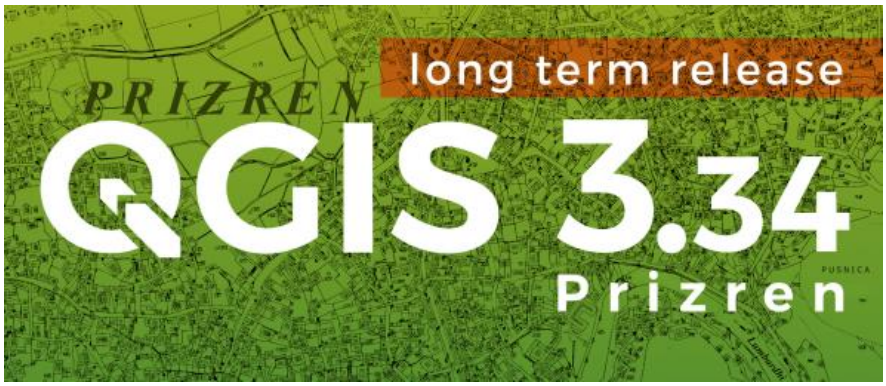
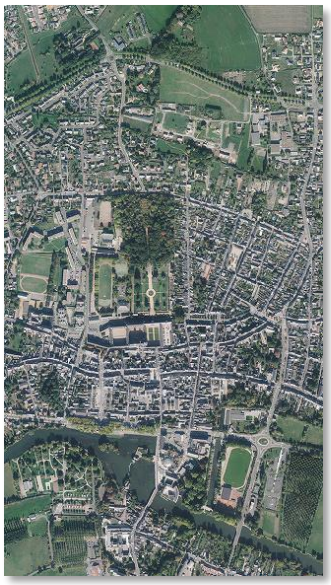
JEAN-MARC GILLIOT
QGIS-tuto.fr

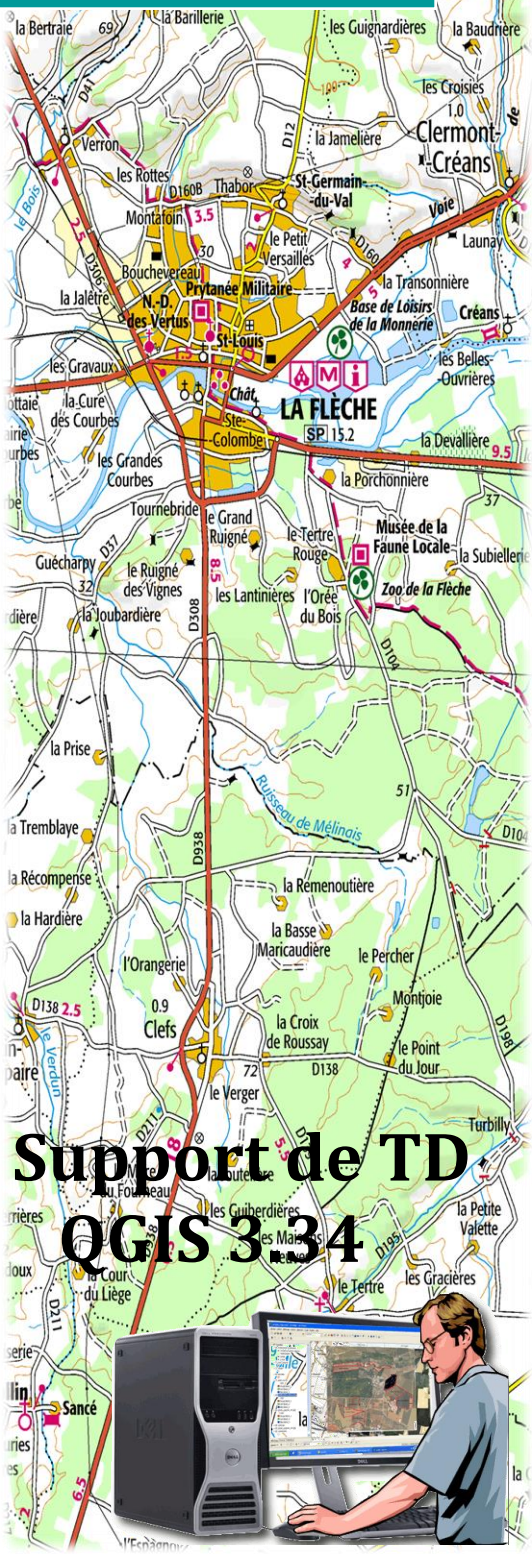
2024-2025

SIG
TD 8-2
QGIS 3.34

Introduction à la Télédétection avec QGIS
Applications aux milieux naturels et agricoles







Support de TD QGIS 3.34

Version janvier 2025

AgroParisTech

Grande école européenne d'ingénieurs et de managers
dans le domaine du vivant et de l'environnement

Jean-marc.gilliot@agroparistech

Table des matières

1. Démarrage de QGIS et téléchargement des données 3
2. Installation de OTB : Orfeo ToolBox du C.N.E.S..... 3
3. Approches par Classification 5
3.1. Classification non supervisée par KMeans 5
3.2. Classification supervisée Par Random Forest (RF) 12
Etape 1 : Apprentissage : 12
Etape 2 : Classification application du modèle.....13
Calculer la matrice de confusion en sortie pour analyser la qualité de la classification. 14
3.3 Analyse texturale des paysages 19
3.3.1. Analyse texturale sur l'image du catalogue Brodatz 19
3.3.2. Classification texturale par analyse d'histogrammes locaux 24
4. Analyse diachronique par télédétection..... 27
4.3. Extraction automatique des cultures sous QGIS à partir d'une analyse diachronique d'images Sentinel et du RPG.....27
4.4. Cartographie par Télédétection diachronique des dégâts de la tempête de 1999 sur les forêts de la Gironde dans le canton « Sud Medoc » 36

Conseil A vous de jouer Attention durée objectif

clic souris : gauche droit YouTube tuto. Vidéo à retenir

1. Démarrage de QGIS et téléchargement des données

Depuis le menu Windows : le menu de QGIS est QGIS 3.34

Lancer ***QGIS Desktop 3.34***

https://qgis.org/downloads/

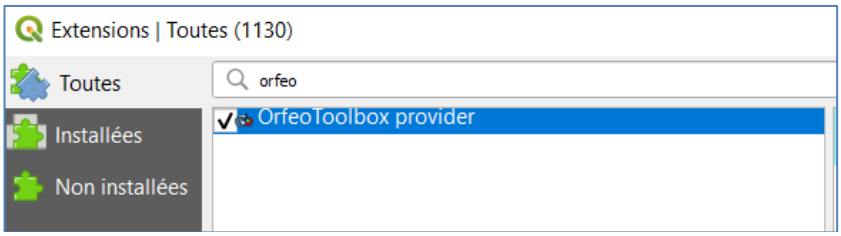
Sur le site de téléchargement de QGIS, privilégier la dernière version dite « LTR » (long time release) qui est la version la plus stable car elle subit des tests plus rigoureux avant sa sortie et reçoit au moins un an de mises à jour de correction de bogues.

Long Term Version for Windows (3.34 LTR)

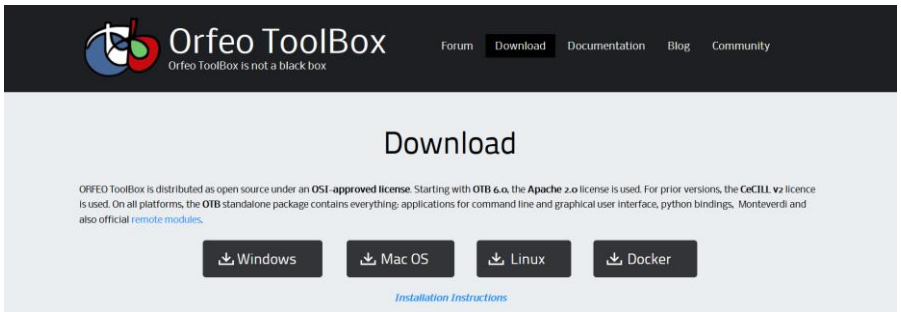
2. Installation de OTB : Orfeo ToolBox du C.N.E.S.

Orfeo ToolBox est un logiciel open source de télédétection initialement développé par le Centre national d'études spatiales (CNES), une boite à outils QGIS de OTB existe pour QGIS, pour l'installer :

Installer l'extension : OrfeoToolbox provider depuis le menu Extension de QGIS :



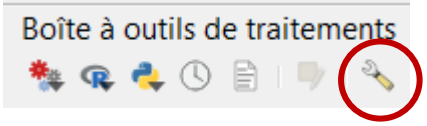
Télécharger OTB :
https://www.orfeo-toolbox.org/download/



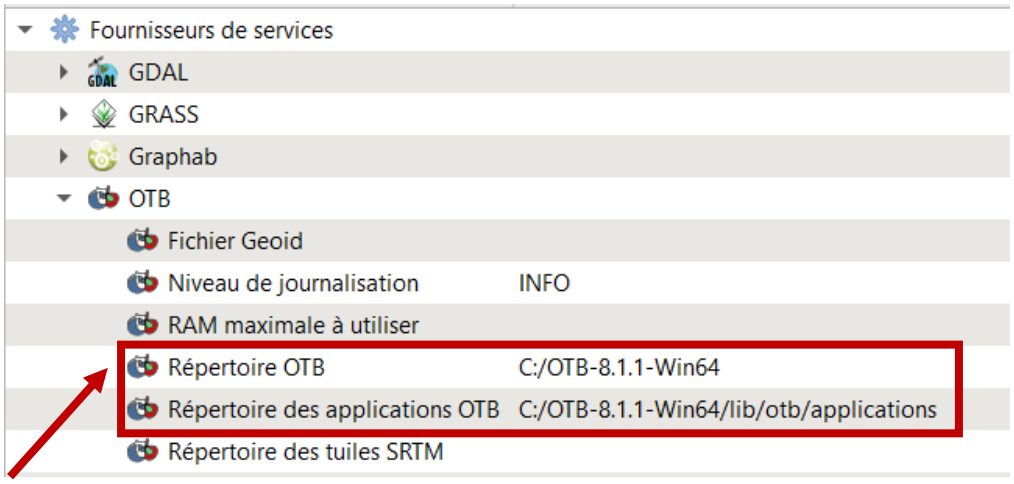
Télécharger OTB-8.1.1-Win64.zip, puis le dézipper dans le dossier téléchargement et copier OTB-8.1.1-Win64 directement à la racine de c : (par exemple) ce qui donne le dossier : c:\OTB-8.1.1-Win64

Sous QGIS , dans la boite à outils de traitements (processing)

Ouvrir les options avec



Dans Fournisseurs de services >> OTB
Fixer les chemins pour **Répertoire OTB** et **Répertoire des applications OTB** comme ci-dessous :



Saisir en premier le répertoire OTB de base

3. Approches par Classification

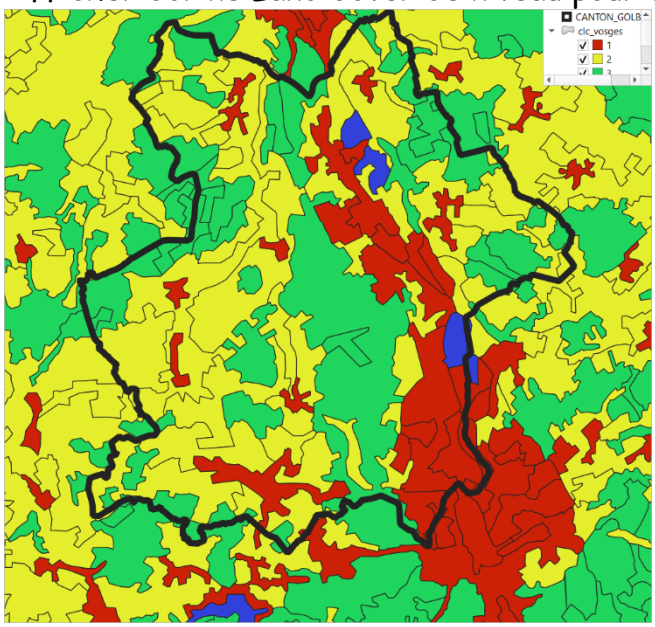
3.1. Classification non supervisée par KMeans

Sur l'image de GOLBEY que vous avez géoréférencé précédemment.

Sentinel_2008_08_16_modifié

Réaliser une classification de l'occupation des sols à partir de l'image Sentinel, le plus proche possible du niveau 1 de Corine Land Cover :

Afficher Corine Land Cover de niveau pour visualiser les grandes zones :

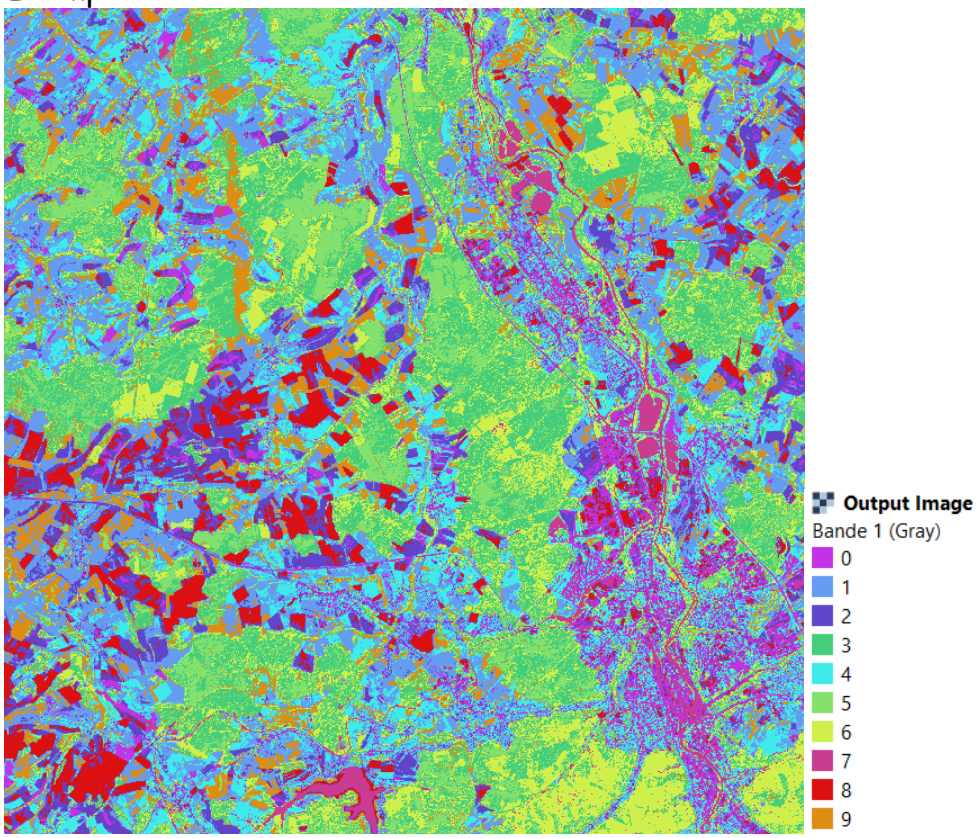


1 : Urbain
2 : Agricole
3 : Forêt
4 et 5 : Surfaces en eau

Pour l'image Sentinel faire une composition colorée 4/3/2 = PIR/R/V

Classification non supervisée par exemple par KMEANS sur l'image Sentinel

Exemple en 10 classes



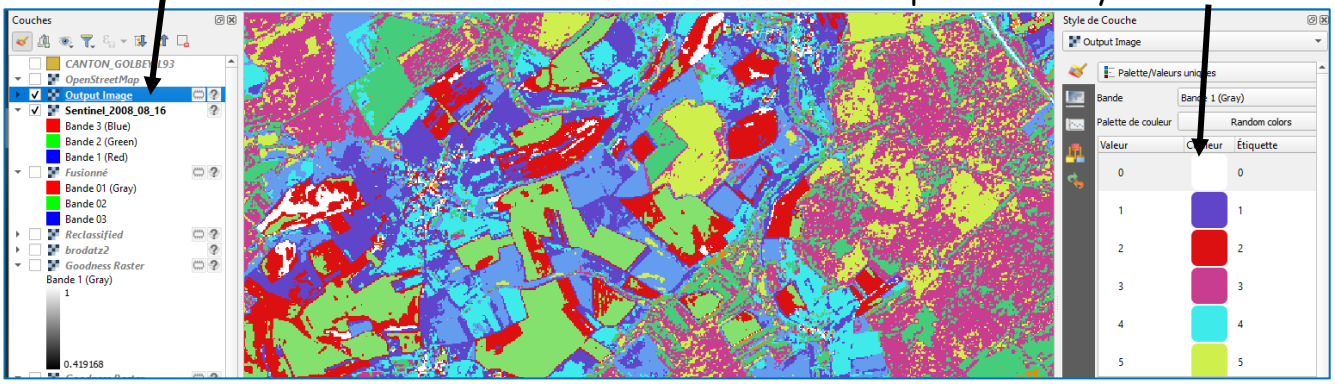
Il est conseillé de faire plus de classes que le nombre recherché, puis de le regrouper ensuite.

