Master Géomatique

UE « Télédétection et traitement d'images »

Travaux Pratiques :

Initiation à la télédétection avec Orfeo ToolBox et QGIS

8 et 22 novembre 2021

Stéphane DUPUY

CIRAD, UMR TETIS, F-34398 Montpellier, France.
TETIS, Univ Montpellier, AgroParisTech, CIRAD, CNRS, INRAE, Montpellier, France

Ce(tte) œuvre est mise à disposition selon les termes de la Licence Creative Commons Attribution 4.0 International.

Ce document a été élaboré pour le Master Géomatique : UE « Traitement d'images, télédétection, satellite »

© Cirad, Département Environnement et Société, Unité Mixte de Recherche Territoire Environnement Télédétection et Information Spatiale, novembre 2021

SOMMAIRE

Table des matières

1. Les images 5

   a. Rappels sur les images 5
   b. Images utilisées dans ces TP 6

2. Les logiciels 8

   a. OrfeoToolBox 8
   b. QGIS 8

3. Visualiser des images (QGIS) 9

   a. Afficher une image multispectrale 9
   b. Utiliser la barre d'outils « raster » 9
   c. Utiliser les outils de navigation cartographique 9
   d. Principe de visualisation 10
   e. Consulter l'histogramme 11
   f. Recherche des informations générales sur l'image 11
   g. Image panchromatique 12

4. Prétraitements d'une image à THRS (OTB + QGIS) 13

   a. Mosaïque des tuiles pour l'image Panchromatique 13
   b. Conversion d'une image en Comptes Numériques (CN) en réflectance apparente / TOA (Top Of Atmosphere) 14
   c. Orthorectification 16
   d. Recalage avec points d'amer avec l'extension de géoréférencement de QGIS 18
   e. Orthorectification avec sélection automatique de points d'amer sur OTB (optionnel) 21
   f. Pansharpening (OTB) 24

5. Prétraitements d'images HRS Sentinel-2 (OTB) 26

   a. Description d'une image Sentinel-2 26
   b. Découper une image 26
   c. Ré-échantillonner les bandes S2 à 20m 27
   d. Compiler les bandes dans une seul image multispectrale (Stack) 27
   e. Exécuter le script 28

6. Création de néo-canaux et détection de changements (OTB + QGIS) 29

   a. Indice de végétation : NDVI 29
   b. Indice de brillance 29
   c. Composition colorée diachronique et détection de changements avec 2 NDVI 30

7. Classification non-supervisée (OTB + QGIS) 31

   a. Classification K-Means 31
   b. Vérification de la classification 31
   c. Utilisation de l'extension de QGIS ThRasE 32

d. Appliquer les traitements à un image S-2 (optionnel) 32

8. Classification supervisée (OTB + QGIS) 33

a. Base de données de référence 33
b. Calcul des statistiques 34
c. Algorithme Bayes (Normal Bayes Classifier) 34
i. Apprentissage 34
ii. Prédiction 35
iii. Analyser la classification 36
d. Algorithme SVM (Support Vector Machine) 36
e. Algorithme RF (Random Forest) 37

9. Validation (OTB) 39

a. Rappel sur la validation d'une classification 39
b. Valider une classification avec OTB (optionnel) 40

10. Remerciements 40

1. LES IMAGES

a. Rappels sur les images

Les images utilisées en télédétection se différencient par leurs résolutions spatiales, radiométriques et spectrales. En fonction des objectifs et des thématiques, le choix se portera sur un ou plusieurs types d'images.

➢ Résolution spatiale

La dimension du pixel définit la limite de résolution spatiale d'un capteur. Exprimée en mètres au niveau du sol, elle définit la plus petite surface pour laquelle les différentes réponses radiométriques des objets pouvant composer cette plage sont confondues.

On distingue les images à :

- Basse résolution spatiale, tel Modis/Terra (pixel de 500m),
- Haute résolution, tels Sentinel-2 (10 à 60m), Landsat (15 à 90m), Spot 5 (10m), RapidEye (5m),
- Très haute résolution, tels Spot 6/7 (1.5m), Pléiades (0.5m), WorldView 3 (0.3m)

L'augmentation de la résolution réduit d'autant le champ de visée lors de l'acquisition d'image.

La taille du pixel constitue une donnée très importante car elle fixe très largement l'échelle d'analyse.

La prise d'image en mode panchromatique permet d'augmenter la résolution au détriment de l'information spectrale (ex Pléiades : 2m en mode multispectral, 50cm en mode panchromatique).

➢ Résolution radiométrique

Elle correspond à la capacité d'un système d'acquisition à séparer de petites différences de niveaux d'énergie à l'intérieur d'une même bande spectrale.

Cette capacité dépend de la sensibilité du capteur (rapport signal/bruit) mais aussi du type de codage retenu. Les comptes numériques des images de télédétection peuvent être échantillonnées sur :

- 8 bits (signés ou non) par pixel > 256 pas d'échantillonnage : images Spot, Landsat 1 à 7,
- 11 bits (signés ou non) > 2 048 niveaux (images WorldView 2, Landsat 8,),
- 16 bits (signés ou non) > 65 536 niveaux (images radar),
- 24 bits (signés ou non) > 16 millions (radar),

La dynamique des images peut être d'autant plus importante que le nombre de niveaux théoriques est élevé.

➢ Résolution spectrale

La résolution spectrale détermine la capacité d'un capteur à utiliser de petites fenêtres de longueur d'onde pour discriminer des signatures spectrales. Plus la résolution spectrale est fine, plus les fenêtres des différents canaux sont étroites.

b. Images utilisées dans ces TP

➢ Images à THRS

DINAMIS - Dispositif Institutionnel National d'Approvisionnement Mutualisé en Imagerie Satellitaire - permet d'accéder gratuitement à un ensemble d'imagerie satellitaire à des fins de recherche scientifique, d'aménagement ou de planification territoriale, de suivi ou de gestion environnementale, d'innovation ou de création de valeur ajoutée.

Pour plus d'informations consulter le site web : https://dinamis.data-terra.org. À partir du moment où une connexion est établie sur le site, il est possible de consulter le catalogue et de télécharger les images disponibles (pour Spot6/7 et/ou Pléiades).

Les images Spot6/7 diffusées par Dinamis sont disponibles en deux différents niveaux de traitement : primary (primaire) et ortho. Tous les produits sont corrigés pour la radiométrie et les distorsions du capteur, en utilisant des paramètres d'étalonnage internes, éphémérides et mesures d'altitude.

Nous allons notamment utiliser dans ces travaux pratiques (informations issues du site de Dinamis : https://dinamis.data-terra.org/produits-spot-6-spot-7) :

- **Une image acquise le 01/09/2018 acquise au niveau « Primary » (capteur parfait).** Il s'agit du niveau de traitement (niveau 1) le plus proche de l'image telle qu'acquise par le capteur, pour garantir une restitution parfaite des conditions d'acquisition : le capteur est mis en géométrie rectiligne et l'image est corrigée de toute distorsion radiométrique. Le produit Primary est conçu pour les utilisateurs familiarisés avec les techniques de traitement d'imagerie satellite qui souhaitent appliquer leurs propres méthodes de production (orthorectification ou modélisation 3D par exemple). Les produits de niveau 1 sont géo référencées ce qui signifie que les métadonnées du produit renseignent sur la localisation exacte de l'image.
- **Une image acquise le 24/05/2020 au niveau « Ortho ».** Il s'agit du traitement de niveau 3 (Ortho) qui consiste à appliquer des corrections radiométriques identiques à celles du niveau 1. L'image géoréférencée et corrigée des effets d'angles et de relief, optimale pour une utilisation simple et directe de l'image et intégration dans un Système d'Information Géographique.

Le format des produits SPOT 6/7 et PLEIADES est DIMAP V2. (Le format DIMAP V2 est une version améliorée du format DIMAP V1 de SPOT)

➢ Série temporelle d'images à HRS

La mission Sentinel-2 a débutée en 2015 dans le cadre du programme Copernicus de l'ESA (agence spatiale européenne). Elle est constituée de 2 satellites : S2A et S2B ce dernier est opérationnel depuis mi 2017. Avec une fauchée de 290 Km, ils réalisent une couverture globale des surfaces continentales tous les 5 jours. Le capteur est multi-résolution (de 10m à 60m) et permet de réaliser des acquisitions avec 12 bandes spectrales (Cf. Figure 1).

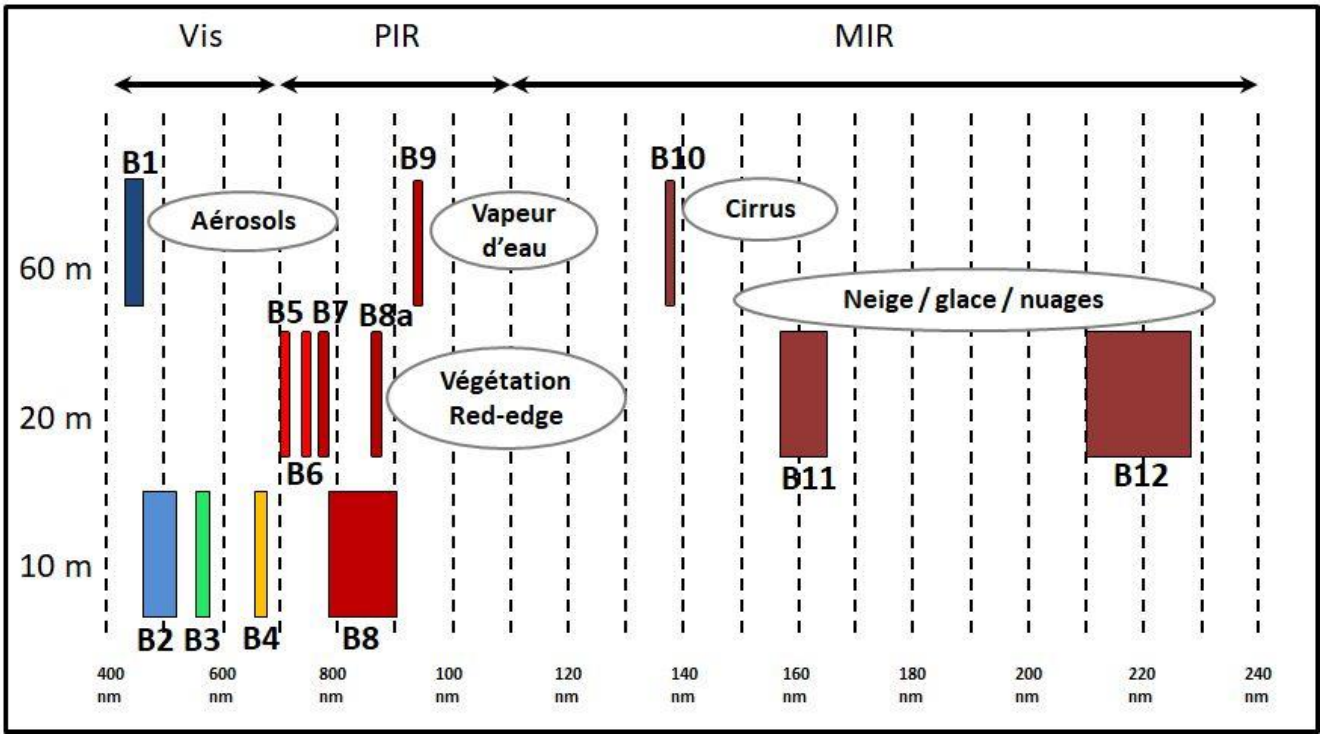

Figure 1 : Bandes spectrales des satellites Sentinel-2 et leurs résolutions spatiales : 4 bandes (B, V, R et PIR) sont à 10 mètres de résolution spatiale, 6 bandes (red-edge et moyen infrarouge) à 20 mètres et 3 bandes (aerosol, vapeur d'eau et cirrus) à 60m

Ces images sont disponibles en libre accès ce qui a bouleversé profondément les usages et les traitements de données satellitaires. En effet, compte tenu du nombre d'images disponibles on parle maintenant d'utiliser des séries temporelles d'images pour produire des classifications.

Le pôle de données et de services surfaces continentales Theia est un des membres actifs de l'infrastructure de recherche sur le système Terre, Data Terra. Il développe des produits à valeur ajoutée pour la communauté scientifique et les acteurs publics français. Le CNES diffuse notamment des données Sentinel-2 de niveau 2A. Ces produits fournissent des réflectances de surfaces corrigées des effets atmosphériques et sont fournies avec un masque des nuages et leurs ombres. Le traitement des données est réalisé avec la chaîne MAJA (autrefois appelée MACCS), par le centre de traitement MUSCATE du CNES, pour le compte du centre de données Theia. Les données sont accessibles via cette plateforme : https://theia.cnes.fr

Au cours de cette formation nous allons mettre en œuvre les traitements permettant d'exploiter des images à Très Haute Résolution Spatiale acquises via Dinamis et des images Sentinel-2 de niveau 2A distribuées par THEIA.

2. LES LOGICIELS

a. OrfeoToolBox

L'Orfeo ToolBox (OTB) est une librairie de traitement d'image de télédétection initié par le CNES en 2006 et qui est constamment en développement. Sa structure est constituée d'une librairie avec une série d'algorithmes construits en langage C++ qui servent au traitement des ressources issues de la télédétection.

L'utilisation des modules d'OTB peut se faire de plusieurs manières :

- Par ligne de commande. Dans ce cas, nous utilisons la fenêtre de commandes OSGeo4W, contenue dans le dossier de QGIS.
- Sur QGIS, accessible depuis la fenêtre de boîte à outils.
- Avec l'interface graphique MAPLA, disponible dans le dossier d'OTB.

