaires pour affiner les données : le premier pour nettoyer les données du département 10, et le second pour permettre l'exploration des métriques LiDAR.

Le premier filtre a été appliqué uniquement au département 10. Ce filtrage visait à retirer les pixels provenant de parcelles dont la date de plantation était très ancienne et qui avaient nécessairement subi des changements entre 2017 et 2022. De plus, avec l'appui de la photo-interprétation, certains pixels ont également été exclus après vérification à l'aide des orthophotos de l'IGN. Ces pixels ont été jugés non pertinents pour notre jeu de données, leur inclusion risquant de biaiser l'analyse. L'image ci-dessous illustre l'une de ces incohérences:



Figure 3 : Capture d'écran montrant l'incohérence entre la date de plantation renseignée dans la table attributaire et ce qui peut être constaté par photo-interprétation sur la BD Ortho de 2021

Dans l'image, la parcelle est identifiée comme ayant été plantée en 2011 (pixels en jaune), mais en observant la BD Ortho de 2021, on se rend compte qu'il y a eu une coupe.

Le deuxième filtre a été réalisé pour permettre l'analyse des métriques LiDAR. Un filtre a été appliqué pour ne conserver que les pixels dont le raster de certitude avait été produit à la même date que le raster contenant la métrique LiDAR. Dans ces données, un critère supplémentaire a également été appliqué : seuls les pixels ayant une couverture de canopée (colonne grid_CC) supérieure à 5 % ont été conservés, si le peuplier avait plus de 5 ans. Ce dernier filtre a aidé à identifier les parcelles où un abattage aurait pu avoir lieu et qui n'avait été détecté par aucun filtre précédent.

2.3.5. Création des boîtes à moustaches avec encoche

Après le filtrage du fichier .csv, des scripts ont été créés pour générer des diagrammes, visant à visualiser et analyser les données. Les fonctions de traçage sont organisées dans le script functions_plots.py et sont appelées par différents scripts selon les besoins.

Pour tous les graphiques, nous avons choisi de représenter les variables à l'aide de boîtes à moustaches avec notch (ou boxnotch pour simplifier). Un boxnotch est une variante du

boxplot traditionnel qui inclut des encoches autour de la médiane pour représenter l'intervalle de confiance. Plus l'encoche est étroite, plus la médiane est statistiquement précise. Cette représentation permet de mieux visualiser la précision des données et de comparer les groupes de manière statistiquement informée.

La fonction `top_cultivars` ne génère pas de graphiques, mais réalise la sélection et le tri des cultivars les plus représentatifs, en se basant sur le nombre de pixels disponibles dans le jeu de données. Cette fonction utilise les arguments suivants :

- `max_age` : Âge maximum des plantations à considérer, avec une valeur par défaut de 12, puisque l'analyse se concentre sur les jeunes cultivars.
- `min_years` : Nombre minimum d'années pendant lesquelles le cultivar doit avoir des données disponibles, avec une valeur par défaut de 6. Ainsi, seuls les cultivars ayant des enregistrements sur au moins 6 ans sont sélectionnés.
- `top_n` : Nombre maximal de cultivars à inclure dans la sélection finale, avec une valeur par défaut de 20.

Le résultat est un dictionnaire contenant les cultivars sélectionnés et le nombre de pixels associés, classés par ordre décroissant.

Après la sélection des cultivars avec `top_cultivars`, les fonctions de traçage sont appelées pour générer des graphiques. Les principales fonctions sont :

1. **boxnotch_confidenceXage** : crée un diagramme en boîte avec encoches pour représenter graphiquement la probabilité d'appartenance à la classe peuplier en fonction de l'âge de la plantation, pour un cultivar spécifique. Inclut des statistiques et des points colorés représentant les origines des données.
2. **grid_boxnotch_confidenceXage_par_anne** : génère une grille de diagrammes en boîte montrant la probabilité d'appartenance en fonction de l'âge de la plantation, segmentée par année d'acquisition des données, pour un cultivar spécifique.
3. **boxnotch_lidar_metrics** : produit un diagramme en boîte avec encoches et un nuage de points pour une métrique LiDAR spécifique en fonction de l'âge de la plantation, mettant en évidence la variation des données pour un cultivar donné.
4. **boxnotch_confidenceXage_lidar_metrics** : génère deux graphiques combinés : l'un pour la probabilité d'appartenance et l'autre pour une métrique LiDAR spécifique,

tous deux en fonction de l'âge de la plantation, mettant en évidence de possibles corrélations entre les variables.

Toutes ces fonctions exportent les graphiques générés au format .png ainsi que dans un fichier consolidé au format .pdf.

2.3.6. Crédit de l'application avec Dash et Plotly

Avec la génération des premiers graphiques, nous avons ressenti le besoin d'explorer les données de manière plus interactive. Pour cela, nous avons créé des applications en utilisant les bibliothèques Dash et Plotly.

La première application se trouve dans le script **7_0_App_un_graphique.py**. Elle permet de choisir un ou plusieurs cultivars simultanément ainsi que les variables X et Y via des listes déroulantes. Les visualisations suivantes peuvent être générées :

- Boxplot avec nuage de points ;
- Scatter plot ;
- Heatmap ;
- Grille de graphiques par année (facets).

De plus, il est possible d'explorer les valeurs de chaque point du nuage de points ou du boxplot à l'aide de la fonction hover ([Figure 4](#)).

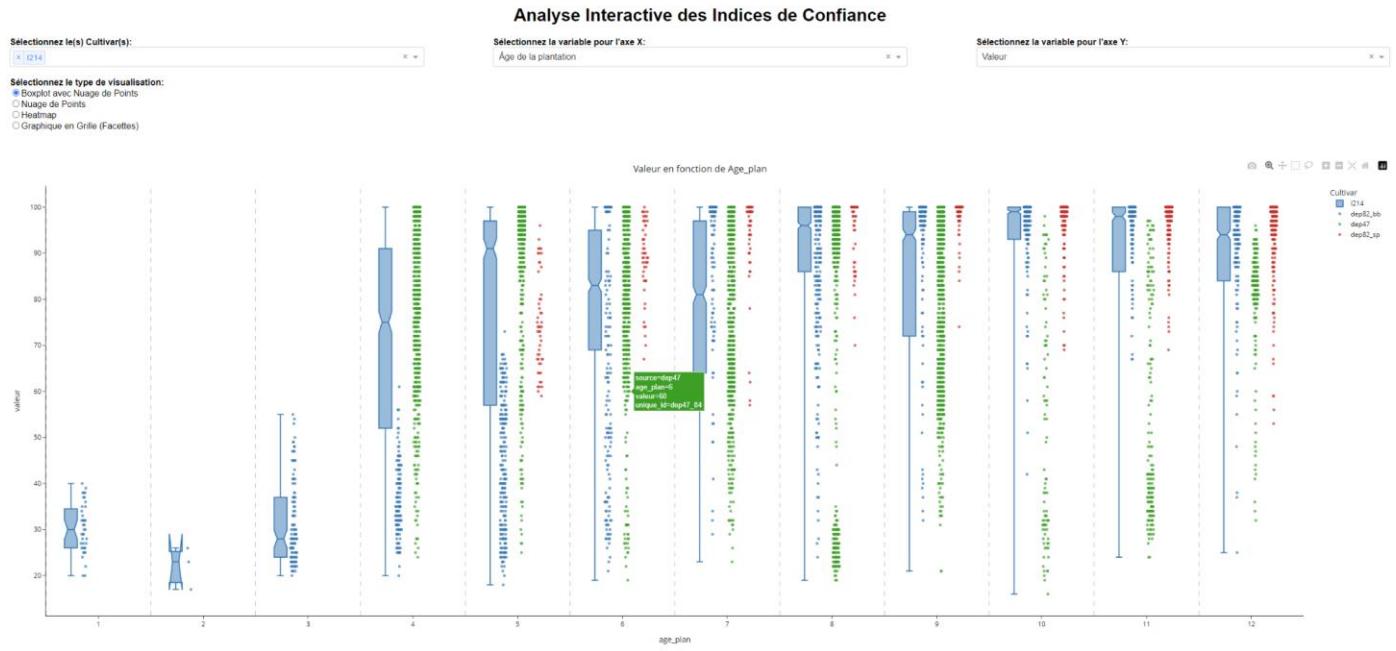


Figure 4 : Capture d'écran de l'application créée pour explorer les graphiques de manière interactive

La deuxième application, dans le script 7_1_App_px_confidenceXage, est une variation de la première. Cependant, celle-ci affiche deux graphiques côté à côté, permettant ainsi une comparaison directe ([Figure 5](#)).

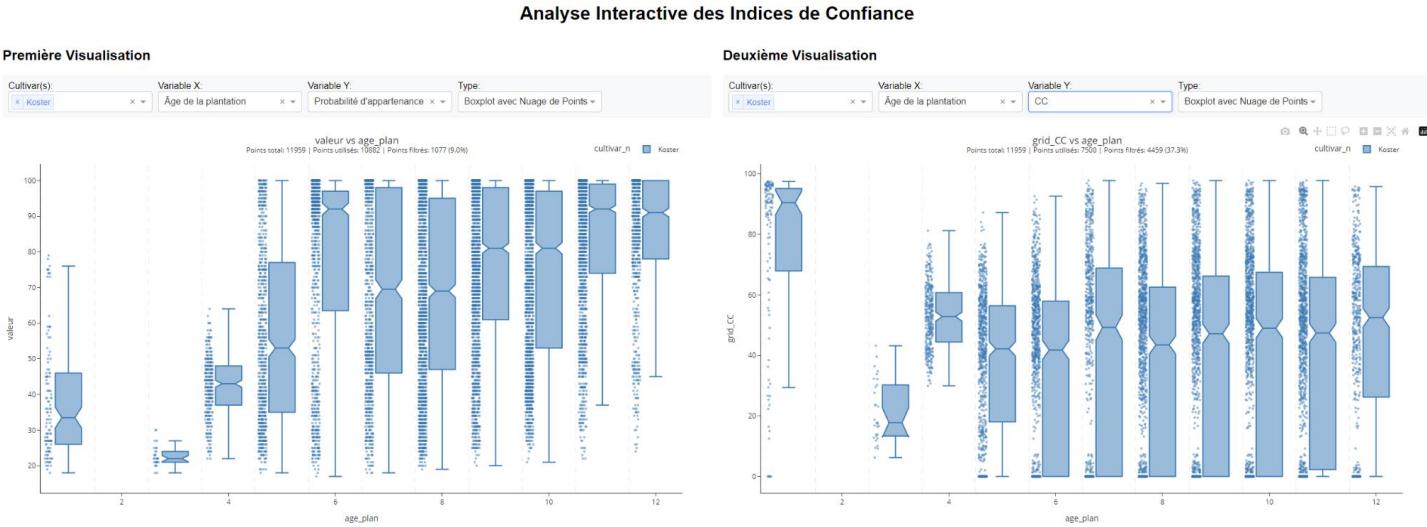


Figure 5 : Capture d'écran de l'application créée pour comparer deux graphiques de manière interactive

3. Résultats

3.1. Variation du nombre des données suivant les différents filtres

Le tableau ci-dessous illustre la variation du nombre de parcelles et de pixels après l'application des différents filtres :

- Parcelles

Tableau 1 : Variation de nombre de parcelles par département tout au long de chaîne de

Localisation (tuile/département)	Nombre initial	Nombre avec cultivar unique	Nombre des parcelles après extraction
82 Borde Basse	54	46	42
82 St Pierre	259	226	138
47	114	72	60
73	73	25	19
10	90	71	64
TOTAL	590	440	323

traitement

- Pixel

Tableau 2 : Quantité de données en pixel tout au long de chaîne de traitement

Données initiales	Filtre 1	Filtre 2
Nombre de pixels extraits des rasters.	Nombre de pixels restants après le filtrage par photo-interprétation pour le département 10.	Nombre de pixels dont la date de tuile Sentinel 2 est identique à celle de l'acquisition des données LiDAR plus le filtre du Canopy Cover
77617	71453	9031

Comme expliqué dans la section 2.2.5, une fonction appelée top_cultivars a été créée pour sélectionner les cultivars les plus représentatifs. Dans les graphiques montrant la relation entre la probabilité d'appartenance et l'âge de plantation, les paramètres de la fonction étaient : max_age = 12 et min_years = 6.