Regroupement des classes :

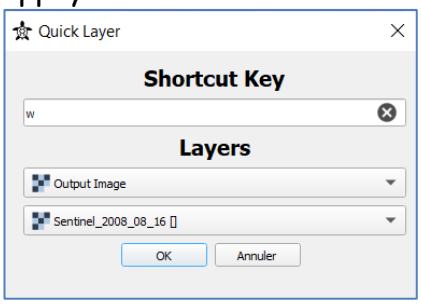
Etape 1 : Fusionner dans un premier temps visuellement les classes, en donnant la même couleur aux classes candidates à la fusion, dans la légende :

Afficher dans le panneau couche
La classif. Au-dessus de Sentinel

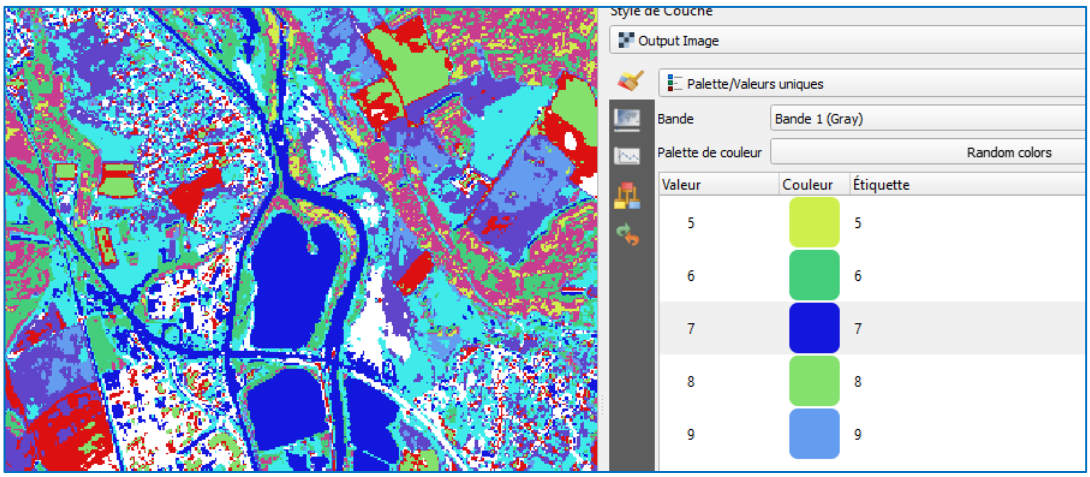
Afficher le panneau « style de couche »



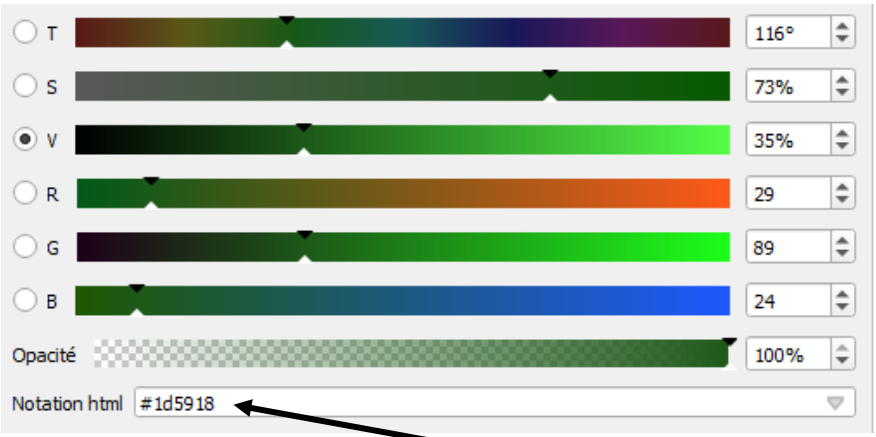
Cocher / Décocher la classif. Pour visualiser alternativement la classif. et l'image originale et repérer ainsi les classes à regrouper ; on peut aussi utiliser pour cela, l'extension « Quick Layer » qui permet de switcher rapidement l'affichage entre deux couches en appuyant sur une touche du clavier.



Remarquez qu'il y peut y avoir **confusion** entre certaines classes (pixels mal classés).

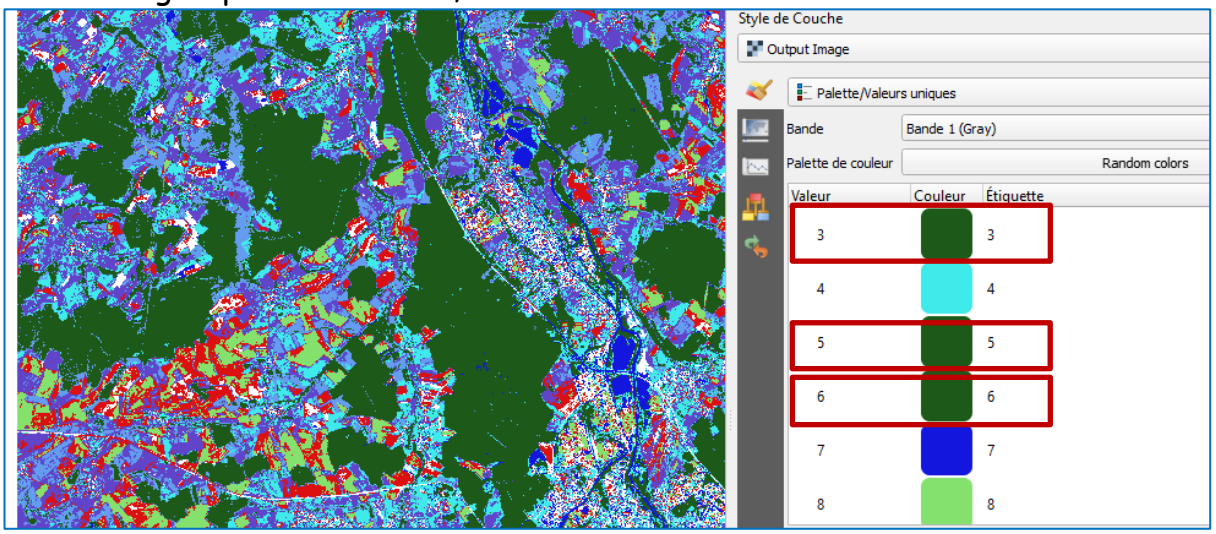


Fusionner les classes de Forêt en les mettant en vert foncé



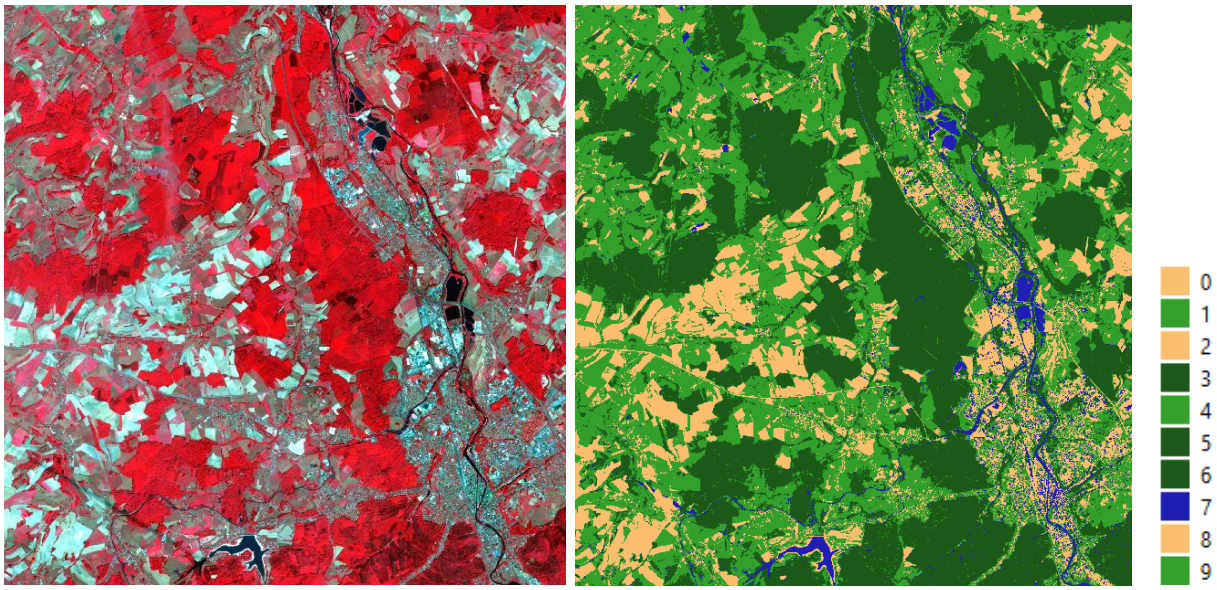
Pour copier exactement la même couleur : copier / coller sa notation html dans le panneau symbologie :

Ici on a regroupé les classes 3, 5 et 6 en vert foncé « forêt »



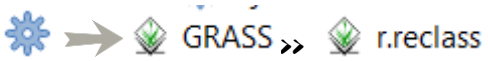
On remarquera qu'il est très difficile de faire ressortir une classe « urbaine » cela est dû à l'hétérogénéité des réponses spectrales de l'urbain (mélange).

Pour les parcelles agricoles on va faire deux classes : une pour les parcelles avec végétation chlorophyllienne, rose ou rouge dans la composition colorée et une autre pour les parcelles sans végétation chlorophyllienne (bleutée ou verdâtre) :



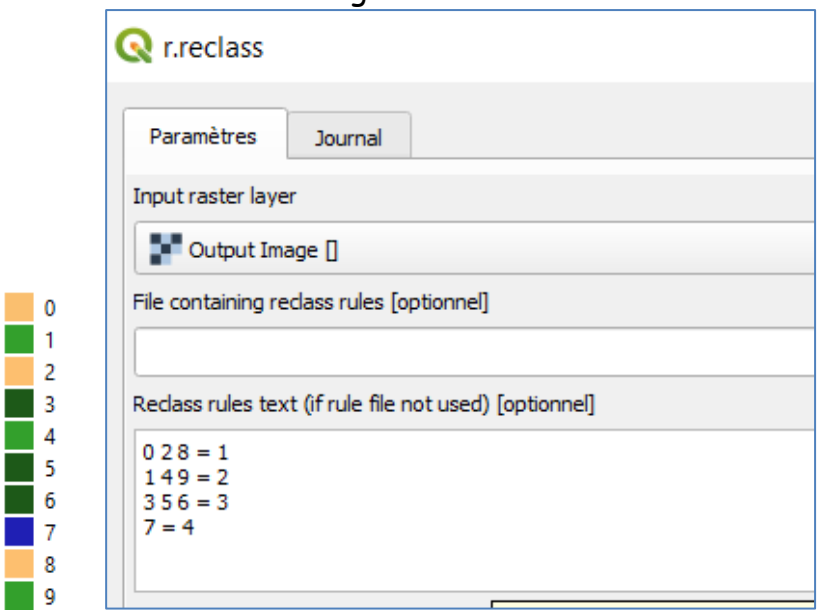
Sauvegarder nos choix de couleurs dans un fichier de style si on veut les retravailler ultérieurement :
Clic droit sur la couche >> Exporter >> Enregistrer en tant que fichier de style .qml
(Bouton style dans Symbologie pour le recharger)

Etape 2 : Création de l'image finale reclassée



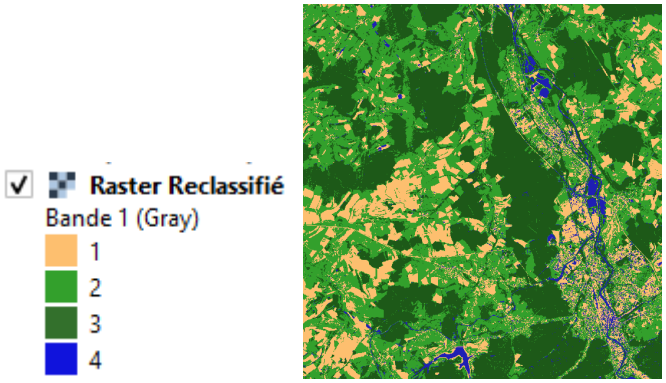
On peut aussi utiliser

Saisir la table de recodage

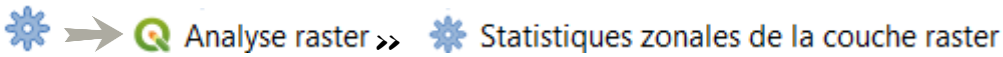


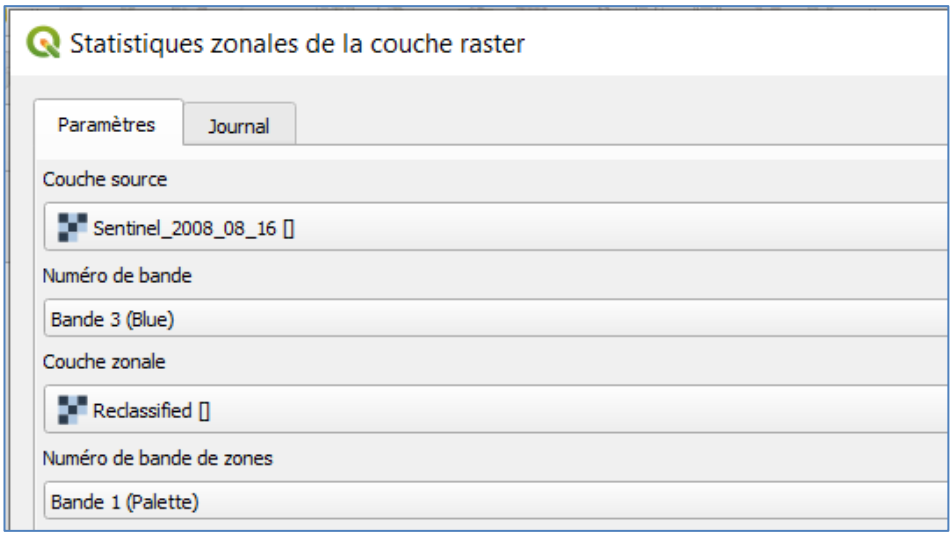
Vous pouvez sauver vos règles de reclassification par un simple copier / coller dans un fichier texte.