Le téléchargement du logiciel est disponible via ce lien : https://www.orfeo-toolbox.org

Les TP sont préparés avec la version (7.4).

Lorsque le téléchargement est terminé, il faut décompresser l'archive zip et copier le dossier sur le disque dur de la machine. Il n'est pas nécessaire d'effectuer une installation.

b. QGIS

QGIS (également appelé Quantum GIS jusqu'à la version 1.9) est un logiciel SIG libre multiplateformes publié sous licence GPL (General Public Licence). Le développement a débuté en mai 2002 et est sorti en tant que projet sur Source Forge en juin 2002. Plus d'information via ce lien : https://www.qgis.org/fr/site

Les TP sont préparés avec la version stable : 3.16 LTR

Il faut installer QGIS sur la machine avant de commencer les traitements.

Au cours de ce TP, nous allons appeler les outils de l'OTB en utilisant MAPLA ou des lignes de commandes via l'outil OSGEOW de QGIS.

Les applications de l'OTB peuvent être lancées à partir de QGIS de la même façon que Mapla. Pour que cela fonctionne, il est nécessaire de procéder à la configuration suivante :

- Démarrer QGIS Desktop
- Installer l'extension OTB. Si l'extension n'est pas disponible, se rendre sur le site suivant et suivre la procédure indiquée : https://gitlab.orfeo-toolbox.org/orfeotoolbox/qgis-otb-plugin
- Cliquer sur le bouton « options » de la boite à outil de traitements
- Sélectionner le fournisseur de traitements de l'OTB
- Entrer les chemins vers les dossiers suivants :
    - OTB application folder : entrer le chemin d'accès « lib/otb/applications » de votre installation OTB ;
    - OTB folder : entrer le chemin d'accès vers le dossier de votre installation OTB.
    - Cocher sur « Activate » pour pouvoir utiliser les outils OTB.

3. VISUALISER DES IMAGES (QGIS)

Objectif : à l'issue de cet exercice, vous serez capable : d'afficher une image à l'écran et de connaître les manipulations de base avec QGIS, de consulter les propriétés d'une couche raster, de manipuler les images numériques et de gérer les différents canaux qui constituent ces fichiers

a. Afficher une image multispectrale

Dans les données disponibles, chercher l'image Spot6/7 acquise en 2020. Décompresser les 2 fichiers pour accéder à l'image multispectrale et l'image panchromatique.

- Ouvrir le logiciel QGIS
- Cliquer sur **« Couche > Ajouter une couche > Ajouter une couche raster »**
- Naviguer dans l'arborescence des données pour ouvrir l'image multispectrale acquise en 2020 : SPOT6_MS_20200524_ORT_TOA.tif
- Cliquer sur Ouvrir puis Ajouter

b. Utiliser la barre d'outils « raster »

La barre d'outil « raster » permet de modifier le contraste, la luminosité (etc…) des images. Si elle n'est pas disponible, il faut la sélectionner :

- clic droit sur le bandeau puis dans **« barres d'outils »** cocher **« raster »**

Figure 2 : Barre d'outil raster de QGIS

- Tester les différents outils disponibles

c. Utiliser les outils de navigation cartographique

La barre d'outil de navigation cartographique permet de se déplacer dans l'image et de faire des zooms avant ou arrière.

Figure 3 : Barre d'outils de navigation cartographique

Tester les différents outils disponibles

d. Principe de visualisation

Rappel : Les images satellitaires ou aériennes numériques sont acquises dans différentes plages de longueurs d'ondes du spectre électromagnétique. L'écran de visualisation est constitué de trois plans colorimétriques : Rouge, Vert et Bleu. Conformément aux principes de la synthèse additive, le chargement de trois bandes spectrales identiques ou différentes restituera soit une image « monocanal » en tons de gris, soit une composition colorée.

Attention à l'ordre de chargement des plans ! Les images Spot et Pléiades livrées par Airbus ont un ordre de bandes adapté pour que la visualisation dans un outil de SIG se fasse par défaut en couleurs naturelles. En effet, pour obtenir une composition colorée en couleurs naturelles, il faut associer au plan rouge le canal rouge (ici 1), au plan vert le canal vert (ici 2) et au plan bleu le canal bleu (ici 3). L'ordre « classique » des bandes spectrales respecte les longueurs d'ondes soit Bleu, Vert, Rouge. Attention donc car en fonction du capteur et du fournisseur il faudra adapter l'ordre des bandes.

Afin d'ouvrir les propriétés de la couche, **sélectionner l'image dans la liste des couches > cliquer sur le bouton droit > Propriétés.** La fenêtre Propriétés de la couche s'ouvre (Cf. figure 4). **Sélectionner l'onglet « Symbologie ».**

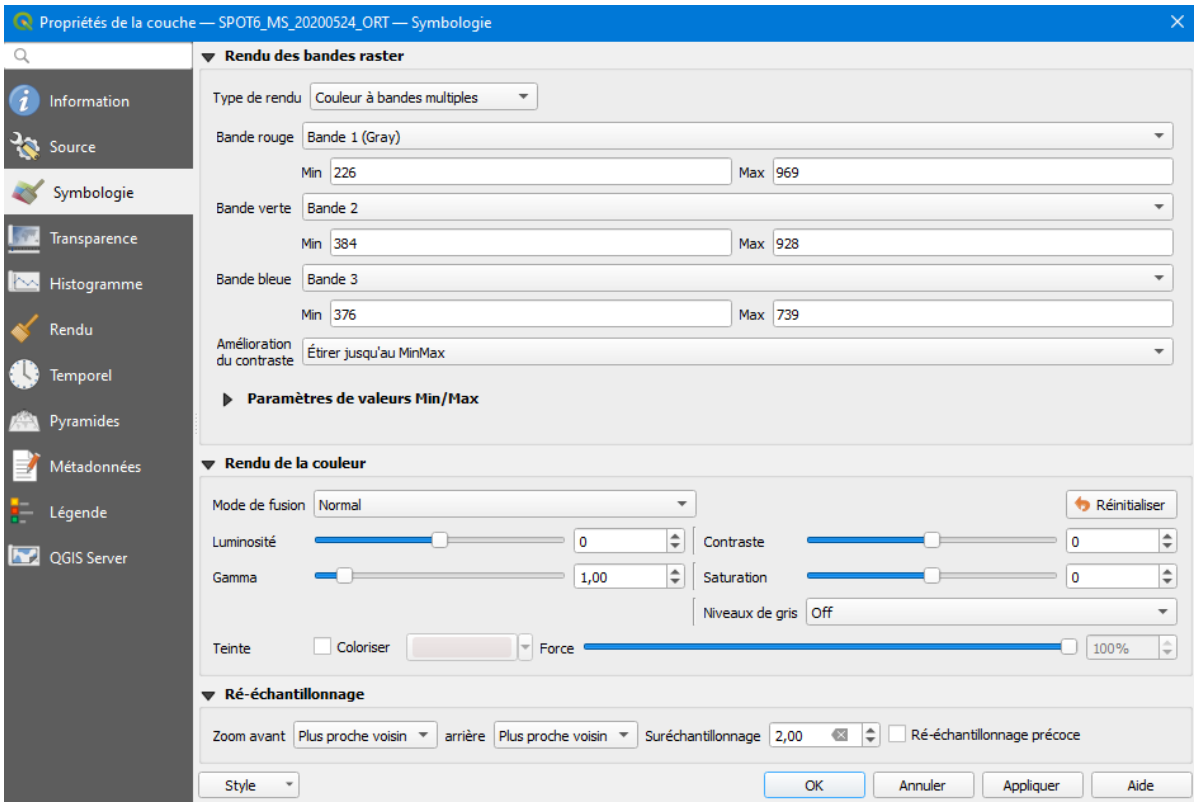

Figure 4 : Fenêtre des Propriété de la couche

Dans l'onglet **« Symbologie »**, sélectionner les trois canaux à afficher pour obtenir une composition colorée fausses couleurs infrarouge (cf. cours), sachant que pour cette image, l'ordre des canaux est le suivant :

- Canal 1 : rouge
- Canal 2 : vert
- Canal 3 : bleu
- Canal 4 : proche Infrarouge (PIR)

Q2 : Observer l'image affichée et comparer la végétation et les autres types d'occupation du sol. Que constatez-vous ?

N.B. Attention à l'ordre de chargement des bandes spectrales ! Pour obtenir une composition colorée infrarouge, les canaux proche infrarouge, rouge et vert doivent être respectivement affectés au Rouge, au Vert et au Bleu.

Afficher une seconde fois la même image afin de comparer une composition colorée en couleurs naturelles et fausses couleurs.

Q1 : La végétation active apparaît dans quelle couleur ? Pourquoi ?

e. Consulter l'histogramme

Ouvrir les propriétés de la couche et sélectionner l'onglet « Histogramme » (Cf. figure 5).

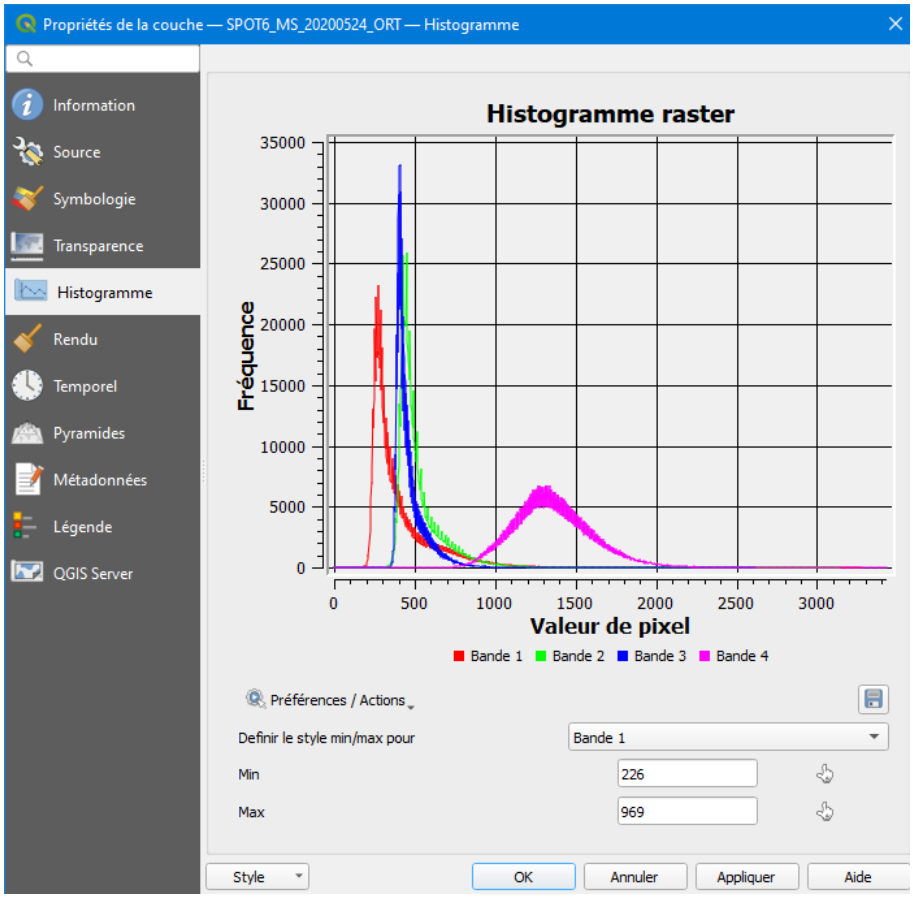

Figure 5 : Histogramme de l'image

Cet outil permet d'accéder à l'information spectrale d'une image. Avec le menu « Préférences », afficher alternativement chaque bande spectrale pour observer la distribution de la valeur des pixels.

Revenez dans l'interface graphique. À l'aide de l'outil , cliquer sur différents types d'occupation du sol (urbain, cultures, forêt, eau, etc.) et observer les différentes valeurs des pixels.

f. Recherche des informations générales sur l'image

Ouvrir les propriétés de la couche et sélectionner l'onglet « Information » (Cf. figure 6).

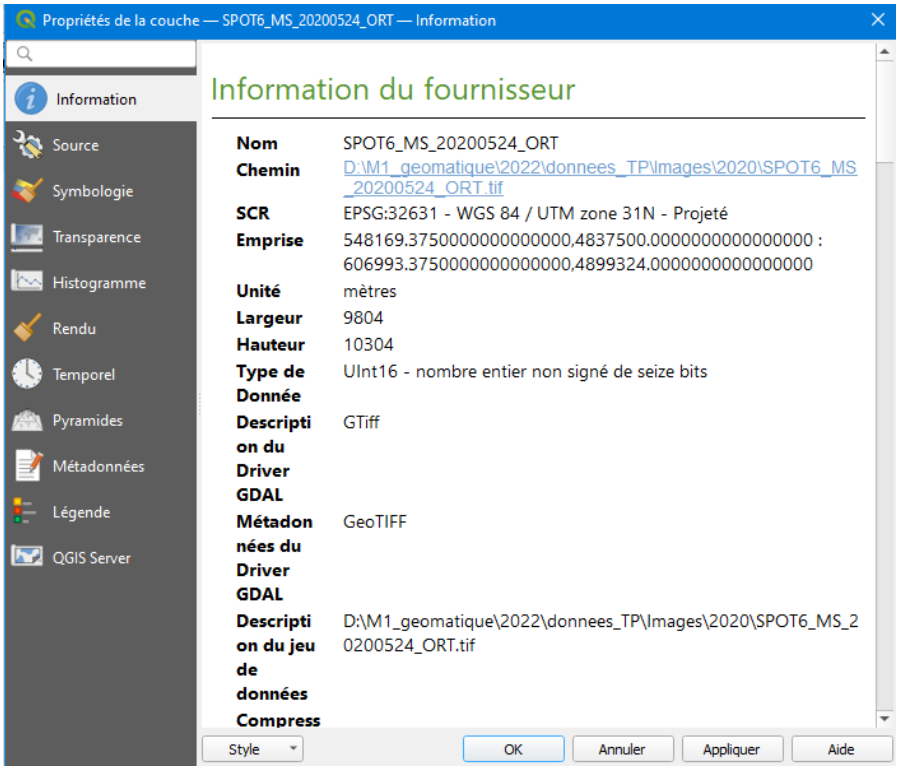

Figure 6 : Information sur l'image

***Q3** : Trouver les informations suivantes :*

- ***Résolution spatiale***
- ***Projection (code EPSG)***
- ***Type de la donnée***
- ***Nombre de lignes et de colonnes***

g. Image panchromatique

Afficher l'image panchromatique associée en la superposant à l'image multispectrale utilisée jusqu'à présent. Il s'agit de cette image : SPOT6_P_20200524_ORT_TOA.tif

Ajuster le contraste de l'image panchromatique

Pour faciliter la comparaison entre les 2images, installer l'extension « Map Swipe Tool » puis à l'aide du bouton de cette extension sélectionner une des 2 images :

Faites glisser une image sur l'autre à l'aide du curseur.

