

Copernicus & biodiversité

Carto physiologique végétation naturelle

©Samuel Alleaume

22/07/2019

Objectif

L'objectif de ce TP est de créer un exemple de fond de cartes physiologiques des végétations en utilisant des données issues de la télédétection. L'approche physiologique de la végétation consiste à caractériser les milieux naturels selon leur structure, leur hauteur, leur densité ou leur productivité.

Méthode générale

L'exercice se base sur l'apport conjoint d'une image mono-date à très haute résolution spatiales (THRS, d'une résolution métrique) et d'une série temporelle annuelle d'images de hautes résolutions spatiales (HRS, d'une résolution décamétrique). Il s'agit de caractériser des polygones issus d'une segmentation (1) selon une approche orientée objets et hiérarchisée.

La première étape consiste à segmenter l'image THRS en objets qui seront par la suite caractérisés selon une démarche orientée objet et hiérarchisée. L'étape suivante consiste à extraire des milieux naturels en se basant sur la série temporelle annuelle des images HRS. Cette approche s'appuie sur le fait que les milieux naturels ont une production végétale toute l'année (en dehors des périodes d'enneigement) contrairement à d'autres occupations du sol (cultures, bâtis ...). En effet, les cultures se retrouvent au moins une fois en sol nu dans l'année après récolte et en période du labour tandis que les zones urbaines ont une production végétale quasiment nulle toute l'année. Par la suite, les images THRS permettent de décliner les végétations des milieux naturels en trois niveaux de densité: herbacées (ouverts), ligneux mixtes (semi-ouverts) et ligneux denses (fermés). Puis, la série temporelle est de nouveau mise à contribution afin de caractériser les zones herbacées (pelouses ou prairies) en niveaux de productivité de biomasse végétale. Finalement, l'ensemble de ces traitements permet d'aboutir à la carte de végétation. Pour mener à bien cet objectif, les différentes parties de ce travail se résument comme suit (Figure 1):

1. Segmentation ou découpage automatique en zones homogènes à partir de l'image THRS mono-date ;
2. Extraction des milieux naturels à partir de la série temporelle d'images HRS ;
3. Extraction du degré d'ouverture des milieux c'est-à-dire de la densité des végétations naturelles ;
4. Extraction du niveau de productivité de la biomasse végétale ;
5. Carte de végétation finalisée.

Figure 1 : Chaîne de traitements pour la cartographie physiologies végétales.1

Zone d'étude

Le secteur analysé dans cet exercice est situé dans le département de l'Isère à environ 10 kilomètres au sud-est de Grenoble entre la commune d'Herbeys au nord et Vaulnareys-le Bas au sud (latitude 45°07N, longitude 5°48E, Figure 1). La zone d'intérêt représente environ 25 km² et couvre un gradient de paysage de plaine agricole au nord vers un paysage semi-montagnard au sud, atteignant jusqu'à 1000 mètres d'altitude.

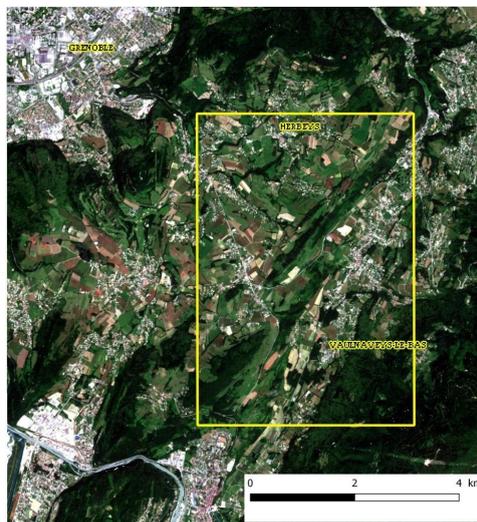


Figure 1. Localisation de la zone d'étude

Logiciels et données

Logiciel requis

Les traitements prévus dans la méthode de cartographie des végétations font appel aux fonctionnalités de base du logiciel QGIS (version 3.18) ainsi que la bibliothèque intégrée de traitement d'image Orfeo ToolBox (OTB) développée par le CNES. Les fonctions d'OTB seront exécutées par le lanceur d'applications MAPLA (Monteverdi Application Launcher).

Données de télédétection

Les images de télédétection mises à contribution pour cet exercice sont fournies par l'infrastructure de données spatiales (IDS) du pôle surfaces continentales Theia (theia-land.fr).

Image à très haute résolution spatiale (THRS)

L'image à très haute résolution spatiale (THRS) utilisée est une image **SPOT 6&7** extraite de **la couverture France entière de l'Equipex Geosud** (<http://ids.equipex-geosud.fr>). Les images bénéficient d'une résolution spatiale de 1,5 mètres en mode panchromatique et de 6 mètres en multispectrale composée des bandes bleue, verte rouge et proche-infra-rouge. L'image utilisée a déjà été prétraitée par la technique du pansharpening, c'est-à-dire que les bandes multispectrales ont été fondues avec l'image panchromatique afin d'obtenir une

image couleur d'une résolution de 1,5 mètres. De plus, l'image a subi une correction radiométrique de type Top of Atmosphere (TOA) en utilisant l'extension de QGIS « Geosud reflectance TOA ». Les valeurs numériques des pixels sont en milli-réfectance (16 bits). L'image SPOT7 date du 9 juillet 2020 et est projetée en Lambert 93. Les bandes sont réparties ainsi : rouge (bande 1), vert (bande 2), bleu (bande 3) et proche-infra-rouge (canal 4). Pour l'exercice cette image est nommée **SPOT_20200709.tif**.