On obtient une image en 4 classes : Faire une légende par valeur unique et remettre les mêmes couleurs que précédemment



Visuellement le résultat est absolument identique à l'étape précédente, mais maintenant la matrice ne contient que les valeurs de classe : 1, 2, 3 et 4, ce qui est indispensable si on veut faire des traitements par classe, par exemple calculez les statistiques de la bande PIR de l'image Sentinel dans chacune des zones des 4 classes :



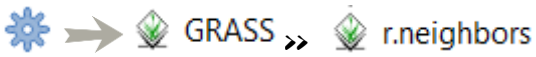


Statistiques — Total des entités: 4, Filtrées: 4, Sélectionnées: 0

	zone	area	sum	count	min	max	mean
1	4.00000000	98085.00000000	3943471.00000...	98085	0	82.00000000	40.20462864
2	2.00000000	1128884.00000...	160667196.000...	1128884	78.00000000	255.00000000	142.32391991
3	1.00000000	503761.00000000	63188933.0000...	503761	0	255.00000000	125.43434883
4	3.00000000	1030785.00000...	172149963.000...	1030785	50.00000000	255.00000000	167.00860315

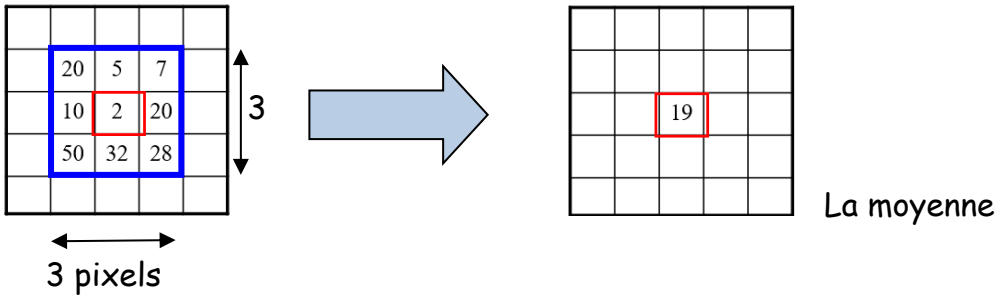
<u>Post-traitement de la classification = simplification géométrique = lissage</u>

On applique très souvent après la classification une étape de post-traitement pour simplifier la forme des zones obtenues et supprimer les très petites zones, on va utiliser une opération de Filtrage par voisinage.

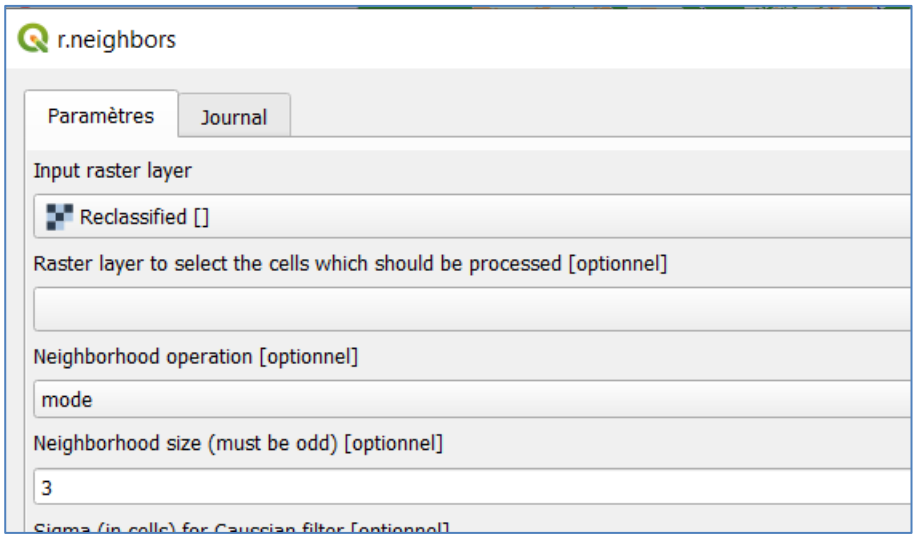


R.neighbors remplace la valeur d'un pixel par le résultat d'une opération statistique entre cette valeur et celles des pixels dans le voisinage immédiat dans l'image.

Exemple d'un Filtrage par la moyenne avec une taille de fenêtre 3 x 3



Dans notre cas on ne peut pas faire une moyenne, car l'image classée contient des numéros de classe qui sont des données non quantitatives, on n'a pas le droit de faire la moyenne de données qualitatives. On va remplacer la moyenne par un calcul de la valeur modale, qui est le numéro de la classe qui est la plus fréquente dans le voisinage (adapté aux données qualitatives comme de numéros de classe)



Calcul de la valeur modale

Taille du voisinage de la fenêtre : 3 x 3

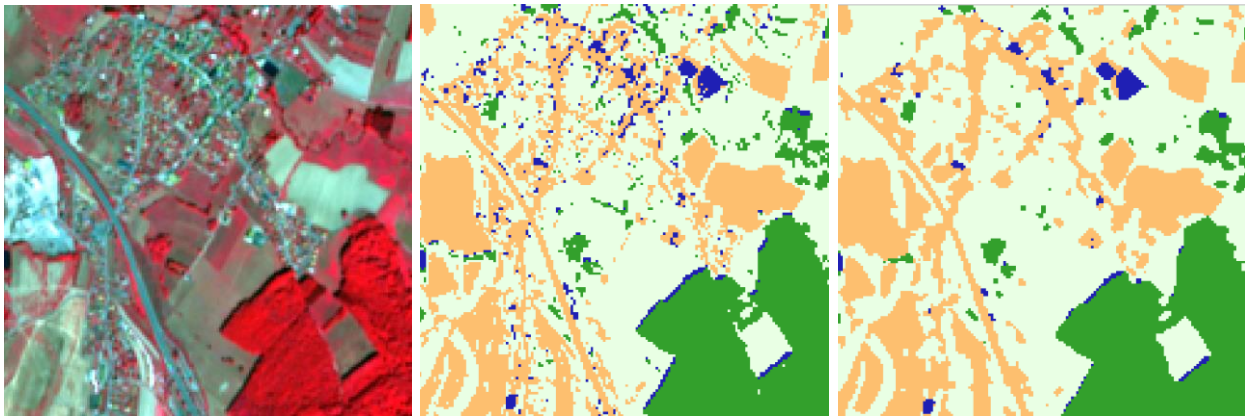


Image de départ classification kmeans filtrage mode 3x3

On voit bien l'effet de simplification ainsi que la perte des toutes petites surfaces, mais aussi une petite perte de précision globale, il y a toujours un effet positif et un effet négatif au filtrage.

Si la classification est « suffisamment » simplifiée on peut éventuellement la vectoriser avec :

GRASS → Raster (r.*) → r.to.vect

3.2. Classification supervisée Par Random Forest (RF)

La classification se déroule en deux étapes :

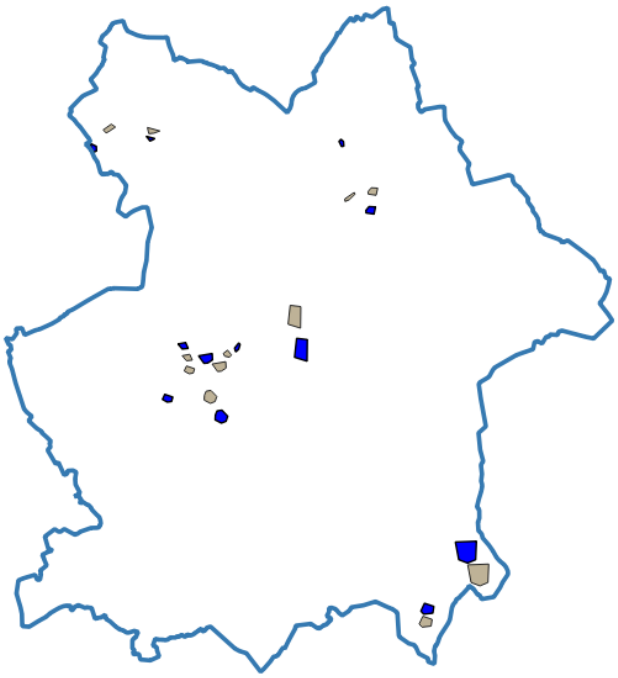
1) On choisit des zones d'apprentissage = Region Of Interest (ROI), soit en les dessinant sur la composition colorée, soit en utilisant des polygones de vérité terrain, ces polygones pouvant être issus d'une base de données de référence comme : BD topo IGN, Corine Land Cover, RPG agricole etc, sur ces polygones des statistiques représentatives de chaque classe sont alors calculées.

2) On classe tous les pixels par rapport aux statistiques de classes calculées à l'étape 1.

Etape 1 : Apprentissage :

Ajouter les couches roi_1.shp et roi_2.shp

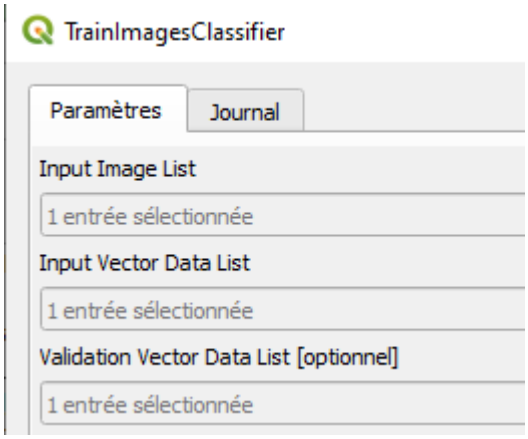
Roi_1 contient 12 polygones d'apprentissage (ROI) saisis visuellement sur l'image, avec les classes ci-dessous
Roi_2 même chose mais sur d'autres positions, pour validation de la classification.



	CLASSE_NAM	CLASSE_NUM
1	AGRI_NU	1
2	AGRI_VEG	2
3	FORET	3
4	URBAIN	5
5	EAU	4

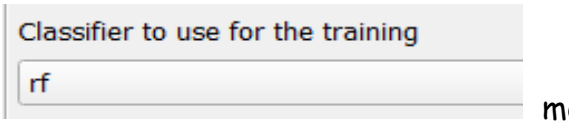
OTB → Learning → TrainImagesClassifier

Cette fonction va calculer les statistiques de classes à partir de ROI 1

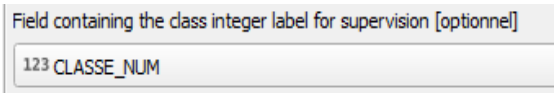


L'image Sentinel

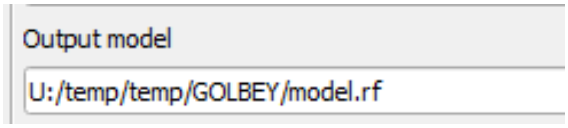
le fichier shape des ROI roi1
Le roi de validaton roi2



méthode de classification ici RF (Random Forest)

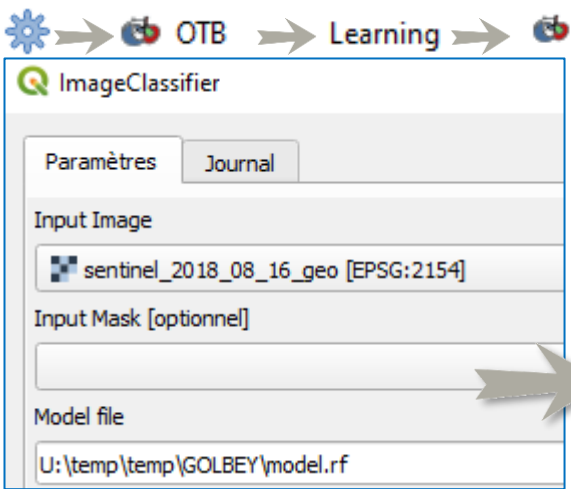


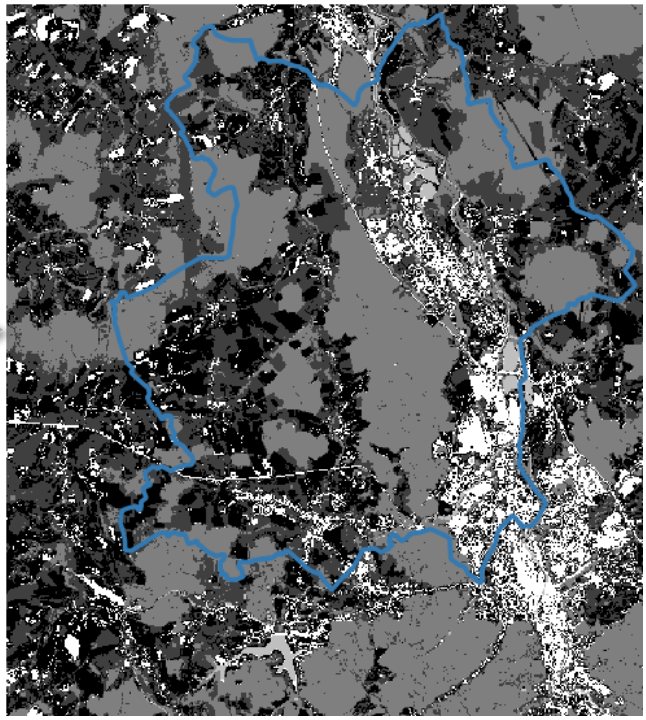
Le champ qui contient le numéro de classe