***Q4** : Trouver les informations sur l'image panchromatique :*

- ***Résolution spatiale***
- ***Nombre de ligne et de colonnes***

Que constatez-vous ?

4. PRÉTRAITEMENTS D'UNE IMAGE À THRS (OTB + QGIS)

Objectif : à l'issue de ces exercices, vous serez capable de réaliser une conversion en réflectance TOA d'une image à THRS Spot6/7, de réaliser l'orthorectification d'une image THRS Spot6/7 puis un recalage en utilisant une image de référence. Ceci sera réalisé avec les logiciels OrfeoToolBox et QGIS

a. Mosaïque des tuiles pour l'image Panchromatique

Les images Spot6/7 panchromatiques sont fournies au format tif et découpées en tuiles. Avant toute autre manipulation, il faut les mosaïquer avec QGIS. Pour qu'ensuite OTB puisse lire les métadonnées il faut renommer le résultat de la mosaïque par le nom de la première tuile (il faut alors supprimer toutes les tuiles du dossier). Attention à bien copier le nom de façon exacte (même l'extension TIF en lettres majuscules !).

Dans les données disponibles, chercher l'image Spot6/7 acquise en 2018. Décompresser les 2 fichiers pour accéder à l'image Multispectrale et l'image panchromatique.

Q5 : Naviguer dans l'arborescence de ces 2 images. Que constatez-vous ?

Dans QGIS, ouvrir l'outil : **Raster > Divers > Fusion**

- Dans **« Couches en entrée »** sélectionner tous les fichiers TIF qui correspondent aux tuiles de l'image panchromatique
- Dans type de données en sortie, choisir **« Uint16 »**
- Dans Fichier en sortie, sélectionner le nom de la première tuile en ajoutant par exemple **« temp »** (Cf. Figure 7).
- Cliquer sur exécuter la commande

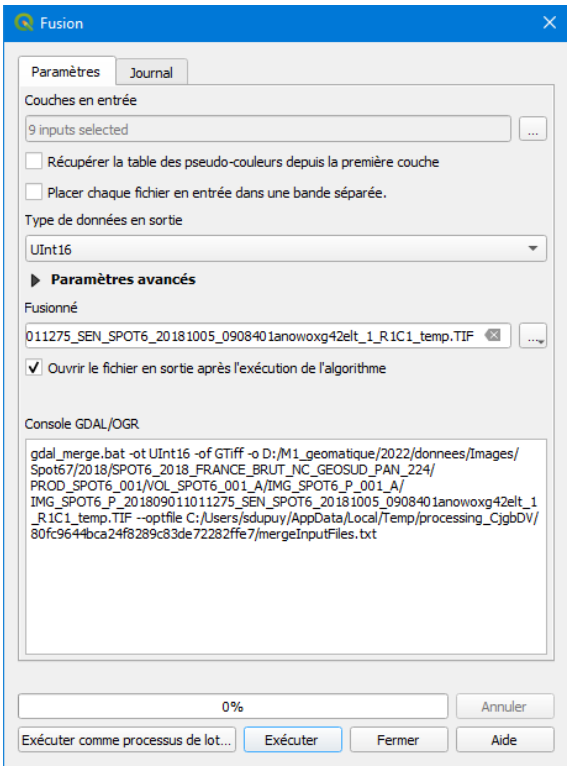

Figure 7 : Commande fusionner prête à être exécutée

Ouvrir l'image dans QGIS pour vérifier le résultat.

Créer un dossier **« tuile »** puis y transférer tous les fichiers tif. Faire la même chose avec les fichiers TFW (sauf pour la tuile RC1C.TFW). Ne pas transférer le résultat de la mosaïque et le renommer en enlevant « temp ». De cette façon OTB connectera automatiquement cette image aux métadonnées pour effectuer les traitements suivants.

b. Conversion d'une image en Comptes Numériques (CN) en réflectance apparente / TOA (Top Of Atmosphere)

La méthode de conversion des comptes numériques (**CN)** en **réflectance apparente** (R_b) dépend souvent des capteurs. Il faut se référer à la documentation des fournisseurs d'images pour trouver l'expression adéquate.

La **réflectance apparente** (propriété intrinsèque d'un objet) étant le ratio de la ***luminance apparente*** (énergie réfléchie) normalisée par l'**éclairement solaire incident** (énergie reçue), il est nécessaire de calculer cette ***luminance apparente*** (***TOA*** *radiance en anglais = sommet de (Top Of) l'atmosphère*) pour chacune des bandes spectrales.

Conversion des comptes numériques en ***luminance apparente*** (*TOA*).

$$\mathbf{L_b} = \frac{\mathrm{CN(b)}}{\mathrm{Gain_b}} + \mathrm{Bias_b}$$

Soit $\mathbf{L_b}$ la ***luminance apparente*** pour une bande spectrale donnée, **CN**(**b**) le compte numérique, ***Gain$_b$*** et ***Bias$_b$*** le gain et le biais pour cette bande spectrale :

Conversion de la ***luminance apparente*** en ***réflectance apparente*** (> valeurs comprises entre 0 et 1)

$$\mathbf{R_b} = \frac{(\boldsymbol{\pi} \,.\, \mathbf{L_b})}{\left(\mathbf{E_b} \,.\, \boldsymbol{\cos}(\boldsymbol{\theta s})\right)} = \frac{(\,\pi \,.\, \mathrm{CN(pb)} + \mathrm{Bias_b}\,)}{(\,\mathrm{E_b} \,.\, \cos(\theta s).\, \mathrm{Gain_b})}$$

Soit $\mathbf{R_b}$ la **réflectance apparente**, E_b la **luminance (irradiance) solaire** et θs l'**angle zénithal** solaire :

N.B. L'utilisateur peut compléter cette expression en ajoutant par exemple la distance entre la Terre et le Soleil, fonction de la date d'acquisition ($\boldsymbol{\pi} \,.\, \mathbf{L_b} . \mathbf{d2}$)*. Cette valeur, exprimée en unité astronomique, est ici supposée constante dans le temps et égale à 1. Pour la calculer plus précisément, se reporter à la page Web suivante : http://www.fourmilab.ch/cgi-bin/Solar*

La conversion en réflectance TOA sera réalisée pour les 2 images SPOT6 (panchromatique et multispectrale). On choisit d'enregistrer l'image en Float puis de la convertir en 16bits

- **Recherche des paramètres**

En se basant sur la métadonnée de l'image Spot 6, recherchez les valeurs des paramètres requis pour chacune des bandes spectrales afin de pouvoir compléter le tableau suivant.

Pour Spot6/7 les métadonnées sont disponibles via le fichier de type « xml » située dans le répertoire suivant :
\..\SPOT6_2018_FRANCE_BRUT_NC_GEOSUD_MS_224\PROD_SPOT6_001\VOL_SPOT6_001_A\IMG_SPOT6_MS_001_A\DIM_SPOT6_MS_201809011011275_SEN_SPOT6_20181005_0908341x1t9ck5pkaep_1.XML

Pour faciliter la lecture de ce fichier, l'outil xmlPad est disponible dans le répertoire « outils » des données du TP. Les données que vous cherchez sont dans : Radiometric Calibration > Band Measurement list. Utilisez l'outil de recherche pour trouver les paramètres.

Tableau 1 : Paramètres requis pour la conversion en réflectance TOA de l'image Spot6 avec <en gras> le nom du paramètre dans les métadonnées

		Valeurs			
		Bleu (layer 3) B0	**vert (layer 2) B1**	**Rouge (layer 1) B2**	**PIR (layer 4) B3**
Paramètres	Gain **<GAIN>**	7.75	9.25	10.34	13.88
	Biais **<BIAS>**	0	0	0	0
	Luminance solaire **<Band_Solar_Irradiance>**	1982.67	1826.09	1540.50	1094.75
	Angle zénithal θs déduit de l'élévation solaire **<SUN_ELEVATION>**	49.267			

Attention, le nom des paramètres diffère entre les satellites. Pour Landsat 8 la métadonnée est un fichier texte :

- RADIANCE_MULT_BAND_N = gain
- RADIANCE_ADD_BAND_N = biais

- **<u>Conversion des images en réflectance TOA</u>**

OTB va chercher automatiquement les paramètres nécessaires, à condition que l'image et les métadonnées correspondent aux standards attendus par le logiciel.

Ouvrir l'interface MAPLA située dans le répertoire **OTB > mapla.bat**

Outil : Optical Calibration (il est également possible d'utiliser la commande otbcli_OpticalCalibration en ligne de commande d'OSGeo4W)

Sur la fenêtre de dialogue :

- Input : **< Image brute en entrée >**
- Output **< Image en sortie : ImageTOA.tif MS ou PAN >** Type : float. Simplifiez le nom de l'image produite par exemple : « 20180901-MS-TOA-float.tif »
- Dans le menu **« Calibration Level »** choisir **« image to TOA »**
- Exécuter la commande

Au moment du choix de l'image en entrée les paramètres d'acquisition de l'image sont automatiquement affichés cela donne une indication que la lecture des métadonnées est correcte (par exemple la valeur **« Sun elevation angle »** relevée dans les métadonnées). **Appliquer aux images panchromatique et multispectrale.**

- **<u>Multiplication des valeurs des pixels par 10000 et conversion en UINT16</u>**

Cette opération permet de diminuer la taille de l'image en modifiant le type d'image : on passe de flottant (float) à 16 bits non signés uint16).

Nous allons utiliser l'outil : otbcli_BandMathX d'OTB sur la fenêtre de commande d'OSGeo4W.

Pour ouvrir la fenêtre OSGeo4W Shell il faut aller dans **démarrer > programmes > OSGeo4W Shell** (souvent un raccourci est aussi disponible sur le bureau).

Se positionner dans le dossier bin de l'OTB. Si le dossier OTB est installé sur le disque D, il faut saisir **« d : »** dans la fenêtre OSGeo4W puis **« cd »** et le chemin jusqu'au dossier bin (on peut s'aider de la touche tabulation). Il faut ensuite saisir la commande ci-dessous.

Pour saisir le nom du fichier d'entré et de sortie, on peut **copier le fichier en tant que chemin d'accès** via l'explorateur Windows en maintenant enfoncée la touche majuscule du clavier et en cliquant sur le nom du fichier avec le bouton de droite de la souris.

Ligne de commande :

otbcli_BandMathX.bat -il < image en entrée : ImageTOA.tif > -exp "10000*im1" -out < Image en sortie : TOAx10000.tif > uint16

Exemple : otbcli_BandMathX.bat -il D:/Images/Spot67/2018/20180901_MS_TOA_float.tif -out D:/Images/Spot67/2018/20180901_MS_TOA_16b.tif uint16 -exp 10000*im1

Il est possible de copier la ligne de commande depuis Mapla avec **« Application > Copy command line to clipboard ».**

Q6 : Comparer la taille des images sur le disque dur puis afficher les 2 images MS dans QGIS. Utiliser l'outil de sondage pour consulter la valeur des pixels des 2 images

Que constatez-vous ?

L'extension Value Tool pour QGIS est utile pour cet exercice.

Appliquer aux images panchromatique et multispectrale.

c. Orthorectification

➢ Préparation du référentiel altimétrique

Ouvrir la page Web : http://earthexplorer.usgs.gov : le catalogue Earth Explorer, géré par l'USGS (U.S. Geological Survey) aux États-Unis, est une infrastructure de données spatiales qui donne accès, entre autres, à des images satellitaires et des modèles numériques de terrain (MNT).

Ces données sont pour la plupart gratuites. L'accès au téléchargement nécessite de créer au préalable un compte d'utilisateur.

- Dans l'onglet ***Search Criteria*** : au niveau de ***Adress/Place***, entrer Montpellier, puis cliquer sur Show et sur l'Adresse trouvée (Montpellier. France)
- Cliquer ensuite sur le bouton ***Data Sets*** »
- Dans l'onglet Data Sets : cocher Digital Elevation > ***SRTM > SRTM 1 Arc-second Global***

Les données SRTM (Shuttle Radar Topography Mission) ont été acquises par la navette spatiale Endeavour. Longtemps distribuées avec une résolution spatiale dégradée de 3 secondes d'arc (environ 90 mètres à l'équateur), celles-ci sont désormais disponibles à une résolution d'1 seconde d'arc (environ 30 mètres à l'équateur). Pour plus d'informations, voir la page Web suivante : https://www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-shuttle-radar-topography-mission-srtm-1-arc?qt-science_center_objects=0#qt-science_center_objects

- Cliquer sur le bouton ***Results*** »
- Dans l'onglet Results apparaissent les résultats de la recherche, soit six dalles SRTM pour couvrir la zone d'étude.
- Cliquer sur :
  - (Show Footprint) pour visualiser l'emprise d'une image sur la fenêtre cartographique,

- (Show Browse Overlay) pour visualiser un aperçu géolocalisé,
- (Show Metadata and Browse) pour consulter les métadonnées ,
- (Download Options) pour télécharger le(s) fichier(s).