Série temporelle *Sentinel-2*

La série temporelle utilisée est issue d'une série d'images **Sentinel-2 de 2020**. La mission Sentinel 2 de l'ESA, lancée dans le cadre du programme Copernicus de l'Union Européenne, se compose de deux satellites, Sentinel 2A et Sentinel 2B : le premier a été lancé en juin 2015 et le second en mars 2017. Chacun de ces satellites permet une observation de toutes les terres tous les 10 jours sur 13 bandes spectrales du visible au moyen infra-rouge. En combinant les données des deux satellites, chaque pixel est observé tous les 5 jours. Les images utilisées sont des données de niveau 3A, c'est-à-dire qu'elle correspondent à des synthèses annuelles de réflectance de surface. Les images ont été corrigées des effets atmosphériques et de relief et sont accompagnées d'un bon masque de nuages et d'ombres de nuages. Pour cet exercice nous utiliserons les quatre bandes d'une résolution de 10 mètres à savoir les bandes B2, B3, B4 et B8 correspondant respectivement au bleu (bande 1), vert (bande 2), rouge (bande 3) et proche-infra-rouge (canal 4). Ces images ont été reprojetées en Lambert 93. Les 12 images constituant la série temporelle sont nommées ainsi : **S2A_2020MMJJ.tif** (MM pour le mois, JJ pour le jour)

Données vectorielles

Quatre fichiers vectoriels sont fournis et correspondent à des sites échantillon pour la calibration et la validation des classifications par seuillage ou supervisée. Ces échantillons proviennent d'une photo-interprétation fondée sur l'image SPOT 7 et de polygones issus d'une segmentation.

echant_anthropique_naturelle.shp : échantillons d'entraînement nécessaire à la classification par SVM pour les classes « Anthropisée » et « végétation Naturelle ». La classe « Anthropisée » correspond à des zones urbanisées et des zones de cultures, la classe « végétation Naturelle » comprend des prairies permanentes, des landes et des forêts.

validation_anthropique_naturelle.shp : échantillons de validation pour évaluer la distinction entre la classe dite « anthropisées » et la classe « végétation naturelle ».

La table attributaire est constituée de 3 champs: ID, ECHANT et CLASS. ID correspondant aux identifiants uniques des polygones. Le champ texte ECHANT contient "H" pour la classe «Herbacées», "M" pour la classe « ligneux Mixtes » et "D" pour la classe « ligneux Denses » correspondant respectivement aux valeurs numériques 3, 4, et 5 du champ numérique CLASS.

echant_densite_veg.shp : échantillons d'entraînement pour les classes de densités des végétations semi-naturelles.

validation_densite_veg.shp : échantillons de validation pour ces mêmes classes de densités.

La table attributaire est constituée de 2 champs principaux : ECHANT et CLASS. 3 modalités ECHANT sont possibles : H pour la classe «Herbacée», M pour la classe « Mixte » et D pour la classe « ligneux denses » correspondant respectivement à la valeur numérique 3, 4, et 5.

Etape 1 : Traitement de l'image de THRS

1.1. Segmentation de l'image

L'image optique THRS est d'abord segmentée en zones homogènes sous forme de polygone (objets).

Une segmentation consiste en un regroupement de pixels suivant les mêmes propriétés prédéfinies. Les pixels sont alors regroupés en régions (encore appelés objets ou segments). Elle crée un lien conceptuel avec une perception humaine en essayant d'imiter la capacité de perception du photo-interprète à délimiter des zones homogènes. Il existe différents algorithmes qui permettent de générer une segmentation. Les segments ainsi créés sont aussi appelés objet et sont la base de la classification orientée objet. Toutes les études s'accordent sur le fait que les méthodes orientées objet permettent une plus grande flexibilité dans la capacité à transférer des stratégies de classification basées sur des images et des relevés de terrain récents vers des classifications sur des images passées où les relevés de terrain sont souvent inexistantes. Par conséquent, la couche de segments ainsi générée servira de référence spatiale pour les traitements et les classifications qui suivent.

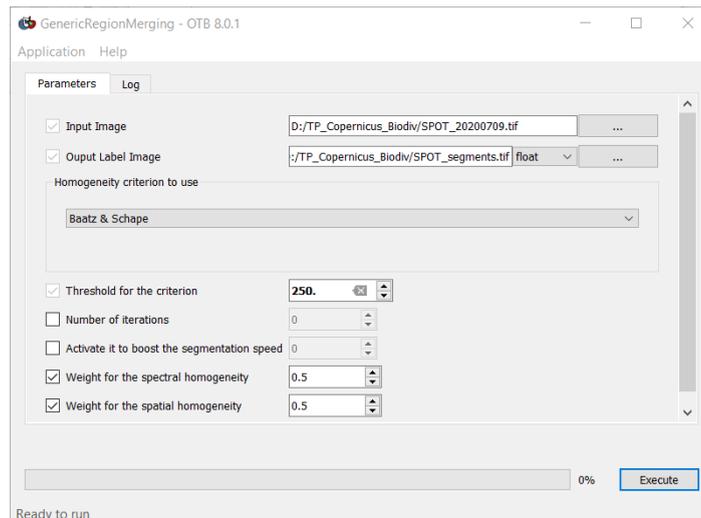
La technique de segmentation retenue est celle de Baatz-Shape. Cette technique de segmentation multi-résolution est une technique ascendante de croissance par paire de régions qui débute par un ensemble d'objets correspondants à un seul pixel de l'image. Les plus petits objets sont ensuite fusionnés de façon itérative dans de plus grands en fonction des critères d'homogénéité utilisant l'information des bandes spectrales de l'image ainsi que la forme des objets courants. Le processus s'arrête lorsque l'homogénéité des objets candidats à la fusion dépasse un seuil d'hétérogénéité maximale acceptée.

Cette méthode de segmentation multi-résolution a été développée à l'origine dans le logiciel propriétaire eCognition, elle a été par la suite implémentée dans OTB via l'application **GenericRegionMerging**.