Seuls 18 cultivars sur 53 ont satisfait ces critères. La quantité de pixels associée à chacun d'eux est présentée ci-dessous :

Tableau 3 : Nombre de cultivars et nombre de pixels restants après l'application de la fonction top_cultivars pour le traçage du graphique de probabilité d'appartenance en fonction de l'âge de la plantation

N	Cultivar	Nombre de pixels	N	Cultivar	Nombre de pixels
1	Koster	10764	10	Albelo	1031
2	Polargo	7728	11	Alcinde	886
3	I214	7655	12	Diva	500
4	I45/51	5073	13	Tucano	300
5	Soligo	3940	14	Senna	294
6	Flevo	3499	15	San Martino	284
7	Dorskamp	2608	16	Aleramo	177
8	Raspalje	2459	17	Moncalvo	171
9	Triplo	1370	18	Rona	160

Pour tracer les métriques LiDAR, les paramètres de la fonction top_cultivars ont été ajustés en raison de la quantité de données beaucoup plus faible. Le même max_age = 12 a été utilisé, mais avec min_years = 2. Le nombre de cultivars répondant à ce critère est de 16.

Tableau 4 : Nombre de cultivars et nombre de pixels restants après l'application de la fonction top_cultivars pour le traçage du graphique des métriques LiDAR

N	Cultivar	Nombre de pixels	N	Cultivar	Nombre de pixels
1	Koster	783	9	Moncalvo	114
2	I45/51	580	10	Senna	114
3	Raspalje	536	11	San Martino	83
4	Polargo	516	12	Rona	71
5	I214	497	13	Aleramo	64
6	Soligo	317	14	Hoogvorst	30
7	Diva	182	15	Delvignac	26

8	Tucano	117		16	Vesten	5
---	--------	-----	--	----	--------	---

3.2. Probabilité d'appartenance des pixels à la classe peuplier en fonction de l'âge de la plantation, tous cultivars confondus

Pour mieux visualiser l'origine des données, nous avons ajouté, à gauche du boxnotch, un nuage de points représentant les données, avec un code couleur permettant de différencier les départements. Une légère séparation entre les années a été ajoutée à l'aide d'une ligne pointillée. Si un cultivar ne possède aucun pixel pour une année donnée, nous avons inséré la mention NA dans l'espace correspondant. Lorsque l'espace le permet, nous avons également ajouté, sous chaque boxnotch, quatre métriques :

- **Nb** = Nombre de pixels (points)
- **Moy** = Moyenne
- **Med** = Médiane
- **σ** = Écart-type

Un graphique de l'incertitude de la classification selon l'âge, intégrant tous ces éléments visuels, peut être observé ci-dessous.

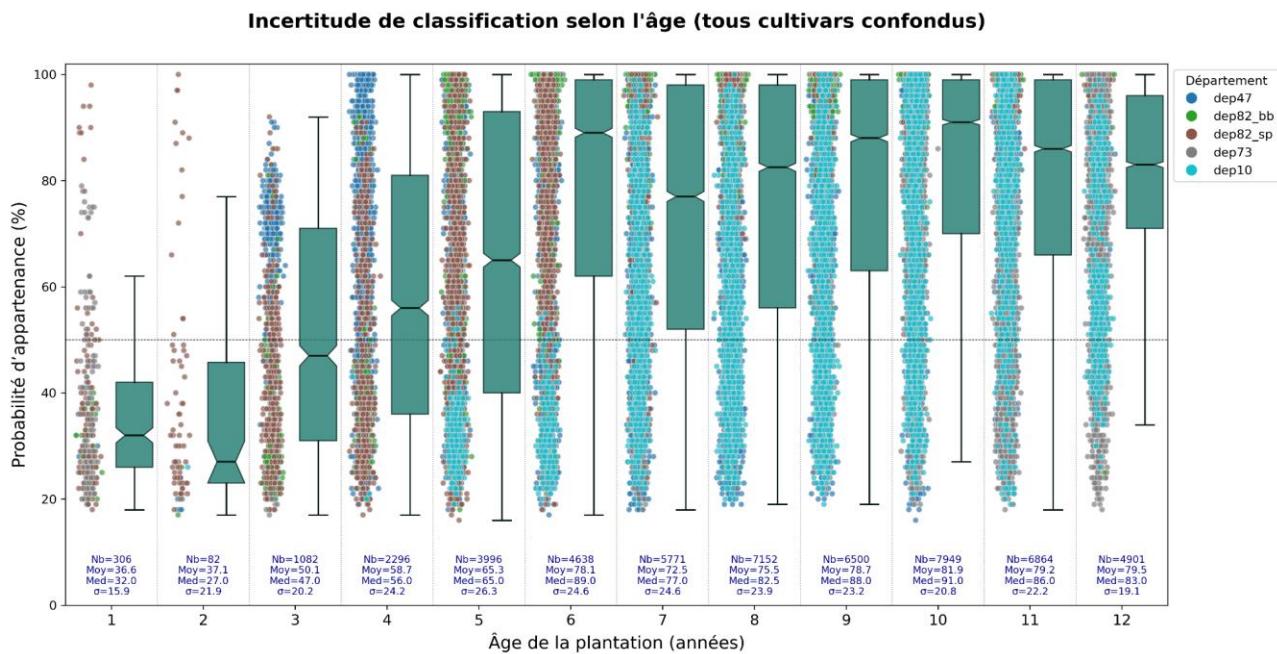


Figure 6 : Probabilité d'appartenance en fonction de l'âge de plantation tous cultivars confondus

3.2.1. Probabilité d'appartenance des pixels à la classe peuplier en fonction de

l'âge de la plantation, par type de cultivars

Seuls les cultivars les plus représentatifs en termes de volume de données (nombre de pixels) sont présentés dans cette partie du rapport.

3.2.1. Probabilité d'appartenance par cultivar sans différenciation des dates d'acquisition de données Sentinel

Cultivar Koster :

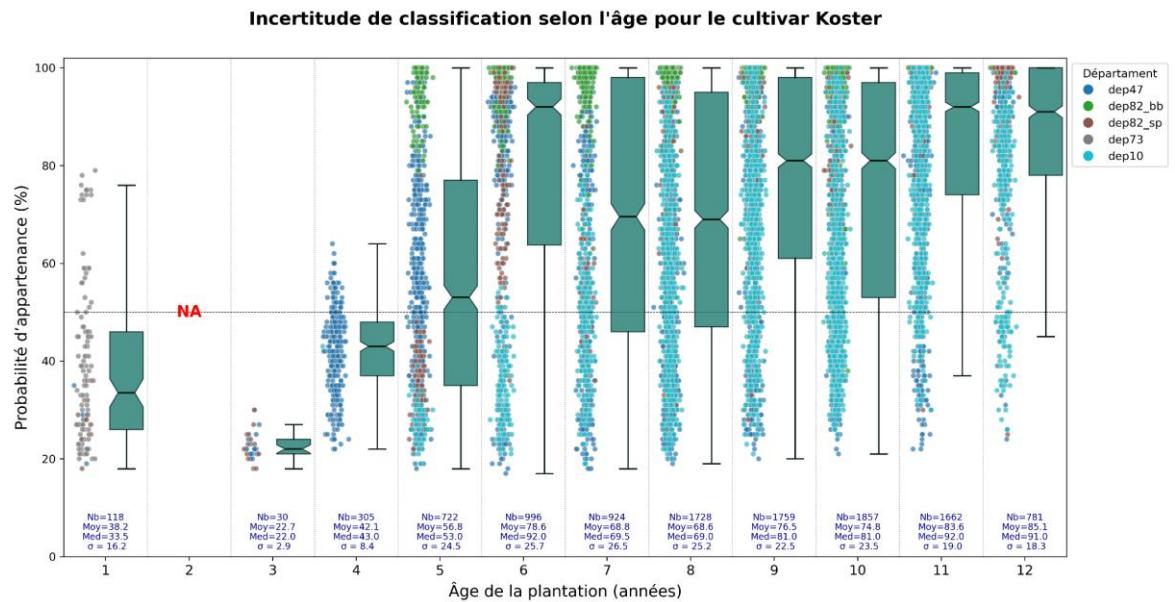


Figure 7 : Probabilité d'appartenance en fonction de l'âge de plantation pour le cultivar Koster

Cultivar Polargo :