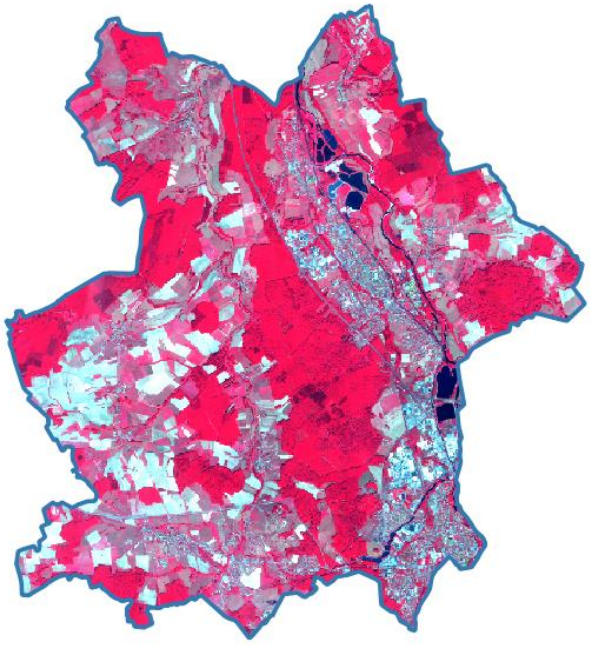
Le fichier sortie avec les stat. du modèle

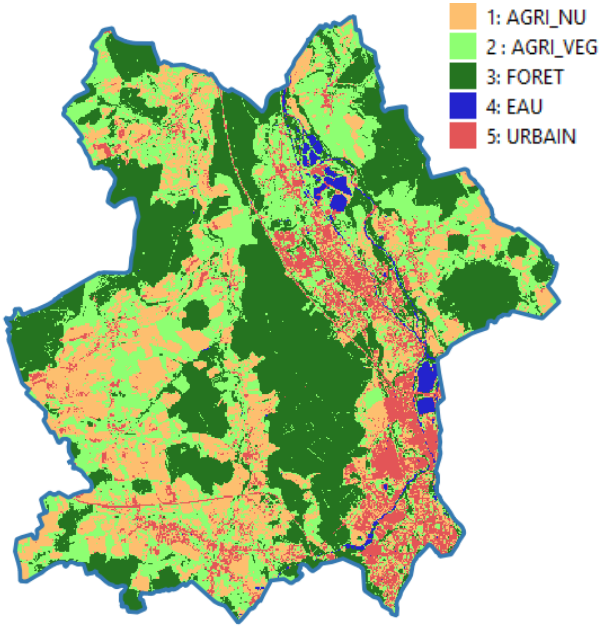
<u>Etape 2 : Classification application du modèle</u>



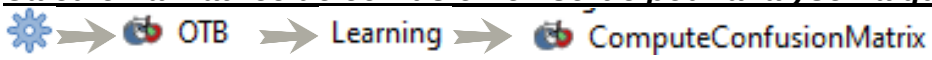


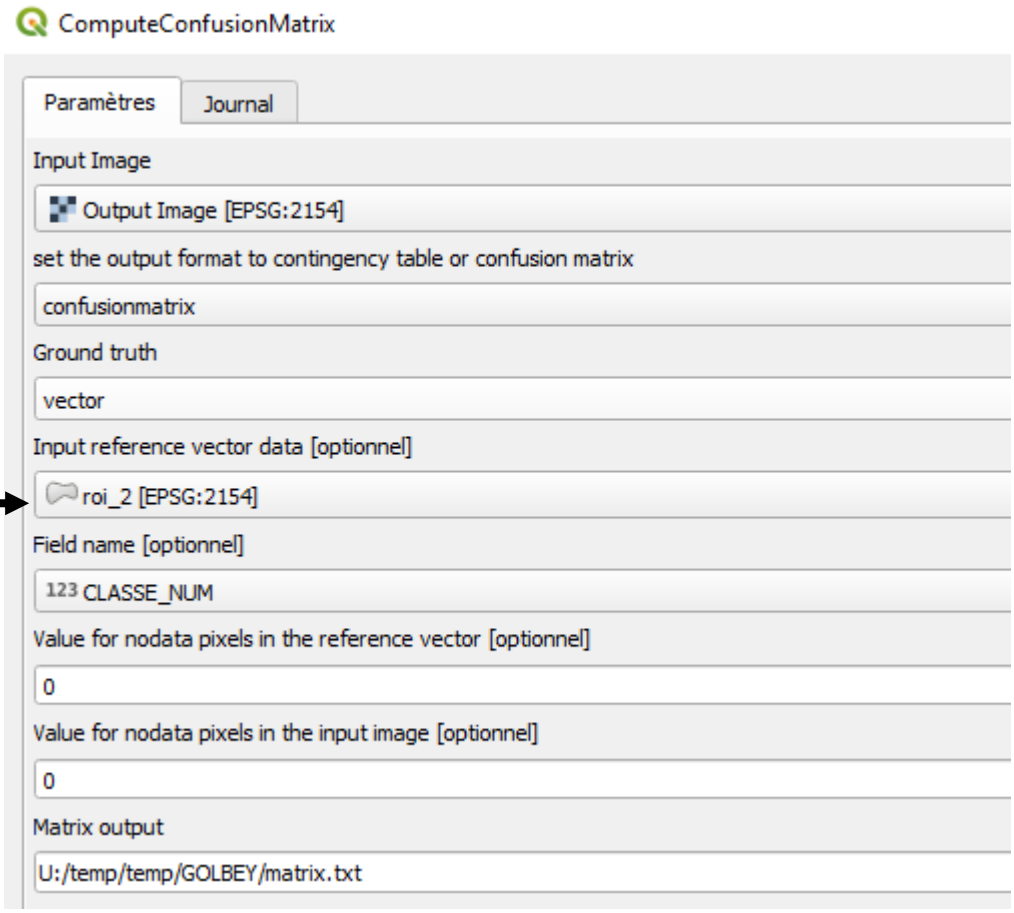
Changer la symbologie du résultat,
<u>Palette / valeurs uniques</u>
Adapter les couleurs et saisir
Les intitulés de la légende
Masquer par les limites du canton





Calculer la matrice de confusion en sortie pour analyser la qualité de la classification.





Le fichier ROI2.shp est utilisé comme zones de validation

Regarder la matrice de confusion produite par la classification :

```
#Reference labels (rows):1,2,3,4,5
#Produced labels (columns):1,2,3,4,5
784,4,0,0,0
10,455,0,0,4
0,26,1717,0,0
0,0,0,379,0
686,286,23,0,1331
```

La mettre en forme dans Excel par exemple de cette façon :

			PRODUCED ou classés				
			AGRI_NU	AGRI_VEG	FORET	EAU	URBAIN
			1	2	3	4	5
REFERENC	AGRI_NU	1	784	4	0	0	0
	AGRI_VEG	2	10	455	0	0	4
	FORET	3	0	26	1717	0	0
	EAU	4	0	0	0	379	0
	URBAIN	5	686	286	23	0	1331

⚠ Attention selon les auteurs, la matrice peut être présentée en inversant lignes et colonnes, à savoir avec en ligne les classés et en colonne les références.

Les nombres représentent des co-occurrences de pixels qui correspondent à une classe produite par la classification (en colonne) et une valeur de référence (vérité terrain) en ligne. Par exemple la première colonne représente tous les pixels testés qui ont été classés comme AGRI_NU, ceux de la première cellule (784) correspondent bien à une référence AGRI_NU, alors que ceux de la seconde cellule (10) sont mal classés car ils correspondaient en fait à la référence AGRI_VEG etc …

Dans l'idéale on devrait avoir une matrice qu'avec des valeurs sur la diagonale ce qui voudrait dire que tous les pixels sont parfaitement classés, mais il y a des confusions entre certaines classes.

On peut calculer différentes métriques à partir de la matrice de confusion, pour estimer la qualité de la classification. Commencer par calculer les totaux par colonne et par ligne (sommes marginales). Ri est la somme d'une ligne pour la classe i, c'est le nombre de pixels qui sont dans les zones références de la classe. Pi (produced ou classés) est la somme d'une colonne pour la classe i, c'est le nombre de pixels qui ont été classés dans cette classe.

N est le nombre total de pixels de roi_2, c'est la somme des totaux soit en ligne soit en colonne (même valeur)
Pour i variant de 1 à nc (nombre de classes)

$n_{i,i}$ est le nombre dans une cellule de la diagonale de la matrice, qui correspond donc aux pixels bien classés pour la classe i.

$$N_T = \sum_{i=1}^{nc} n_{i,i}$$

N_T est le nombre total de pixels bien classés (True) pour l'ensemble des classes, c'est la somme de la diagonale.

			PRODUCED ou classés							
			AGRI_NU	AGRI_VEG	FORET	EAU	URBAIN			
REFERENCE			1	2	3	4	5	Ri	PP	ED
	AGRI_NU	1	784	4	0	0	0	788	0.995	0.005
	AGRI_VEG	2	10	455	0	0	4	469	0.970	0.030
	FORET	3	0	26	1717	0	0	1743	0.985	0.015
	EAU	4	0	0	0	379	0	379	1.000	0.000
	URBAIN	5	686	286	23	0	1331	2326	0.572	0.428
		Pi	1480	771	1740	379	1335	5705		
		PU	0.530	0.590	0.987	1.000	0.997		4666	
		EE	0.470	0.410	0.013	0.000	0.003			

N = 5705 pixels le nombre total de pixels qui ont été testés par la validation
N_T = 4666 le nombre total de pixels bien classés (True) c'est la somme de la diagonale
N_F : Nombre de pixels mal classés (False) = N - N_T = 5705 – 4666 = 1039
La précision Globale P_G de la classification ou Overall Accuracy A_O est donnée par :

$$P_G = \frac{N_T}{N} = \frac{nombre\ de\ pixels\ bien\ classés}{nombre\ total\ de\ pixels}$$

Dans notre cas P_G = 4666 / 5705 = 0.82 soit 82 %

<u>82% des pixels de la validation sont donc bien classés</u>, il faut toutefois être prudent avec la précision globale car elle peut masquer des problèmes particuliers à certaines classes, surtout dans le cas où les effectifs testés pour les différentes classes ne sont pas les mêmes, une classe surreprésentée qui est facile à classer peut entrainer une surestimation de P_G.

La précision utilisateur P_U : Si un utilisateur de la classification regarde par exemple un pixel classé « forêt » dans l'image, quelle est la probabilité que ce soit effectivement de la forêt dans la référence (vérité terrain) ? Dans la matrice si on regarde la colonne forêt (pixels classés) il y a 1740 pixels qui ont été classés forêt (somme de la colonne) parmi lesquels 1717 étaient effectivement de la forêt, la probabilité est donc de 1717 / 1740 = 0.987 soit 98.7% pour la classe forêt.

L'erreur d'excédent E_E = 1 – P_U : c'est le complémentaire de P_U, elle représente les pixels « sur détecter » pour une classe.

La précision producteur P_P : Si une position est réellement de la forêt (vérité) quelle est sa probabilité d'être bien classée en forêt ? Dans la matrice on regarde cette fois la ligne forêt, au bout de la ligne forêt la somme est de 1743 pixels de référence « forêt » parmi ces 1743, 1717 ont effectivement été classés en forêt, la probabilité est donc de 1717 / 1743 = 0.985 soit 98.5% pour la classe forêt.

L'erreur de déficits $E_D = 1 - P_P$: c'est le complémentaire de P_P

Calculer PU et PP pour toutes les classes, on remarque que cela revient à chaque fois à faire le rapport de la cellule en diagonale sur la somme soit de la ligne ou de la colonne.
On peut remarquer que certaines classes sont plus facile à classer que d'autres, ainsi pour l'eau $P_U = P_P = 100\%$ il n'y a aucune confusion avec les autres classes ; pour la forêt on a $P_U = 98.7\%$ et $P_P = 98.5\%$ donc comme l'eau très peu de confusion. Cependant pour la classe AGRI_NU $P_P = 99.5\%$ mais $P_U = 53$ % seulement, l'erreur de déficit est très faible cela signifie que toutes les zones de référence AGRI_NU ont très bien été classées par contre l'erreur d'excédent est de 47% cela signifie qu'il y a 47% de pixels classés à tort en AGRI_NU (en trop) et on peut constater que cela provient essentiellement de la classe URBAIN qui représente à elle seule 46.3% de cet excédent. L'urbain est difficile à classer avec ces classifications « pixel à pixel » car c'est milieu hétérogène par nature à l'échelle d'un pixel.

<u>Calcul du coefficient Kappa</u>, comme indicateur global de la matrice de confusion, il vari entre 0 et 1, plus il se rapproche de 1, meilleur est la classification :

$$Kappa = \frac{N\sum_{i=1}^{nc} n_{i,i} - \sum_{i=1}^{nc}(R_i P_i)}{N^2 - \sum_{i=1}^{nc}(R_i P_i)}$$

Avec N le nombre total de pixels testés, $n_{i,i}$ le nombre de pixels de la classe i qui sont bien classés en i, R_i le nombre total de pixels de référence pour la classe i et Pi le nombre total de pixels produits (classés) dans la classe i (les sommes marginales).
Dans notre exemple :
N = 5705

$$\sum_{i=1}^{nc} n_{i,i} = N_T = 4666$$

$$\sum_{i=1}^{nc}(R_i P_i) = 1480 * 788 + 771 * 469 + 1740 * 1743 + 379 * 379 + 1335 * 2326 = 7809510$$

$$Kappa = \frac{5705 * 4666 - 7809510}{5705^2 - 7809510} = 0.76$$

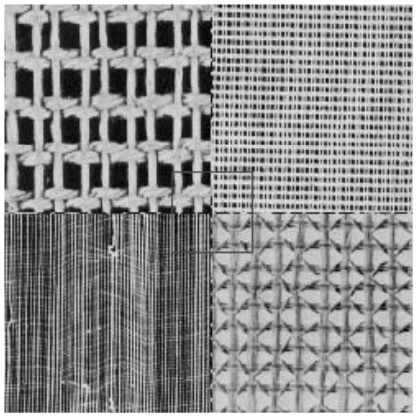
On considère généralement que le Kappa doit être supérieur à 0.6, un Kappa > 0.75 la classification est bonne, Kappa > 0.8 Forte et Kappa > 0.9 presque parfaite.