- Si vous possédez un compte, se connecter puis cliquer pour chaque image sur l'icône
- Dans la fenêtre Download Options, télécharger la version DTED 1 Arc-second

N.B. Les fichiers sont également disponibles dans le répertoire SRTM

➢ Orthorectification avec Mapla

Pour cette étape, nous utiliserons l'outil Mapla dans le dossier de l'OTB. Ouvrir le menu OrthoRectification (Cf. Figure 8).

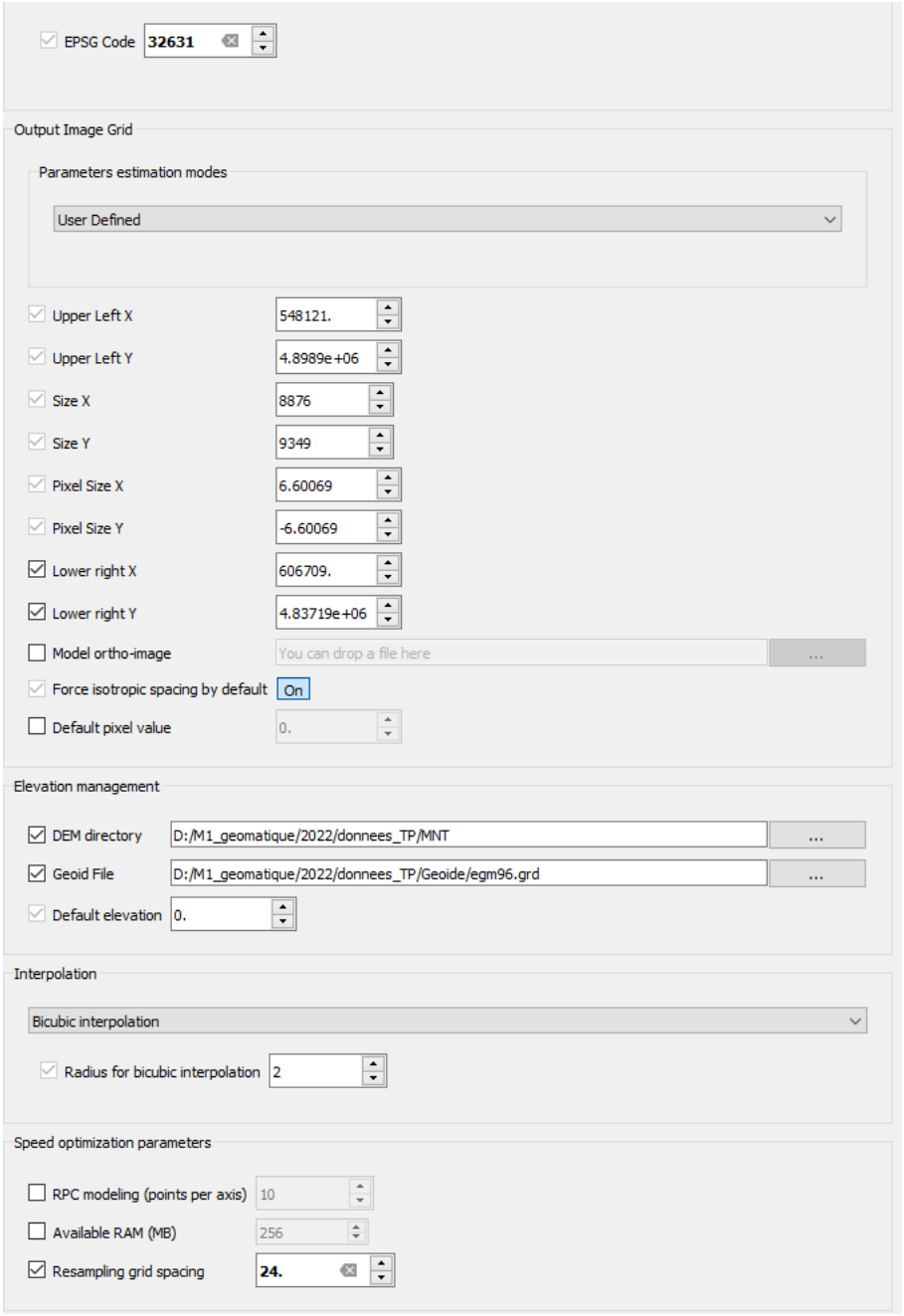

Figure 8 : commande d'orthorectification

Il faut renseigner l'image en entrée (multispectrale ou panchromatique), l'image en sortie et le dossier dans lequel se trouvent les tuiles du MNT SRTM 30m. Penser à corriger la taille des pixels de l'image

en sortie (6m pour l'image multispectrale et 1,5m pour la panchromatique. **Attention à bien enregistrer la nouvelle image en Unsigned Integer 16**.

Attention, après l'extension .tif de l'image en entrée, il faut ajouter la mention **?&skipcarto=true** Si on ne respecte pas cela le nord et le sud sont inversés dans l'image en sortie. Cette extension est nécessaire pour qu'OTB puisse orthorectifier tous les produits de type « OrthoReady ». En effet ces produits ont un pseudo système de coordonnées non standard permettant leur exploitation approximative dans un SIG sans prétraitement préalable. L'option **?&skipcarto=true** permet d'ignorer cette partie des métadonnées inutile pour ce traitement.

Q7 : Superposez dans QGIS, l'image orthorectifiée à celle de 2020

Que constatez-vous ?

L'extension Map Swipe Tool pour QGIS est utile pour cet exercice.

Les options à paramétrer :

- **Map Projection :** indique le système de projection dans lequel la donnée sera produite. On choisit ici EPSG dans la liste déroulante puis le code 32631 correspondant au système de coordonnées WGS84 – UTM31 Nord.
- **Resampling grid spacing :** permet de définir la valeur de la grille de déformation pour le ré-échantillonage de l'image. La norme est de choisir pour ce paramètre une valeur égale à 4 fois la résolution spatiale de l'image en sortie soit pour l'image multispectrale 6x4 = 24 et pour la panchromatique 1,5x4=6
- **DEM directory :** pointe vers le dossier contenant les tuiles SRTM
- **Geoid File :** pointe vers le fichier de géoïde (disponible dans le répertoire Geoide)

➢ Orthorectification en ligne de commande (optionnel)

Outil : otbcli_OrthoRectification d'OTB sur la fenêtre de commande d'OSGeo4W.

Ligne de commande :

otbcli_orthorectification.bat -io.in ""<OrthoExRoiTOAx10000.tif>"?&skipcarto=true" -io.out <Fichier en sortie : OrthoFinalExRoiTOAx10000.tif> uint16 -map epsg -map.epsg.code 32631 -opt.gridspacing < 24 ou 6 > -elev.dem <dossier SRTM> -elev.geoid <fichier du géoïde>

Exemple : otbcli_OrthoRectification.bat -io.in ""D:/donnees/Images/Spot67/2018/20180901_MS_TOA_16b.tif"?&skipcarto=true" -io.out D:/donnees/Images/Spot67/2018/20180901_MS_TOA_16b_ortho.tif uint16 -map epsg -map.epsg.code 32631 -outputs.mode auto -elev.dem D:/donnees_TP/MNT -elev.geoid D:/donnees_TP/Geoide/egm96.grd -opt.gridspacing 24

Appliquer aux images panchromatiques et multispectrales.

d. Recalage avec points d'amer avec l'extension de géoréférencement de QGIS

Les modèles de capteurs permettent de positionner les images avec une erreur relative de quelques mètres mais cette erreur n'est parfois pas acceptable. C'est le cas par exemple si on dispose d'une image de référence sur laquelle on souhaite superposer de nouvelles données. Dans ce cas on utilise des points d'amer (de référence) pour améliorer le modèle de capteur prédéfini. L'image sera donc corrigée avec une planimétrie proche de la référence.

➢ Effectuer un recalage en utilisant l'algorithme « Polynomial »

Les algorithmes polynomiaux de degrés 1 à 3 font partie des algorithmes les plus utilisés pour faire correspondre les points de contrôle au sol aux points saisis sur l'image à corriger. L'algorithme polynomial de degrés 2 et 3 autorise une déformation de l'image. La transformation polynomiale du premier ordre (affine) permet uniquement la mise à l'échelle, la translation et la rotation.

- Dans QGIS, dans le menu « Raster » sélectionner « Géoréférencer ».
- Dans l'outil « Géoréférenceur », afficher l'image de 2018 sur laquelle vous avez appliqué l'orthorectification.
- Dans « Paramètres », modifier les « Paramètres de la transformation » :
  - Type de transformation : Polynomiale de degré 3
  - Méthode de ré-échantillonnage : Cubique
  - SCR Cible : code EPSG 32631
  - Nom du raster de sortie : donner un nom et un emplacement au fichier qui sera créé
- Dans la fenêtre principale de QGIS, afficher l'image de référence (celle acquise en 2020)
- Chercher dans le « Géoréférenceur » et sur l'image de référence un invariant pour sélectionner un point d'appui.
- Pour sélectionner ce point, cliquer sur le bouton « Ajouter un point » dans le « Géoréférenceur »
- La fenêtre « Saisir les coordonnées de la carte » s'ouvre. Cliquer sur le bouton « Depuis le canevas de carte » pour saisir le point homologue sur l'image de référence (Cf. figure 9).
- Cliquer sur OK. Les points sont positionnés sur les 2 images.

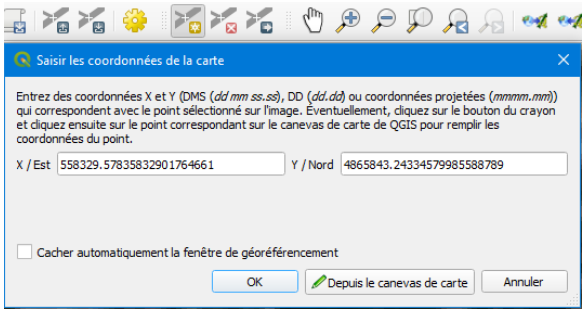

Figure 9 : Saisie de points homologues avec dans le bandeau les outils de saisie et de modification des point et la fenêtre de saisie d'un nouveau point

- Il est obligatoire de saisir un minimum d'une dizaine de points avec un polynôme de degré 3.
- Lorsque le nombre de points et suffisant, l'erreur moyenne s'affiche. Vérifier aussi l'erreur de chaque point pour les modifier si besoin
- Enregistrer les points régulièrement

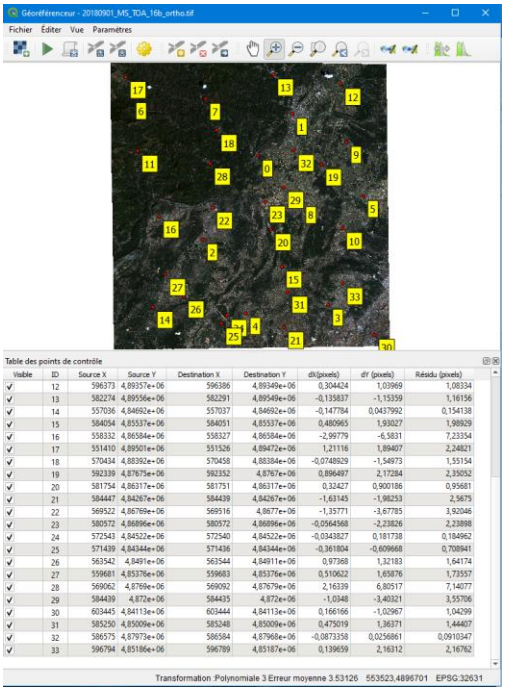

Figure 10 : Outil "Géoréférenceur" de QGIS

- Exécuter le géoréférencement
- Vérifier le nouveau calage en superposant l'image à la référence.
- Modifier éventuellement les points puis relancer le traitement.

Q8 : Superposez dans QGIS, l'image corrigée à celle de 2020

Que constatez-vous ?

L'extension Map Swipe Tool pour QGIS est utile pour cet exercice.

➢ Effectuer un recalage en utilisant l'algorithme « Thin Plate Spline » (TPS)

L'algorithme Thin Plate Spline (TPS) est une méthode plus moderne que « Polynomial ». Il est capable d'introduire des déformations sur des secteurs précis de l'image. Il est très pratique quand des sources de faible qualité sont utilisées.

- Sans modifier les points, modifier les paramètres de transformation pour remplacer « Polynomiale 3 » par « Thin Plate Spline »
- Modifier le nom de l'image en sortie
- Exécuter le géoréférencement

➢ Effectuer le même travail avec l'image panchromatique. Il est possible de charger les points identifiés sur les images multispectrales mais dans ce cas il faut tous les vérifier et les déplacer si besoin. Pour modifier le contraste de l'image dans l'outil géoréférenceur, il faut d'abord modifier la symbologie dans la fenêtre principale de QGIS et l'enregistrer par défaut.