Dans Mapla, sélectionner la fonction **Segmentation > GenericRegionMerging**

3 paramètres sont à prendre en compte :

- **th : Seuil du critère**, valeur de seuil (sans unité de mesure) pour le critère du maximum degré d'hétérogénéité des objets produits par la segmentation. Une valeur plus grande produit des plus grands segments.
- **cw : Poids pour l'homogénéité spectrale**, d'une valeur entre 0 et 1, il définit l'importance relative de la radiométrie (couleur) dans le critère d'agglomération. Une valeur égale à 0 impliquera que les objets seront optimisés exclusivement par leur homogénéité spectrale.
- **sw : Poids pour l'homogénéité spatiale**, d'une valeur entre 0 et 1, il définit l'importance relative de l'information de forme. Une valeur de 0 correspond à des objets de forme plus sinueuse. Une valeur plus élevée implique une majeure compacité. Dans la théorie, un cercle est la forme plus compacte possible.



La segmentation est un processus long qui peut prendre quelques minutes.

La segmentation ainsi réalisé est un raster, il doit être maintenant transformé en raster dans **QGIS** en utilisant l'outil *Polygonize*. Sauver la vectorisation sous le nom de *segments.shp*.

Afficher la segmentation en faisant apparaître l'image SPOT_20200709.tif en fond, comme ceci :



1.2. Calcul de l'indice de végétation NDVI

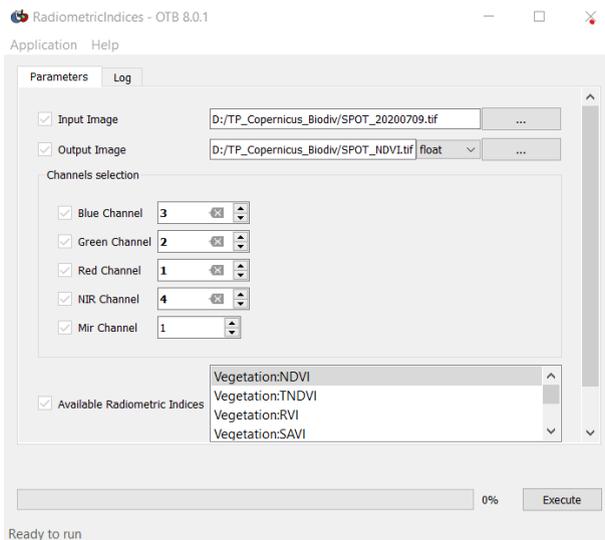
L'indice de végétation normalisé (NDVI) permet d'observer et d'analyser la couverture végétale sur les images de télédétection optiques. Cet indice traduit la signature spectrale très caractéristique de la végétation active. Dans la bande spectrale du visible, et en particulier dans le rouge (0.6 à 0.7 μm), la végétation absorbe fortement le rayonnement solaire permettant l'activité chlorophylle, mais cet effet cesse dans le proche infrarouge (0.7 à 0.9 μm) où, au contraire, la végétation réfléchit fortement. Comme les sols nus présentent un faible écart de réflectance entre le rouge et le proche infrarouge, la distinction entre un sol nu et un sol couvert ne pose généralement aucun problème :

$$NDVI = \frac{\rho_{PIR} - \rho_{Rouge}}{\rho_{PIR} + \rho_{Rouge}} \quad [1.1]$$

où ρ_{PIR} est la réflectance de la bande proche infrarouge et ρ_{Rouge} celle de la bande rouge.

Le dénominateur est un facteur de normalisation qui compense partiellement la différence de réflectance de la surface liée à la hauteur du soleil ou aux angles de prise de vue des satellites. Par construction, le NDVI est compris entre -1 et 1. Un NDVI proche de 1 traduit une végétation active, proche de 0 un sol nu et, en-dessous de 0 des essentiellement des surfaces en eau.

Avec Mapla, sélectionner la fonction **Feature Extraction> RadiometricIndices**. Attribuer le numéro du canal à chaque bande spectrale (Bleu 3, Vert 2, Rouge 1 et Proche infrarouge 4) et sélectionner l'indice radiométrique NDVI dans la liste déroulante.



1.3. Calcul des indices de texture

La très haute résolution spatiale permet la reconnaissance de textures discriminantes pour les mosaïques végétales. La texture fait référence à la variation de ton d'une image et peut être une bonne approximation de la structure de la végétation. L'analyse de texture doit permettre non seulement de distinguer les zones herbacées (prairies et pelouses) des strates ligneuses, mais également de séparer les strates ligneuses mixtes (ouvertes) des strates ligneuses denses. Deux types de texture sont mises à contribution : les indices d'Haralick et les indices SFS.

Les indices d'Haralick

Les indices de texture contiennent des informations relatives à la distribution spatiale des variations de tons pour une bande. Le concept de ton est basé sur la variation des nuances de gris alors que celui de la texture concerne la distribution spatiale de ces tons (Haralick, 1979). Un premier groupe d'indices de texture d'Haralick est extrait de la matrice des niveaux de gris appelée matrice de cooccurrence (GLCM pour Grey Level Cooccurrence Matrix). Elle permet de capturer la dépendance spatiale des valeurs de niveaux de gris dans une image. La matrice de cooccurrence mesure le nombre de fois qu'un pixel d'un certain niveau de gris apparaît dans une direction définie et selon une distance spécifiée de ses pixels voisins ayant un niveau de gris spécifique. Sept indices simples d'haralick sont ainsi calculés en utilisant l'Orfeo Toolbox : Cluster Shade, Correlation, Energy, Entropy, Haralick Correlation, Inverse Difference Moment (IDM) et Inertie.

Dans MAPLA sélectionner la fonction **Feature Extraction> HaralickTextureExtraction**

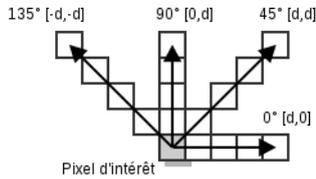
Les indices de texture simples d'Haralick sont calculés sur la bande rouge (canal 1) de l'image SPOT.

Tester les paramètres de bases (selon la figure ci-dessous), en les faisant varier.