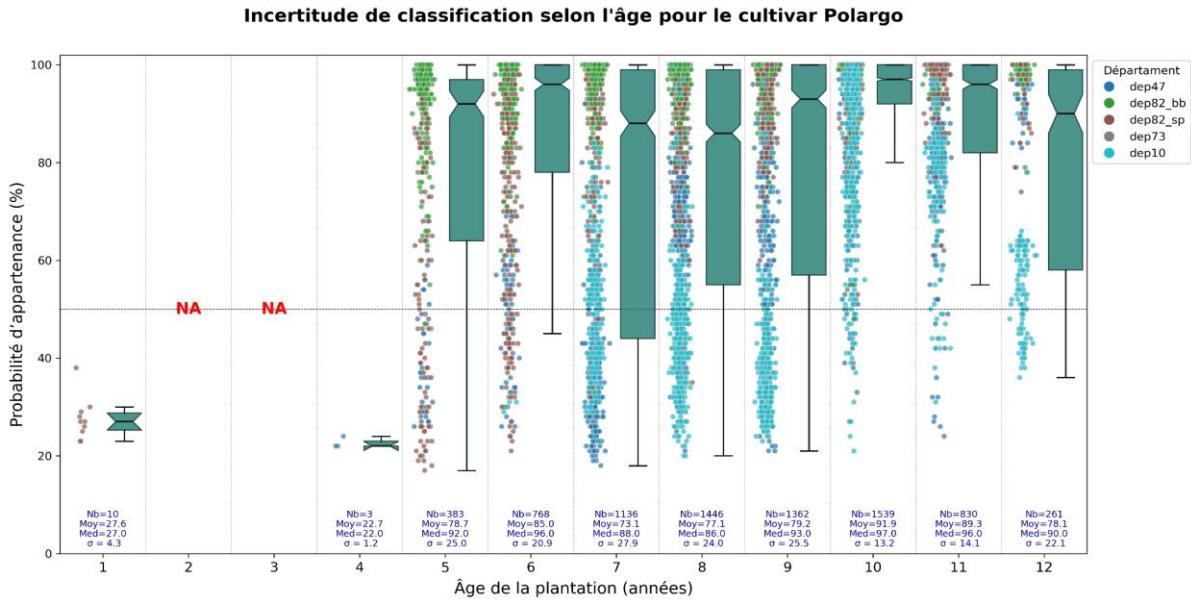


Figure 8 : Probabilité d'appartenance en fonction de l'âge pour le cultivar Polargo

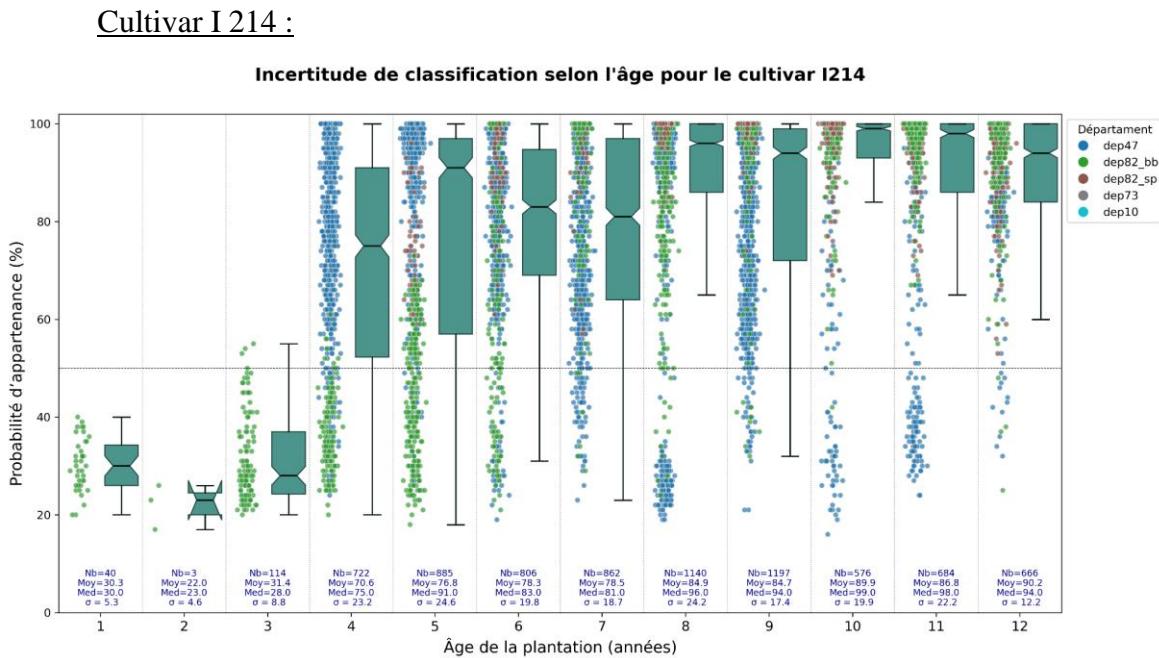


Figure 9 : Probabilité d'appartenance en fonction de l'âge pour le cultivar I214

3.2.2. Probabilité d'appartenance des pixels à la classe peuplier en fonction de l'âge de la plantation, par cultivar et par années d'acquisition de Sentinel

L'objectif est d'identifier s'il existe une influence potentielle de la date d'acquisition des données Sentinel sur la probabilité d'appartenance à la classe peuplier.

Les codes couleurs sont appliqués pour distinguer les différentes années et faciliter l’interprétation visuelle des résultats.

Cultivar Koster :

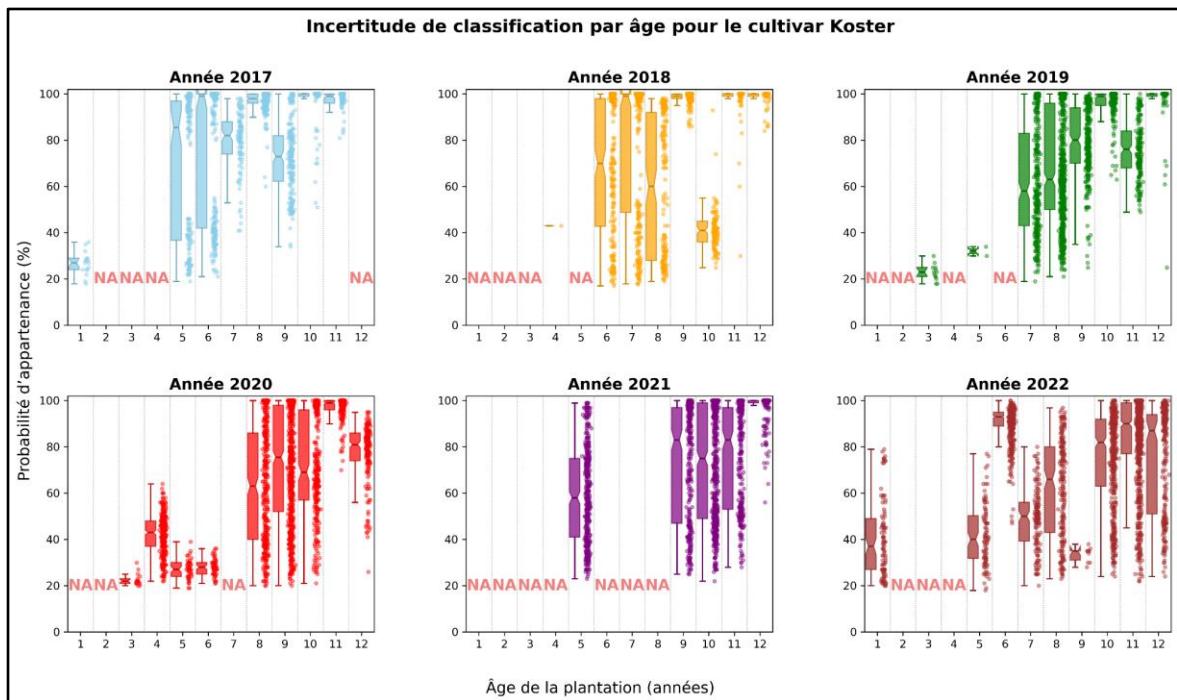


Figure 10 : Probabilité d'appartenance en fonction de l'âge entre les années de 2017 à 2020 pour le cultivar Koster

Cultivar Polargo :

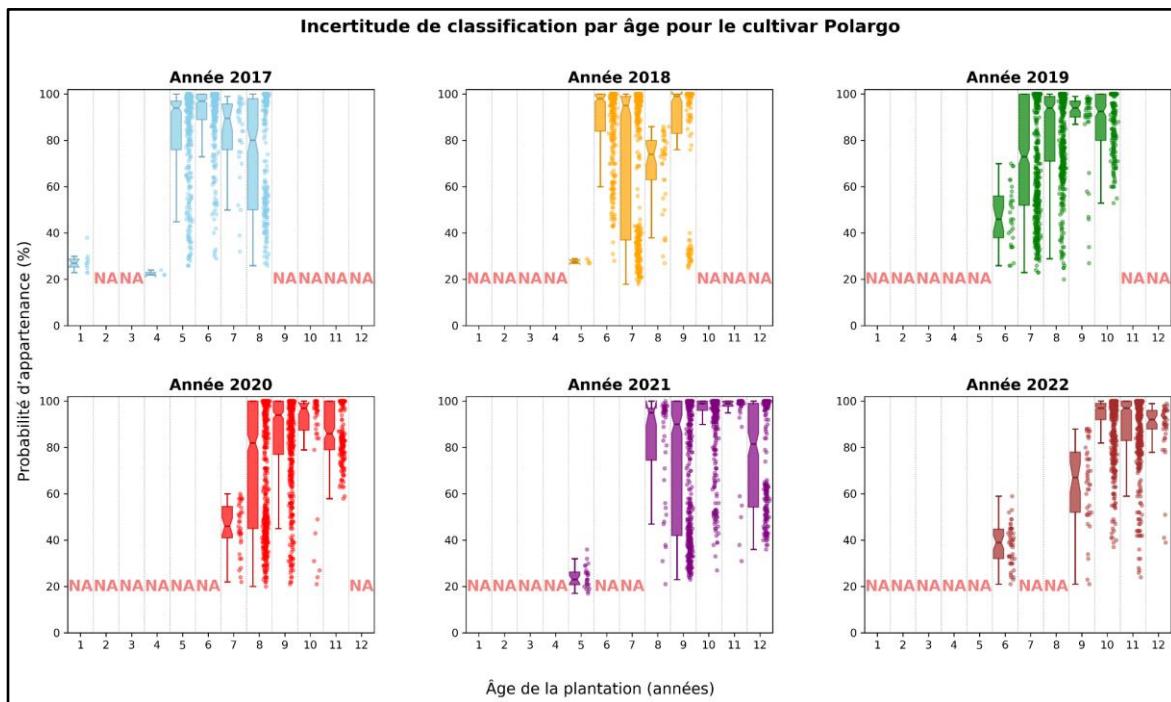


Figure 11 : Probabilité d'appartenance en fonction de l'âge entre les années de 2017 à 2020 pour le cultivar Polargo

Cultivar I 214 :