Q9 : Superposez dans QGIS, l'image corrigée à celle de 2020

Quel algorithme est le mieux adapté pour cette transformation ? Pourquoi ?

e. Orthorectification avec sélection automatique de points d'amer sur OTB (optionnel)

Les modèles de capteurs permettent de positionner les images avec une erreur relative de quelques mètres mais cette erreur n'est parfois pas acceptable. C'est le cas par exemple si on dispose d'une image de référence sur laquelle on souhaite superposer de nouvelles données. Dans ce cas on utilise des points d'amer (de référence) pour améliorer le modèle de capteur prédéfini. L'image sera donc corrigée avec une planimétrie proche de la référence.

La première étape est d'extraire des paires de points invariants. Chaque paire étant constituée d'un point sur l'image de référence et son équivalent sur l'image à rectifier. OTB propose des méthodes de sélection automatique de ces paires de points dans l'outil **HomologousPointsExtraction** : SIFT (Scale-Invariant Feature Transform) et SURF (Speeded Up Robust Features). Pour les 2 outils on distingue 2 étapes :

- Extraction de descripteurs à partir des images. Ces descripteurs sont invariants à l'orientation et à la résolution de l'image et peu sensibles à sa dynamique et netteté.
- Mise en correspondance des descripteurs fondée sur la distance euclidienne entre les descripteurs.

Lorsqu'on dispose d'un nombre suffisant de paires de points, il est possible de modifier légèrement les paramètres du modèle de capteur pour qu'il permette de produire une image superposable à la référence. Pour modifier le modèle de capteur on utilisera l'outil **RefineSensorModel.**

Pour mettre en œuvre ces outils on va orthorectifier l'image acquise en 2018 en utilisant comme référence l'image de 2020 téléchargée sur Dinamis en version orthorectifiée. La recherche de points homologues sera effectuée entre les 2 images panchromatiques puis entre les 2 images multispectrales.

- **Extraire le modèle de capteur :** comme nous avons appliqué un traitement radiométrique, nous disposons d'un modèle de capteur (fichier avec l'extension .geom). Dans le cas où il ne serait pas disponible, il est possible de le générer en utilisant l'outil ReadImageInfo avec l'option outkwl, suivie d'un nom de fichier avec l'extension .geom.
- **Recherche des points Homologues :** avec les paramètres suivants :
  - **Image 1 :** image de 2018 à orthorectifier sans oublier l'option **?&skipcarto=true**
  - **Image 2 :** Image de 2020 de référence
  - **Output file:** donner un nom au fichier texte contenant les points
  - **Output vector :** pour exporter les points en format vecteur pour les contrôler visuellement
  - **Algorithme :** choisir SIFT.
  - **Distance threshold for matching :** 0.6 (*Il s'agit du seuil de détection compris entre 0 et 1. Lorsque le seuil est proche de 1 un grand nombre de paires de points sera créées.).*
  - **Use back-matching :** off *(Si ON, on minimise le nombre de mauvaises correspondances.)*.
  - **Keypoints mode :** choisir « search keypoints in smal spatial bins… » (*Pour ne pas utiliser la totalité de l'image.*).
    - **Size of bin :** *taille des sous-régions*
    - **Steps between bins :** *pas entre chaque sous-région*
  - **Estimated precision of the colocation function :** saisir 60 (*Lorsque le filtre ci-dessous est actif, cette valeur définit la distance max entre les points de l'image 1 et ceux de l'image 2. Les paires de points dont la distance est supérieure à ce seuil sont rejetées.*).
  - **Filter point according to geographical or sensor based colocalisation :** on
  - **If enabled, points from second image will be exported en WGS84 :** on (pour exporter les coordonnées des points en latitude/longitude)
  - **Renseigner le dossier contenant le MNT et la localisation du fichier Géoide**

Ligne de commande :

otbcli_HomologousPointsExtraction.bat -in1 ""<image à orthorectifier>"?&skipcarto=true" -in2 <image de référence> -out <fichier texte des points en sortie -outvector <fichier shape des points en sortie> -algorithm sift -mode geobins -precision 60 -mfilter true -2wgs84 true -elev.dem <dossier contenant les tuiles du MNT> -elev.geoid <elipsoide>

Exemple : otbcli_HomologousPointsExtraction.bat
-in1 ""D:/Images/Spot67/2018/20180901_MS_TOA_16b.tif"?&skipcarto=true"
-in2 D:/Images/2020/SPOT6_MS_20200524_ORT.tif
-out D:/Images/Spot67/2018/20180901_MS_TOA_16b_points.txt
-outvector D:/Images/Spot67/2018/20180901_MS_TOA_16b_points.shp
-algorithm sift -mode geobins -precision 60 -mfilter true -2wgs84 true
-elev.dem D:/SRTM
-elev.geoid D:/Geoide/egm96.grd

Attention la recherche de points homologues est un traitement qui prend du temps pour l'image panchromatique ! Un fichier de points est disponible dans les données pour ne pas bloquer la suite du TP.

- **Améliorer le modèle de capteur :** Cette opération est réalisée avec l'outil RefineSensorModel. A appliquer pour les images panchromatique et multispectrale de 2018.
  - **Input geom :** fichier .geom du modèle de capteur
  - **Output geom :** nouveau modèle de capteur
  - **Input file containing tie points :** fichier texte avec les coordonnées des points produit ci-dessus
  - **Statistiques (optionnel) :** production d'un fichier de statistiques de la précision des points
  - **DEM :** dossier contenant les tuiles du MNT
  - **Geoid :** fichier géoïde

Ligne de commande :

otbcli_RefineSensorModel.bat -ingeom <fichier geom de l'image à orthorectifier> -outgeom <nouveau fichier geom> -inpoints <fichier de points homologues> -elev.dem <dossier contenant les tuiles du MNT> -elev.geoid <elipsoide>

Exemple : otbcli_RefineSensorModel.bat -ingeom
D:/Images/Spot67/2018/20180901_MS_TOA_16b.tif -outgeom
D:/Images/Spot67/2018/20180901_MS_TOA_16b_new.geom -inpoints
D:/Images/Spot67/2018/20180901_MS_TOA_16b_points.txt -elev.dem D:/donnees_TP/SRTM -elev.geoid D:/donnees_TP/Geoide/egm96.grd

- **Orthorectification avec le nouveau modèle de capteur** (ne pas détruire les images orthorectifiées précédemment) **:** on utilise le même outil d'orthorectification. La différence est qu'à la suite de la mention ***?&skipcarto=true*** on ajoute ***&geom=<fichier geom avec le chemin d'accès>***

Ligne de commande :

otbcli_OrthoRectification.bat -io.in ""<image à orthorectifier>"?&skipcarto=true&geom=<nouveau fichier geom>" -io.out <image en sortie> uint16 -map epsg -map.epsg.code <code EPSG de la projection> -elev.dem <dossier contenant les tuiles du MNT> -elev.geoid <elipsoide>

Exemple : otbcli_OrthoRectification.bat -io.in
""D:/donnees/Images/Spot67/2018/20180901_MS_TOA_16b.tif"?&skipcarto=true&geom=D:\donnees\Images\Spot67\2018\20180901_MS_TOA_16b_new.geom" -io.out
D:/donnees/Images/Spot67/2018/20180901_MS_TOA_16b_ortho_v2.tif uint16 -map epsg -map.epsg.code 32631 -outputs.spacingx 6 -outputs.spacingy -6 -elev.dem D:/donnees_TP/SRTM -

elev.geoid D:/donnees_TP/Geoide/egm96.grd -interpolator bco -interpolator.bco.radius 2 -opt.ram 256 -opt.gridspacing 24

Appliquer aux images panchromatiques et multispectrales.

f. Pansharpening (OTB)

Objectif : à l'issue de cet exercice, vous serez capable découper une image puis de combiner un bundle Spot6/7 (image multispectrale et panchromatique acquises simultanément) avec OTB pour produire une image couleur avec la résolution de l'image panchromatique.

La plupart des satellites à THRS acquièrent simultanément une image panchromatique et une image multispectrale (plusieurs bandes mais de résolution spatiale plus basse). Afin d'exploiter au mieux les deux sources d'information, des techniques ont été élaborées pour produire une image multispectrale ayant la résolution spatiale de l'image panchromatique. Cette opération s'appelle le pansharpening. OTB propose principalement les méthodes suivantes :

- **Relative Component Substitution (RCS) :** avec cette méthode on considère que l'image panchromatique est la somme des canaux de l'image multispectrale. Cette fusion consiste à appliquer un filtre passe-bas à l'image panchromatique pour lui donner un contenu spectral (dans le domaine de Fourrier) équivalent à l'image multispectrale. Puis l'image multispectrale est normalisée par l'image panchromatique filtrée et multipliée par l'image panchromatique d'origine. Le seul paramètre de cette méthode est le gabarit du filtre passe-bas appliqué à l'image panchromatique.
- **Bayesian Data Fusion (Bayesian) :** cette méthode est plus complexe et nécessite d'effectuer des paramétrages différents en fonction de l'information spectrale de la zone étudiée. Une publication[1] expliquant cette méthode propose des valeurs de pondération en fonction des contextes (urbain, forestiers, agricoles…)

L'outil **Bundle To perfect Sensor** combine dans un même outil les étapes de Superimpose (réechantillonage) et de pansharpening que nous allons décomposer dans cette partie du TP.

Nous allons produire l'image fusionnée en utilisant le bundle acquis en 2018 et en utilisant les images que vous venez de corriger géographiquement.

➢ Extraire un morceau de l'image

Le processus de fusion d'image prend beaucoup de temps pour une image entière. Il est donc préférable de découper un morceau des 2 images (MS et PAN) et de réaliser le traitement sur ces extraits.

Pour cette opération on utilise l'outil ExtractROI. On va extraire le ¼ de l'image situé en bas à droite :

- Renseigner les images en entrée et sortie
- Dans Extraction mode choisir « Extent » puis saisir (en nombre de pixels):
  - Pour l'image multispectrale :
    - en X Upper left corner : 4458
    - en Y Upper left corner : 4690
  - Pour l'image panchromatique :
    - en X Upper left corner : 17829
    - en Y Upper left corner : 18760
- Executer la commande

otbcli_ExtractROI.bat -in <image à découper au format .tif > -out <image découpée au format .tif > uint16 -mode extent -mode.extent.ulx 4458 -mode.extent.uly 4690

[1] D. Fasbender, J. Radoux and P. Bogaert, "Bayesian Data Fusion for Adaptable Image Pansharpening," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 6, pp. 1847-1857, June 2008, doi: 10.1109/TGRS.2008.917131

Exemple : otbcli_ExtractROI.bat -in D:/donnees/Images/Spot67/2018/20180901_MS_TOA_16b_ortho_georef_TPS.tif -out D:/donnees/Images/Spot67/2018/20180901_MS_TOA_16b_ortho_georef_TPS_extract.tif uint16 -mode extent -mode.extent.ulx 4458 -mode.extent.uly 4690

- Superimpose

La fonction superimpose sensor permet d'aligner la grille des pixels d'une image par rapport à une image de référence. Ici cette fonction permet de passer l'image multispectrale à la résolution de l'image panchromatique.

otbcli_Superimpose.bat -inr <image panchromatique orthorectifiée au format .tif >-inm <image multispectrale orthorectifiée au format .tif >--out <image multispectrale en sortie au format .tif > uint16

Exemple : otbcli_Superimpose.bat -inr D:/donnees/Images/Spot67/2018/20180901_P_TOA_16b_ortho_georef_TPS_extract.tif -inm D:/donnees/Images/Spot67/2018/20180901_MS_TOA_16b_ortho_georef_TPS.tif -out D:/donnees/Images/Spot67/2018/20180901_MS_TOA_16b_ortho_georef_TPS_extract_superimpose.tif uint16

- PANSHARPENING

Nous utiliserons la méthode Bayes en laissant les valeurs par défaut (sauf le type de donnée qu'il faut toujours choisir en uint16).

otbcli_Pansharpening.bat -inp <image panchromatique orthorectifiée avec le nouveau modèle de capteur au format .tif > -inxs <image multispectrale issue du superimpose au format .tif > -out <image fusionnée en sortie au format .tif > uint16 -method.bayes.lambda 0.99990 -method.bayes.s 1

Exemple : otbcli_Pansharpening.bat -inp D:/donnees/Images/Spot67/2018/20180901_P_TOA_16b_ortho_georef_TPS_extract.tif -inxs D:/donnees/Images/Spot67/2018/20180901_MS_TOA_16b_ortho_georef_TPS_extract_superimpose.tif -out D:/donnees/Images/Spot67/2018/20180901_PMS_TOA_16b_ortho_georef_TPS_bayes.tif uint16 -method bayes

Tester les algorithmes de pancharpening, RCS et Bayesian.

Q10 : Comparer les images des différents tests de pansharpening. Afficher les images en couleurs naturelles et fausses couleurs.

Que constatez-vous ?

5. PRÉTRAITEMENTS D'IMAGES HRS SENTINEL-2 (OTB)

Dans ce TP nous allons mettre en œuvre des outils permettant de rendre exploitables une série temporelle avec OTB.

Les données Sentinel-2 sont distribuées au niveau L1C correspondant à une image calibrée radiométriquement (TOA) et orthorectifiée (projection WGS84 de la zone UTM couverte par l'image). Il est aussi possible de trouver des produits traités avec des niveaux différents qui sont aussi en accès libre. C'est le cas sur les hubs de l'ESA qui proposent des images de niveau 2A (corrections atmosphériques réalisés avec la chaine de traitement Sen2Cor) sur l'Europe. En France, le Pôle Surfaces Continentales THEIA distribue des images prétraitées avec l'algorithme MAJA (plus performant que Sen2Cor). Dans ce TP nous allons utiliser ces images S2 de THEIA.