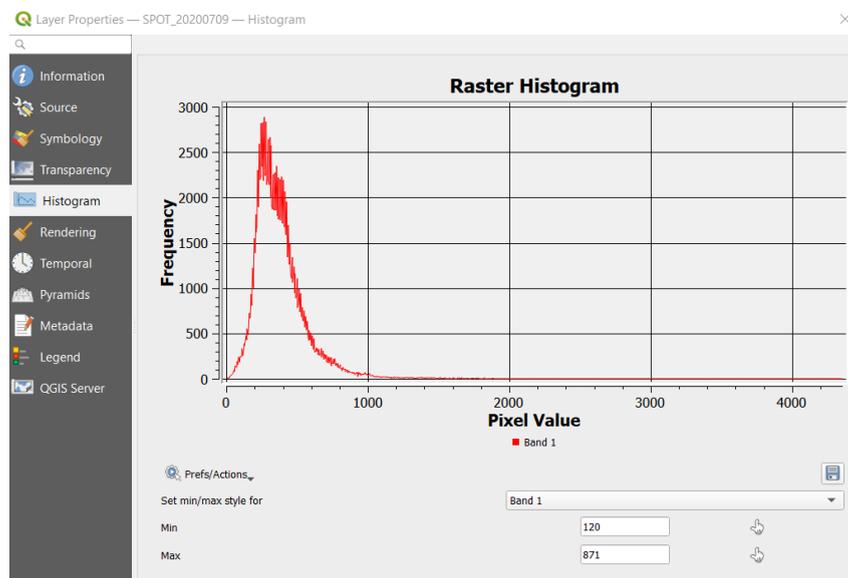
Les principaux paramètres à prendre en compte sont :

X/Y Radius : « rayon » en distance pixel pour indiquer la taille de la fenêtre d'analyse (par exemple si X/Y radius = 2,2 alors la taille de la fenêtre sera de 5x5 pixels)

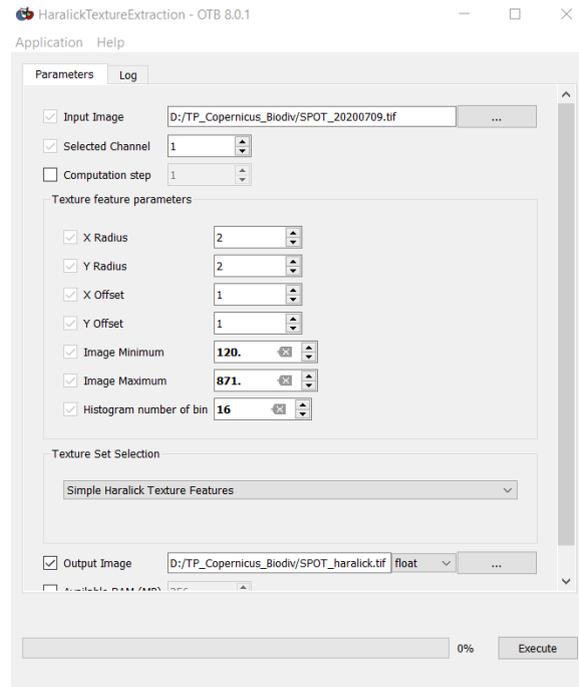
- **X/Y Offset** : ces paramètres indiquent une direction selon une droite dont un premier point serait [0,0] et le deuxième [X offset , Y offset]. Par exemple [1,1] crée une direction d'analyse à 45°.



- **Image Minimum / Image maximum** : bornes de gris correspondant à la dynamique de l'image et conservés pour l'analyse. Pour cela, dans QGIS ouvrir les propriétés de l'image à analyser (clic droit), ouvrir l'onglet histogramme et afficher l'histogramme de la bande rouge (1). Noter les valeurs minimales et maximum affichées.



- **Histogram number of bins** : nombre d'intervalles de quantification. Indiquer 16.



L'image produite est image en 8 bandes avec un indice par bande. Puisque chaque bande est indépendante, afficher une bande à la fois dans QGIS.

Les indices SFS

SFS (Structural Feature Set) est une approche directionnelle de calcul de texture qui consiste à appliquer une analyse directionnelle à partir du pixel central de la fenêtre d'analyse. Ce sont des mesures statistiques employées pour réduire la dimension et extraire les entités d'un histogramme de lignes de direction. Les lignes de direction peuvent être définies comme étant un nombre prédéterminé de lignes également espacées et partant du pixel central. L'extension de ces lignes est basée sur la similarité du niveau du gris du voisin et les lignes partant du pixel central dans différentes directions. Orfeo Toolbox permet de calculer six mesures statistiques SFS: longueur, largeur, Pixel Shape Index (PSI), moyenne pondérée (W-Mean), Ratio et l'écart type (SD).

Dans MAPLA sélectionner la fonction **Feature Extraction> SFSTextureExtraction**

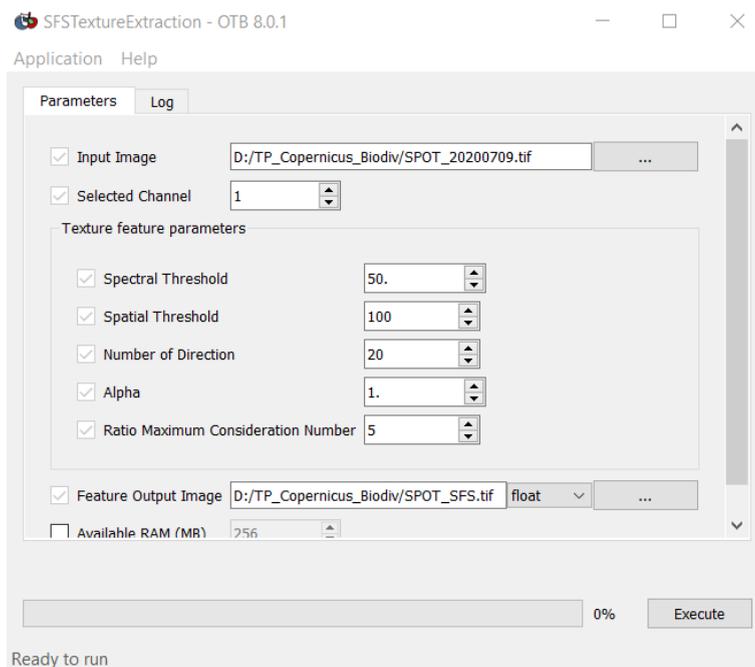
Les paramètres utilisés sont

Spectral threshold : la différence maximale autorisée entre la valeur du pixel central et celle d'un pixel de la ligne de direction.