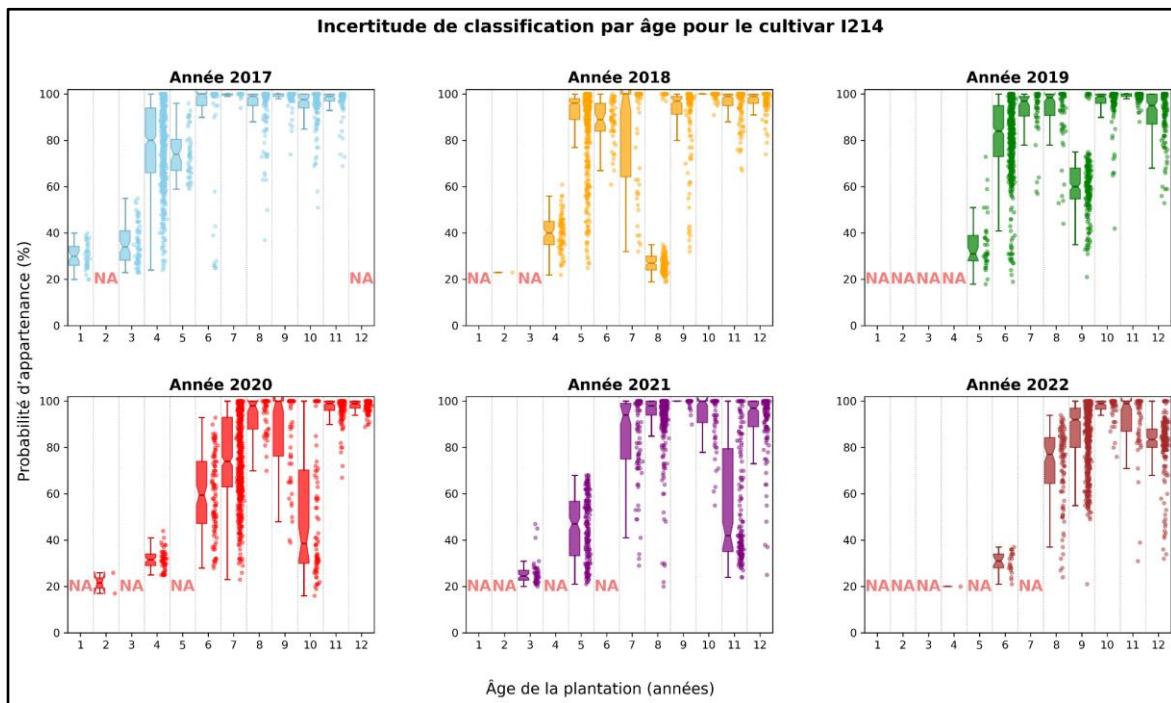


Figure 12 : Probabilité d'appartenance en fonction de l'âge entre les années de 2017 à 2020 pour le cultivar I214

3.3. Probabilité d'appartenance des pixels à la classe peuplier, densité de végétation (PAI) et l'ouverture de la canopée (Canopy cover) en fonction de l'âge

Les graphiques de cette section se réfèrent à l'analyse des métriques LiDAR. Comme l'acquisition des données LiDAR a été réalisée entre 2021 et 2023, les données analysées ici ont fait l'objet d'un **second filtrage**. Autrement dit, seuls les pixels dont la date de la tuile Sentinel-2 correspond à la date d'acquisition des données LiDAR ont été pris en compte, en plus du critère supplémentaire de Canopy Cover expliqué dans la section 2.2.4.

3.3.1. Visualisation Globale : tous cultivars confondus

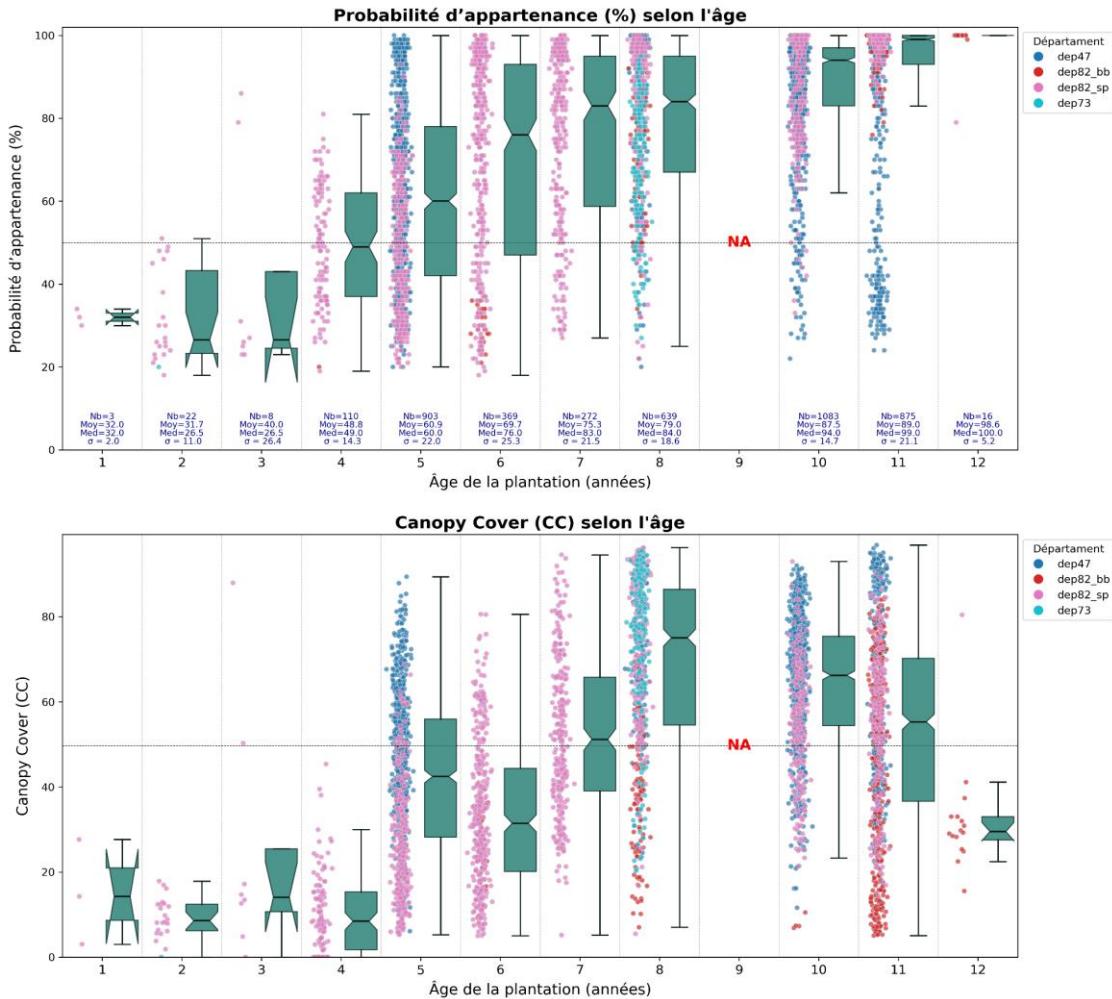


Figure 13 : Canopy cover en fonction de l'âge tous cultivars confondus

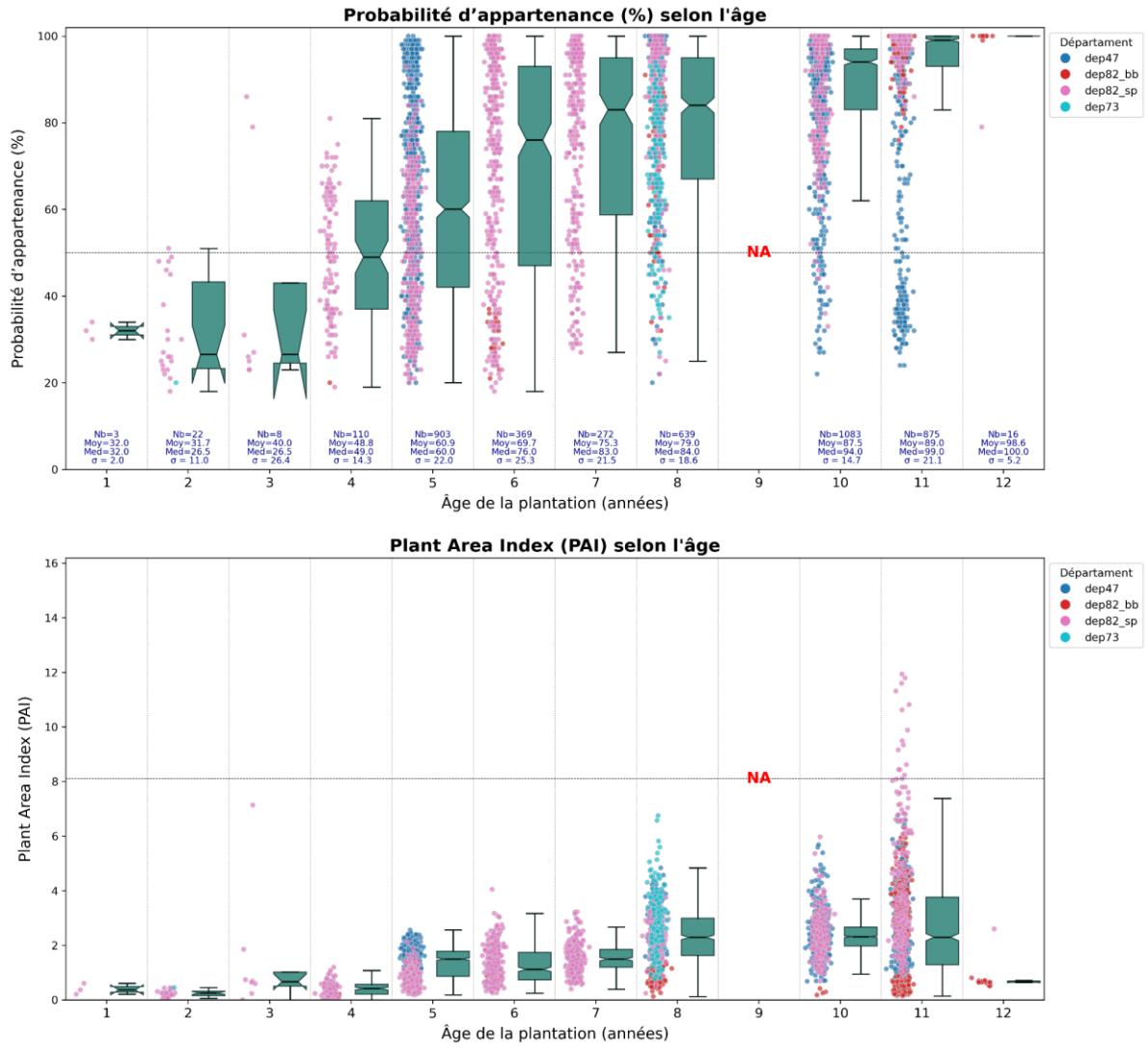


Figure 14 : Densité de végétation (PAI) en fonction de l'âge tous cultivars confondus

3.3.2. Visualisation détaillée: par type de Cultivar

La quantité de données étant trop faible, seuls le cultivar Koster et I 45 / 51 ont été analysés pour explorer la corrélation entre les variables LiDAR (CC et PAI) et l'âge des cultivars.