Les images S2 sont fournies sous la forme d'un ensemble de fichiers contenant les images décomposées bande par bande et des méta données (informations sur l'acquisition, qualité, masques de nuages…). Nous allons dans ce TP mettre en œuvre les applications de l'OTB permettant de produire un jeu de données de type série temporelle exploitable pour réaliser une classification automatique pour produire une carte d'occupation du sol.

Pour limiter la taille des données composant ce TP, une seule image est disponible dans le dossier S2.

a. Description d'une image Sentinel-2

Dans les données du TP, aller dans le dossier de la première image S2 :

S2/SENTINEL2A_20200308-104850-422_L2A_T31TEJ_C_V2-2

Avec le nom du dossier qui est repris pour chaque fichier, on comprend qu'il s'agit d'une image acquise le 8 mars 2020, par Sentinel 2A, de niveau L2A, sur la tuile 31TEJ (région de Montpellier).

Les fichiers dont les noms se terminent par SRE_Bxx.tif correspondent aux réflectances de surface pour les différentes bandes.

Les fichiers dont les noms se terminent par FRE_Bxx.tif correspondent à un produit plus élaboré prenant en compte les corrections radiométriques de pente et d'environnement. Nous allons utiliser ces fichiers.

Dans le dossier MASKS, de nombreux fichiers sont disponibles mais nous pourrions utiliser ceux dont le nom se termine par CLM_X qui sont les masques de nuages à 10m de résolution (R1) et 20m (R2). Toutefois, les images choisies n'ont pas (ou peu) de nuage et nous n'aborderons pas dans ce TP la gestion des masques de nuages.

b. Découper une image

L'image Sentinel2 couvre une grande superficie, il est donc nécessaire de travailler sur un morceau plus petit. Nous allons donc découper en utilisant l'outil ExtractROI avec comme modèle d'emprise l'extrait d'image Spot6/7 multispectrale produite au cours du TP précèdent (pansharpening) : 20180901_MS_TOA_16b_ortho_georef_TPS_extract.tif

La commande pour chaque bande serait de ce type :

otbcli_ExtractROI.bat -in <image à découper au format .tif > -out <image découpée au format .tif > uint16 -mode extent -mode.extent.ulx 4458 -mode.fit.im <image du TP pansharpening>

Exemple : otbcli_ExtractROI.bat -in D:/donnees_TP/Images/S2/SENTINEL2A_20200308-104850-422_L2A_T31TEJ_C_V2-2/SENTINEL2A_20200308-104850-422_L2A_T31TEJ_C_V2-2_SRE_B2.tif -out D:/donnees/Images/S2/20200308_ S2A_B2.tif uint16 -mode fit -mode.fit.im D:/donnees/Images/Spot67/2018/20180901_MS_TOA_16b_ortho_georef_TPS_extract.tif

Mais attention ! : au lieu de lancer la commande sur une seule bande S2, nous allons enchainer les traitements via un script de système (batch pour Windows). Pour cela utiliser un nouveau fichier texte appelé par exemple « S2proces.bat » en ajoutant avant chaque commande le mot clé « call ». Copier/coller la première ligne et modifier les noms des images pour chaque bande spectrale (soit 10).

c. Ré-échantillonner les bandes S2 à 20m

Il faut ré-échantillonner les bandes à 20m sur la grille de celles à 10m pour avoir un ensemble de bande homogène à la résolution la plus élevée. Pour faire cela on utilise, comme pour le pansharpening, l'outil Superimpose en choisissant comme référence une des bandes à 10m. L'appliquer sur toutes les bandes à 20m en ajoutant ces commandes dans le script (6 commande au total).

otbcli_Superimpose.bat -inr <bande à 10m au format .tif >-inm <bande à 20m au format .tif >--out <bande20m ré-échantillonnée à 10m en sortie au format .tif > uint16

Exemple : -inr D:/M1_geomatique/2022/donnees/Images/S2/20200308_ S2A _B2.tif -inm D:/donnees/Images/S2/20200308_ S2A _B5.tif -out D:/donnees/Images/S2/20200308_ S2A _B5_10m.tif uint16

d. Compiler les bandes dans une seul image multispectrale (Stack)

La dernière étape sera de compiler dans une même image l'ensemble des bandes spectrales disponibles. Cette opération permet de constituer une image multispectrale en utilisant l'outil ConcatenateImages. Dans la liste des images, veillez à respecter l'ordre des bandes afin de faciliter ensuite l'affichage dans les différentes compositions colorées.

otbcli_ConcatenateImages.bat -il <bandes à 10m au format .tif > -out <image multispectrale en sortie au format .tif > uint16

Exemple : otbcli_ConcatenateImages.bat -il D:/ donnees/Images/S2/ 20200308_B2.tif D:/donnees/Images/S2/20200308_ S2A_B3.tif D://donnees/Images/S2/20200308_B4.tif D:/donnees/Images/S2/ 20200308_ S2A_B5_10m.tif D:/donnees/Images/S2/20200308_ S2A_B6_10m.tif D:/donnees/Images/S2/20200308_ S2A_B7_10m.tif D:/donnees/Images/S2/20200308_ S2A_B8.tif D:/donnees/Images/S2/20200308_ S2A_B8A_10m.tif D:/donnees/Images/S2/20200308_ S2A_B11_10m.tif D:/donnees/Images/S2/20200308_ S2A_B12_10m.tif -out D:/donnees/Images/S2/20200308_S2A.tif uint16

e. Exécuter le script

Pour exécuter le script, copier le fichier « S2process.bat » dans le dossier « bin » de l'OTB et double cliquer dessus.

Pour eviter cette manipulation, il est également possible de saisir le chemin complet d'accès aux outils OTB.

<u>Q11</u> : Vérifier l'image produite en l'affichant dans QGIS. Tester différentes combinaisons de bandes en insérant plusieurs fois l'image dans la liste des couches

Que constatez-vous ?

L'extension Map Swipe Tool pour QGIS est utile pour cet exercice.

6. CRÉATION DE NÉO-CANAUX ET DÉTECTION DE CHANGEMENTS (OTB + QGIS)

a. Indice de végétation : NDVI

Objectif : à l'issue de cet exercice, vous serez capable de calculer un indice de végétation.

Le NDVI (Normalized Difference Vegetation Index) est un indice de végétation basé sur la différence entre les valeurs du rouge et du proche infrarouge (pIR). Sa formule est la suivante :

$$\frac{(\boldsymbol{pIR - Rouge})}{(\boldsymbol{pIR + Rouge})}$$

Cet indice exploite la signature spectrale de la végétation qui est très particulière, car elle montre un pic très marqué dans le proche infrarouge, et une réflectance moindre dans le rouge. Cet indice est très efficace pour déterminer la présence de végétation, mais il peut également servir à évaluer l'importance de la biomasse végétale ainsi que l'intensité de l'activité de photosynthèse.

Concrètement, le calcul du NDVI est une opération arithmétique pixel à pixel entre deux rasters (bandes spectrales). OTB propose un outil pour calculer automatiquement plusieurs indices de végétation.

Calculer les indices de végétation normalisés pour les images de 2018 et 2020 en utilisant les images multispectrales.

Outil : RadiometricIndices d'OTB sur la fenêtre de commande d'OSGeo4W

otbcli_RadiometricIndices.bat -in <image pansharp> -list Vegetation:NDVI -out <Images en sortie au format tif> -channels.blue 3 -channels.green 2 -channels.red 1 -channels.nir 4

Exemple : otbcli_RadiometricIndices.bat -in D:/donnees_TP/Images/2020/SPOT6_MS_20200524_ORT_TOA.tif -out D:/donnees/Images/Spot67/2020/20200524_NDVI.tif float -channels.blue 3 -channels.green 2 -channels.nir 4 -list Vegetation:NDVI

Q12 : Les valeurs du NDVI sont comprises entre -1 et +1. En analysant ces valeurs (traduites à l'écran par des teintes grisées) par ordre décroissant, associer les quatre ou cinq catégories d'occupation du sol.

b. Indice de brillance

Objectif : à l'issue de cet exercice, vous serez capable de calculer un indice de brillance.

L'indice de brillance le plus connu est exprimé par la relation suivante :

$$\boldsymbol{IB = \sqrt{R^2 + pIR^2}}$$

Calculer l'indice de brillance pour les images de 2018 et 2020 en utilisant les images multispectrales.

Outil : RadiometricIndices d'OTB. Vous avez le choix entre BI-1 (utilisation des bandes V,R) et BI-2 (utilisation des bandes V,R,PIR).

c. Composition colorée diachronique et détection de changements avec 2 NDVI

Objectif : à l'issue de cet exercice, vous serez capable de combiner avec OTB 2 indices de végétation dans une même image puis de visualiser dans QGIS les changements.

La combinaison de 2 NDVI calculés sur 2 images différentes est intéressant pour visualiser de façon simple les changements de végétation intervenus entre les 2 images (par exemple les défrichements).

- **Étape 1 découper avec une emprise :** Il s'agit de découper les 2 NDVI avec une emprise afin qu'ils couvrent la même emprise spatiale

Pour cette étape on utilisera l'outil Extract ROI en utilisant une emprise au format shape déjà créée.

Outil : ExtractROI d'OTB sur la fenêtre de commande d'OSGeo4W

otbcli_ExtractROI.bat -in <image à découper> -out <image en sortie> float -mode.fit.vect <emprise de découpage au format shape>

Exemple : otbcli_ExtractROI.bat -in D:/donnees/Images/Spot67/2018/20180901_NDVI.tif -out D:/donnees/Images/Spot67/2018/20180901_NDVI_extract.tif float -mode fit -mode.fit.vect D:/2022/donnees/Emprise_Spot6.shp

- **Étape 2 Superimpose :** Il s'agit d'aligner les pixels du NDVI de 2018 sur ceux du NDVI de 2020 afin d'avoir une correspondance parfaite

Pour cette étape on utilise l'outil Superimpose déjà utilisé précédemment lors du pansharpening.

otbcli_Superimpose.bat -inr <NDVI découpé de 2020 utilisé comme référence >-inm <NDVI de 2018 à aligner sur 2020>--out <image en sortie au format .tif > float

Exemple : otbcli_Superimpose.bat -inr D:/donnees/Images/Spot67/2020/20200524_NDVI_extract.tif -inm D:/donnees/Images/Spot67/2018/20180901_NDVI_extract.tif -out D:/donnees/Images/Spot67/2018/20180901_NDVI_extract_superimpose.tif float

- **Étape 3 Concaténation des NDVI :** il s'agit de combiner les 2 NDVI dans un même fichier pour obtenir une image multispectrale

otbcli_ConcatenateImages -il <ndvi 2018 issu du superimpose > <ndvi 2020 > -out <Image en sortie au format tif>

Exemple : otbcli_ConcatenateImages.bat -il D:/Images/Spot67/2018/20180901_NDVI_extract_superimpose.tif D:/Images/Spot67/2020/20200524_NDVI_extract.tif -out D:/Images/Spot67/ndvi_2018_2020.tif float

- Afficher dans QGIS les 2 images multispectrales et l'image issue de la combinaison des 2 NDVI. Modifier la symbologie pour afficher dans le rouge la bande 1 (NDVI 2018), dans le vert la bande 2 (NDVI 2020) et dans le bleu la bande 1. Adapter le contraste et visualiser le résultat.

Q13 : En utilisant la symbologie préconisée, comment apparaissent les changements ?

7. CLASSIFICATION NON-SUPERVISÉE (OTB + QGIS)

Objectif : à l'issue de cet exercice, vous serez capable de classifier une image multispectrale en utilisant une méthode non-supervisée.

Ce type de classification ne demande pas de connaissances a priori sur les thèmes à différencier. Aucune vérité-terrain n'est utilisée ; les classes spectrales sont créées automatiquement par le logiciel (analyse statistique des données en entrée).

Les classes sont étiquetées a posteriori en établissant des relations entre les classes spectrales et des classes d'intérêt (classes thématiques).

Nous allons utiliser l'image Spot6/7 de 2020 issue du traitement de pansharpening. Il s'agit d'un extrait de l'image déjà utilisée précédemment qui est disponible dans les données fournies.

a. Classification K-Means

Les principaux paramètres à régler sont les suivants :

- **Input Image (in):** image à classer.
- **Output Image (out):** image classée en sortie codée sur 8 bits non signés
- **Number of classes (nc):** Nombre de modes, qui seront utilisés pour générer l'appartenance à une classe (et donc le nombre de classes).
- **Training set size (ts):** taille du jeu d'apprentissage : plus il sera représentatif, plus la classification sera de bonne qualité
- **Maximum number of iterations (maxit):** Nombre maximal d'itérations pour l'étape d'apprentissage.
- **Sampler type (sampler):** Type d'échantillonnage (periodic, pattern based, random)
    - **Periodic sampler (periodic):** prend des échantillons régulièrement espacés
        - **Jitter amplitude (sampler.periodic.jitter):** amplitude ajoutée lors de la sélection de l'échantillon (0 = no jitter)
    - **Random sampler (random):** les positions pour sélectionner les échantillons sont mélangées de façon aléatoire.

Avec l'outil KMeans Classification d'OTB vous allez calculer une classification en demandant la création de 20 classes, un Training set size (ts) de 90 000, un Maximum number of iterations (maxit) de 10 000 et en masquant les pixels à 0.

otbcli_KMeansClassification.bat -in <image à classer> -out <classification en sortie> uint8 -nc 20 -ts 90000 -maxit 10000 -sampler periodic

Exemple : otbcli_KMeansClassification.bat -in D:/donnees/Images/Spot67/2020/SPOT6_PS_20200524_ORT_TOA_extract.tif -out D:/donnees/Images/Spot67/2020/SPOT6_PS_20200524_ORT_TOA_extract_classif_KMEANS_20cl_PC_v10.tif uint8 -nc 20 -ts 90000 -maxit 10000 -sampler periodic

b. Vérification de la classification

Afficher la classification K-Means sur QGIS. Dans la symbologie, choisir un type de rendu « palette / valeurs uniques » et classer pour afficher toutes les classes.