Spatial threshold : la longueur maximale de la ligne de direction.

Alpha et **Ratio Maximum Consideration Number** : constantes pour ajuster la valeur de la moyenne pondérée.

Laisser l'ensemble des valeurs par défaut.



Etape 2 : Calcul du DHI (Dynamique Habitat Index)

L'exploitation d'une série temporelle annuelle d'images permet de prendre en compte la variabilité temporelle des végétations par conséquent la phénologie des habitats potentiels. Le DHI proposé initialement par Coops et al. (2008) est un indice défini par trois composantes. Il peut se calculer à partir de n'importe quel indice radiométrique. Il donne des indications sur la nature et la dynamique des habitats naturels. De plus, les végétations naturelles se distinguent des végétations culturelles de par leur plus grande stabilité phénologique dans le temps.

Trois indices composent le DHI :

- DHI cum : il correspond au cumul sur l'année (soit l'aire sous la courbe phénologique d'une année). Il traduit un niveau de productivité annuelle de la végétation.
- DHI min : il correspond à la valeur minimale de la courbe phénologique d'une année. Il traduit le niveau le plus bas de productivité de la végétation.
- DHI max : le maximum de la courbe phénologique d'une année
- DHI cv : il correspond au coefficient de variation de la courbe phénologique d'une année. Il traduit la saisonnalité de l'habitat.

2.1. Calcul des indices de végétation NDVI

Extraire le NDVI de chacune des 12 images Sentinel 2

Dans MAPLA sélectionner la fonction **Feature Extraction > RadiometricIndices**

2.2. Calcul des minimum et maximum de la série

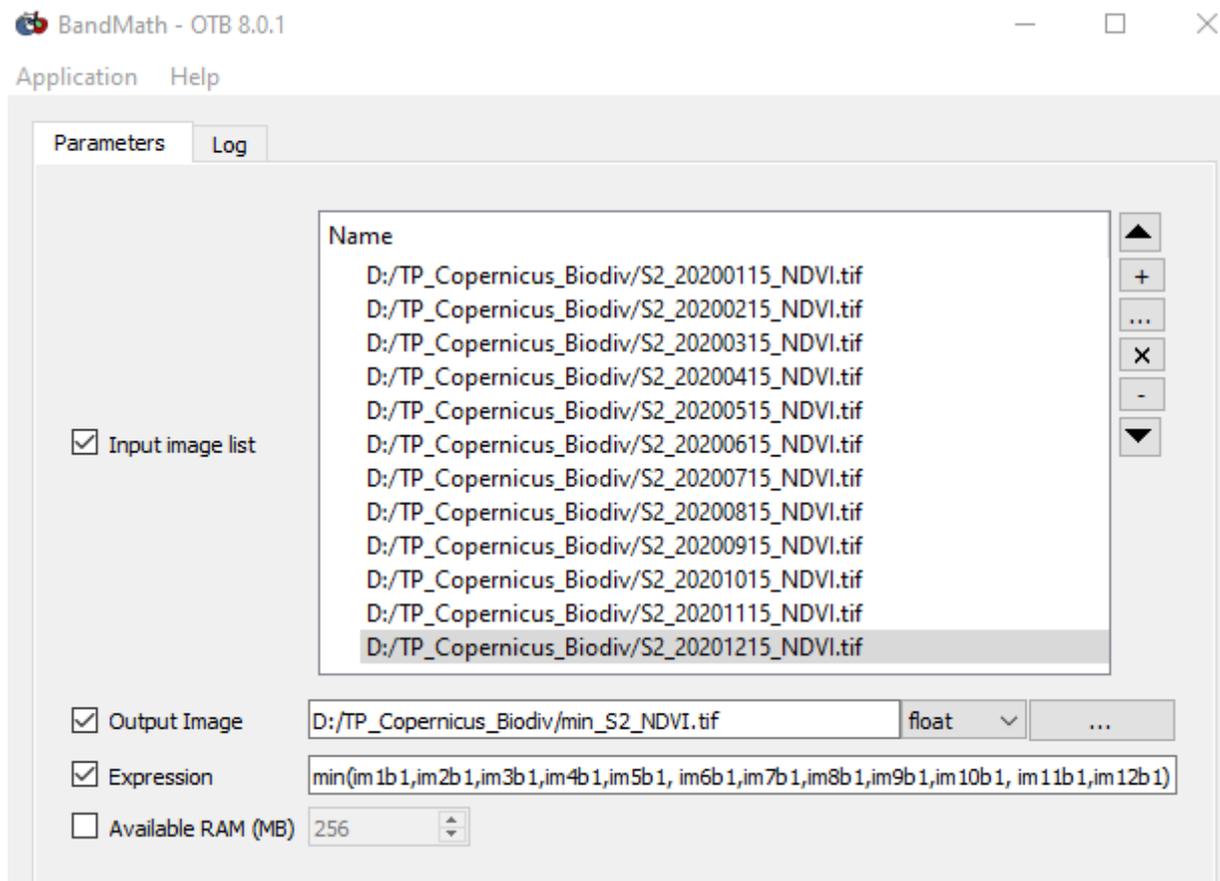
Dans MAPLA sélectionner la fonction **Image Manipulation > Band Math**

Calcul du maximum de la série des NDVI de Sentinel 2.

Dans la liste, sélectionner les 12 images NDVI générées à l'étape précédente.