Un autre graphique illustre la variation de la probabilité d'appartenance en fonction de l'âge. Cette approche améliore la visualisation et permet une analyse plus approfondie de l'évolution des variables étudiées.

Cultivar Koster :

- Variable CC en fonction de l'âge de plantation:

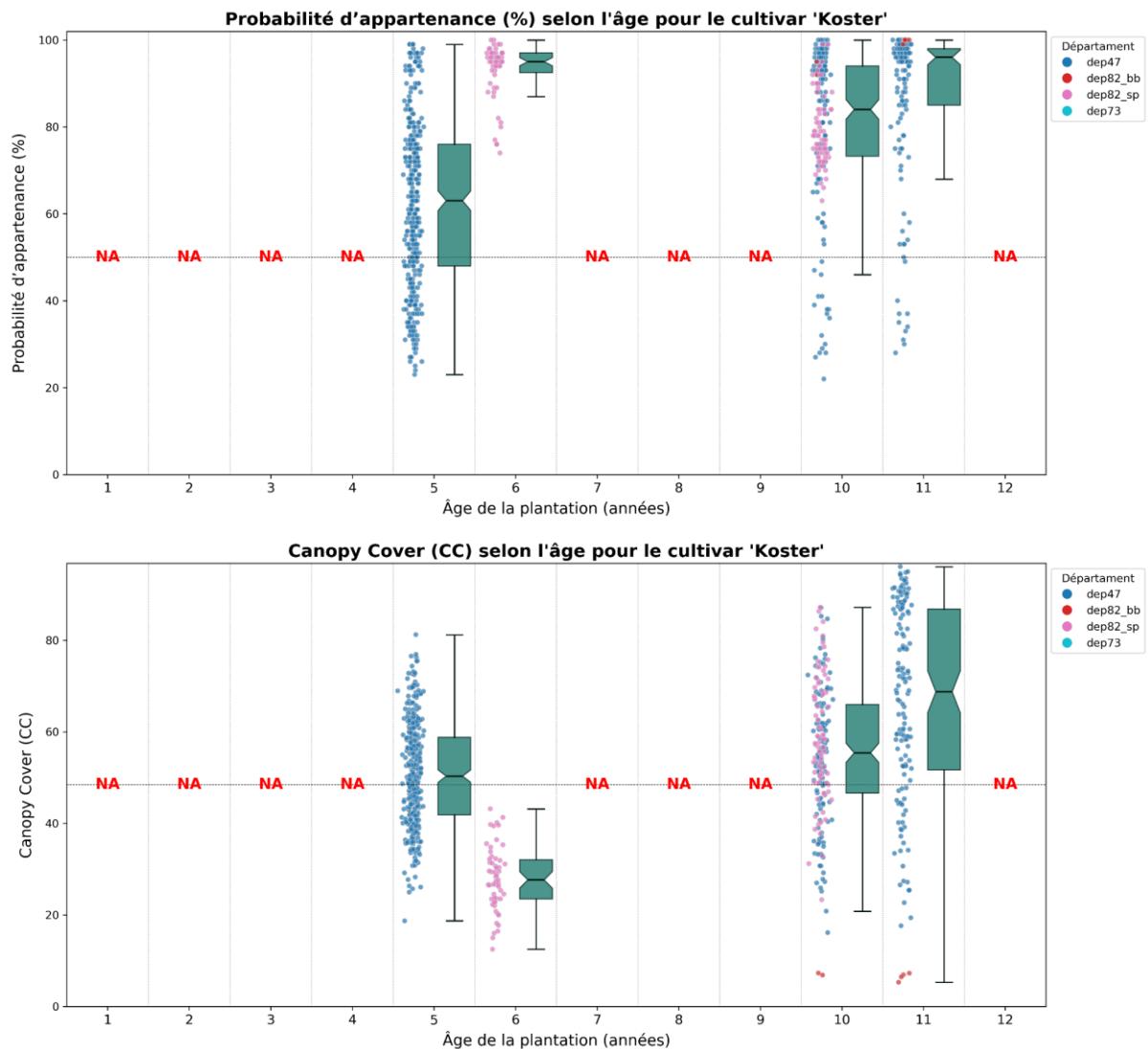


Figure 15 : Canopy cover et probabilité d'appartenance en fonction de l'âge de plantation pour le cultivar Koster

- Variable PAI en fonction de l'âge de plantation

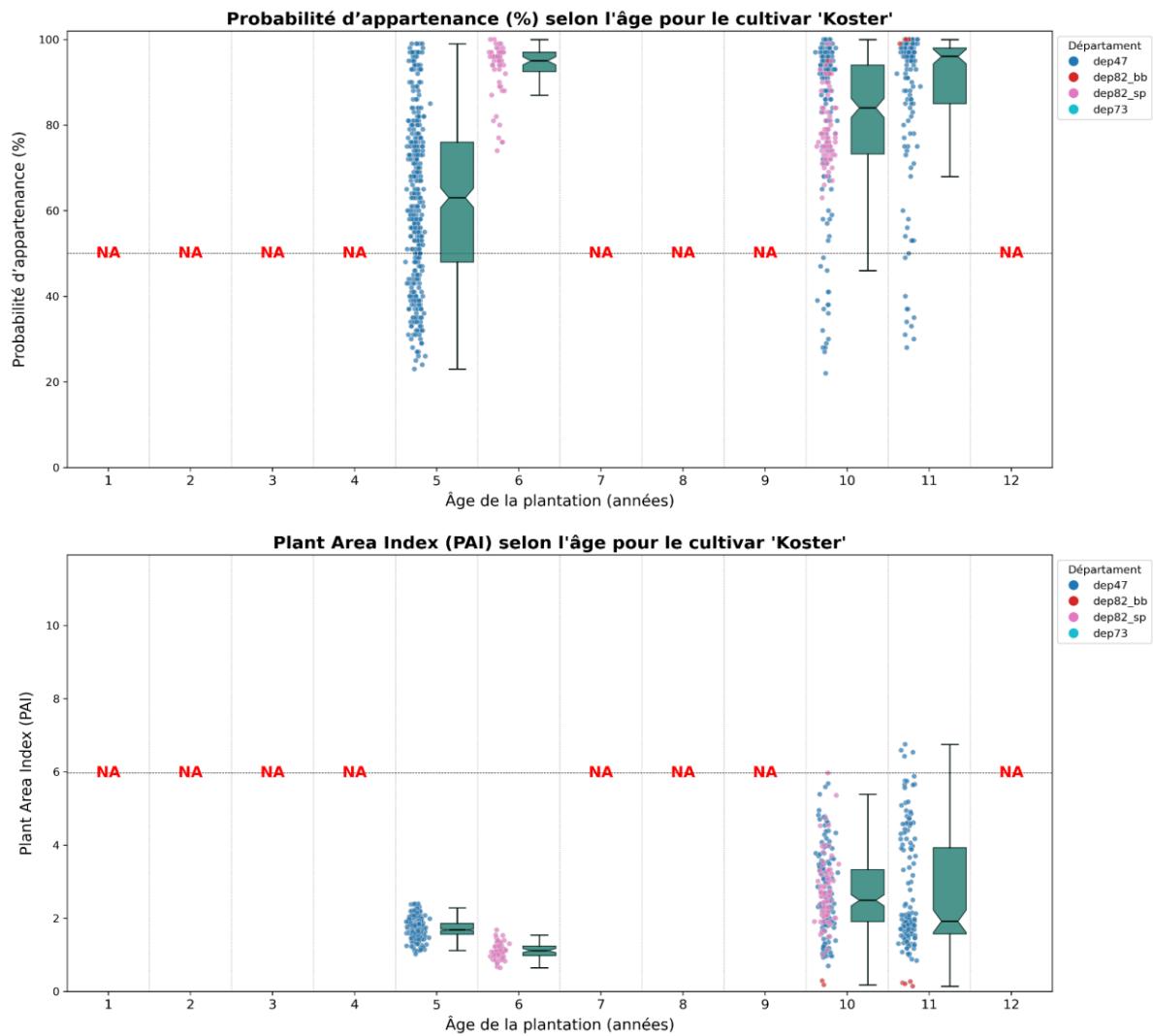


Figure 16 : Densité de végétation (PAI) et probabilité d'appartenance en fonction de l'âge de plantation pour le cultivar Koster

Cultivar I 45 / 51 :

- Variable CC en fonction de l'âge de plantation:

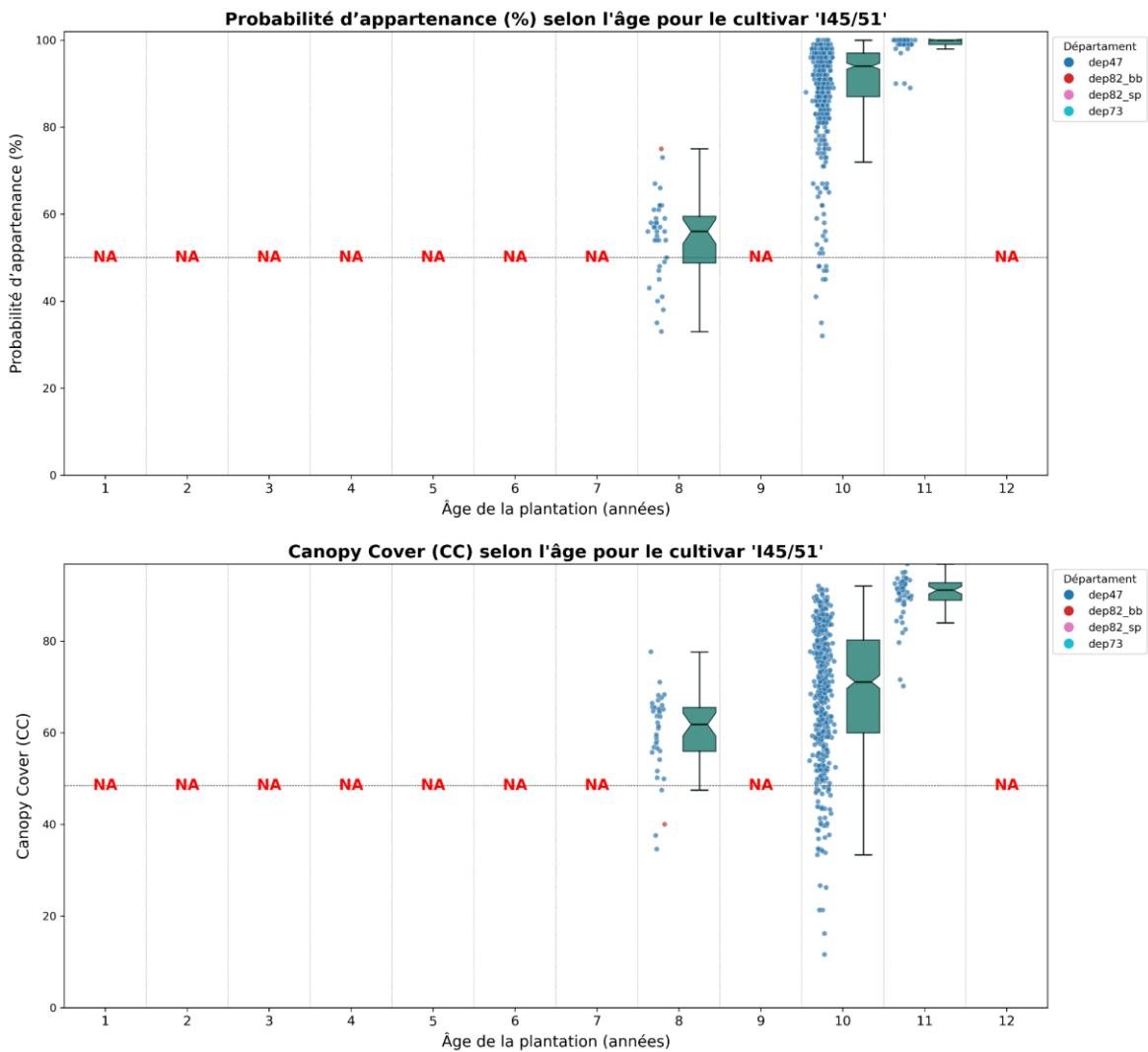


Figure 17 : Canopy cover et probabilité d'appartenance en fonction de l'âge de plantation pour le cultivar I45/51

