

Université de Toulouse

MASTER GEOMATIQUE

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

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

Projet de télédétection



**ABRAHAM Pierre
JOUANNO Léo
LAMGHARI Mohamed
RIGAUD Axel**

Introduction	3
I. Description de la méthode	3
I.1 Construction d'un masque à partir de la BD forêt V2	3
Figure 1 : Diagramme de flux construction du masque	3
I.2 Sélection des échantillons	3
I.3 Analyse des échantillons	4
I.3.1 Nombre d'échantillons	4
Figure 2 : Diagramme de flux général de l'analyse du nombre d'échantillon	4
I.3.2 Phénologie des peuplements purs	5
Figure 3 : Diagramme de flux analyse des peuplements	5
I.3.3 Analyse de la variabilité spectrale de la BD forêt	6
Figure 4 : Diagramme de flux analyse de la variabilité spectrale	6
I.3 Pré-traitement des images	7
Figure 5 : Diagramme de flux pré-traitement des images	7
Figure 6 : Diagramme de flux création image NDVI	8
I.4 Production d'une essence forestière à l'échelle du pixel	9
Figure 7 : Diagramme de flux production d'une essence forestière à l'échelle du pixel	9
II. Résultats	11
II.1 Analyse des échantillons	11
II.1.1 Distribution générale des essences	11
Figure 9 : Histogramme du nombre de polygones par classe	11
Figure 10 : Histogramme du nombre de pixels par essences d'arbres	12
Figure 11 : Violin plot de la distribution du nombre de pixel par polygone pour chaque essence d'arbres	13
II.1.2 Phénologie des essences pures	13
II.1.3 Comparatif de la variabilité spectrale	14
II.2 Analyse de la carte des essences	16
II.2.1 Analyse cartographique de la classification	16
II.2.2 Analyse des scores de précision	17
III. Discussions	18
III.1 Retour sur la méthode de classification et de validation	18
III.2 Performances du modèle	18
III.3 Analyse des caractéristiques des essences	19
Conclusion	19

Introduction

La BD Forêt V2 produite par l'IGN permet de renseigner les essences forestières sur le territoire métropolitain. Élaborée entre 2007 et 2018 par photo-interprétation d'images en infrarouge couleurs de la BD ORTHO, cette dernière n'est donc pas mise à jour régulièrement et il est intéressant d'étudier si elle peut être utilisée dans le cadre d'une classification supervisée. Pour cela, nous avons utilisé une série temporelle d'image Sentinel-2, datant de début 2022 à début 2023, téléchargées préalablement sur le site de Théia. Il y a 6 images composées de 10 bandes chacune, une en hiver, deux au printemps, deux en automne et enfin une en été, représentant un total de 60 fichiers différents. L'objet de ce projet est donc d'évaluer l'employabilité de produite par l'IGN en tant que données de référence dans le but d'identifier les différentes essences végétales d'une série temporelle d'images. Une des consignes était d'utiliser le moins de bibliothèques python possible, c'est pour cela que seulement les bibliothèques Numpy, Pandas, Geopandas, Os, Gdal, RasterStats, Scikit-Learn et Matplotlib seront utilisées, en plus de fonctions déjà fournies. Nous verrons dans un premier temps la méthode que nous avons utilisée, suivie de nos résultats. Enfin, nous analyserons ces derniers.

I. Description de la méthode

I.1 Construction d'un masque à partir de la BD forêt V2

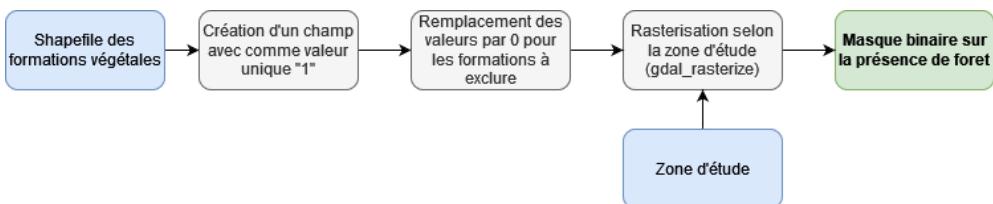


Figure 1 : Diagramme de flux construction du masque

Avant d'analyser les images Sentinel 2, il est nécessaire d'élaborer un masque arboré afin de masquer tout pixel n'étant pas considéré comme de la forêt. Pour cela, nous avons ajouté un nouveau champ "binaire" avec comme valeur unique "1" au shapefile des essences végétales de la BD Forêt V2 à l'aide de la librairie "GeoPandas". Une liste répertoriant les codes TFV propres à chaque formation végétale à exclure a ensuite été créée et la valeur du champ "binaire" des polygones concerné par ces codes a été remplacée par 0. Ces polygones ont ensuite été rasterisés avec une fonction basée sur la commande Gdal "gdal_rasterize", en conservant uniquement les pixels situés dans l'emprise de notre zone d'étude.

I.2 Sélection des échantillons

Pour la sélection des échantillons, nous avons construit deux nouveaux champs, "code" et "nom" à partir du shapefile de la BD Forêt V2 (ESSENCE_VEGETALES.shp), chargé dans un géodataframe. Nous avons ensuite construit un dictionnaire manuellement, en associant chaque code TFV aux code et nom correspondant renseignés dans l'objectif de ce projet.

Nous avons ensuite mappé chaque valeur du dictionnaire au code TFV correspondant. Le geodataframe a ensuite été enregistré afin d'être mobilisé pour la suite du ce projet (Sample_BD_foret_T31TCJ.shp).

I.3 Analyse des échantillons

I.3.1 Nombre d'échantillons

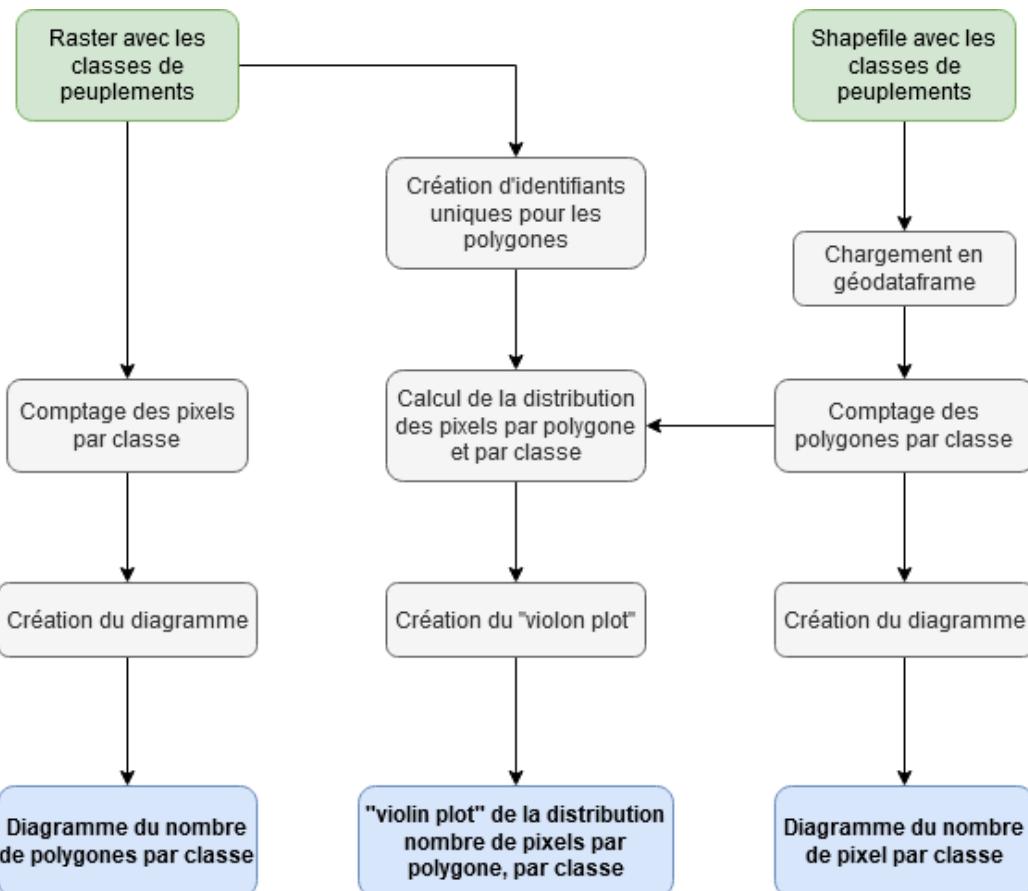


Figure 2 : Diagramme de flux général de l'analyse du nombre d'échantillon

Pour l'analyse et la visualisation des statistiques des polygones et pixels des échantillons par classe, nous avons tout d'abord chargé les données vectorielles des essences (Sample_BD_foret_T31TCJ.shp) dans un géodataframe. Nous avons ensuite compté le nombre de polygones par classe à l'aide de la fonction Pandas ".value_counts()". Puis nous avons créé un histogramme du nombre de polygones par classe avec la bibliothèque python "MatPlotLib".