Utiliser l'expression :

min(im1b1,im2b1,im3b1,im4b1,im5b1, im6b1,im7b1,im8b1,im9b1,im10b1, im11b1,im12b1)



puis l'expression :

max(im1b1,im2b1,im3b1,im4b1,im5b1, im6b1,im7b1,im8b1,im9b1,im10b1, im11b1,im12b1)

Etape 3 : Séparation des milieux naturels des cultures à partir de la série par méthode des seuillages

La valeur minimale du NDVI annuelle (NDVImin) permet de détecter les valeurs les plus faibles d'activités chlorophylliennes considérées comme fortement anthropisées tels que les sols nus, les zones urbaines ou les carrières. Les milieux cultureux sont particuliers puisque durant l'année ils présentent une phase végétalisée et une phase de sol nu liée à la récolte et/ou aux labours. Parallèlement, les milieux naturels maintiennent une activité chlorophyllienne sur toute l'année.

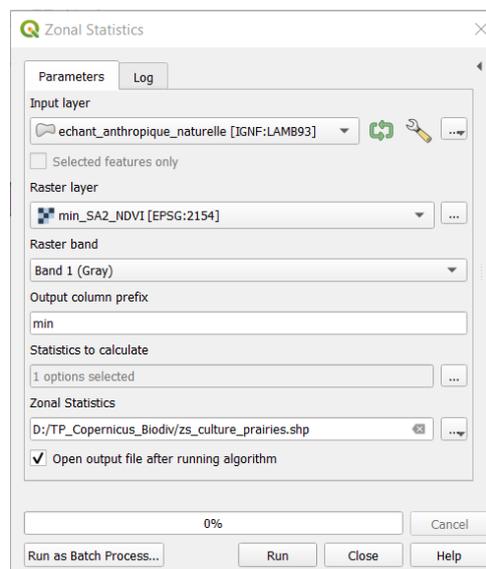
Par conséquent, on utilise des sites de calibration représentant deux classes à distinguer : les "cultures" et les "végétations naturelles" pour trouver le seuil entre des milieux toujours en végétation et des milieux anthropisés. Les polygones de calibration font l'objet d'un croisement avec l'image du NDVImin au travers d'une statistique zonale. La moyenne du NDVImin est ainsi calculée pour chaque polygone. Une recherche des seuils de séparabilité entre les deux classes s'effectue par la méthode SEaTH (Separability and Treshold) détaillée ci-dessous. Par la suite, ce seuil est appliqué pour attribuer aux polygones de la segmentation soit la classe « milieux naturels » soit la classe « anthropisés ».

3.1. Statistique de zone sur le minimum du NDVI de la série temporelle sentinelle

A ce niveau, il s'agit d'extraire les valeurs moyennes du minimum de NDVI annuel, d'une part pour chaque polygone d'échantillon, et d'autre part pour chaque polygone de la segmentation.

Dans la toolbox de QGIS sélectionner la fonction **Raster analysis/Zonal Statistics** (Cette fonction existe aussi dans Mapla sous ZonalStatistics, mais ne permet pas de sélectionner la statistique d'intérêt, ni d'éditer un préfixe dans la colonne.)

Choisir le fichier *min_SA2_NDVI.tif* comme couche raster et le fichier d'échantillons *echant_anthropique_naturelle.shp* comme couche de polygone contenant les zones.



3.2. Réaliser la méthode Seath à l'aide d'un script

La méthode SEaTH (Separability and Treshold) a été développée par Nussbaum et al en 2006. Cette méthode permet de calculer pour chaque couple de classes le degré de séparabilité des deux classes ainsi que la valeur du seuil optimale à utiliser pour une variable testée. Cette méthode est applicable pour une classification d'image orientée objet. Deux classes sont comparées à travers des sites échantillons (objets) selon la « distance de Bhattacharyya » B (1). Pour une meilleure comparabilité, la valeur de B est transformée en distance de Jeffries-Matusita ou J (2).

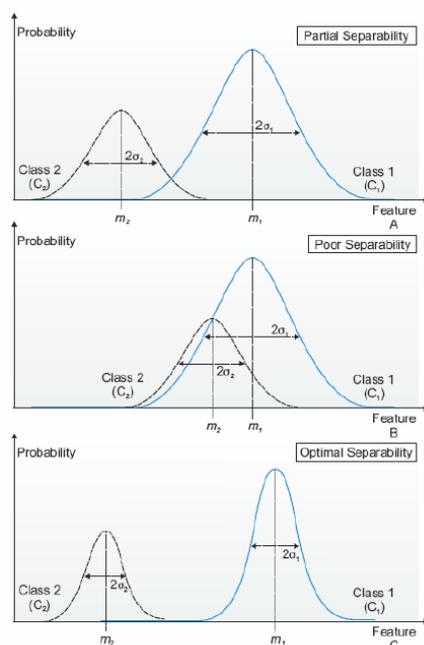
$$B = \frac{1}{8} (m_1 - m_2)^2 \frac{2}{\sigma_1^2 + \sigma_2^2} + \frac{1}{2} \ln[] \quad (1)$$

où m et σ sont la moyenne et l'écart-type pour la distribution de la variable pour les deux classes 1 et 2.

$$J = 2(1 - e^{-B}) \quad (2)$$

La valeur de J varie entre 0 et 2. Plus cette valeur de J est proche de 2, plus la séparabilité est optimale. La Erreur : source de la référence non trouvée 4 montre trois exemples de séparabilité. Ceci pour des distributions de probabilité de variables pour deux classes. La valeur du seuil optimal est estimée pour une forte valeur de J . Ce seuil optimal est calculé par une approche statistique bayésienne résolvant l'équation (3) pour résoudre la valeur de x . Cette équation est sous la forme d'un modèle mixte de probabilité gaussienne.

$$p(x) = p(1)p(C_1) + p(2)p(C_2) \quad (3)$$



Séparabilité partielle

$$0 < J < 2$$

Séparabilité faible

$$J = 0$$

Séparabilité optimale

$$J = 2$$

Figure 2: Probabilité de distribution et valeurs de J , coefficient de séparabilité de Jeffries-Matusita

Certaines méthodes utilisées n'existent pas dans Qgis. Il est possible d'en créer en écrivant un script python formalisé qui sera intégré comme une boîte à outils.

Un script déjà conçu en Python et qui représente la méthode Seath est fourni sous forme du fichier **seath_qgis.py**.