- Variable PAI en fonction de l'âge de plantation :

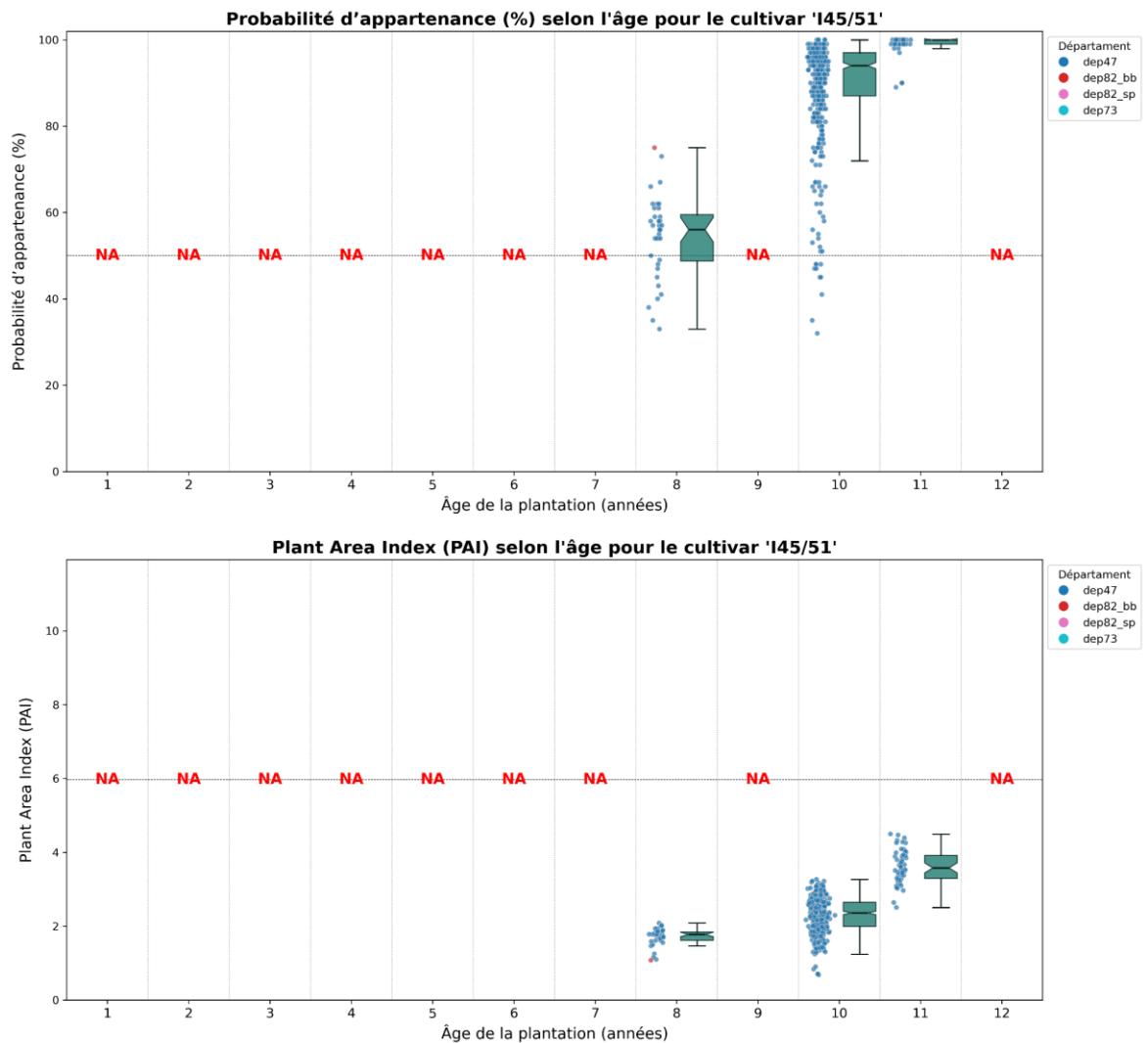


Figure 18 : Densité de végétation (PAI) et probabilité d'appartenance en fonction de l'âge de plantation pour le cultivar I45/51

4. Discussion

4.1. Diminution du nombre de données exploitables

Les [Tableaux 1 et 2](#) mettent en évidence une diminution progressive et significative du nombre de données exploitables à mesure de l'application de différents filtres. Cette réduction s'explique principalement par l'élimination :

- Des parcelles présentant deux cultivars ou plus
- Des parcelles qui ne sont pas couvertes par les tuiles du Sentinel-2 contenant les informations sur la probabilité d'appartenance.
- Des parcelles avec des peupleraies dont l'âge de la plantation est négatif : la date de plantation est postérieure à la date de tuiles de sentinel-2
- Des pixels dont la date de tuiles de Sentinel-2 n'a pas de correspondance à la date d'acquisition des données LiDAR
- Des pixels correspondant à des cultivars âgés de 5 ans ou plus et dont la couverture de canopée est inférieure à 5 %.

Ces données plus faibles réduisent la représentativité des échantillons. Cette perte concerne surtout les données métriques et a été particulièrement marquée dans certains départements, notamment le département 10, où les données LiDAR étaient indisponibles, ainsi que dans les départements 47 et 73, où plus de la moitié des données initiales ont été éliminées.

Sur les graphiques illustrant la variation des valeurs métriques “Canopy cover” et “Plante Area Index” en fonction de l'âge pour les cultivars Koster et I 45/51 par exemple ([Figure 15, 16, 17 et 18](#)) seules quelques classes d'âge disposent des données. Cette limitation complique une analyse fine de la corrélation entre ces variables pour ces cultivars.

4.2. Variation de la probabilité d'appartenance en fonction de l'âge tous cultivars confondus

Les probabilités d'appartenance semblent augmenter avec l'âge des plantations, en particulier après 4 ans, où les boîtes à moustaches montrent un pourcentage de probabilité d'appartenance supérieur à 50% ([Figure 6](#)).

Au-delà de 5 ans, la probabilité d'appartenance demeure élevée, bien que sa variabilité interannuelle soit perceptible. Cette dernière est probablement influencée par les différences de la quantité de données disponible d'une année à l'autre.

Une tendance claire d'amélioration de la probabilité d'appartenance avec l'âge se dessine, ce qui suggère une meilleure classification des plantations plus anciennes. Cela pourrait être lié à des traits ou caractéristiques plus marqués et développés chez les plantations âgées (biomasse, surface foliaire,...).

4.3. Variation des valeurs du Canopy Cover (CC) et du Plant Area Index (PAI) en fonction de l'âge tous cultivars confondus

Les valeurs des variables CC et PAI augmentent globalement avec l'âge des cultivars. Pour le PAI, une forte augmentation est observée après la 4^e année, suivie d'une stagnation avec un léger effet "yo-yo" jusqu'à 11 ans, âge auquel il atteint son maximum (aux alentours de 2). En revanche, le CC progresse de manière continue et marquée jusqu'à 11 ans, atteignant une moyenne de 15 %.

Ce scénario s'explique par le déploiement progressif du couvert végétal accompagnant le développement des jeunes arbres. La valeur de ces deux métriques évolue en parallèle.

Les valeurs de CC et PAI des parcelles du département 82-Borde Basse semblent rester faibles dès leurs apparitions dans les graphiques (6, 8, 11 et 12 ans). Ces données, issues du cultivar I214, peuvent influencer la moyenne annuelle de CC et PAI de l'ensemble des données en tirant vers le bas. Toutefois, l'outil mis en place dans le cadre de cette étude pour visualiser les données montre bien que ces valeurs augmentent de manière nette chaque année jusqu'à l'année 10 avant de chuter significativement à partir de l'année 11.

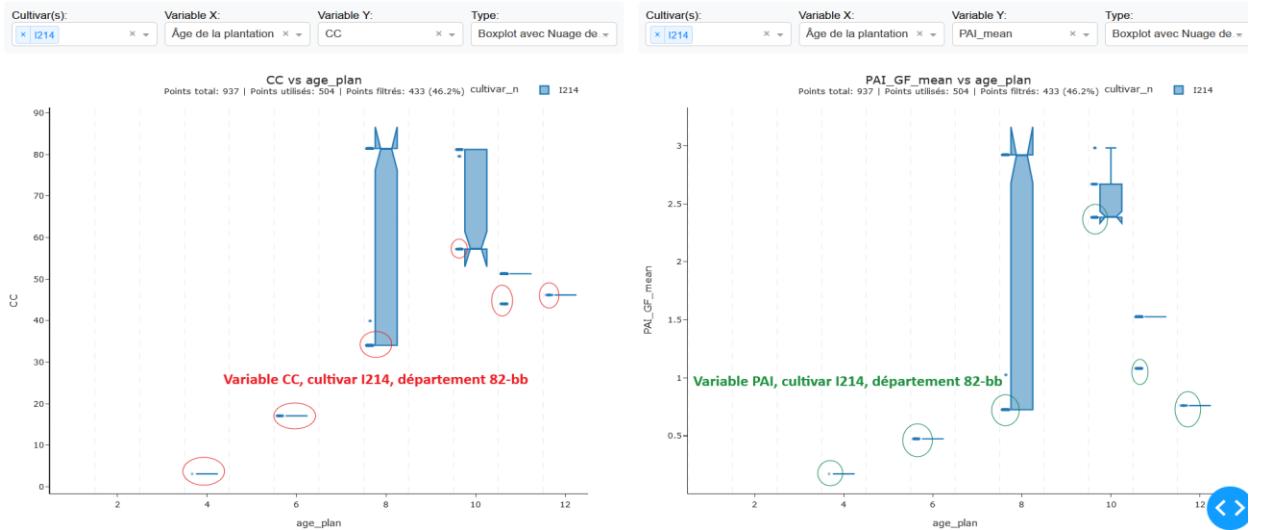


Figure 19 : Canopy cover et Plant Area Index (PAI) en fonction de l'âge pour le cultivar I214

L’analyse visuelle de ces plantations, réalisée avec QGIS et l’ortho-photo de l’IGN, n’a pas permis d’identifier la raison de la chute des valeurs de CC et PAI observée à partir de 11 ans.