Après cela nous avons effectué l'analyse du nombre de pixels par classe en rasterisant tout d'abord les échantillons obtenus au point précédent (Sample_BD_foret_T31TCJ.tif) puis en les chargeant dans un tableau numpy. Cela nous permet de compter pour chaque valeur de classe unique le nombre de pixel associé, afin d'être stocké dans un array avec "np.unique()". Nous avons ensuite créé dataframe contenant chaque classe de peuplement

et leur nombre de pixels respectifs afin de représenter la distribution des pixels par polygone et par classe à l'aide de Matplotlib.

Enfin pour pouvoir créer le violon plot, nous avons créé une colonne contenant les identifiants unique (unique_id) dans un géodataframe pour chaque polygone échantillon, et nous l'avons rastérisé. Nous avons ensuite repris la même méthode que pour le calcul des pixels par classe, en comptant d'abord le nombre de pixels par identifiant unique, puis en les associant ensuite aux classes respectives de chaque polygone. Cela nous permet enfin de créer le “violon plot” du nombre de pixels par polygone par classe.

I.3.2 Phénologie des peuplements purs

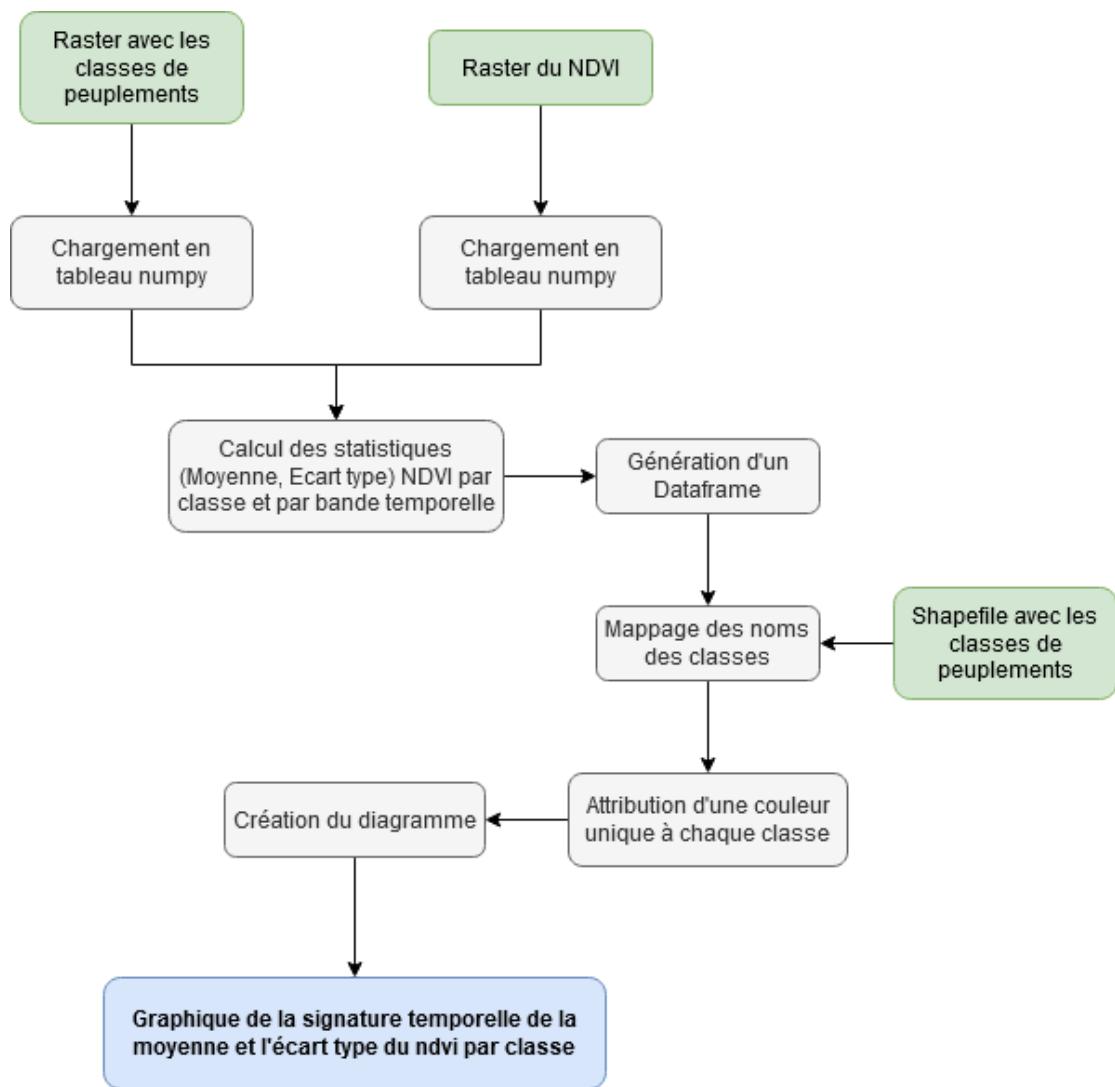


Figure 3 : Diagramme de flux analyse des peuplements

Pour l'analyse des phénologies des peuplements purs, nous avons tout d'abord chargé les rasters dans un tableau numpy. Puis nous avons calculé les statistiques (Moyenne et écart type) du NDVI par classe et pour chacune des 6 bandes temporelles. Par la suite nous avons mappé les codes de classe avec leur nom à l'aide du shapefile, et distinguer chaque classe d'une couleur différente pour une meilleure visibilité sur le graphique. A l'aide du

dataframe des statistiques, nous avons pu ensuite créer le graphique avec les moyennes et écarts type pour chacune des classes et en fonction des 6 dates.

I.3.3 Analyse de la variabilité spectrale de la BD forêt

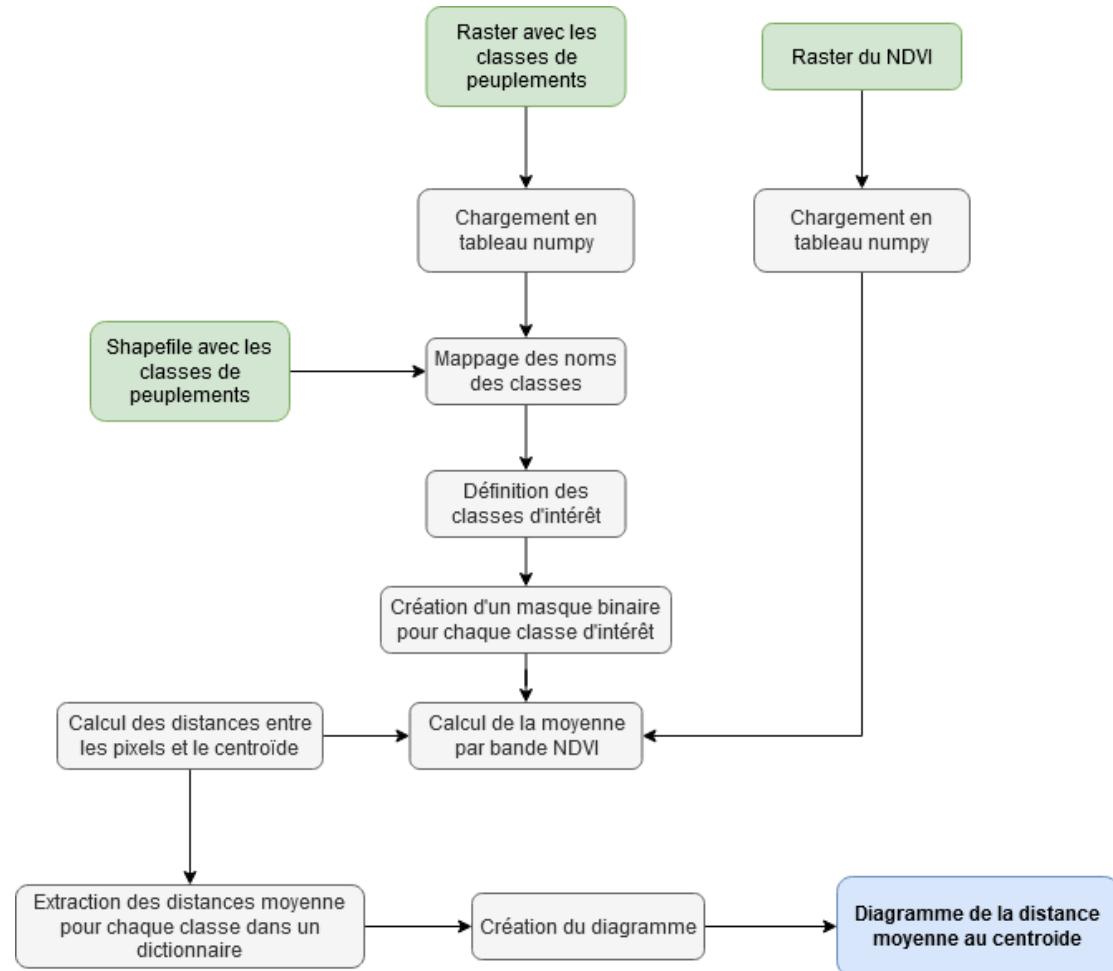


Figure 4 : Diagramme de flux analyse de la variabilité spectrale

Pour l'analyse de la variabilité spectrale de la BD forêt, nous devons calculer des distances moyennes au centroïde par classe. Pour ce faire, nous avons tout d'abord chargé la série temporelle NDVI et le raster des échantillons dans un tableau numpy. Après cela nous avons pu calculer les distances moyennes au centroïde. Pour ce faire, nous avons d'abord défini les codes des classes des peuplements purs et les classes des peuplement en mélange dans deux listes. A partir de leurs codes, nous avons créé un masque binaire basé sur l'ensemble des essences, et calculé la moyenne par bande du NDVI masquée. Enfin nous avons extrait les distances moyennes pour chaque classe dans un dictionnaire. Pour cela, nous avons soustrait la valeur de chaque pixel à la valeur moyenne du NDVI, élevé au carré, pour chaque classe, et crée un histogramme comparatif des deux types de formations végétales en différenciant les couleurs selon l'appartenance à l'une des deux listes afin de faire une analyse comparative entre les peuplements purs et en mélange.