- Pour ajouter un script :

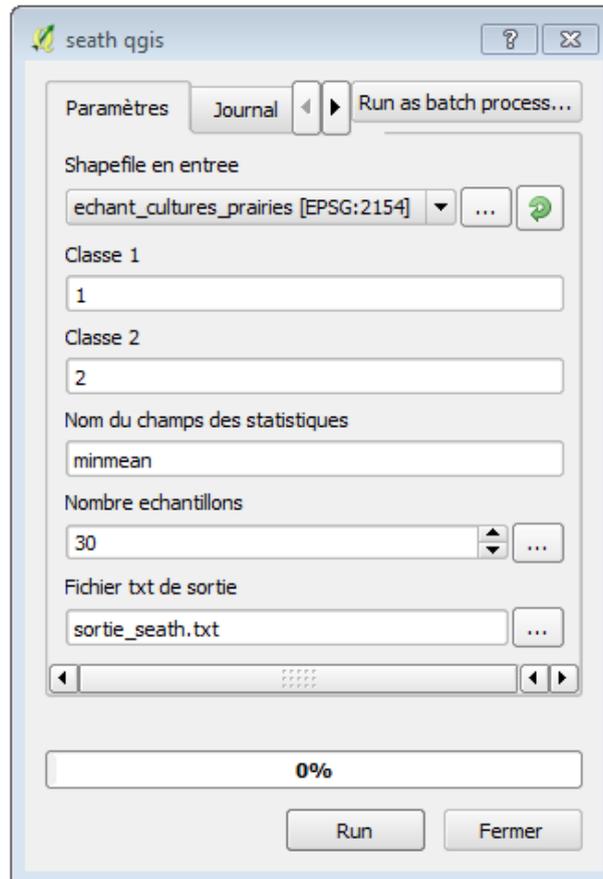
Dans la **Boîte à outils de traitements**

Scripts  > **Add script to toolbox** > Sélectionner le script **seath_qgis.py**

-Lancer le script

Dans la **Boite à outils de traitements**

Scripts > Scripts utilisateur > seath qgis



Ce script crée un fichier sortie (sortie_seath.txt) qui fournit la valeur du seuil optimale pour discriminer les 2 classes d'entrée : classe 1, les zones anthropiques, classe 2 les zones de végétation naturelles. Ouvrir ce fichier de sortie et noter cette **valeur seuil**.

"To extract the class 1, we need to this index minmean < 0.4112557799"

3.3. Appliquer le seuil sur la segmentation

Dans MAPLA sélectionner la fonction Image **Analysis/ZonalStatistics**

Choisir le fichier **min_SA2_NDVI.tif** comme couche raster et le fichier **segments.shp** comme couche de polygone contenant les zones.

Appliquer la valeur seuil déterminée précédemment sur les segments : Il suffit de sélectionner les polygones dont la valeur minmean est inférieure à la valeur seuil (ici 0.411) et de leur affecter la classe A pour anthropique (la classe 1 était attribuée à la classe anthropique dans la méthode Seath).

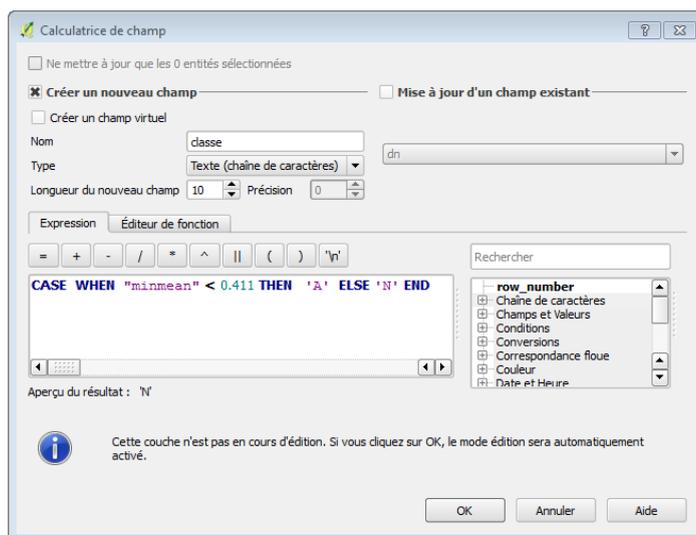
Dans QGIS, ouvrir la table attributaire (clic droit sur la couche, **Ouvrir la table d'attributs**) ;

Activer la **calculatrice de champ** 

Créer un nouveau champ texte appelé « classe »

Utiliser l'expression

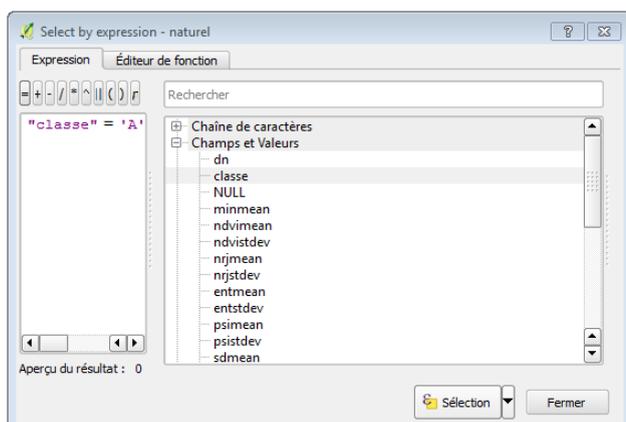
CASE WHEN "minmean" < 0.411 THEN 'A' ELSE 'N' END