3.3 Analyse texturale des paysages

3.3.1. Analyse texturale sur l'image du catalogue Brodatz

DOSSIER : « texture »

Ajouter sous QGIS l'image Brodatz2.tif

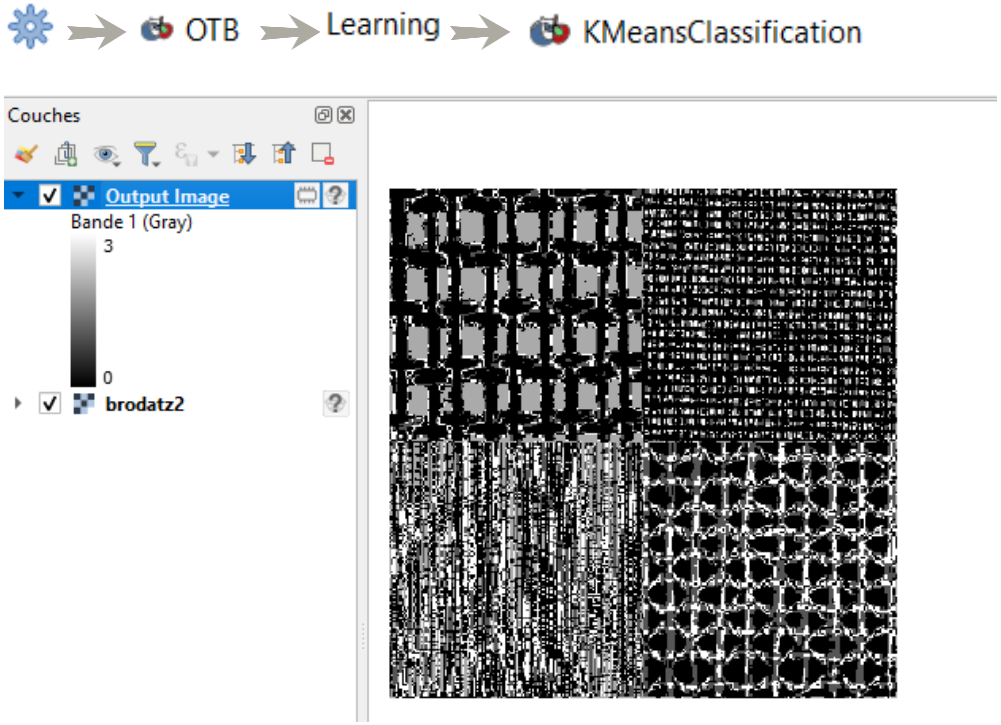


Il s'agit d'une image de textures naturelles extraites du catalogue de Brodatz.

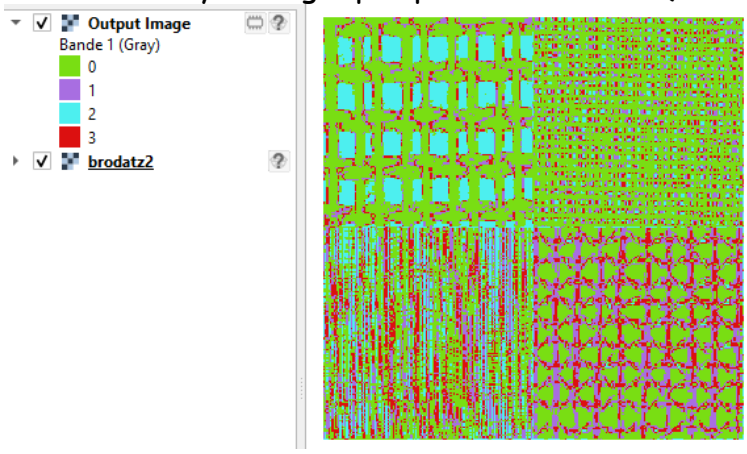
Sur l'image Brodatz2 essayez de segmenter les 4 zones représentants les 4 textures différentes.

a) Classification colorimétrique / radiométrique non supervisée par K-MEANS

Utiliser la classification non supervisée de K-MEANS en 4 classes



Mettre une symbologie par pseudo-couleurs (couleurs aléatoires)



Ce type de classification qui ne tient pas compte du voisinage n'est pas adaptée à la classification d'images texturées : on ne voit pas ressortir les 4 zones texturées comme 4 classes.

b) Classification texturale par matrice de cooccurrence (GLCM)

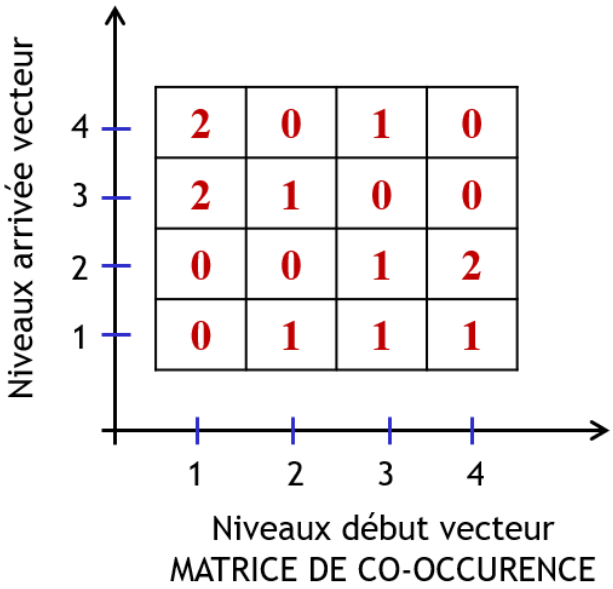
GLCM = Grey Level Cooccurrence Matrix

Principe du calcul d'une matrice de co-occurrence de niveaux de gris :

début ⟶ arrivée
Vecteur direction
pour calcul des co-occurences

1	4	2	3
3	1	4	2
3	2	1	3
1	3	4	1

IMAGE
quantification à 4 niveaux: 1,2,3,4



En rouge = effectif des cooccurrences

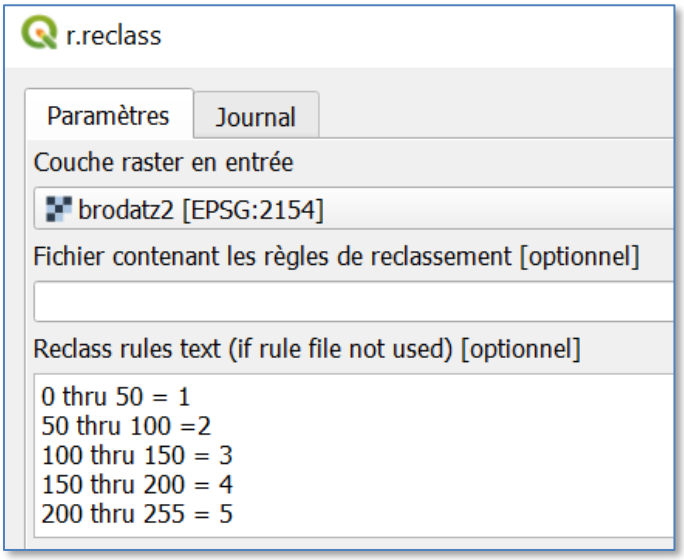
La quantification est le nombre de niveaux de gris existant, ici dans l'exemple 4, le plus souvent 256 (0-255). Cela détermine la taille de la matrice de co-occurrence qui peut être une limite pour les calculs.

Réduire le nombre de niveaux de quantification de l'image brodatz2 à 5 niveaux

Eventuellement Histogramme d'une couche raster pour analyser la répartition des niveaux dans l'histogramme et fixer les intervalles sinon cf ci-dessous.

r.reclass fonction Grass pour reclasser les niveaux initiaux en 5 niveaux.

(plusieurs autres fonctions qgis pourraient aussi être utilisées pour faire cela).

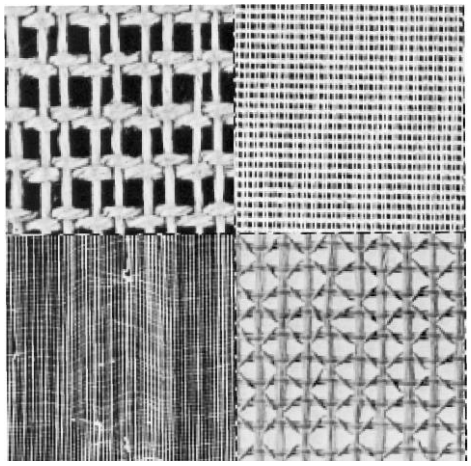


Il faut définir des règles de recodage des valeurs de niveaux de gris.

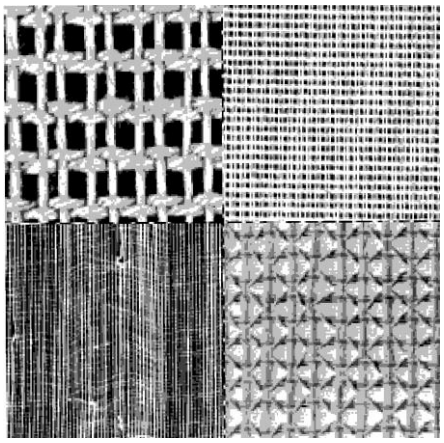
Soit depuis un fichier texte soit dans la zone ci-dessous

(*Consulter l'aide en ligne*)
On peut transformer une valeur unique en une nouvelle valeur, une liste de valeurs uniques en une nouvelle valeur ou un intervalle de valeurs grâce au mot clé « thru »

valeur_min ***thru*** valeur_max = nouvelle valeur



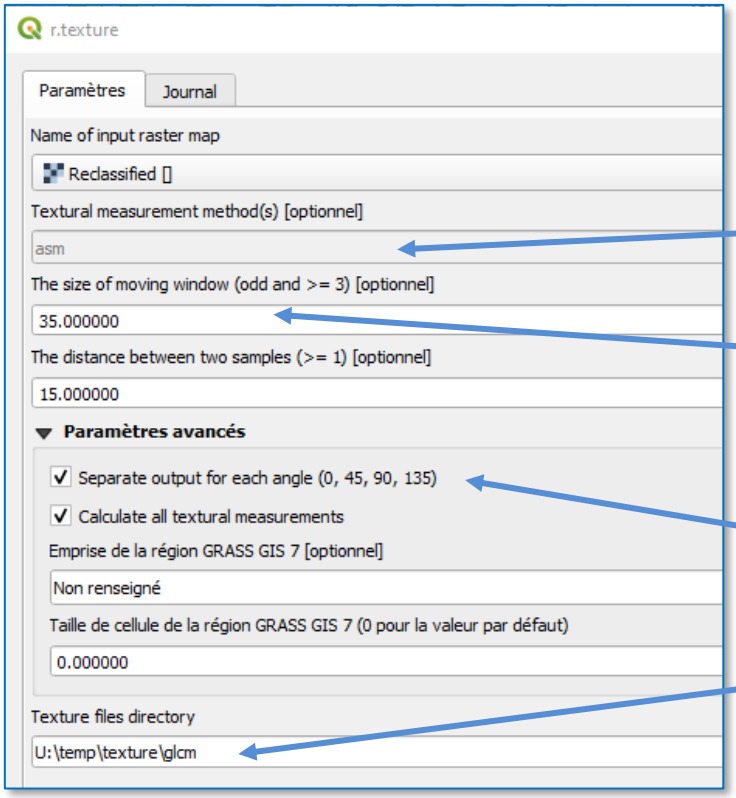
Brodatz2



quantifiée en 5 niveaux

Calculer les paramètres texturaux de la GLCM avec la fonction Grass : r.texture

Chaque paramètre choisit donne lieu à un néocanal (nouvelle image)
A partir de l'image reclassifiée précédente



Méthode = 13 calculs tous cochés *(Consulter l'aide en ligne)*

A définir en fonction de la « taille élémentaire » des textures à analyser

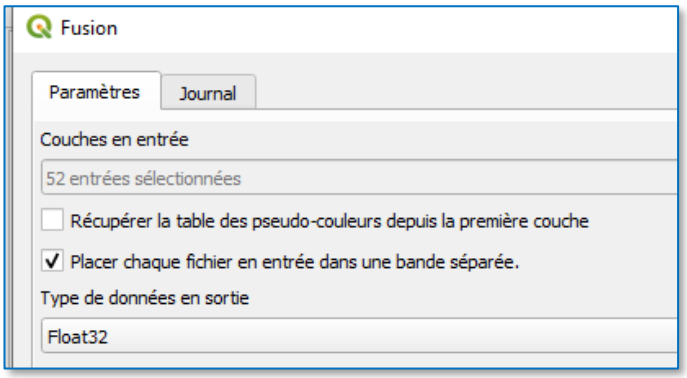
On calcule tous les 13 paramètres texturaux, selon toutes les directions (4 angles)

Définir un dossier de sortie pour bien localiser les néocanaux :

Ajouter les néocanaux au projet après calcul.

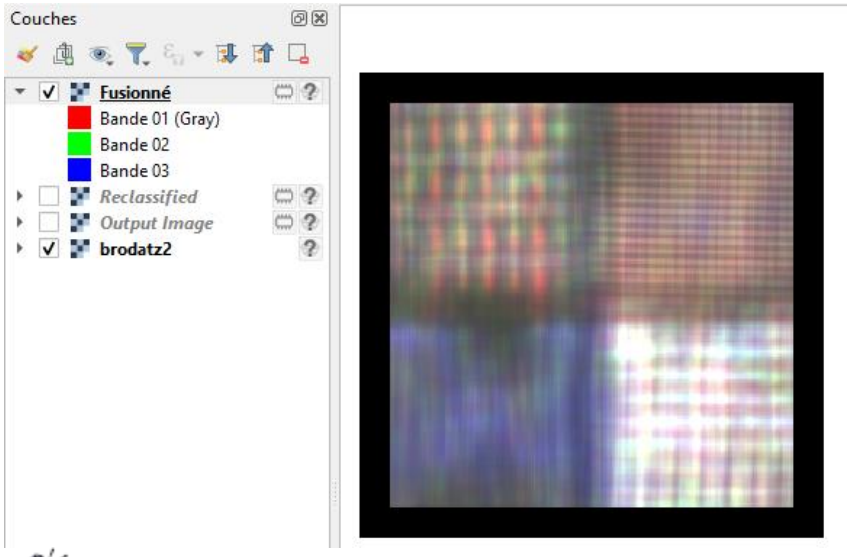
52 néocanaux sont calculés = 13 calculs x 4 angles = 52 néocanaux

Créer une image fusionnée des 52 canaux



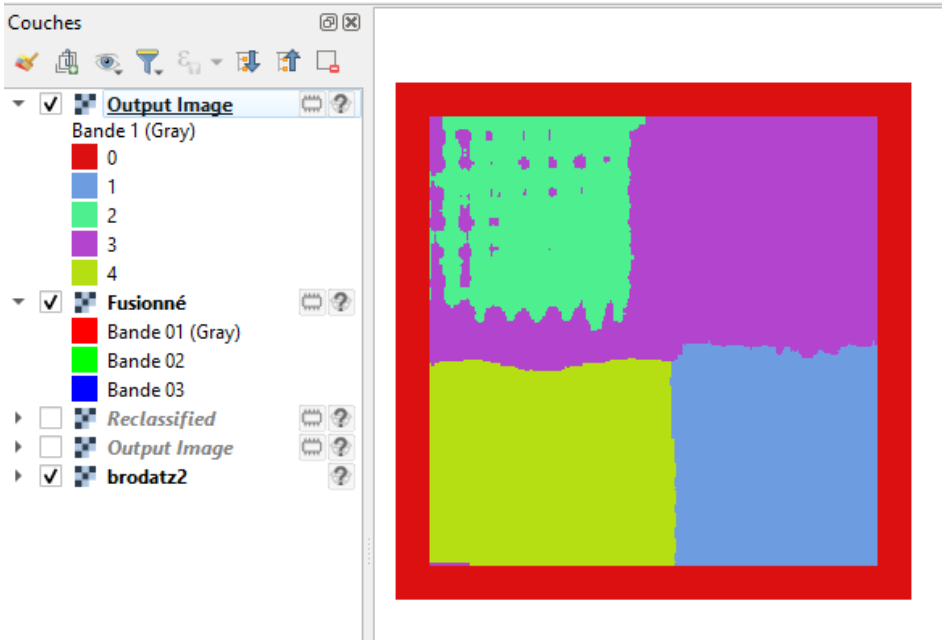
Choisir sur le disque les 52 neo-canaux (tous les fichiers .tif)

Bien cocher cette option sinon Gdal fait une fusion géographique et non des canaux.