Pour produire la distribution des distances moyennes des pixels par polygones par classes, nous avons appliqué une méthode similaire, mais en ajoutant un second masque, basé sur le raster des identifiant unique des polygones, afin d'appliquer un masque par polygone. Cela permet de stocker dans un dictionnaire les distances au centroïde de chaque polygone pour chacune des classes, afin de réaliser un violin plot.

I.3 Pré-traitement des images

Pour le prétraitement des images, nous avons privilégié une approche fonctionnant avec les chemins de fichier de sorte à automatiser la récupération des 60 images que nous avons téléchargées. Cependant, le code ne fonctionnera pas avec d'autres chemins.

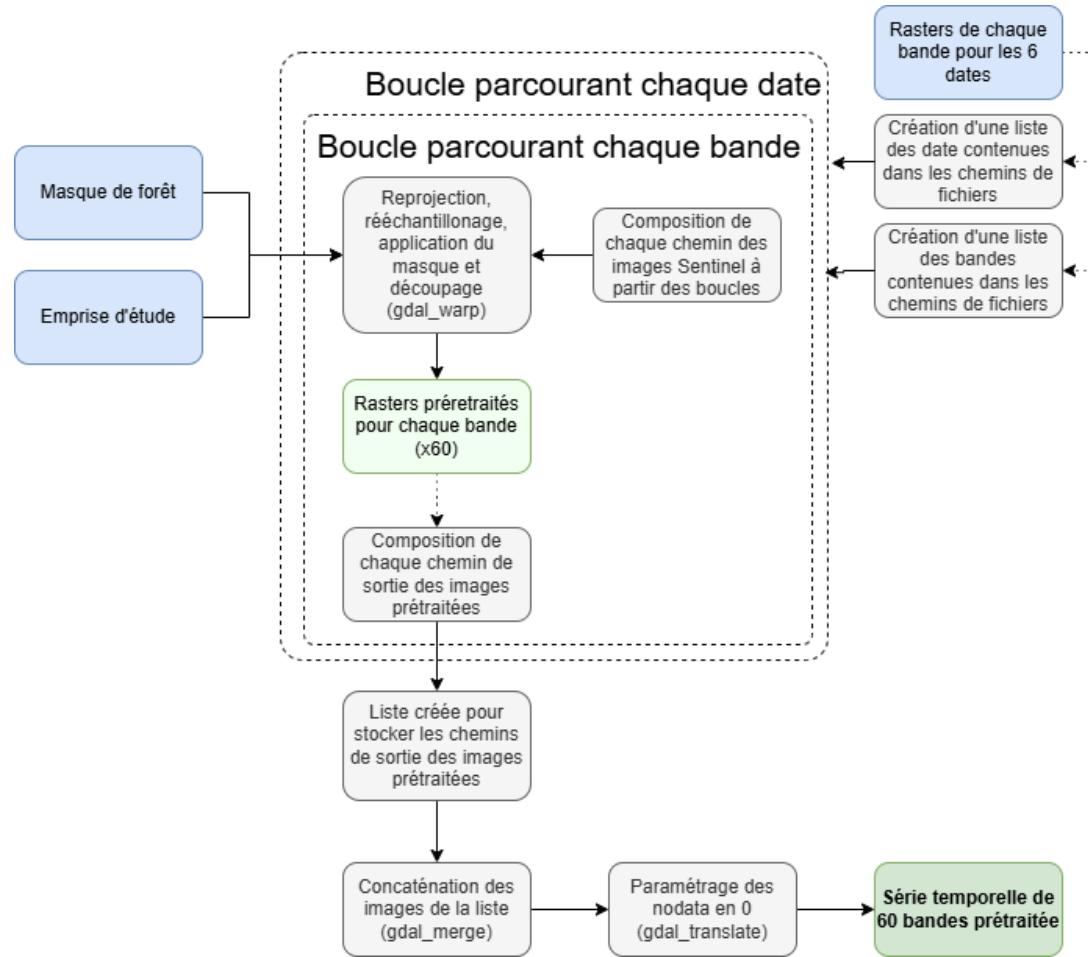


Figure 5 : Diagramme de flux pré-traitement des images

Deux listes sont créées, une contenant la première partie des chemins des fichiers correspondant aux 6 dates différentes et une autre au niveau de chacune des 10 bandes. En faisant une boucle imbriquée, il est alors possible d'obtenir les 60 chemins de fichier. A partir de ces chemins, chaque image est ensuite reprojétée, rééchantillonnée, découpée selon l'emprise et filtrée à partir du masque à l'aide de la commande Gdal "gdal_warp", qui permet de prendre en compte tous ses paramètres. A l'issue des deux boucles, 60 nouveaux raster sont donc créés correspondant à chaque bande prétraitées au format GEOTIFF, avec leur chemin renseigné dans une liste.

Une fonction basée sur la commande “gdal_merge” permet ensuite de concaténer les images à partir de la liste des chemins des images prétraitées puis une commande “gdal_translate” permet enfin de configurer les nodata en 0. Nous avons opté pour cette solution au lieu d’utiliser OTB notamment car nous ne sommes pas parvenus à l’utiliser dans nos script. A l’issue de cette partie, nous obtenons donc une série temporelle à 60 bandes sur notre zone d’études, filtrée par les polygones des essences végétales.

La série temporelle du NDVI est ensuite calculée à partir de cette image. Pour ce faire, l’image prétraitées est convertie en matrice Numpy. Les indices pour les 6 dates des bandes rouges et infrarouges sont ensuite récupérés sous forme de deux listes. Une boucle sur chaque date est ensuite faite pour calculer les 6 NDVI, en récupérant les valeurs des pixels des bandes rouges et infrarouges depuis les listes des indices créés précédemment. Chaque NDVI calculé est ensuite ajouté à une liste, contenant à l’issue de ces opérations 6 matrices numpy. Ces dernières sont ensuite concaténées à l’aide de la fonction dstack de numpy afin d’obtenir la série temporelle du NDVI. Comme il s’agit de calculs sur des rasters, nous avons privilégié cette approche plutôt que d’utiliser la commande “gdal_merge”, qui fonctionne à partir de fichier déjà créé.