Parcourez la classification en la superposant à l'image pour affecter chaque classe à un comportement que vous pouvez identifier : sol nu, foret, culture, bâti, eau… Lorsque la classe est identifiée, affectez-lui une couleur et donnez-lui un nom dans la légende. Le but est de regrouper les classes similaires sans provoquer de confusion entre les classes d'intérêt.

Q14 : Constatez-vous des confusions entre des classes ? À votre avis comment pourrait-on améliorer le résultat ?

Enregistrer la symbologie de la classification en choisissant « enregistrer par défaut ».

c. Utilisation de l'extension de QGIS ThRasE

La classification contient plusieurs classes identifiées pour une même occupation. Avec l'extension de QGIS ThRasE (Thematic Raster Editor), vous allez regrouper les classes similaires pour en diminuer le nombre. Elle permet également de procéder à des opérations de post classification pour corriger des erreurs ponctuelles de classification.

- Installer l'extension puis cliquer sur le bouton suivant pour l'utiliser :
- Dans la fenêtre ThRasE – Init dialog, choisissez dans l'onglet New : 2 colonnes et 1 ligne
- Dans l'outil ThRasE, afficher à gauche l'image satellite et au centre la classification que vous avez analysée précédemment
- Dans la partie droite, choisissez la classification à recoder (Cf. Figure 11)

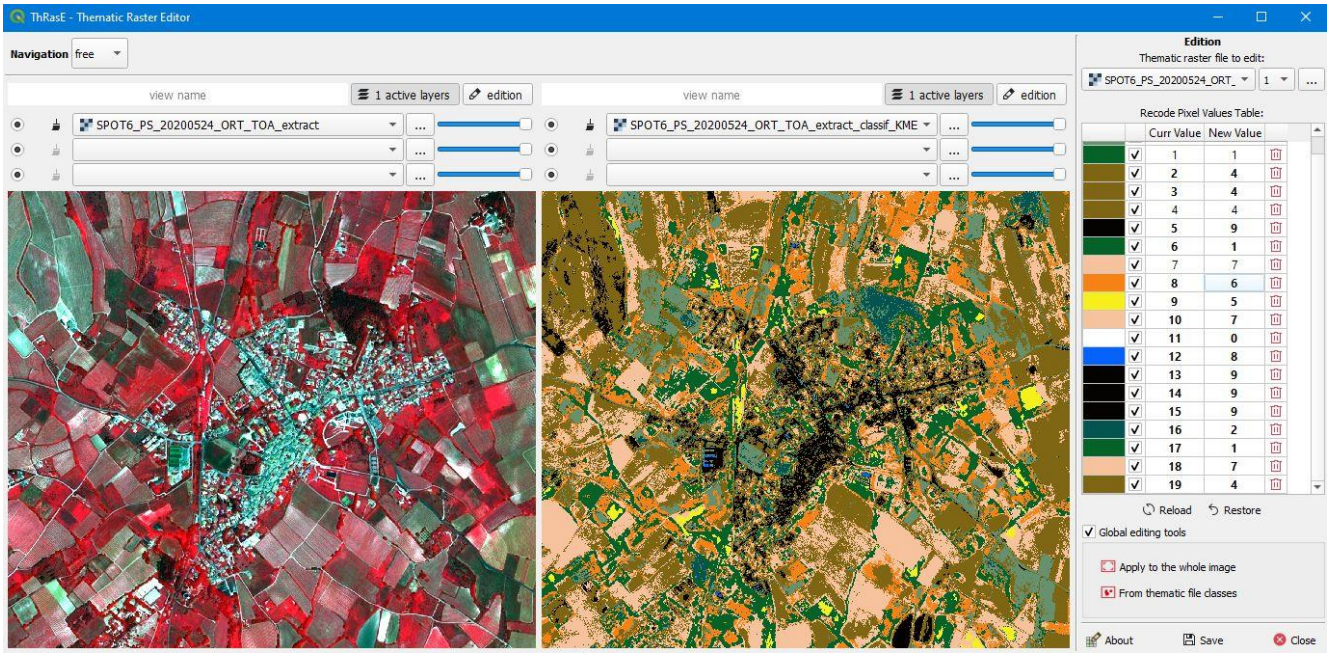

Figure 11: Interface de l'extension ThRaseE avec à gauche l'image satellite, au centre la classification et à droite la table de correspondance pour renuméroter et regrouper les classes

- Naviguer dans l'image pour étudier les fonctionnalités qu'il propose puis saisissez les nouvelles valeurs des classes dans la colonne « New Value ».
- Cocher la case « Global editing tools » puis sur « Apply to the Whole image » pour appliquer la modification à l'ensemble de l'image

d. Appliquer les traitements à un image S-2 (optionnel)

En utilisant les mêmes paramètres, appliquez une classification non supervisée à l'image Sentinel-2 du 22 mai 2020.

8. CLASSIFICATION SUPERVISÉE (OTB + QGIS)

Objectif : à l'issue de cet exercice, vous serez capable de classifier une image multispectrale en utilisant une méthode supervisée.

Dans ce type de classification les connaissances a priori sont utilisées pour la création des classes et la saisie des échantillons de chaque classe thématique recherchée. Les polygones (zones) d'entraînement permettent d'établir une relation entre des classes spectrales et des classes thématiques (recherche de signatures spectrales). À l'issue de la classification il est possible d'évaluer sa validité à l'aide d'une matrice d'erreur et d'une analyse statistique.

Sous OTB, ce processus de classification supervisé se déroule en deux étapes : l'apprentissage d'un modèle qui permet ensuite de prédire sur l'ensemble de l'image une classification.

Nous allons utiliser l'image Spot6/7 de 2020 multispectrale. Nous n'utilisons pas celle issue du pansharpening car les temps de traitement sont trop longs. Il s'agit d'un extrait de l'image d'origine qui est disponible dans les données fournies.

a. Base de données de référence

Une base de données de référence est disponible dans les données du TP. Elle est constituée des données IGN suivantes (disponible ici : https://geoservices.ign.fr) :

- BD TOPO® pour le bâti
- BD Foret pour les espaces naturels
- Registre Parcellaire Graphique : pour les espaces agricoles

La géométrie de ces bases de données est de bonne qualité mais des traitements spécifiques ont été réalisés :

- Sélection aléatoire de 70 des polygones de chaque BD
- Suppression des polygones superposés. En effet, les couches RPG et BD Foret ont parfois des limites qui se chevauchent. En utilisant l'outil de sélection de QGIS, ces polygones ont été repérés puis supprimés.
- Fusion des couches dans un même fichier shape puis création d'attributs uniques : un champ numérique « Code2 » et un champ textuel « Nom2 ».
- Le champ Code1 est un regroupement en 4 classes du niveau 2 qui lui contient 12 classes
- Division de la BD en 2 fichiers : 80% des polygones constituent la BD d'apprentissage et les 20% restants sont la BD de validation. Le seuillage a été effectué sur le champs Code2.

La nomenclature finale retenue est présentée dans le tableau 2.

Tableau 2 : Nomenclature de la base de donnée de référence

Code 1	Nom 1	Code 2	Nom 2
1	agriculture	1	culture annuelle
		2	prairie et fourrage
		3	oliveraie
		4	verger
		5	vigne

2	végétation naturelle	6	foret de feuillus
		7	foret de conifères
		8	foret mixte
		9	formation herbacée
		10	lande
3	bâti	11	bâti
4	eau	12	eau

b. Calcul des statistiques

Le calcul préalable des statistiques de chaque bande spectrale utilisée est nécessaire pour certains algorithmes de classification (comme les Support Vector Machines) qui sont sensibles à la dynamique des données. Il s'agit de calculer les moyennes et écarts types des canaux de l'image. Cette opération est réalisée avec OTB en utilisant : Compute Images Statistics.

otbcli_ ComputeImagesStatistics.bat -il <image à classer> -out <statistiques au format xml>

Exemple : otbcli_ComputeImagesStatistics.bat -il D://donnees/Images/Spot67/2020/SPOT6_MS_20200524_ORT_TOA_extract.tif -out D://donnees/Images/Spot67/2020/SPOT6_MS_20200524_ORT_TOA_extract.xml

c. Algorithme Bayes (Normal Bayes Classifier)

i. Apprentissage

Cette étape est réalisée avec OTB et l'outil « TrainImagesClassifier ». Les principaux paramètres à régler sont les suivants :

- **Input Image (io.il) :** image à classer sur laquelle sera réalisé l'apprentissage.
- **Input Vector Data (io.vd) :** la base de données d'apprentissage au format vecteur
- **Validation Vector Data (io.valid) :** la base de données de validation au format vecteur
- **Input XML image statistics File (io.imstat) :** les statistiques de l'image calculées lors de l'étape précédente
- **Output model (io.out):** le fichier texte contenant le modèle calculé à l'issue de cette étape d'apprentissage. Ce fichier sera ensuite utilisé pour l'étape de prédiction
- **Output confusion matrix (io.confmatout) :** le fichier au format csv contenant la matrice de confusion produite lors de la validation interne
- **Maximum training sample size per class (sample.mt) :** le nombre maximum d'échantillons d exprimé en nombre de pixel par classe qui sera nécessaire pour éviter d'avoir un jeu de données trop déséquilibré entre les classes. S'il y a un déséquilibre, il y a un risque de biaiser l'apprentissage. On laissera 1000 pixels pour que le calcul soit réalisé rapidement. Si on ne veut pas imposer de limite, il faudrait mettre -1
- **Maximum validation sample size per class (sample.mv) :** le nombre maximum d'échantillons de validation exprimé en nombre de pixel par classe utilisé pour l'estimation interne du potentiel de la méthode. On laissera 1000 pixels pour que le calcul soit réalisé rapidement. Si on ne veut pas imposer de limite, il faudrait mettre -1
- **Bound sample number by minimum (sample.bm) :** option permettant de limiter les échantillons de chaque classe en fonction du nombre d'échantillons de la classe la moins bien représentée. On mettra 0 pour cette option.
- **Training and validation sample ratio (sample.vtr) :** fraction d'échantillon à utiliser pour la validation si le jeu de validation n'est pas donné. On mettra 0.5 pour cette option
- **Field containing the class integer label for supervision (sample.vfn) :** Le champ de la table du vecteur contenant les identifiants des classes. Ce doit être obligatoire un champ de type numérique. On choisira le champ « Code1 »
- **Classifier to use for the training (classifier) :** Classifier utilisé. Ici on choisira Normal Bayes classifier

otbcli_TrainImagesClassifier -io.il <image à classer>-io.vd <BD apprentissage au format shp> -io.valid <BD validation au format shp>-io.imstat <statistiques au format xml> -io.out <modèle au format txt> -io.confmatout <matrice de confusion au format csv>-sample.bm 0 -sample.vfn Code1 -classifier bayes

Exemple : otbcli_TrainImagesClassifier -io.il D:/donnees/Images/Spot67/2020/SPOT6_MS_20200524_ORT_TOA_extract.tif -io.vd D:/donnees_TP/BD_Ref/BD_apprentissage_v2.shp -io.valid D:/donnees_TP/BD_Ref/BD_validation_v2.shp -io.imstat D:/donnees/Images/Spot67/2020/SPOT6_MS_20200524_ORT_TOA_extract.xml -io.out D:/donnees/Images/Spot67/2020/model_classif_MS_bayes.txt -io.confmatout D:/donnees/Images/Spot67/2020/matrice_classif_MS_bayes.csv -sample.bm 0 -sample.vfn Code1 -ram 256 -classifier bayes

Cliquez sur l'onglet Log de Mapla pour analyser les résultats et vérifier si l'apprentissage est correct.

(INFO) TrainImagesClassifier: Precision of class [1] vs all: 0.840426
(INFO) TrainImagesClassifier: Recall of class [1] vs all: 0.835979
(INFO) TrainImagesClassifier: F-score of class [1] vs all: 0.838196

(INFO) TrainImagesClassifier: Precision of class [2] vs all: 0.891541
(INFO) TrainImagesClassifier: Recall of class [2] vs all: 0.908995
(INFO) TrainImagesClassifier: F-score of class [2] vs all: 0.900183

(INFO) TrainImagesClassifier: Precision of class [3] vs all: 0.899087
(INFO) TrainImagesClassifier: Recall of class [3] vs all: 0.886243
(INFO) TrainImagesClassifier: F-score of class [3] vs all: 0.892619

(INFO) TrainImagesClassifier: Precision of class [4] vs all: 0.999471
(INFO) TrainImagesClassifier: Recall of class [4] vs all: 0.999471
(INFO) TrainImagesClassifier: F-score of class [4] vs all: 0.999471

(INFO) TrainImagesClassifier: Global performance, Kappa index: 0.876896

L'apprentissage a été effectué puis testé avec le jeu de validation fourni. On constate que les résultats sont assez bons avec un indice de Kappa de 0.87. Les indicateurs par classe (précision, recall et f-score) sont aussi corrects puisque supérieurs à 80% dans la plupart des cas.