Faire une classification non supervisées K-Means en 5 classes (à cause du bord noir) à partir de l'image fusionnée à 52 canaux

OTB → Learning → KMeansClassification



Le résultat est beaucoup plus satisfaisant que la classification colorimétrique, le point sensible est le choix de la taille du vecteur GLCM, pour bien prendre en compte les différentes textures, cela peut être compliquée quand une des textures a une taille assez différente des autres, comme dans l'exemple ci-dessus la texture en haut à gauche et on voit que c'est celle qui sort le moins bien.

Pourquoi le bord de l'image n'a-t-elle pas été traitée ? quelle est la taille de cet espace non traité ?

3.3.2. Classification texturale par analyse d'histogrammes locaux

Partir de l'image de Brodatz classée en 5 classes au point précédent

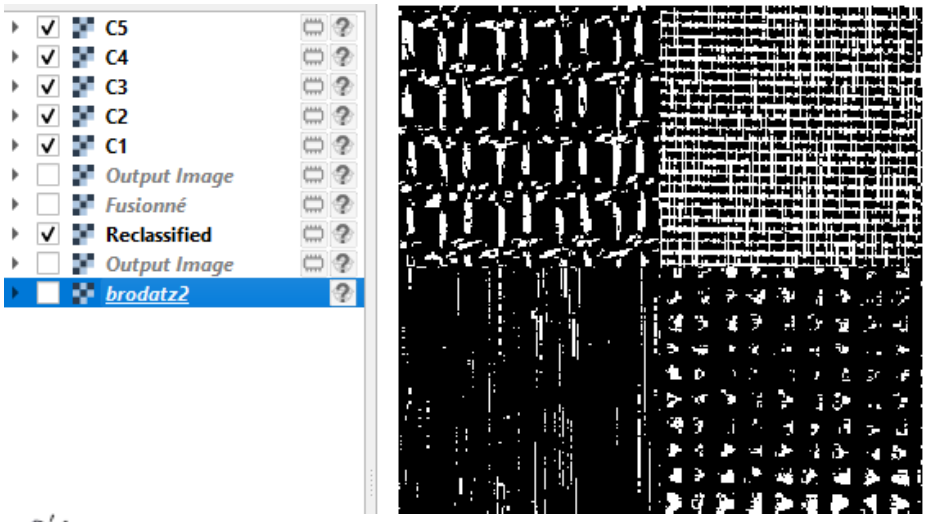
Créer un néocanal binaire pour chacune des classes (5 néocanaux) de telle façon que pour chaque néocanal la valeur du pixel sera :

1 = si le pixel appartient à la classe et 0 si pas dans la classe (masque binaire de la classe)

Utiliser la calculatrice raster : Raster calculator

Renommer les néocanaux au fur et à mesure : C1 … C5

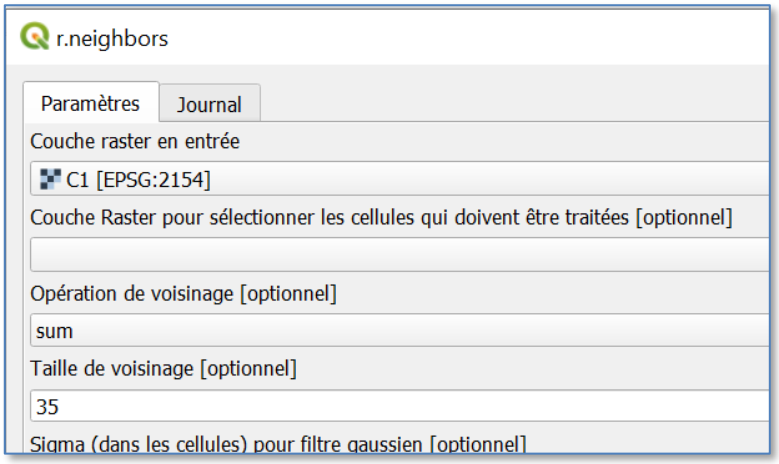
Ne pas refermer la boite de dialogue de la calculatrice entre chaque classe : appliquer la formule, renommer et passer à la classe suivante sans fermer la fenêtre.



Calculer un néocanal « histogramme local » (fenêtre 35x35) pour chacun des néocanaux Cx précédents en utilisant r.neighbors

Les néocanaux étant binaire le nombre de zéro est le complément du nombre de 1 dans la fenêtre de traitement, on va donc considérer comme histogramme simplifié uniquement le nombre de 1, pour cela on va calculer la somme des pixels dans la fenêtre de traitement avec la fonction (on pourrait aussi prendre la moyenne):

r.neighbors

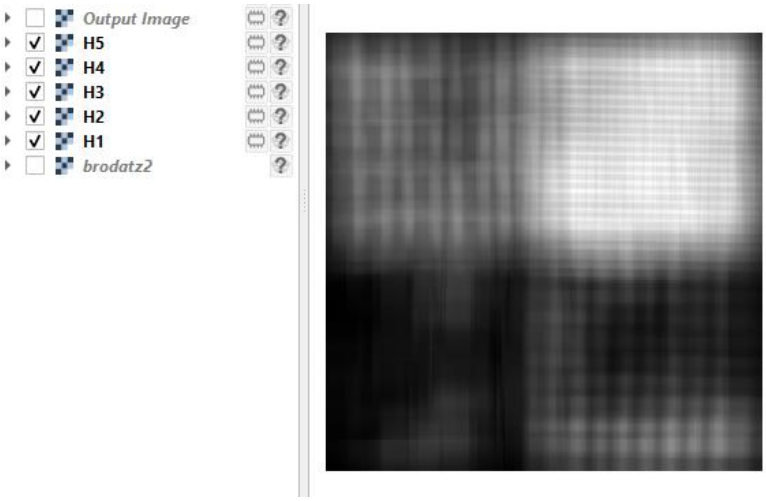


Fonction sum (somme des pixels)

Voisinage 35x35 pixels

Renommer les néocanaux : H1 … H5

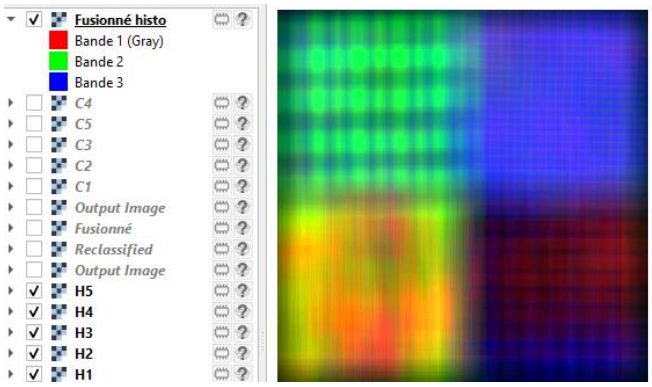
Ne pas refermer la boite de dialogue comme précédemment.



Donc en chaque pixel la valeur dans un canal i représente le nombre d'occurrence de pixels de la classe i qui est présent dans un voisinage de 35x35 (histogramme)

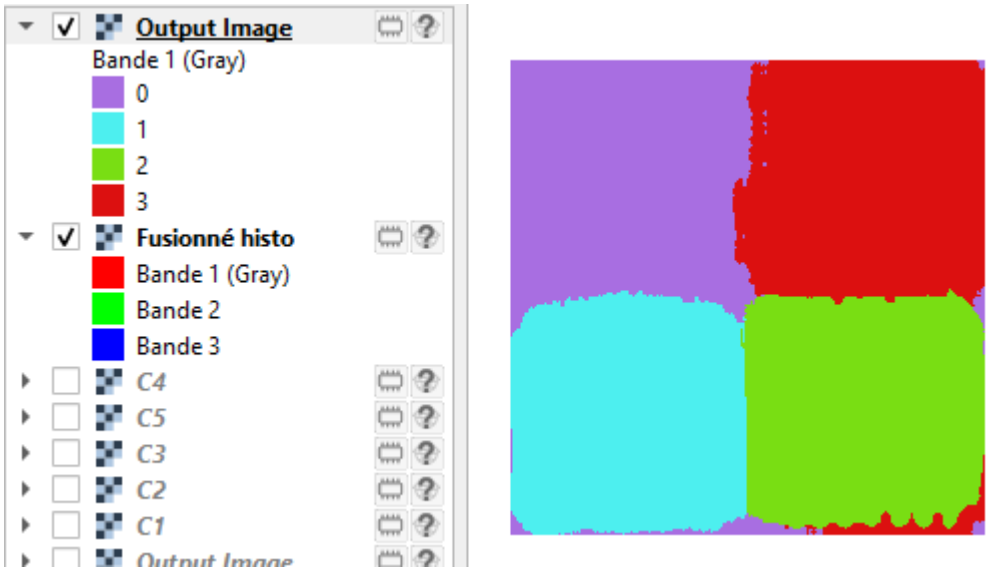
Fusionner les canaux histogrammes dans une image multibandes.

GDAL → Divers raster → Fusion



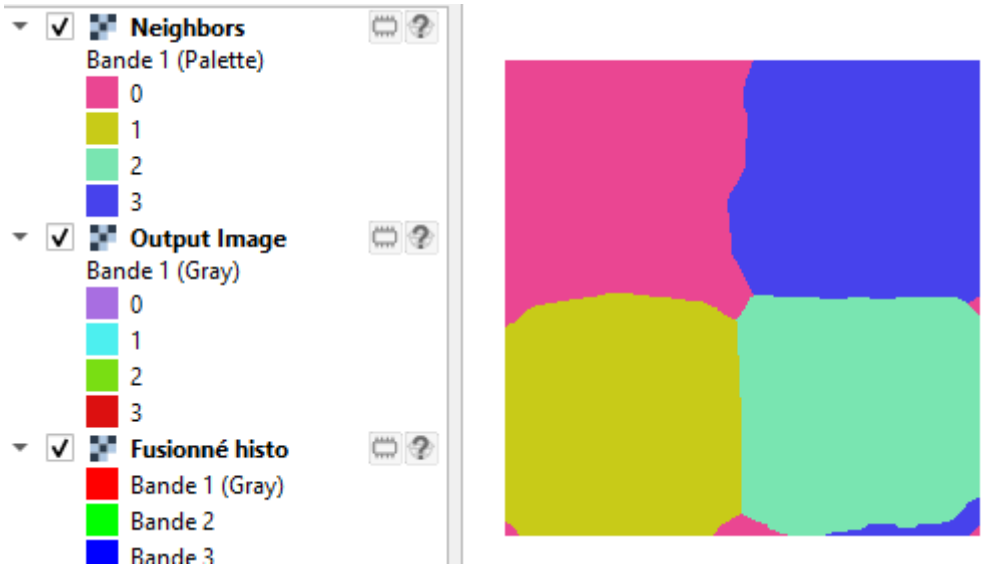
Faire une classification non supervisées K-Means en 4 classes à partir de l'image fusionnée des histogrammes

OTB → Learning → KMeansClassification



Faire un post-traitement pour lisser les zones par un filtrage par le MODE par exemple avec un voisinage d'une taille de 15 :

r.neighbors



On peut réaliser une approche analogue par classification supervisée avec création de ROI.

4. Analyse diachronique par télédétection

4.3. Extraction automatique des cultures sous QGIS à partir d'une analyse diachronique d'images Sentinel et du RPG

WISSEMBOURG Dossier des données du canton de Wissembourg en Alsace.

Le dossier SENTINEL2 contient 7 images satellites Sentinel 2 :

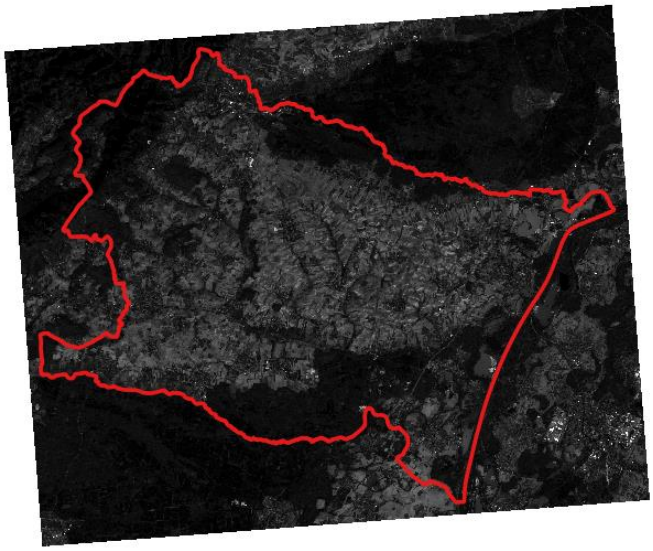
- S2_MSIL1C_2022_02_28_
- S2_MSIL2A_2022_03_25_
- S2_MSIL2A_2022_04_19_
- S2_MSIL2A_2022_05_09_
- S2_MSIL2A_2022_06_28_
- S2_MSIL2A_2022_07_18_
- S2_MSIL2A_2022_08_12_

Pour les mois de : février, mars, avril, mai, juin, juillet et août.
Pour chaque image le dossier IMG_DATA contient les bandes spectrales à 10m de résolution.

- SENTINEL2
  - S2_MSIL1C_2022_02_28_T102849_N0400_R108_T32
    - AUX_DATA
    - DATASTRIP
    - GRANULE
      - L1C_T32UMV_A026017_20220228T103305
        - AUX_DATA
        - IMG_DATA

Les images ont été téléchargées sur https://dataspace.copernicus.eu/

Seules les bandes B02 (bleu), B03 (vert), B04 (rouge) et B08 (proche infrarouge) à 10m de résolution, ont été conservées et uniquement sur l'emprise du canton de Wissembourg.

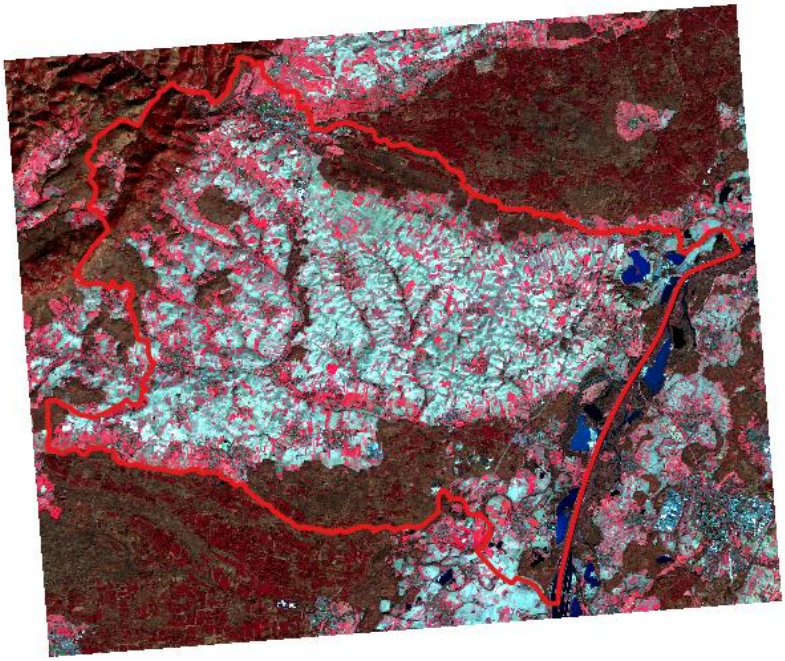


RPG_2022.shp Les parcelles agricoles du RPG (Registre Parcellaire Graphique) de 2022.