4.4. Âge minimum pour une meilleure classification, variable selon le cultivar

Comme développé précédemment, nous nous attendons à une augmentation de la probabilité d’appartenance à la classe « peuplier » avec l’âge des plantations. Dans cette optique, un âge minimum de plantation semble nécessaire pour garantir une bonne classification. D’après nos résultats, ce seuil se situe autour de 5 ans si l’on considère qu’une valeur de probabilité d’appartenance supérieure à 60% est jugée fiable. Cependant, cet âge minimum est susceptible de varier notamment selon les cultivars :

Pour le cultivar I 45/51, par exemple, l’algorithme commence à bien le détecter à partir de l’âge de 7 ans où la valeur moyenne de probabilité d’appartenance est de 65,7 %. D’après [la fiche technique du CNPF](#) relative à ce cultivar : la densité de plantation influe sur son rythme de croissance. Lorsque la densité est inférieure à 204 tiges/ha, la croissance soutenue ne se manifeste qu’après 8 à 9 ans. Dans notre cas, les données relatives à ce cultivar proviennent principalement du département 47 ([Annexe 1.b](#)) et, dans une moindre mesure, du département 82 (Bord Basse), où la densité de plantation varie entre 160 et 204 pieds/ha. Ces observations concordent avec les résultats observés.

Pour le cultivar Flevo, la détection arrive tardivement : aux alentours de 10 ans après la plantation, avec une moyenne de confiance avoisinant le 80 % ([Annexe 1.b](#)). D’après la [fiche](#)

[REP 1999](#), la phase d'installation de ce cultivar est globalement lente et varie suivant les stations : elle est rapide dans les stations sans excès d'eau mais beaucoup plus lente dans les environnements plus humides où elle peut aller jusqu'à 5 ans. La phase de croissance, quant à elle, est progressive et se maintient à un rythme soutenu jusqu'à environ 10 ans. Nos données concernant ce cultivar proviennent majoritairement des départements 73 et 10, situés plus au nord, où le stress hydrique est moins prononcé. Ces caractéristiques pourraient expliquer la détection tardive de ce cultivar.

4.5. Incohérences de la valeur de probabilité d'appartenance chez les plantations jeunes et âgées

Pour certains cultivars, des incohérences dans les valeurs de probabilité d'appartenance ont été observées tant chez les plantations jeunes que chez les plantations âgées.

Pour les plantations très jeunes affichant une probabilité d'appartenance déjà élevée, le cas le plus notable est celui du cultivar San Martino, où cette valeur dépasse le 80% dès la 1ère et 2ème année de plantation ([Annexe1.b](#)). Le même constat est observé chez les jeunes plantations de cultivars Diva, I45/51, Raspalje et Koster, bien que leurs valeurs de probabilité d'appartenance soient moins élevées que celles du cultivar San Martino.

Pour les plantations âgées affichant une probabilité d'appartenance faible, une explication possible pourrait être un changement de l'état de couverture végétale survenu entre la date d'acquisition des données Sentinel et la date de plantation indiquée: souvent dû à une coupe intervenue entre ces deux dates.

A titre d'exemple, la valeur de probabilité d'appartenance du cultivar Albelo est faible à l'âge de 10 et 11 ans avec une moyenne inférieure à 40 %. Une photo-interprétation utilisant la BD Ortho de 2019 a révélé que, pour certaines parcelles situées dans le département 10, des coupes auraient probablement eu lieu entre 2017 et 2019.

Un autre exemple avec le cultivar I214 dans le département 47 : l'examen du graphique ([Figure 20](#)) montrant la variation de la probabilité d'appartenance, suggère des valeurs inférieures à 40 % pour une plantation âgée de 8 ans.

Incertitude de classification selon l'âge (tous cultivars confondus) pour le cultivar I214

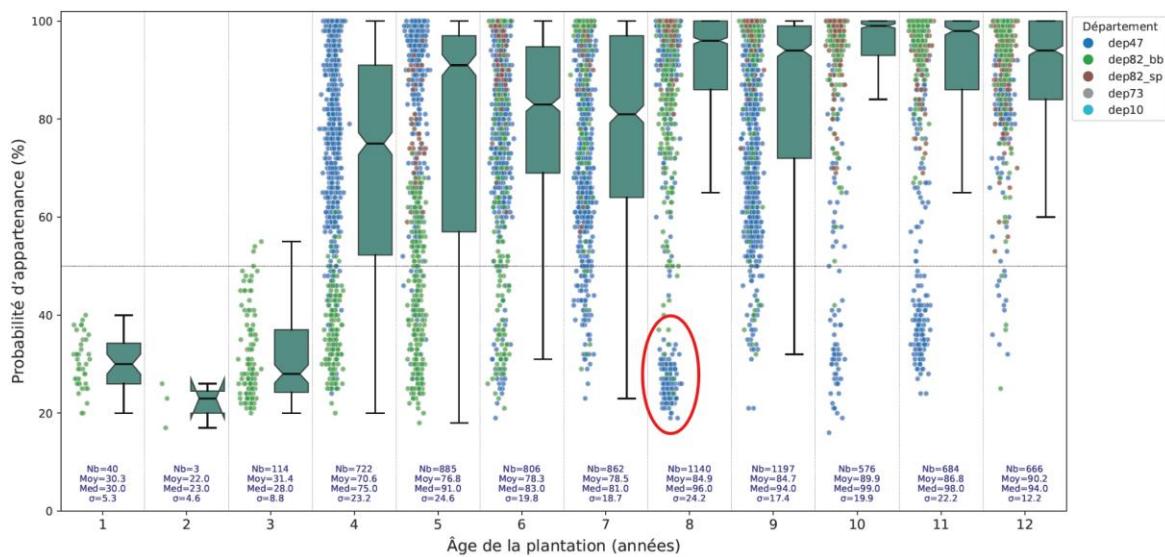


Figure 20 : Pixels à faibles valeurs de probabilité de classification à l'âge de 8 ans pour le cultivar I214

Pour vérifier l'origine de ces pixels, nous les avons visualisé avec l'application ([Figure 21](#)) et ensuite analysé leur apparence à l'aide de la photo-interprétation de 2017.

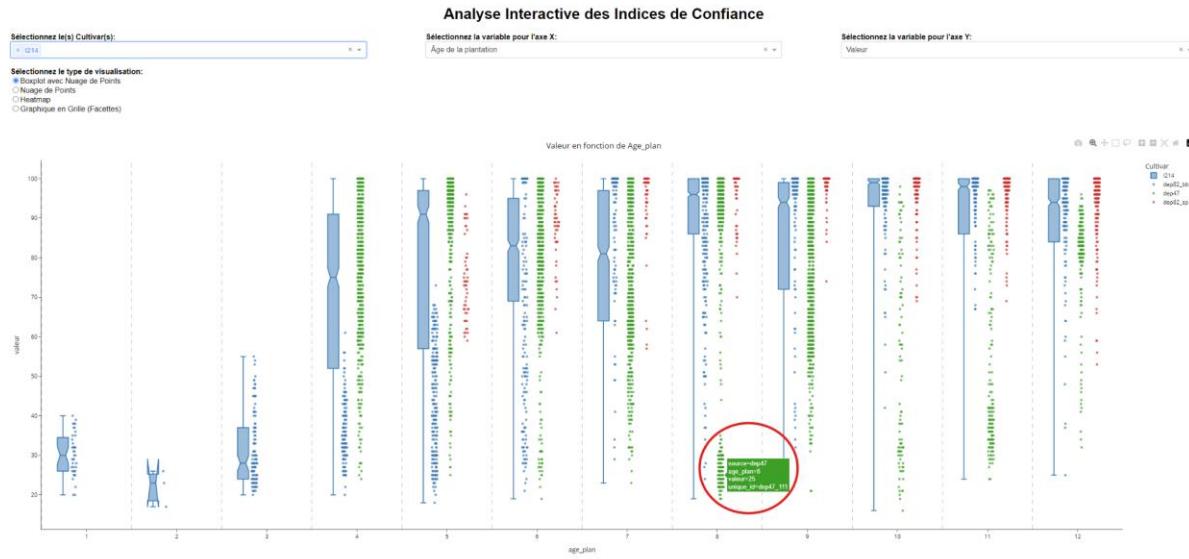


Figure 21 : Capture d'écran de l'affichage sur l'application de pixels à faibles valeurs de probabilité de classification pour le cultivar I214 à l'âge de 8 ans

La photo-interprétation ([Figure 22](#)) a révélé l'existence d'une parcelle (dep 47-111) implantée depuis 2010, mais présentant des plantations qui semblent beaucoup plus jeunes, avec de très faibles valeurs de probabilité de classification.