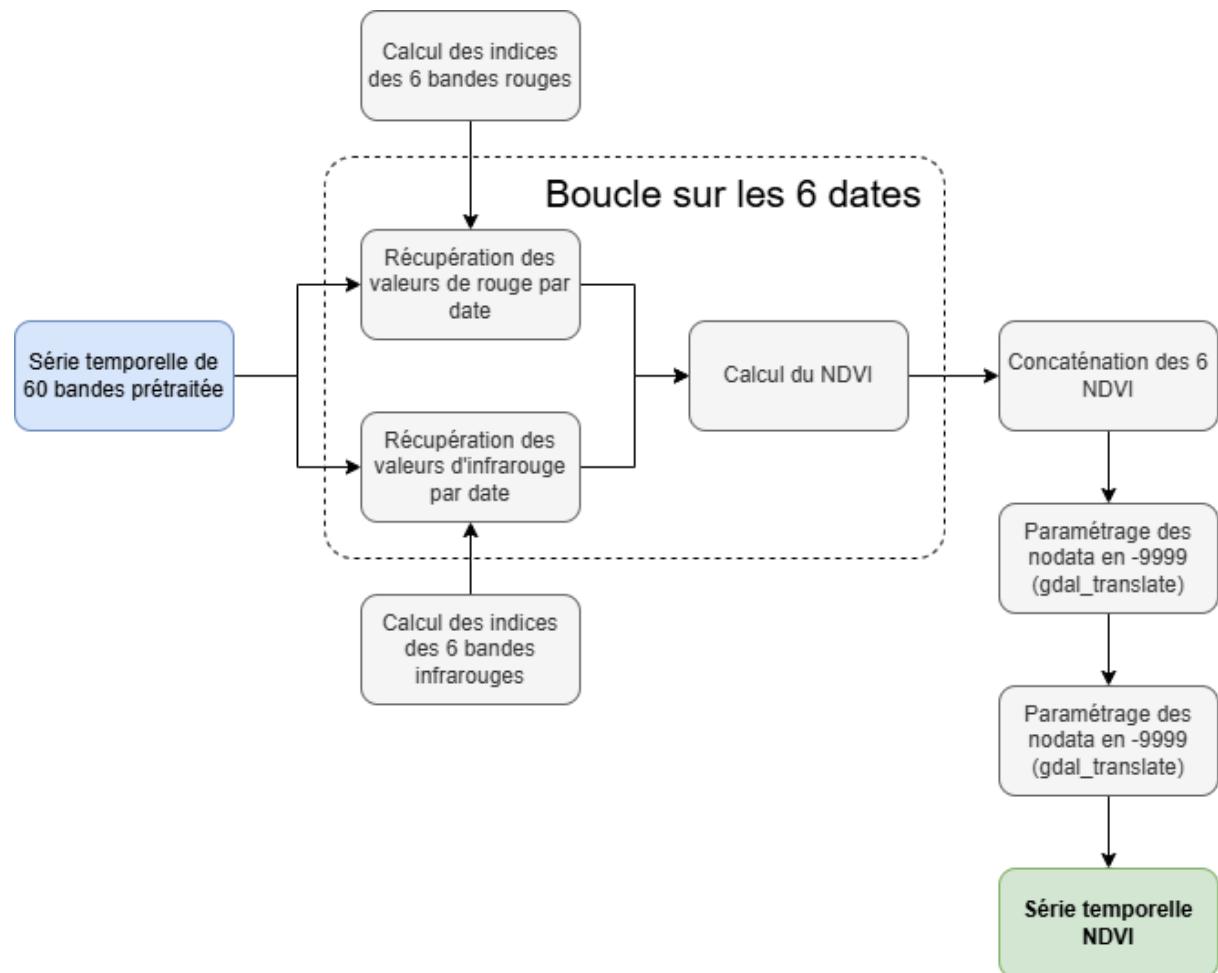


Figure 6 : Diagramme de flux création image NDVI

I.4 Production d'une essence forestière à l'échelle du pixel

Nous avons effectué une classification en utilisant une validation croisée stratifiée en prenant en compte l'appartenance d'un pixel à un polygone, mise en place avec la fonction StratifiedGroupKFold de Scikit-Learn. Cette dernière permet de diviser les données en un nombre de groupes (folds) déterminés (5 dans le cas de ce projet) tout en assurant une répartition représentative de l'ensemble des données selon le type d'essence dans chaque groupe, en prenant en compte le polygone d'appartenance de chaque pixel. Cela permet à des pixels se trouvant dans un même polygone de ne pas être à la fois dans le jeu d'entraînement et de validation car ils risquent d'avoir sensiblement les mêmes valeurs. Dans ce projet, les classes correspondent aux codes des différents peuplements.

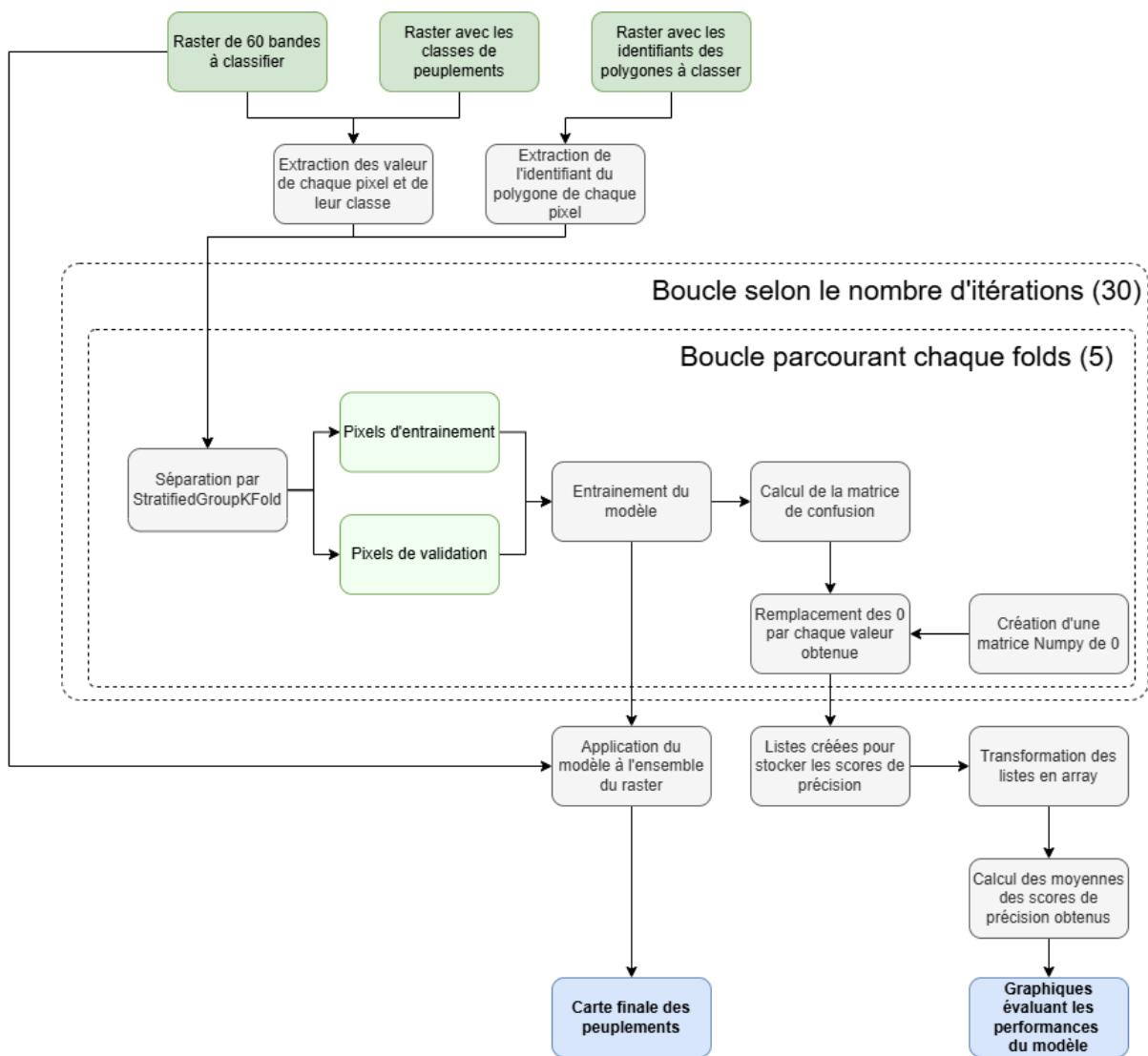


Figure 7 : Diagramme de flux production d'une essence forestière à l'échelle du pixel

Nous avons tout d'abord utilisé la fonction `get_samples_from_roi`. Cette fonction convertit une image dans un format mobilisable par la bibliothèque scikit-learn afin d'effectuer un apprentissage supervisé. Elle permet de renvoyer une matrice où les valeurs d'une image sont retranscrites à travers chaque ligne, correspondant à chaque pixel et chaque colonnes,

correspondant aux bandes de l'image à traiter. Elle renvoie également un array Numpy contenant la classe associée à chaque pixel, extraite depuis un fichier d'échantillon et un tuple contenant les coordonnées des pixels. Pour cela, cette dernière ouvre une image à classifier et une image contenant la région d'intérêt (ROI) tout en vérifiant si leurs dimensions sont identiques avant de poursuivre le traitement. Si ce n'est pas le cas, un message d'erreur est affiché et le programme s'arrête. Si les tailles sont similaires, la fonction extrait les valeurs des pixels de l'image à classifier là où les pixels la ROI sont différents de 0.

Cette fonction nous permet donc de récupérer dans un premier temps sous forme d'une matrice Numpy les pixels à classifier et leur classe correspondant aux codes des différents peuplements sous forme d'array Numpy. Cette fonction est de nouveau exécutée pour récupérer cette fois les identifiant des polygones d'appartenance des pixels à la place des classes de peuplement.

Une boucle s'effectue selon le nombre d'itérations choisi (30 dans notre cas) et répète ainsi la fonction StratifiedGroupKFold. Nous avons initié en amont des listes vides afin d'y stocker les matrices de confusions et le score de précisions pour chaque itération. Une seconde boucle s'effectue sur le nombre de groupes et permet d'entraîner le modèle RandomForestClassifier de la bibliothèque python Scikit-Learn avec les hyperparamètres spécifiés dans l'énoncé du projet. Il s'agit d'une technique d'apprentissage supervisée basée sur un ensemble d'arbres de décision. Une matrice de confusion et le score de précision sont également calculés à ce moment. Cependant, les distributions des classes étant très hétérogène, certaines ne figuraient pas toujours dans les folds, notamment la classe 22 composée de moins de 4000 pixels contre plus de 2200000 pour la classe 12. Nous avons alors généré une matrice Numpy de 0 pour chaque fold avec comme dimension le nombre de classes uniques, au sein de laquelle les véritables valeurs ayant pu être calculées par la matrice de confusion sont intégrées. Cela permet dans le cas où une classe n'apparaît pas dans un fold de garder des matrices de confusion de tailles identiques pour le calcul de la moyenne.

Une fois les itérations effectuées, les listes contenant les matrices de confusion et les scores de précision sont transformées en array Numpy afin d'y calculer les valeurs moyennes puis différents graphiques sont réalisées. Cela permet d'avoir une appréciation plus globale de la qualité du modèle qu'avec une unique itération.

Enfin, le modèle est appliqué à l'ensemble de l'image, en renseignant les valeur de la prédiction aux coordonnées de chaque pixel contenu dans le tuple que renvoie la fonction get_samples_from_roi.