NOTICE_rpg-version2-0.pdf Contient les codes des cultures du RPG

canton_Wissembourg.shp Les limites du canton de Wissembourg

Afficher les images en composition colorée infra-rouge couleur (IRC) : 08 / 04 / 03

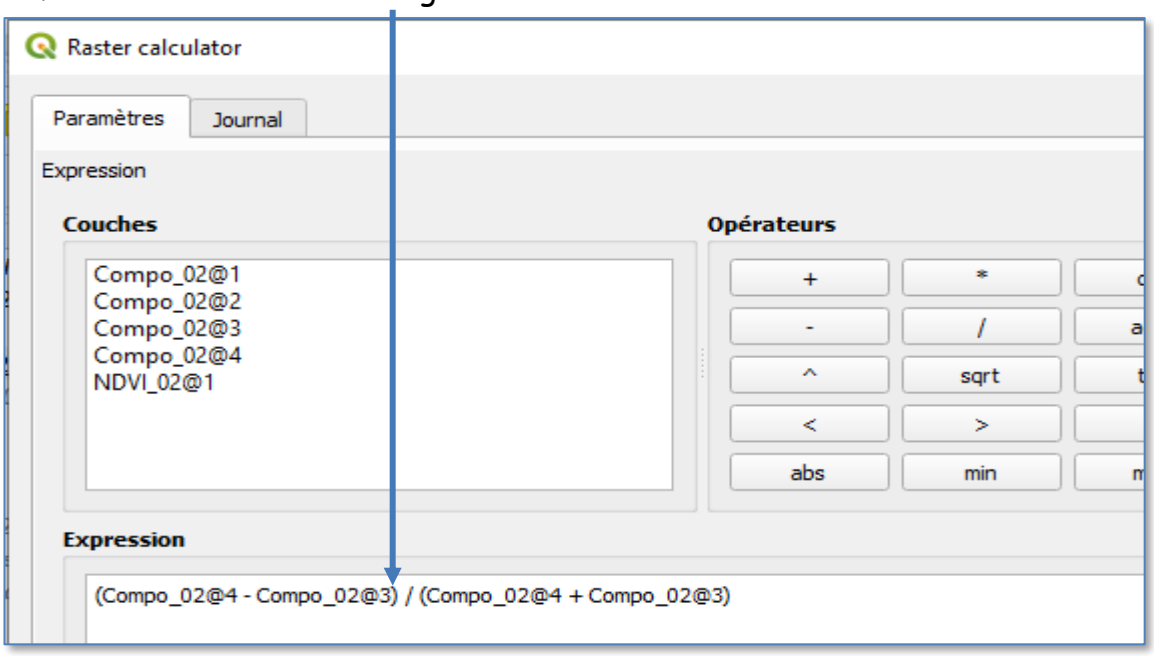


Composition colorée IRC de février

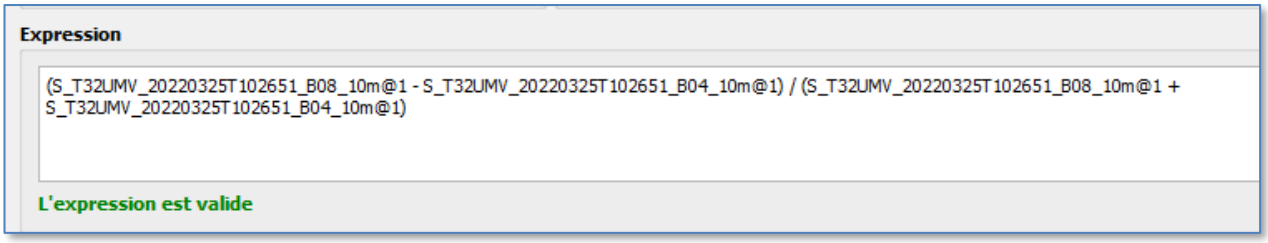
Calculer le NDVI aux différentes dates à l'aide de la calculatrice Raster

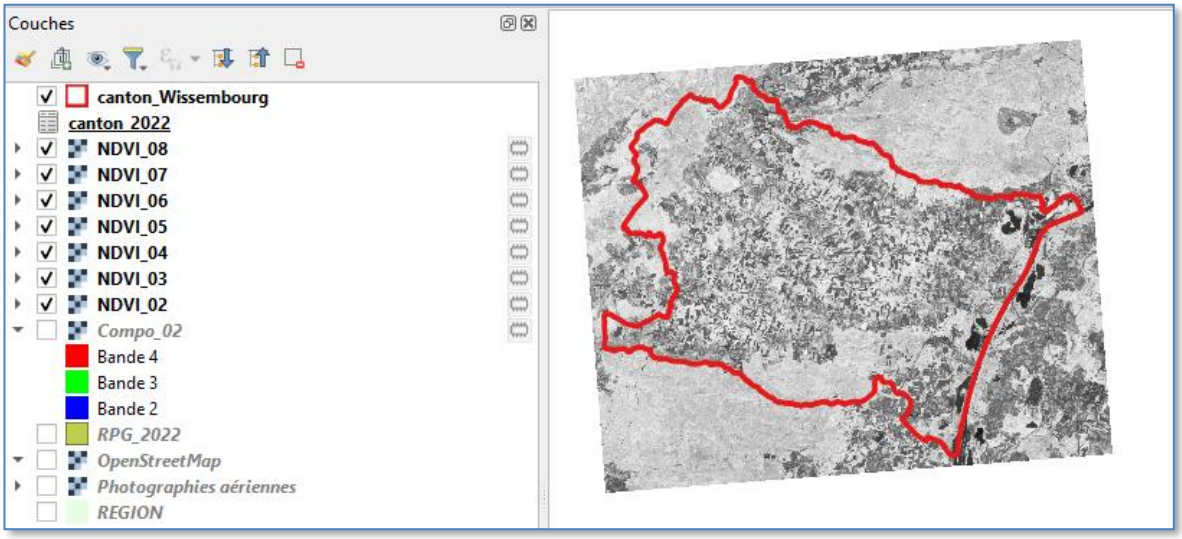
Analyse raster → Raster calculator

La formule du NDVI est intégrée dans la calculatrice Raster



On peut aussi calculer les NDVI directement à partir des fichiers des deux bandes (08 et 04) sans passer par une composition colorée, par ex pour l'image de mars :

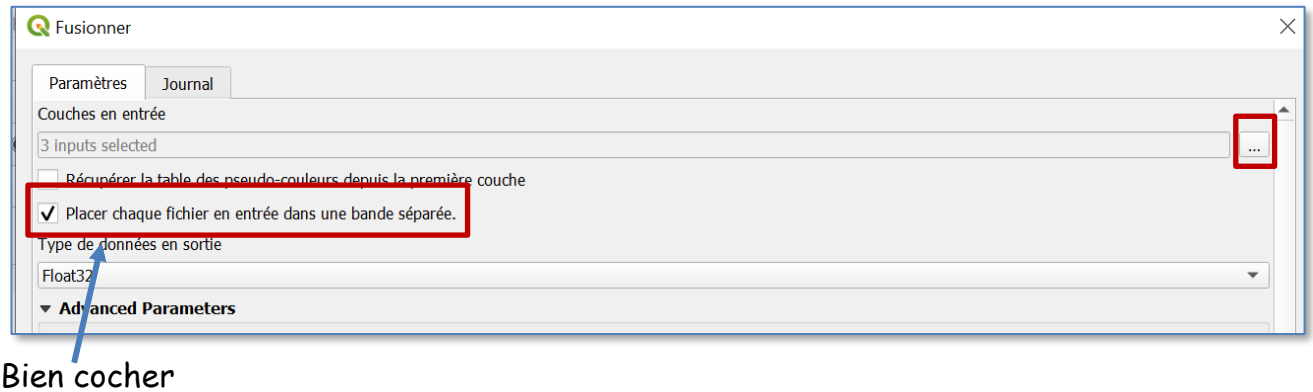




Regrouper les NDVI dans une image multicanal par fusion (ou en image virtuelle)

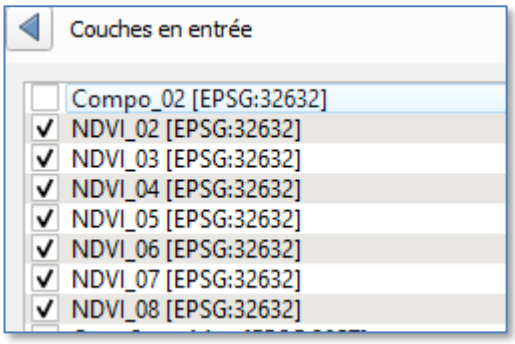
GDAL → Divers raster → Fusionner

Cliquer sur ... pour choisir les bandes à fusionner



Bien cocher

Ordonner les NDVI dans l'ordre des mois, pour que l'ordre des bandes créées respecte l'ordre chronologique.



C'est sur cette image que vous ferez la classification diachronique

Recodages des cultures

Nous allons analyser des classifications à différents niveaux de regroupement de type de cultures, pour se faire créer les codes suivants :

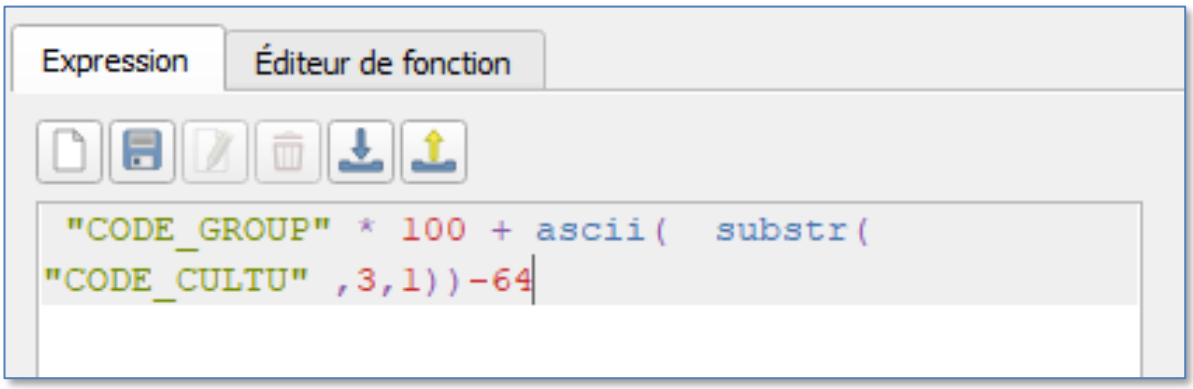
Le champs « CODE_CULTU » contient le code de la culture principale mais avec un code texte de 3 lettres, (cf PDF), pour les traitements rasters qui vont suivre il sera pratique d'avoir une version numérique de ce code, créer un nouveau champ entier « CODE » :

Proposez un recodage numérique avec des codes proches, **MAIS UNIQUE**, pour des cultures proches. Les deux premiers caractères de CODE_CULTU donne la culture et le 3e une variante par ex :

BTH = blé tendre d'hivers (groupe de culture 1)
BTP = blé tendre de printemps (groupe de culture 1)

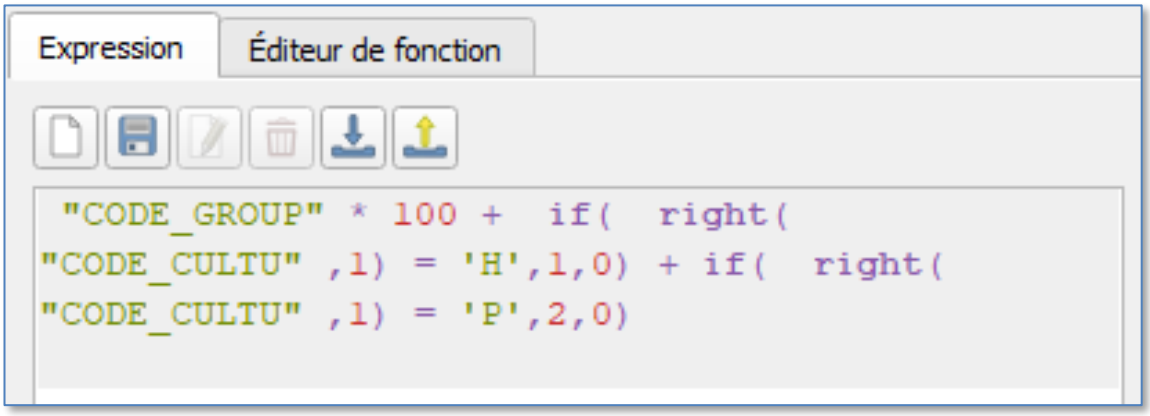
Le champ « CODE_GROUP » contient un code numérique du groupe de la culture

Exemple de recodage :



64 est le code ascci de la lettre 'A'

Créer aussi un autre code « CODE_HP » pour les groupes de cultures en séparant cultures de printemps et cultures d'hivers quand cela existe par exemple :



Créer aussi un code de groupe de cultures « CODE_G2 » plus simple que les groupes de départ :

CODE_G2	cultures
1	Blé et toutes les céréales
2	Maïs
5	Oléagineux : colza, tournesol, autres
8	Protéagineux
18	Prairies
20	Vergers et fruits à coques
21	Vignes
25	Légumes et fleurs
99	Autres : tous le reste

Avec la calculatrice de champ :

Expression | Éditeur de fonction

```
CASE
WHEN  "CODE_GROUP" in(1,3,4) THEN 1
WHEN  "CODE_GROUP" in(2) THEN 2
WHEN  "CODE_GROUP" in(5,6,7) THEN 5
WHEN  "CODE_GROUP" in(8) THEN 8
WHEN  "CODE_GROUP" in(18,19) THEN 18
WHEN  "CODE_GROUP" in(20,22) THEN 20
WHEN  "CODE_GROUP" in(21) THEN 21
WHEN  "CODE_GROUP" in(25) THEN 25
ELSE 99
END
```

Culture	CODE_CULTU	CODE_GROUP	CODE_G2	CODE_HP	CODE
Blé tendre d'hivers	BTH	1	1	101	108
Blé tendre de printemps	BTP	1	1	102	116
Maïs ensilage	MIE	2	2	200	205
Maïs	MIS	2	2	200	219