La matrice de confusion est aussi disponible dans le Log et a été produite au format CSV. Consultez là.

ii. Prédiction

La prédiction sera réalisée avec l'outil ImageClassifier. Les principaux paramètres sont les suivants :

- **Input Image (in) :** image à classer.
- **Model file (model) :** le modèle obtenu lors de l'apprentissage
- **Statistics File (imstat) :** les statistiques de l'image
- **Output Image (out) :** classification au format tif
- **Confidence map (confmap) :** une carte de confiance disponible pour certains classifieurs.

otbcli_ImageClassifier -in <image à classer> -model <modèle au format txt> -imstat <statistiques au format xml> -out <classification au format tif> uint8

Exemple : otbcli_ImageClassifier -in D:/donnees/Images/Spot67/2020/SPOT6_MS_20200524_ORT_TOA_extract.tif -model

D:/donnees/Images/Spot67/2020/model_classif_MS_bayes.txt -imstat
D:/donnees/Images/Spot67/2020/SPOT6_MS_20200524_ORT_TOA_extract.xml -out
D:/donnees/Images/Spot67/2020/SPOT6_MS_20200524_ORT_TOA_extract_classif_bayes.tif uint8

iii. Analyser la classification

Analyser le résultat de cette classification supervisée en l'affichant sur QGIS et en la comparant avec l'image Spot 6/7.

d. Algorithme SVM (Support Vector Machine)

Calculer le modèle pour l'algorithme SVM en choisissant les paramètres suivants dans le cadre « Classifier » :

- Algorithme : LibSVM classifier
- Type de noyau SVM : il s'agit d'un paramètre dont le choix est lié à la complexité des données. Choisir « linear »
- Paramètre de coût : il s'agit d'un réglage laissant plus ou moins de marge par rapport à l'apprentissage : si C<1 on donne plus de marge entre les classes mais il y a plus d'erreurs alors que si C>1, on minimise les erreurs vis-à-vis du jeu d'apprentissage. Laisser la valeur par défaut : 1

otbcli_TrainImagesClassifier -io.il <image à classer> -io.vd <BD apprentissage au format shp> -io.valid <BD validation au format shp>-io.imstat <statistiques au format xml> -io.out <modèle au format txt> -io.confmatout <matrice de confusion au format csv> -sample.vfn Code1 -classifier.libsvm.k linear

Exemple : otbcli_TrainImagesClassifier -io.il
D:/donnees/Images/Spot67/2020/SPOT6_MS_20200524_ORT_TOA_extract.tif -io.vd
D:/donnees/BD_Ref/BD_apprentissage_v2.shp -io.valid
D:/2022/donnees/BD_Ref/BD_validation_v2.shp -io.imstat
D:/donnees/Images/Spot67/2020/SPOT6_MS_20200524_ORT_TOA_extract.xml -io.out
D:/2022/donnees/Classifications/Spot67/model_classif_MS_svm_linear.txt -io.confmatout
D:/2022/donnees/Classifications/Spot67/matrice_classif_MS_svm.csv -sample.vfn Code1 -classifier.libsvm.k linear

Calculer la classification à partir du modèle SVM.

Cliquez sur l'onglet Log de Mapla pour analyser les résultats et vérifier si l'apprentissage est correct.

(INFO) TrainImagesClassifier: Precision of class [1] vs all: 0.841006
(INFO) TrainImagesClassifier: Recall of class [1] vs all: 0.831217
(INFO) TrainImagesClassifier: F-score of class [1] vs all: 0.836083

(INFO) TrainImagesClassifier: Precision of class [2] vs all: 0.885598
(INFO) TrainImagesClassifier: Recall of class [2] vs all: 0.91746
(INFO) TrainImagesClassifier: F-score of class [2] vs all: 0.901247

(INFO) TrainImagesClassifier: Precision of class [3] vs all: 0.895879
(INFO) TrainImagesClassifier: Recall of class [3] vs all: 0.874074
(INFO) TrainImagesClassifier: F-score of class [3] vs all: 0.884842

(INFO) TrainImagesClassifier: Precision of class [4] vs all: 0.999471
(INFO) TrainImagesClassifier: Recall of class [4] vs all: 0.999471
(INFO) TrainImagesClassifier: F-score of class [4] vs all: 0.999471

(INFO) TrainImagesClassifier: Global performance, Kappa index: 0.874074

L'apprentissage a été effectué puis testé avec le jeu de validation fourni. On constate que les résultats sont assez bons avec un indice de Kappa de 0.85. Les indicateurs par classe (précision, recall et f-score) sont aussi corrects puisque supérieurs à 80% dans la plupart des cas.

La matrice de confusion est aussi disponible dans le Log et a été produite au format CSV. Consultez là.

Passez à l'étape de prédiction pour produire la classification puis comparer les résultats.

Tester à nouveau SVM mais en choisissant RBF (Gaussian radial basis function) à la place de Linear.

otbcli_TrainImagesClassifier -io.il <image à classer> -io.vd <BD apprentissage au format shp> -io.valid <BD validation au format shp>-io.imstat <statistiques au format xml> -io.out <modèle au format txt> -io.confmatout <matrice de confusion au format csv> -sample.vfn Code1 -classifier.libsvm.k rbf

Exemple : otbcli_TrainImagesClassifier -io.il D:/donnees/Images/Spot67/2020/SPOT6_MS_20200524_ORT_TOA_extract.tif -io.vd D:/donnees/BD_Ref/BD_apprentissage_v2.shp -io.valid D:/2022/donnees/BD_Ref/BD_validation_v2.shp -io.imstat D:/donnees/Images/Spot67/2020/SPOT6_MS_20200524_ORT_TOA_extract.xml -io.out D:/2022/donnees/Classifications/Spot67/model_classif_MS_svm_rbf.txt -io.confmatout D:/2022/donnees/Classifications/Spot67/matrice_classif_MS_svm_rbf.csv -sample.vfn Code1 -classifier.libsvm.k rbf

e. Algorithme RF (Random Forest)

Calculer le modèle pour l'algorithme Random Forest en choisissant les paramètres suivants dans le cadre « Classifier » :

- Algorithme : Random Forest classifier
- Maximum depth of the tree : il s'agit de la profondeur de l'arbre. Dans RF, les classes sont associées aux feuilles d'un arbre de décision. Si on augmente la profondeur, on autorise la prise en compte de plusieurs variable mais cela peut se faire avec le risque de sur-représenter les données d'apprentissage
- Minimum number of samples in each node : nombre minimum d'échantillon dans chaque nœud. Il s'agit de définir le seuil minimum d'échantillons à chaque nœud de l'arbre pour pouvoir poursuivre la progression dans la profondeur de l'arbre.
- Termination criteria for regression tree : critère d'arrêt de l'arbre de régression : permet de définir une condition d'arrêt en comparant pour chaque classe la distance d'un nouvel échantillon par rapport aux autres échantillons déjà utilisés.
- Cluster possible… :

otbcli_TrainImagesClassifier -io.il <image à classer> -io.vd <BD apprentissage au format shp> -io.valid <BD validation au format shp>-io.imstat <statistiques au format xml> -io.out <modèle au format txt> -io.confmatout <matrice de confusion au format csv> -sample.vfn Code1 –classifier rf

Exemple : otbcli_TrainImagesClassifier.bat -io.il D:/donnees/Images/Spot67/2020/SPOT6_MS_20200524_ORT_TOA_extract.tif -io.vd D:/donnees/BD_Ref/BD_apprentissage_v2.shp -io.valid D:/donnees/BD_Ref/BD_validation_v2.shp -io.imstat D:/donnees/Images/Spot67/2020/SPOT6_MS_20200524_ORT_TOA_extract.xml -io.out

D:/donnees/Classifications/Spot67/model_classif_MS_RF.txt -io.confmatout
D:/2022/donnees/Classifications/Spot67/matrice_classif_MS_RF.csv -sample.vfn Code1 -classifier rf

(INFO) TrainImagesClassifier: Precision of class [1] vs all: 0.753132
(INFO) TrainImagesClassifier: Recall of class [1] vs all: 0.85873
(INFO) TrainImagesClassifier: F-score of class [1] vs all: 0.802472

(INFO) TrainImagesClassifier: Precision of class [2] vs all: 0.92252
(INFO) TrainImagesClassifier: Recall of class [2] vs all: 0.875661
(INFO) TrainImagesClassifier: F-score of class [2] vs all: 0.89848

(INFO) TrainImagesClassifier: Precision of class [3] vs all: 0.897734
(INFO) TrainImagesClassifier: Recall of class [3] vs all: 0.81746
(INFO) TrainImagesClassifier: F-score of class [3] vs all: 0.855719

(INFO) TrainImagesClassifier: Precision of class [4] vs all: 0.998942
(INFO) TrainImagesClassifier: Recall of class [4] vs all: 0.998942
(INFO) TrainImagesClassifier: F-score of class [4] vs all: 0.998942

(INFO) TrainImagesClassifier: Global performance, Kappa index: 0.850265

Q15 : Comparer les classifications obtenues visuellement et avec les statistiques de validation

Quelle est celle que vous préférez ? Pourquoi ?

9. VALIDATION (OTB)

a. Rappel sur la validation d'une classification

La comparaison statistique entre la vérité terrain et le résultat de la classification se présente sous la forme d'une matrice de confusion qui présente plusieurs indicateurs de la fiabilité pour une évaluation de la classification. Dans la matrice, la BD de référence est disposée en ligne et la classification en colonne. Les valeurs dans la matrice représentent des nombres de pixels, d'objets ou des superficies Pour simplifier nous parlerons ici d'individus. Ci-dessous les éléments nécessaires pour bien lire les informations fournies par la matrice :

- La diagonale de la matrice (cellules colorées en vert dans le Tableau 3) correspond aux individus bien classés (accord entre la classification et la vérité terrain). Donc, si on considère la classe 1, la valeur 52839 correspond au nombre d'individus bien classés.
- Si on considère la classe 1 en ligne : les valeurs 22830 et 1583 correspondent à des individus identifiés comme classe 1 dans la BD de référence mais qui ont été incorrectement attribués aux classes 2 et 3 dans la classification (erreur d'omission).
- Si on considère la classe 1 en colonne les valeurs 6943 et 3762 correspondent au nombre d'individus attribués à la classe 1 par la procédure de classification mais appartenant en réalité aux classes 2 et 3 dans la BD terrain (erreur de commission)

Tableau 3 : Exemple de matrice de confusion

		Classification					Précision producteur
		Classe 1	Classe 2	Classe 3	Classe 4	**Total**	
BD Validation	Classe 1	**52839**	22830	1583	0	**77252**	**68%**
	Classe 2	6943	**851378**	1461	0	**859782**	**99%**
	Classe 3	3762	17401	**69988**	0	**91151**	**77%**
	Classe 4	0	0	0	**60336**	**60336**	**100%**
	Total	**63544**	**891609**	**87389**	**60336**	**1088521**	
Précision utilisateur		**83%**	**95%**	**80%**	**100%**		

À partir de cette matrice, on distingue plusieurs mesures ou indicateurs de fiabilité estimant par classe la qualité de la classification :

- **La précision pour l'utilisateur** est le pourcentage d'individus d'une classe issue de la classification qui sont correctement classés vis-à-vis de la référence. Cette précision pour l'utilisateur indique la probabilité pour qu'un individu qui a été classé dans une classe soit effectivement, dans la réalité un individu de cette classe. Elle se lit en colonne dans la matrice de confusion.
- **La précision pour le réalisateur** (ou producteur) est le pourcentage d'individus d'une classe de référence qui sont correctement classés par la procédure de classification. Cette précision pour le réalisateur indique le taux de réussite avec lequel il a réussi à classer les individus appartenant en réalité à la classe. Elle se lit en ligne dans la matrice de confusion.
- **Le f1-score** est la moyenne harmonique des précisions de l'utilisateur et du réalisateur.

On calcule également des estimateurs globaux qui donnent une appréciation globale de la qualité des résultats de la classification :

- **La précision globale :** désigne le nombre d'individus bien classés rapporté au nombre total d'individus. Elle correspond donc à la proportion d'individus bien classés mais ne donne pas d'information sur les classes possédant la meilleure fiabilité.
- **L'indice de Kappa** (ou KIA) tient compte des erreurs en lignes et en colonnes. Il exprime la réduction d'erreur par rapport à celle obtenue par une classification qui serait réalisée de façon

aléatoire. Il varie de 0 à 1. Par exemple une valeur de 0,75 exprime que la méthode de classification utilisée évite 75% des erreurs obtenues par une procédure travaillant complètement au hasard.

b. Valider une classification avec OTB (optionnel)

Outil : ComputeConfusionMatrix d'OTB sur la fenêtre de commande d'OSGeo4W

otbcli_ComputeConfusionMatrix -in **<vrt de la classification par exemple …/classif_cod_1.vrt>** -out **<fichier .csv en sortie qui contiendra la matrice>** -ref vector -ref.vector.in **<jeu de validation>** -ref.vector.field **<Champ, par exemple Cod_1>** > Fichier.log

L'instruction " > Fichier.log" dans cette étape est très importante parce qu'elle crée le fichier texte avec le résultat des précisions des matrices. Le fichier .csv contient seulement les valeurs de la matrice de confusion.

10. REMERCIEMENTS

Merci à Dinamis pour la mise à disposition des images Spot6/7

Merci à THEIA pour la mise à disposition des images Sentinel 2 de niveau 2A. Données à valeur ajoutée traitées par le CNES pour le pôle de données Theia www.theia.land.fr à partir de données Copernicus. Les traitements utilisent les algorithmes développés par les Centres d'Expertise Scientifique de Theia

Département Environnement et Société, Unité Mixte de Recherche Territoire Environnement Télédétection et Information Spatiale, direction régionale Réunion Mayotte

Maison de la Télédétection, 500 rue Jean-François Breton, 34093 Montpellier cedex 5

www.cirad.fr