II. Résultats

II.1 Analyse des échantillons

II.1.1 Distribution générale des essences

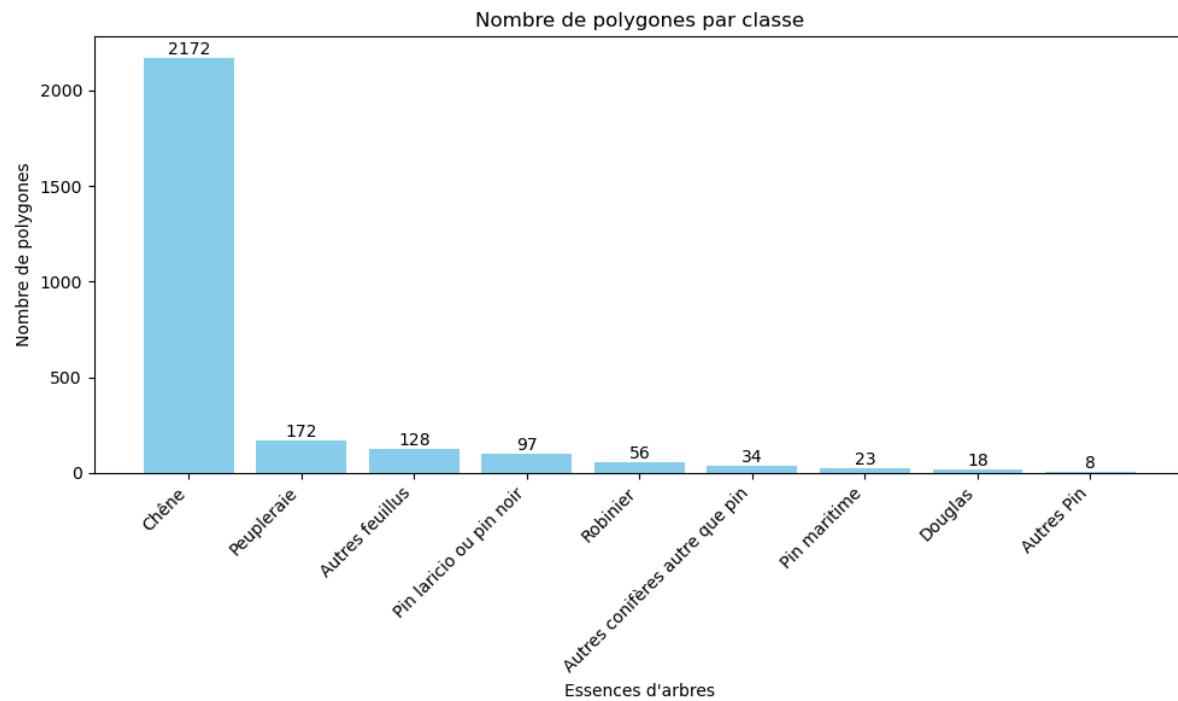


Figure 9 : Histogramme du nombre de polygones par classe

La figure 9 ci-dessus montre le nombre de polygones par classe d'essences d'arbres. Nous pouvons remarquer que la classe chêne est surreprésentée par rapport aux autres classes avec 2172 polygones contre seulement 8 pour la classe "Autres Pin".

La figure 8 précise ce phénomène en indiquant le nombre de pixels par classe d'essences d'arbres. Comme pour la Figure 9, la classe chêne est surreprésentée par rapport aux autres classes, mais de manière encore plus flagrante dans cette figure, avec plus de 2 millions de pixels contre près de 71000 pour la deuxième classe la plus représentée, à savoir les "Autres feuillus". Les "Autres Pins" quant-à eux représentent moins de 3900 pixels, faisant de cette catégorie de peuplements la classe la moins bien représentée.

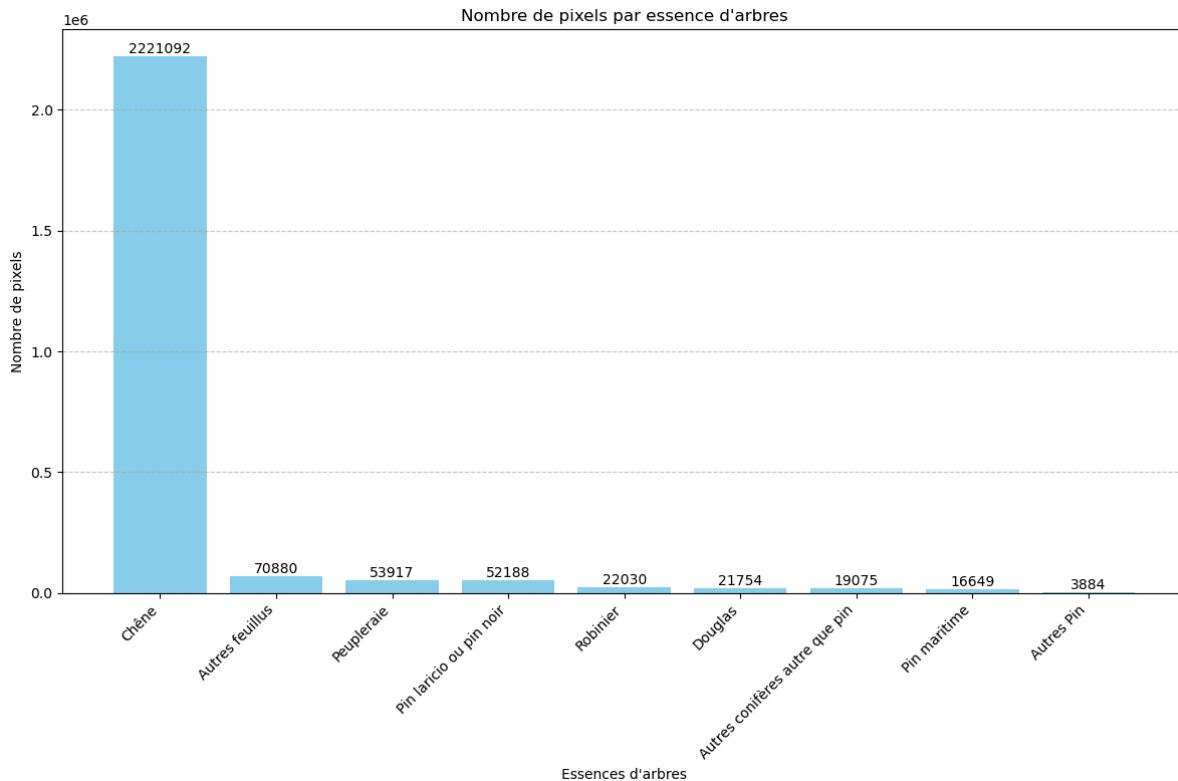


Figure 10 : Histogramme du nombre de pixels par essences d'arbres

Nous avons ensuite produit pour la figure 11 un violin plot qui représente la distribution du nombre de pixel par polygone pour chaque essence d'arbres. Nous avons représenté la moyenne pour chaque classe avec un trait noir et la médiane avec un trait gris pour pouvoir les distinguer. Ce diagramme permet pour chaque classe d'identifier la distribution (à la verticale) de la taille des polygones qui la compose. Plus le violon plot d'une classe sera large en haut, plus elle contiendra de grands polygones, composés de nombreux pixels.

Nous avons fait le choix de représenter ces valeurs sur une échelle logarithmique puisqu'un polygone de la classe 12 est composé de 120000 pixels et "écrase" le reste des classes. Nous pouvons constater graphiquement que les essences de chênes et de douglas ont les plus hautes moyennes, supérieures à 1000 pixels. A l'inverse, les robiniers et les autres pins ont des polygones composés d'en moyenne 600 pixels.

Le violin plot met aussi en évidence une grande variabilité intra-classe par rapport à la taille de chaque forêt. En effet, le chêne présente des forêts hétérogènes comme en témoigne la hauteur de son violin plot, en prenant toujours en compte l'échelle logarithmique, avec des polygones allant de quelques dizaines à plusieurs milliers de pixels allant de petites forêts fragmentées à de vastes massifs homogènes. En revanche, les classes plus rares ont des distributions beaucoup plus compactes, comme les autres pin.