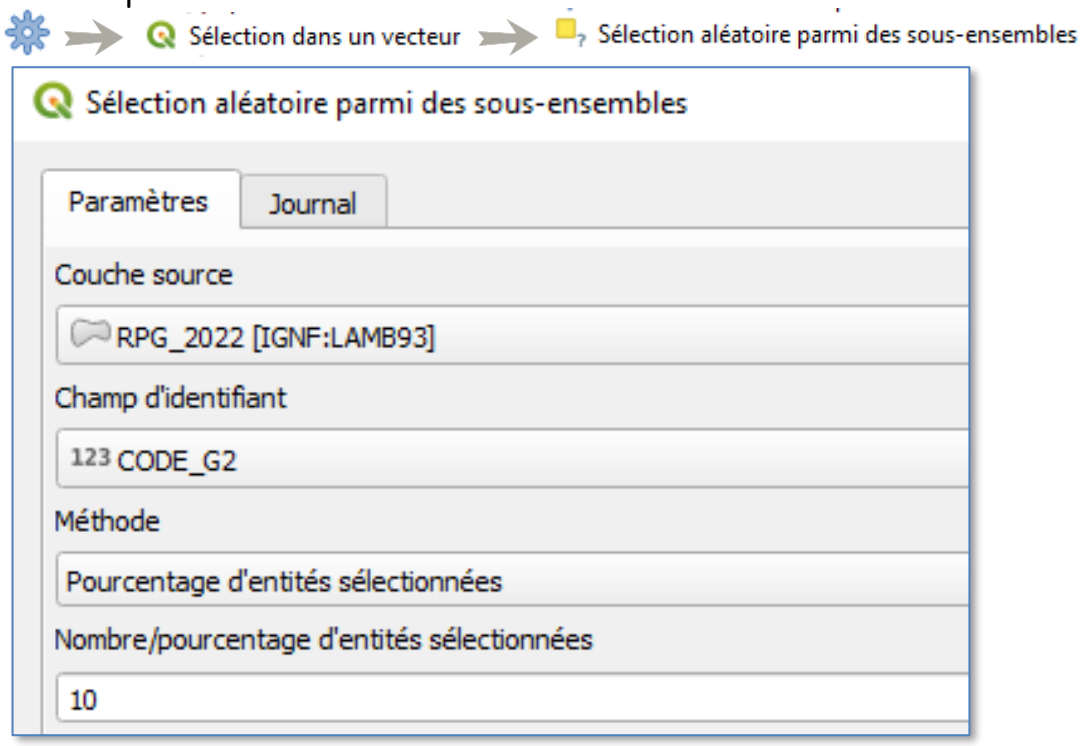
Séparer le fichier rpg_sample_L93 en 2 fichiers : un pour faire des ROI pour la classification, le second pour validation pour calculer par exemple une matrice de confusion.

Par exemple prendre 10% des lignes pour le jeu de données d'apprentissage :

Le champ « no » contient un numéro d'ordre arbitraire des parcelles (numéro de ligne dans la table, on pourrait sélectionner sur la base du reste de la division entière :

%10 donne le reste de la division entière par dix, si cela vaut zéro le numéro (no) est un multiple de dix. Mais pour des cultures faiblement représentées on risque de ne pas les retenir.

Utiliser plutôt :

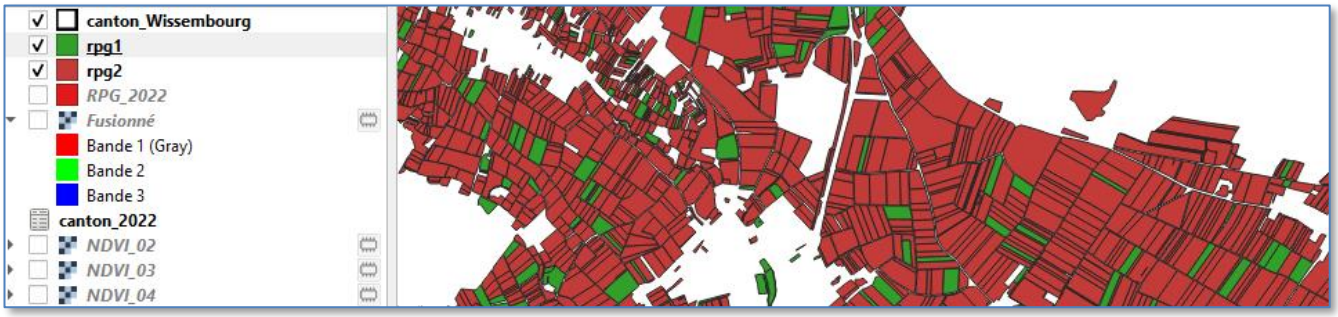


Exporter la sélection dans RPG1

Pour Créer RPG2 (sous ensemble de validation)

Inverser la sélection avec

Puis exporter la sélection dans RPG2



Classer les cultures

OTB → Learning → TrainImagesClassifier

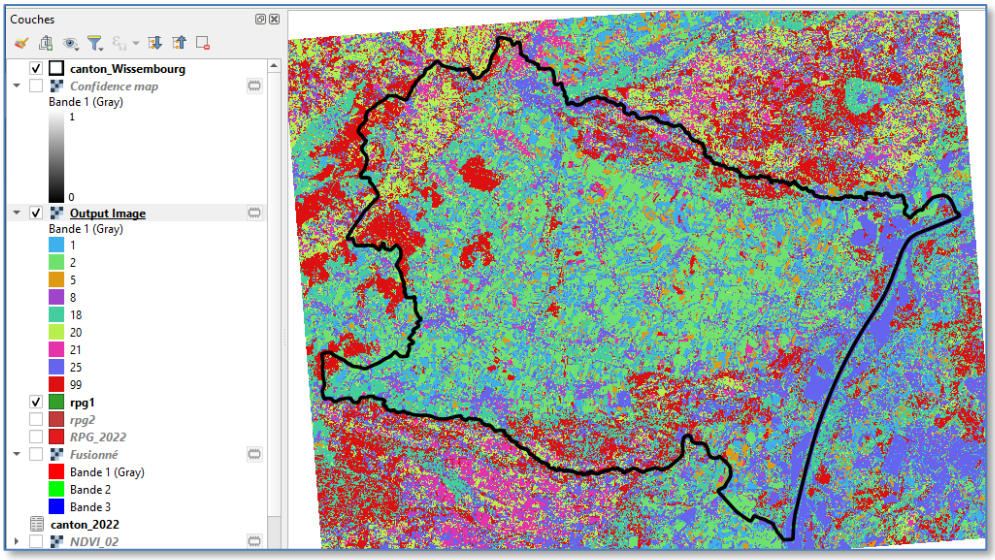
Cette fonction va calculer les statistiques de classes à partir du shapefile RPG1 sur cette image composite (fusion)

Prendre la méthode RF : Random Forest

Appliquer ensuite la classification avec :

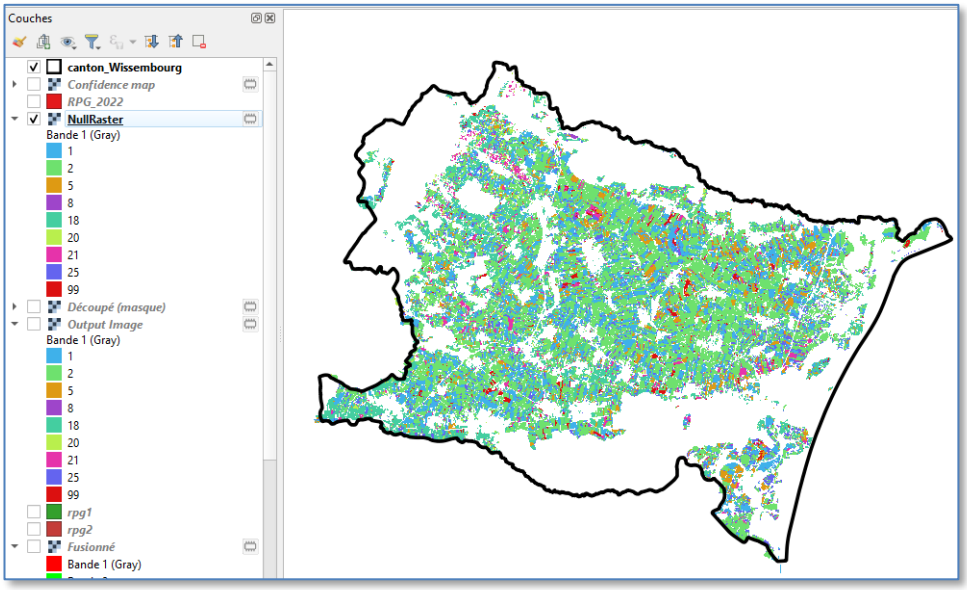
OTB → Learning → ImageClassifier

Classer successivement avec : CODE_G2, CODE_group, CODE_HP et CODE



Classification RF sur CODE_G2

Masquer le résultat avec les parcelles agricoles (rpg_2022.shp)



Classification RF sur CODE_G2 masquée

Calcul de la matrice de confusion avec le fichier RPG2

OTB → Learning → ComputeConfusionMatrix

matricel10.txt - Bloc-notes

Fichier Edition Format Affichage Aide

```
#Reference labels (rows):1,2,5,8,18,20,21,25,99
#Produced labels (columns):1,2,5,8,18,20,21,25,99,255
319408,2298,45103,104,11495,1395,1643,35660,1386,0
2097,757971,2850,4208,7675,3664,22718,79295,18264,0
9026,8706,57949,139,4746,824,6529,3890,1579,0
994,41,2,57,11,3,13,1107,1,0
2317,1160,5056,152,223613,22375,16002,11970,2549,199
126,135,64,51,5526,7561,3131,2305,118,0
13,6,10,124,1371,2849,8359,305,41,0
283,1185,68,76,156,337,530,2982,34,0
2685,26712,5777,649,29384,7711,6881,13024,6627,0
```

Mettre en forme dans Excel

			Céréales	Maïs	Oléagineux	Protéagineux	Prairies	Vergers	Vignes	Légumes	Autres		
			1	2	5	8	18	20	21	25	99		PP
Céréales		1	319408	2298	45103	104	11495	1395	1643	35660	1386	418492	0.76
Maïs		2	2097	757971	2850	4208	7675	3664	22718	79295	18264	898742	0.84
Oléagineux		5	9026	8706	57949	139	4746	824	6529	3890	1579	93388	0.62
Protéagineux		8	994	41	2	57	11	3	13	1107	1	2229	0.03
Prairies		18	2317	1160	5056	152	223613	22375	16002	11970	2549	285194	0.78
Vergers		20	126	135	64	51	5526	7561	3131	2305	118	19017	0.40
Vignes		21	13	6	10	124	1371	2849	8359	305	41	13078	0.64
Légumes		25	283	1185	68	76	156	337	530	2982	34	5651	0.53
Autres		99	2685	26712	5777	649	29384	7711	6881	13024	6627	99450	0.07
			336949	798214	116879	5560	283977	46719	65806	150538	30599	1835241	1.00
		PU=	0.95	0.95	0.50	0.01	0.79	0.16	0.13	0.02	0.05		

Analyser et discuter La matrice de confusion

Calculer pour toutes les parcelles la culture majoritaire par statistique zonale.

Analyse raster → Zonal statistics

Demander le calcul des statistiques « majorité » pour avoir le code de classe qui est majoritaire en surface dans la parcelle.

Construire une matrice de confusion à partir du calcul de stat. Zonale précédent.

4.4. Cartographie par Télédétection diachronique des dégâts de la tempête de 1999 sur les forêts de la Gironde dans le canton « Sud Medoc »

SUD_MEDOC_1999 est le dossier du jeu de données

ADMIN-EXPRESS Sous-dossier des limites administratives

- canton_sud_medoc.shp
- cantons_gironde.shp
- communes_canton_sud_medoc.shp
- DEPARTEMENT.SHP

Dans le sous-dossier LANDSAT il y a deux sous-dossiers : 1999 et 2000
2 extraits d'images des satellites Landsat : une du 08/10/1999 et une du 20/01/2000
Avant et après la tempête de décembre 1999.
Seules les 4 premières bandes ont été conservées :

Band1 : 0,45 - 0,52 micromètres (bleu-vert)
Band2 : 0,52 - 0,60 micromètres (vert)
Band3 : 0,63 - 0,69 micromètres (rouge)
Band4 : 0,76 - 0,90 micromètres (proche infrarouge)

Il s'agit d'images en réflectance de surface fournies par le service EarthExplorer de l'USGS. Il y a un fichier Tif par bande.
Consultez le document descriptif des données « ledaps_product_guide.pdf » pour voir le codage des données, vous masquerez les données hors domaine.

CORINE_CLC_2000 Sous-dossier de Corine Land Cover qui donne l'usage des terres

ETUDE_INVENTAIRE_FORESTIER_NAT Sous-dossier des résultats pour la Gironde, de l'étude par télédétection des dégâts de la tempête menée par l'inventaire forestier national.

Rappel de la formule de l'indice de végétation $\mathrm{NDVI} = \frac{PIR-R}{PIR+R}$

Formule d'estimation simplifiée du cubage de bois d'une forêt :
A l'hectare :

$$V = G * H * F$$

Avec :

V = Volume de bois / ha m^3 / ha
G = surface terrière à l'hectare du peuplement en m^2/ha
H = hauteur totale ou hauteur découpe moyenne du peuplement en m
F = coefficient de forme du peuplement

Pour le Pin maritime on prendra (données inventaire forestier sur les Landes) :

G = 8.3 m^2/ha

H = 20 m

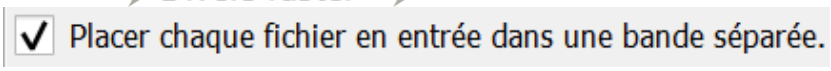
F = 0.46

Déf : La surface terrière d'un arbre (noté g) correspond à la surface de la section transversale de cet arbre à 1.30m de hauteur. La surface terrière d'un peuplement (G) est la somme des g sur un hectare.

Vous allez estimer la destruction de la forêt (essentiellement des pins maritime) par une diminution du NDVI entre 1999 (avant tempête) et 2000 (après tempête). Vous considérerez qu'une diminution de au moins 0.2 du NDVI correspond potentiellement à une destruction partielle ou totale de la forêt en un point.

Les étapes :

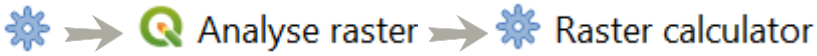
● Créer les compositions colorées « vrai couleurs » et « infrarouge couleurs » pour les images LANDSAT de 1999 et 2000.

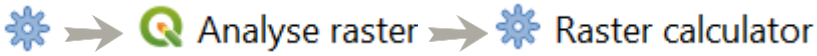
Avec l'option 

permet de regrouper des bandes dans une même image multispectrale.

● Extraire par masquage avec Corine Land Cover les zones forestières, (code de niveau 1 de Corine pour la forêt est 3)

● Cartographier les zones potentiellement détruites par analyse de l'évolution du NDVI entre les deux dates d'images LANDSAT. (seuil de 0.2 d'évolution du NDVI)



● Calculez et cartographier les surfaces détruites en ha dans le canton et par commune.

● Calculer et cartographier le cubage de bois détruit ? pour simplifier on considérera que la forêt n'est composée que de pin maritime.