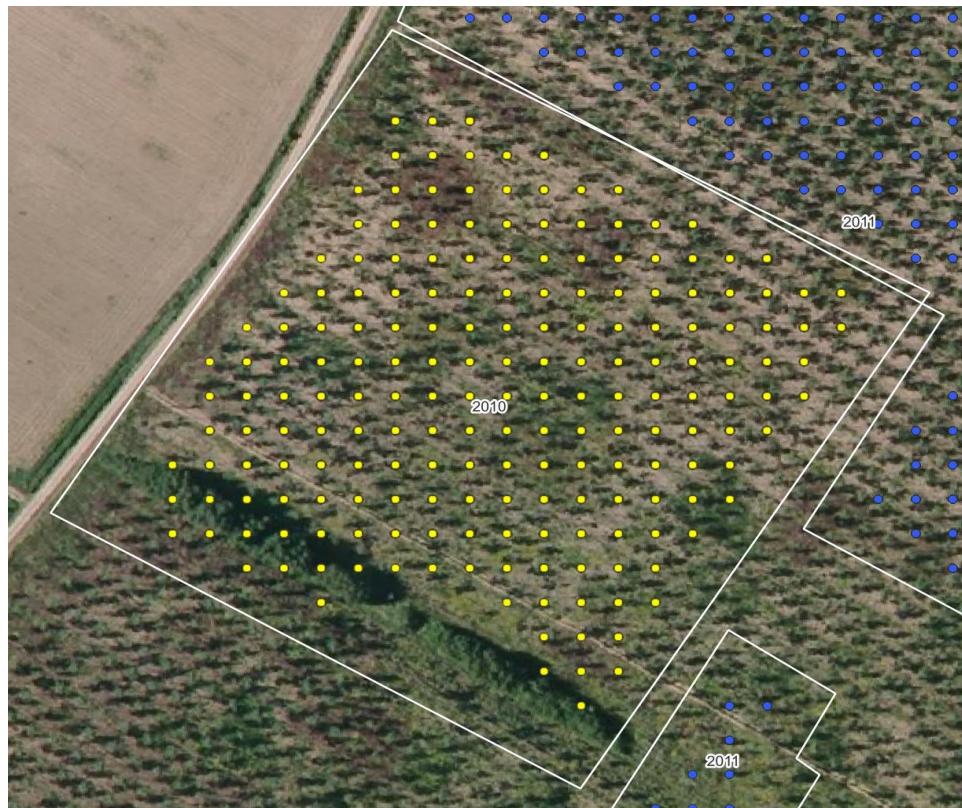


Figure 22 : Capture d'écran de la parcelle dep47-111 avec des faibles valeurs de probabilité d'appartenance à l'âge de 7ans, Ortho Photo 2017

4.6. Corrélation entre les variables métriques CC, PAI et la probabilité d'appartenance

Une corrélation positive se manifeste visuellement entre ces trois variables, indiquant que les variations des variables **CC** et **PAI** influencent directement la probabilité d'appartenance. A titre d'exemple, cinq ans après la plantation, une nette augmentation des valeurs des variables **CC** et **PAI** est observée, entraînant une progression significative de la probabilité d'appartenance. Dans la même optique, une légère régression des deux variables métriques est à relever en année 7 entraînant aussi une baisse modérée de la probabilité d'appartenance ([Figure 9](#)).

Toutefois, ce scénario est attendu : plus l'âge de la plantation avance, plus la plante se développe en hauteur et largeur, meilleure est la détection.

Les tests de corrélations sur les données, dont l'âge de plantation varie entre 4 et 11 ans, confirment l'analyse visuelle : les variables présentent une corrélation positive significative avec une p-value inférieure à 2,2e-16. Les tests ont été effectués sur ces données car nous considérons que c'est surtout pendant la période de croissance active de la plante que ces variables évoluent de manière simultanée:

- Les variables CC et PAI sont fortement corrélées avec un coefficient de corrélation estimé à **0,64** avec la méthode de Kendall car les données n'étant pas normalement distribuées.
- Pour la corrélation entre la probabilité d'appartenance et le Canopy cover, ainsi qu'entre la probabilité d'appartenance et la variable métrique PAI, les coefficients tau de Kendall sont respectivement estimés à **0,24** et **0,34**. Ces résultats suggèrent une corrélation positive modérée, bien que statistiquement significative.
- La corrélation entre âge de plantation et le Canopy cover, ainsi qu'entre âge de plantation et la variable métrique PAI, présente des coefficients tau de Kendall de **0,24** et **0,36** respectivement.

Ces résultats sont cohérents compte tenu de la faible quantité des données, de l'hétérogénéité spatiale des échantillons étudiés et de l'absence d'informations précises concernant d'éventuels changements survenus dans certaines parcelles, ce qui pourrait avoir entraîné des données erronées.

4.7. Limites de l'étude

Cette étude présente des limites qu'il est important de soulever:

- La taille de l'échantillon :

Les données retenues pour répondre à nos objectifs sont plutôt faibles comparées à la quantité de données initiales. C'est surtout le cas des métriques LiDAR où certaines tranches d'âge manquent de données. Cela nécessite une attention particulière pour l'interprétation des résultats parce que:

- La variation de la moyenne des variables (métrique ou probabilité d'appartenance) est plus sensible.
- L'évolution annuelle des variables présente un manque de continuité en raison de l'absence de données.

- La qualité des données :

L'absence d'information sur le calendrier de la rotation culturale (coupe et de replantation) sur la parcelle entraîne une incohérence entre l'âge de plantation, la valeur de la probabilité d'appartenance et les différentes valeurs des variables LiDAR. Cette situation complique le nettoyage des données et peut avoir une influence sur les résultats. Malgré nos efforts pour éliminer ces erreurs, certaines données erronées ont pu subsister et n'ont pas pu être entièrement écartées. Toutefois, ces limitations n'affectent pas la robustesse des résultats globaux, qui demeurent concluants.

- Facteurs exogènes et endogènes :

Des facteurs extérieurs, tels que le climat et le type de sol n'ont pas été pris en compte dans cette étude, bien qu'ils affectent considérablement le rythme de croissance des végétaux. Par ailleurs, les traits fonctionnels spécifiques à chaque type de cultivar n'ont pas été intégrés dans l'analyse. Pourtant, ces éléments influencent les variations observées dans la détection précoce (ou tardive) de certaines essences.

5. Conclusion

Cette étude a permis de mieux comprendre la variation de la probabilité d'appartenance des pixels à la classe peuplier, en lien avec l'âge des plantations, les types de cultivars et leurs caractéristiques structurelles. Elle révèle une amélioration progressive de la classification des plantations avec l'âge, notamment après 5 ans, âge à partir duquel la classification est jugée fiable parce que les traits structurels tels que le développement de la canopée et la surface foliaire deviennent plus marqués contribuant ainsi à une bonne détection.

Les métriques LiDAR, comme le Canopy Cover (CC) et le Plant Area Index (PAI), ont également montré une corrélation positive avec l'âge des plantations. Le PAI connaît une forte augmentation après 4 ans avant de stagner, tandis que le CC affiche une progression continue jusqu'à 11 ans. Ces progressions contribuent à une meilleure classification pour les plantations plus âgées.

L'étude a également mis en évidence des variations selon les cultivars : certains nécessitent une maturité plus avancée pour garantir une détection optimale, tandis que d'autres affichent des bons résultats beaucoup plus tôt.

L'analyse statistique a confirmé une corrélation positive modérée entre les métriques LiDAR et les probabilités d'appartenance. Les facteurs non pris en compte dans l'étude, comme les caractéristiques spécifiques propres à chaque cultivar, cadre climatique, pédologique et changements non documentés dans la gestion des parcelles (coupes, plantation, désherbage) ont engendré des incohérences qui ont potentiellement influencé nos résultats.

Malgré des limites liées à la qualité et à la quantité des données (manque de données pour certaines tranches d'âge et absence d'informations sur les rotations culturales), les résultats obtenus sont robustes et apportent des perspectives intéressantes pour l'amélioration des modèles de classification. Ils ouvrent la voie à des recherches futures intégrant des facteurs exogènes tels que le climat et les types de sols, ainsi que des traits spécifiques aux cultivars, afin d'optimiser la détection des peupleraies.

6. Annexes

6.1. Liens d'accès vers les fichiers PDF contenant tous les graphiques

Ci-dessous vous trouverez le lien vers le pdf combinant l'ensemble des graphiques générés

- a. Boxnotch de la probabilité d'appartenance selon l'âge pour tous cultivars confondus.

https://github.com/caslumali/UE1002_Peupleraie/blob/main/results/1_confidenceXage/Boxnotch_confidenceXage_all.pdf

- b. Boxnotch de la probabilité d'appartenance selon l'âge par cultivar.

https://github.com/caslumali/UE1002_Peupleraie/blob/main/results/1_confidenceXage/Boxnotch_confidenceXage_top_cultivars.pdf

- c. Six facettes présentant les box notches de la probabilité d'appartenance selon l'âge, par année et par cultivar.

https://github.com/caslumali/UE1002_Peupleraie/blob/main/results/2_confidenceXage_par_annee/Boxnotch_confidenceXage_par_annee.pdf

- d. Grid vertical avec le boxnotch des métriques LiDAR en fonction de l'âge pour tous cultivars confondus

https://github.com/caslumali/UE1002_Peupleraie/blob/main/results/4_confidenceXage_lidar_metrics/Boxnotch_confidenceXage_lidar_metrics_all.pdf

- e. Boxnotch des métriques LiDAR en fonction de l'âge par cultivar

https://github.com/caslumali/UE1002_Peupleraie/blob/main/results/3_lidar_metrics/Boxnotch_lidar_metrics_top_cultivars.pdf

- f. Grid vertical avec le boxnotch de la probabilité d'appartenance et des métrique LiDAR selon l'âge de la plantation par cultivar

https://github.com/caslumali/UE1002_Peupleraie/blob/main/results/4_confidenceXage_lidar_metrics/Boxnotch_confidenceXage_lidar_metrics_top_cultivars.pdf

6.2. Liens d'accès vers les scripts

Vous pouvez accéder via le lien GitHub ci-dessous à l'ensemble de scripts créés pour réaliser les différentes étapes décrites dans la partie méthodes et matériels :

https://github.com/caslumali/UE1002_Peupleraie/tree/main/scripts

6.3. Structure finale de l'arborescence du projet

```
UE1002_703_ProjetPeupleraie/
├── data_brut/
│   ├── raster/
│   │   ├── 2017/
│   │   ├── 2018/
│   │   ├── 2019/
│   │   ├── 2020/
│   │   ├── 2021/
│   │   ├── 2022/
│   └── stacks/
        ├── confidence/
        └── lidar/
    └── vector/
├── data_final/
    ├── raster/
        ├── confidence/
        └── lidar/
    ├── tableaux/
        └── analyses/
    └── vector/
└── rapport/
    ├── diagramme/
    └── images/
└── results/
    ├── 1_confidenceXage/
    ├── 2_confidenceXage_par_annee/
    ├── 3_lidar_metrics/
    └── 4_confidenceXage_lidar_metrics/
└── scripts/
```

7. Références bibliographiques

Dawson T.P., Curran P.J., North P.R.J., Plummer S.E. (1999). The propagation of foliar biochemical absorption features in forest canopy reflectance: a theoretical analysis. *Remote Sensing of Environment* 67(2): 147–159. [http://dx.doi.org/10.1016/S0034-4257\(98\)00081-9](http://dx.doi.org/10.1016/S0034-4257(98)00081-9).

Pimmasarn, S.; Tripathi, N.K.; Ninsawat, S.; Sasaki, N. Applying LiDAR to Quantify the Plant Area Index Along a Successional Gradient in a Tropical Forest of Thailand. *Forests* 2020, 11, 520. <https://doi.org/10.3390/f11050520>

Peupliers de France. (Août, 1999). *Cultivar REP Flevo*. [FE130P51_56.PDF](#)

Peupliers de France. (Décembre, 1998). *Cultivar REP I 45/51*. [fe125p44.PDF](#)