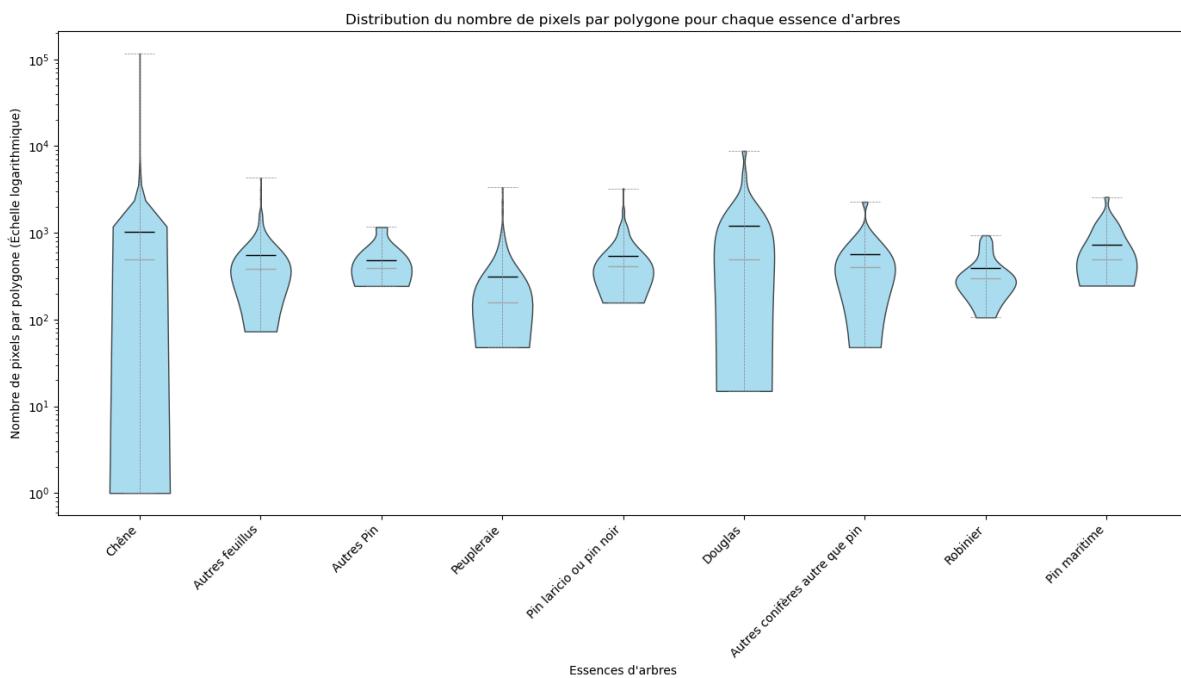


Figure 11 : Violin plot de la distribution du nombre de pixel par polygone pour chaque essence d'arbres

II.1.2 Phénologie des essences pures

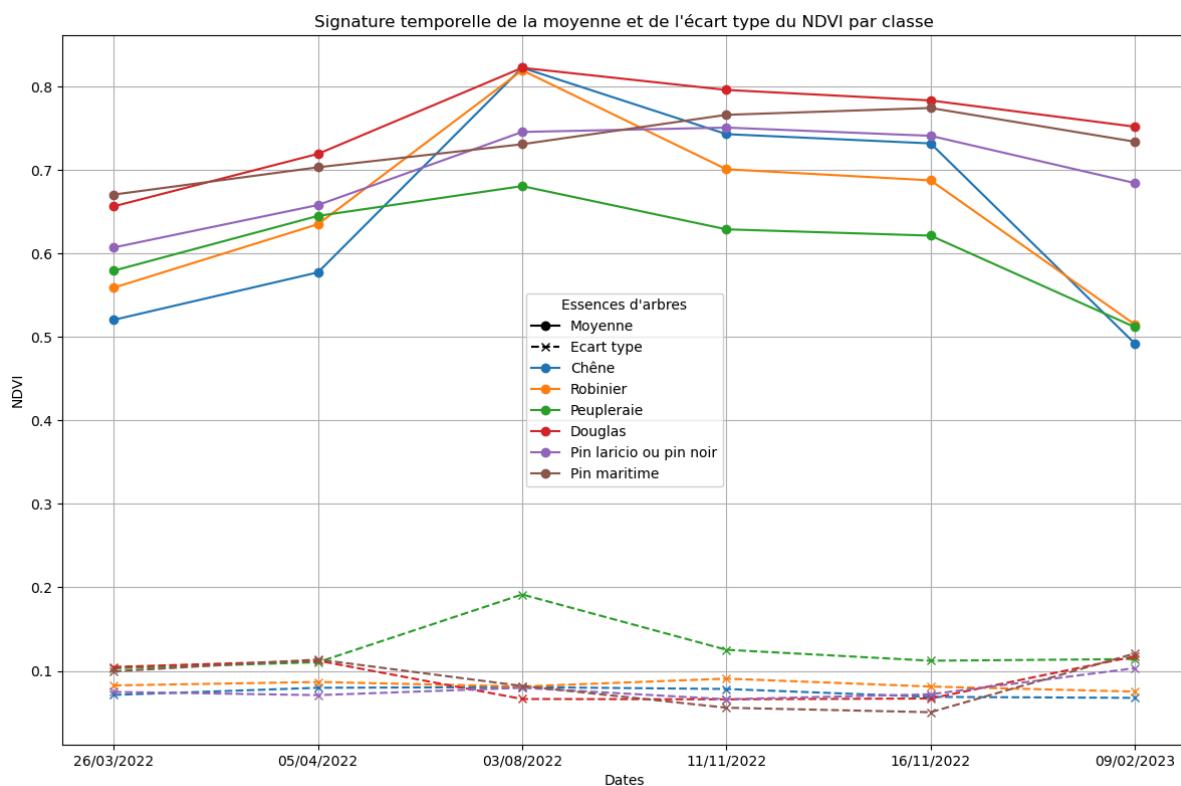


Figure 12 : Signature temporelle de la moyenne et de l'écart type du NDVI par classe

Nous avons ensuite produit pour la figure 12 un graphique qui représente la Signature temporelle de la moyenne et de l'écart type du NDVI par essences d'arbres pour chacune des 6 dates. Nous pouvons tout d'abord constater une croissance générale du NDVI jusqu'à la troisième date, en été, suivie d'une diminution jusqu'à la dernière date, en hiver. On observe ainsi que le NDVI atteint son maximum en août 2022, avec des valeurs autour de 0,82 pour le Douglas. À l'inverse, les valeurs minimales apparaissent en février 2023, où certaines essences chutent à des valeurs proches de 0,5, comme le chêne (0,55) et le robinier (0,58).

Ensuite, nous pouvons voir que la classe Douglas est dans la globalité la classe avec la moyenne de NDVI la plus importante, contrairement à la classe Peupleraie. En termes d'écart type, c'est cette fois-ci c'est la classe Peupleraie qui a globalement les plus fortes valeurs d'écart type et enregistre un pic en été, marquant une certaine hétérogénéité au sein de cette classe. Toutes les autres classes ont sensiblement les mêmes valeurs dans chacune des 6 dates.

II.1.3 Comparatif de la variabilité spectrale

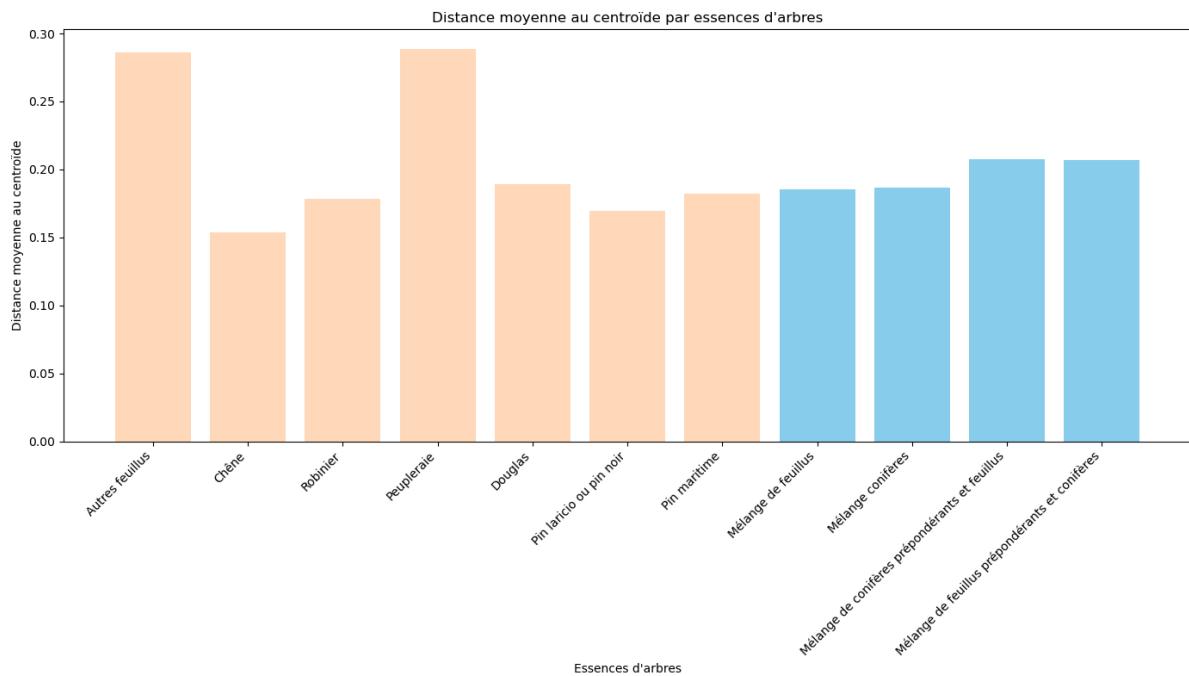


Figure 13 : Diagramme en bâton de la distance moyenne au centroïde par essences d'arbres

Nous avons ensuite produit pour la figure 13 un diagramme en bâton de la distance moyenne au centroïde par essence d'arbres avec les barres jaunes correspondant aux peuplements purs et les barres bleues correspondantes aux peuplements en mélange. Les valeurs représentent la variabilité des valeurs de NDVI par rapport à la moyenne de chacune des classes d'essences d'arbres. Nous pouvons remarquer 2 classes qui se démarquent par rapport aux autres, la classe "Autres feuillus" et la classe "Peupleraie", avec une plus grande variabilité des valeurs (environ 0,28). Toutes les autres classes sont entre 0,15 et 0,20, pour les deux types de peuplements, avec le chêne ayant la variabilité spectrale la plus faible.

Le violin plot de la figure 14 permet de préciser cette analyse. En effet, nous pouvons voir une distribution des écarts au centroïde de chaque polygone par classe sensiblement identique pour les peuplements en mélange, avec une moyenne relativement basse malgré des valeurs extrêmes plus importantes que les peuplements purs. En revanche, la distance au centroïde pour chaque polygone de peuplements purs est hétérogène d'une classe à l'autres, avec une distribution bien plus étalée pour les polygones des autres feuillus par exemple par rapport aux polygones de chênes.

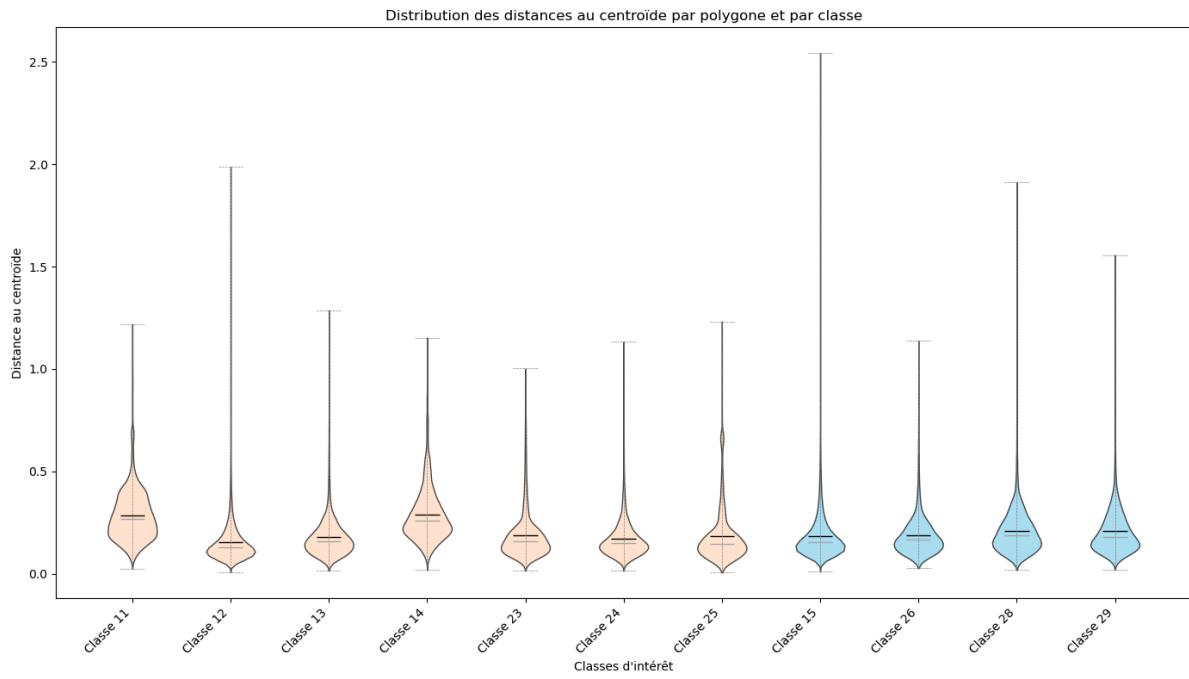


Figure 14 : Violin plot de la distance moyenne au centroïde par polygone par essences d'arbres

II.2 Analyse de la carte des essences

II.2.1 Analyse cartographique de la classification

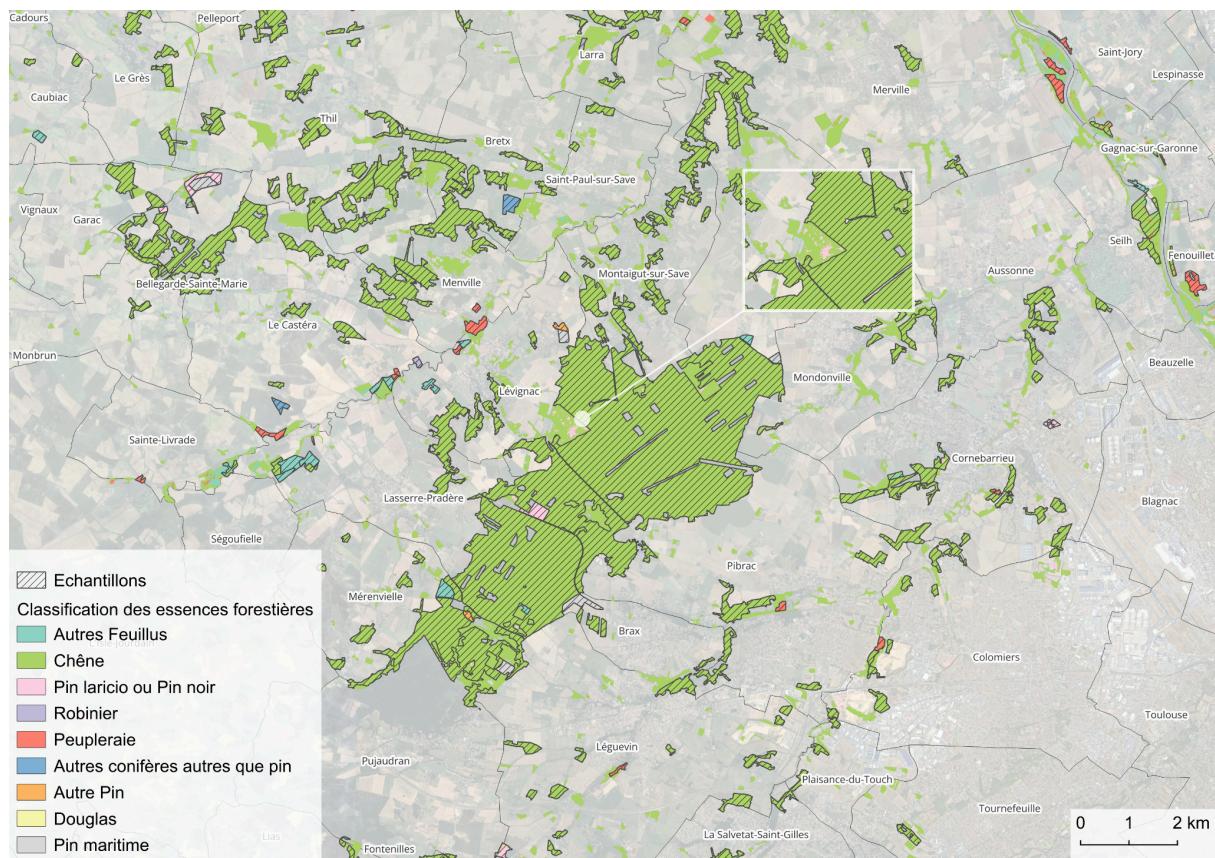


Figure 14 : Carte des essences forestières issue de la classification

La figure 15 montre un extrait de la cartographie des essences forestières, issu de fichier geotiff que génère notre classification. Nous pouvons y voir les différentes essences représentées chacune par une couleur différente, avec une forêt de chênes centrale contrastant avec les taches éparses des autres formations forestières. Nous avons également représenté les échantillons ayant servi pour l'apprentissage du modèle par des hachures car ils englobe moins de classe que le masque arboré appliqué à l'image classifiée tels que les feuillus en îlots, non pris en compte dans la classification à l'échelle du pixel. Nous pouvons alors remarquer qu'une zone au nord-ouest de la forêt centrale, non couverte par les échantillon, présente du bruit, avec des pixels identifiés en tant que "Pin laricio ou pin noir".

II.2.2 Analyse des scores de précision

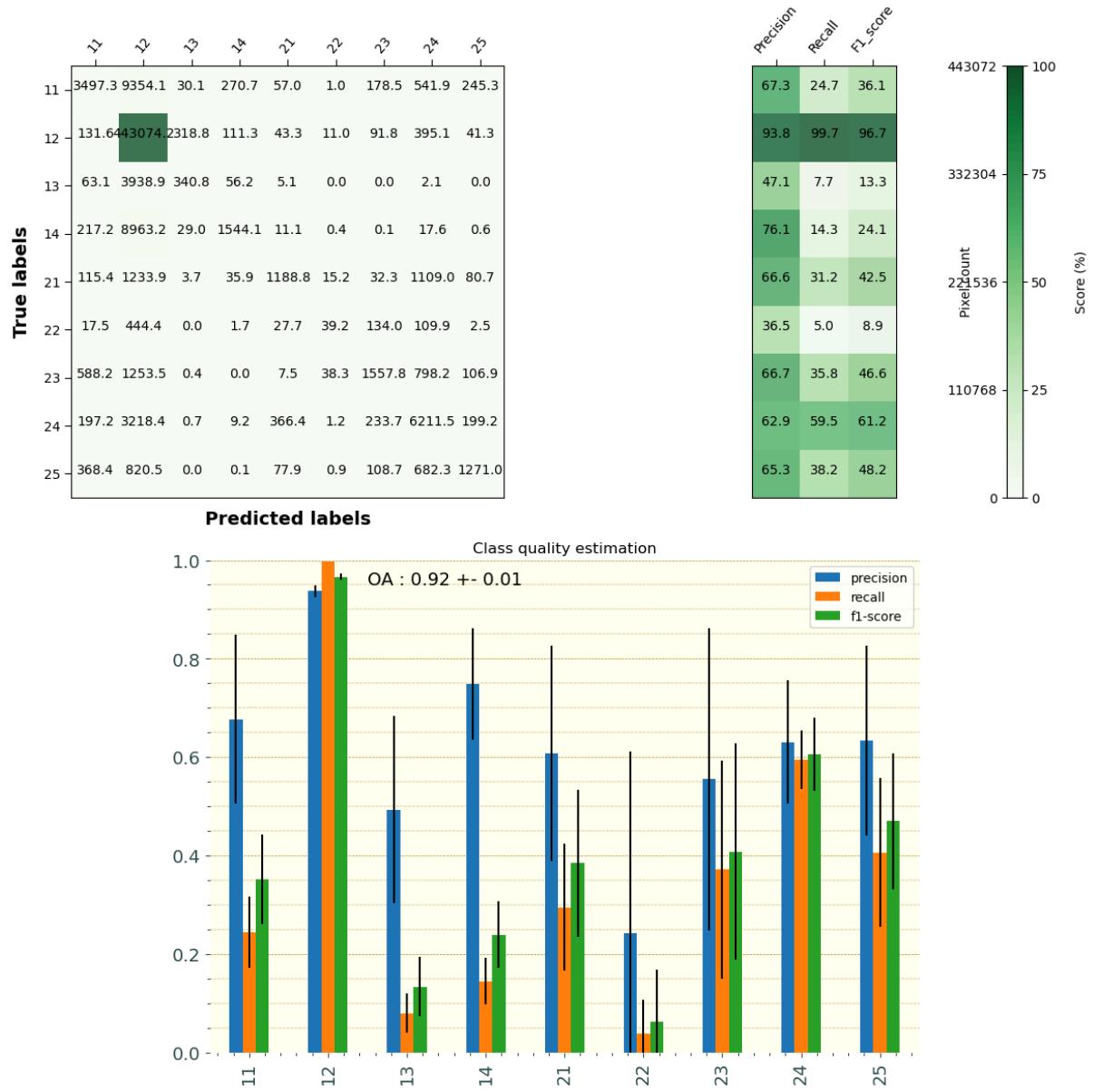


Figure 15 : Matrice de confusion et score de précision issus de la classification

La figure 15 montre la matrice de confusion et les différents scores associés issue de la classification avec une validation croisée stratifiée en prenant en compte l'appartenance d'un pixel à un polygone. Il ressort tout d'abord que seulement la classe 12 et la classe 24 disposent de f1-score dépassant les 50%. A l'inverse, la classe 22 par exemple correspondant au "Autres Pins" a un F1-Score de 8.9%. Ensuite, nous pouvons constater que le rappel de l'ensemble des classes excepté celui de la classe 12 est inférieur à la précision. Enfin, malgré des scores de précision et de rappel bas, l'accord moyen est élevé, de l'ordre de 0.92, mais qu'il faut nuancer en raison de la très grande quantité de pixel de chêne bien classé.

III. Discussions

III.1 Retour sur la méthode de classification et de validation

Comme nous l'avons vu dans nos résultats, la distribution du jeu de données est très hétérogène, avec une classe 12 prédominante, correspondant au peuplements de chênes. De ce fait, nous avons dû faire des choix dans notre stratégie de validation puisque certaines classes sous représentées n'étaient pas prises en compte dans les folds malgré le fait que la fonction "StratifiedGroupKFold" doit conserver une certaine proportion au sein de ces derniers. Cela générât des erreurs lors du calcul des moyennes des scores de précision puisque les matrices de confusion n'avaient pas toutes les mêmes dimensions, ce qui empêchait d'estimer les performances du modèle.

Nous avons alors choisi de conserver les itérations avec des classes manquantes dans les folds, remplaçant alors leur score par 0 au lieu de les éviter puisque cela concernait plusieurs itérations. Nous avons essayé un calcul manuel des performances en générant des matrices de confusion pour chaque itération concernée mais le traitement s'arrêta pour manque de ressource. Ce choix implique que les scores de précision sont faussement bas pour les classes sous représentées puisqu'ils n'ont pu être calculées pour certaines boucles, et donc remplacés par des 0. Il serait alors intéressant d'évaluer la qualité du modèle sur un jeu de données ayant une distribution plus homogène.

Ensuite, nous n'avons pas pu réaliser une classification à l'échelle des peuplements. Il serait intéressant de poursuivre ce travail afin de pouvoir comparer les deux classification et leur score de précision associés.

III.2 Performances du modèle

Comme nous l'avons vu avec les figures 9, 10, il y a une distribution des essence forestière issues de la BD Forêt V2 très hétérogène. En effet, le chêne (classe 12) domine largement les données, avec 2 221 272 pixels (2 172 polygones), soit près de 75 % du total. À l'inverse, certaines classes, comme le Douglas (21 780 pixels - 18 polygones) ou le Pin maritime (16 637 pixels - 23 polygones), sont nettement sous-représentées. Cette disproportion est susceptible d'avoir des conséquences directes sur la classification. En effet, les scores de précision sont très élevés uniquement pour la classe 12, pouvant s'expliquer par le fait que cette dernière concerne la majeure partie des échantillons. Nous pouvions donc nous attendre à ce que le modèle classe efficacement ces pixels, correspondant à des chênes.

Cette distribution peut également expliquer le fait que seule la classe 12 a une précision inférieure au rappel. En effet, la distribution des échantillons est susceptible de favoriser une classification des pixels vers la classe 12, favorisant de nombreux faux positifs à l'intérieur de cette dernière et expliquerait donc aussi le faible score de rappel des autres types d'essences. Il est cependant intéressant de noter que d'autres facteurs influencent la classification car malgré des distributions relativement similaires, les classes "Autres feuillus" et "Pin laricio ou pin noir" ont des rappels très différents (respectivement 24.7 et 59.5). Une poursuite de ce travail pourrait être d'analyser ce qui est susceptible de provoquer de tels écarts. La classe "Peupleraie" a également

Nous pouvons aussi noter que la résolution spatiale des images Sentinel-2 (10m) peut éventuellement limiter la précision de l'analyse, notamment si plusieurs espèces coexistent dans un même pixel. Cela pourrait expliquer certaines des confusions observées.

III.3 Analyse des caractéristiques des essences

La phénologie des peuplements purs montre une évolution qui est attendue. En effet, cette dernière reflète la phénologie des espèces forestières : les feuillus connaissent une forte augmentation du NDVI au printemps avec la pousse des feuilles, tandis que leur signature spectrale décroît en hiver en raison de la chute du feuillage. En revanche, certaines essences comme le Douglas ou le Pin maritime conservent des valeurs élevées toute l'année, ce qui est cohérent avec leur caractère de conifères. Les résultats sont donc en accord avec la réalité phénologique.

L'écart type du NDVI reste globalement faible et homogène sur l'ensemble des périodes observées, avec des valeurs souvent inférieures à 0,1, ce qui indique une certaine homogénéité au sein des classes. Cependant, on note des hausses ponctuelles, notamment en mars et avril 2022, où l'écart type atteint jusqu'à 0,2 pour le peuplier, ce qui est synonyme d'une hétérogénéité au sein de cette classe. Il serait intéressant de cartographier les essences avec un NDVI élevé et celles avec un NDVI faible pour étudier plus en détails les éventuels facteurs de cette hétérogénéité.

Nous pouvons émettre l'hypothèse que cela est en lien avec les résultats de la variabilité spectrale, qui montre notamment une plus grande variabilité spectrale pour la classe "Peupleraie". Ces résultats montrent également une plus forte hétérogénéité de la variabilité au sein des différentes classes de peuplements purs que pour les peuplements en mélange. En effet, certaines classes comme le chêne présentent une faible variabilité, proche de 0.15, contre 0.30 pour les peupleraies par exemple.

Conclusion

Pour conclure, la classification supervisée basée sur les essences de la BD Forêt V2 de l'IGN montre à première vue une performance très satisfaisante de manière générale. Cependant, la distribution des pixels et des polygones révèle une hétérogénéité marquée, avec une forte dominance du chêne et une sous-représentation d'autres essences dans notre zone d'étude. En raison de cette distribution, ce projet ne permet pas de déterminer si cette base de données peut être utilisée en temps que données de référence pour une classification supervisée puisque la majorité des types de peuplement n'ont pas obtenu de scores de précision satisfaisants. Cela pose la question de la pertinence de l'utilisation de la BD Forêt. En effet, cette dernière regroupe dans la classe des chênes les chênes pédonculés, les chênes pubescents, les chênes sessiles et les chênes rouges. Il serait intéressant d'adopter une méthode similaire avec un autre jeu de données plus précis, comme des données issues de l'ONF renseignant précisément les essences à l'échelle des parcelles subdivisées en unités de gestion.