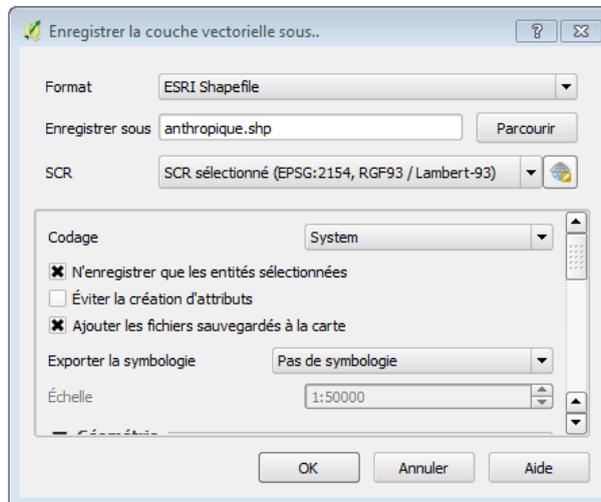
3.4. Séparer la couche segments en couche anthropique.shp et une couche naturel.shp

- Ouvrir la table attributaire de la couche des segments classés

- Faire une sélection par expression 



- Puis enregistrer cette sélection sous **anthropique.shp** en faisant : clic droit sur la couche **segments** / Enregistrer sous.../ **anthropique.shp** en cochant l'option N'enregistrer que les entités sélectionnées



De la même manière créer la couche **naturel.shp** en faisant une sélection inverse des objets déjà sélectionnés  ou en faisant une sélection par expression  « classe » = 'N'

Etape 4 : Classification des densités des végétations par classification supervisée SVM

Cette étape consiste à faire une classification supervisée sur les niveaux de densité des milieux naturels basée sur une approche SVM. Les primitives utilisées pour la classification correspondent aux données de NDVI et de texture extraites de l'image THRS (SPOT).

La techniques *Support Vector Machines* (SVM) est encore appelée technique des *Machines à Vecteur de Support* ou *Séparateurs à Vastes Marges*. Elle consiste à résoudre un problème de classification en plaçant un hyperplan dans l'espace des données comme frontière de décision. L'hyperplan maximise la distance entre deux classes des points les plus proches (vastes marges).

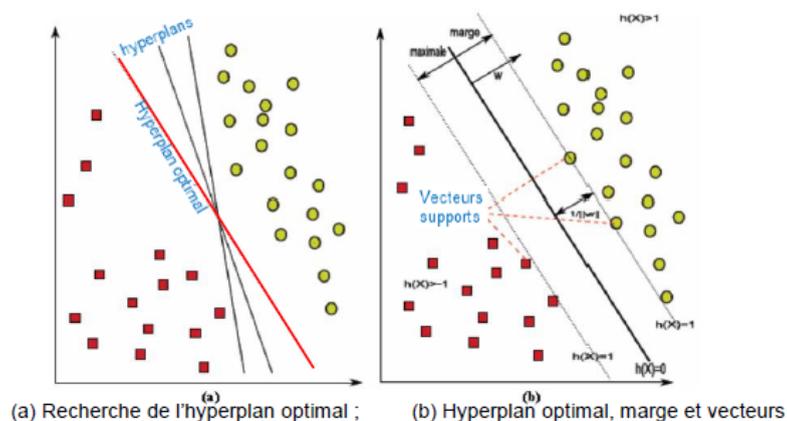


Figure 3: Recherche de l'hyperplan optimal par la technique SVM.

4.1. Zonal statistique à la fois sur les segments à classer et sur les zones d'entraînement

Calculer les statistiques de zones (moyennes et écart-types) sur la couche vectorielle à classer **naturel.shp**. puis sur la couche échantillon **echant_densite_veg.shp**.

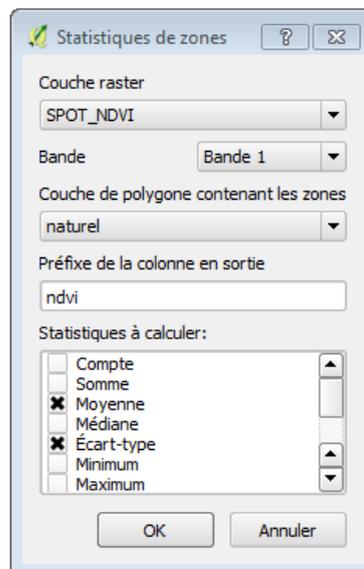
Dans la **barre de menu**:

- **Raster > Statistique de zones > Statistique de zones**

Réaliser cette opération sur SPOT_NDVI.tif ; utiliser **ndvi** comme préfixe de colonne et sélectionner Moyenne et Ecart-type.

Puis sur les indices de texture :

- SPOT_haralick.tif , bande 1 , préfixe **nrj** (pour énergie)
- SPOT_haralick.tif , bande 2 , préfixe **ent** (pour entropie)
- SPOT_SFS.tif , bande 3 , préfixe **psi** (pour PSI)
- SPOT_SFS.tif , bande 6 , préfixe **sd** (pour SD)



Les deux tables attributaires des fichiers ainsi traités ont maintenant 10 nouveaux champs contenant les variables prédictives (**features**) qui vont être utilisées pour la classification SVM : "ndvimean" "ndvistdev" "nrjmean" "nrjstdev" "entmean" "entstdev" "psimean" "psistdev" "sdmean" "sdstdev"

Attention :

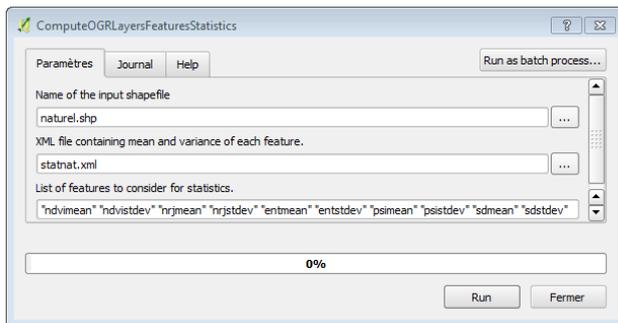
- Utiliser un préfixe en minuscule
- Ne pas oublier de le faire sur les deux couches vectorielles **echant_densite_veg.shp** et **naturel.shp** .

4.2. Création d'un fichier statistique sur les variables

- Calculer les statistiques globales de la couche vectorielle à classer

Dans la **Boîte à outils de traitements**

- **barre à outils Orfeo : Segmentation > ComputeOGRLayersFeaturesStatistics**



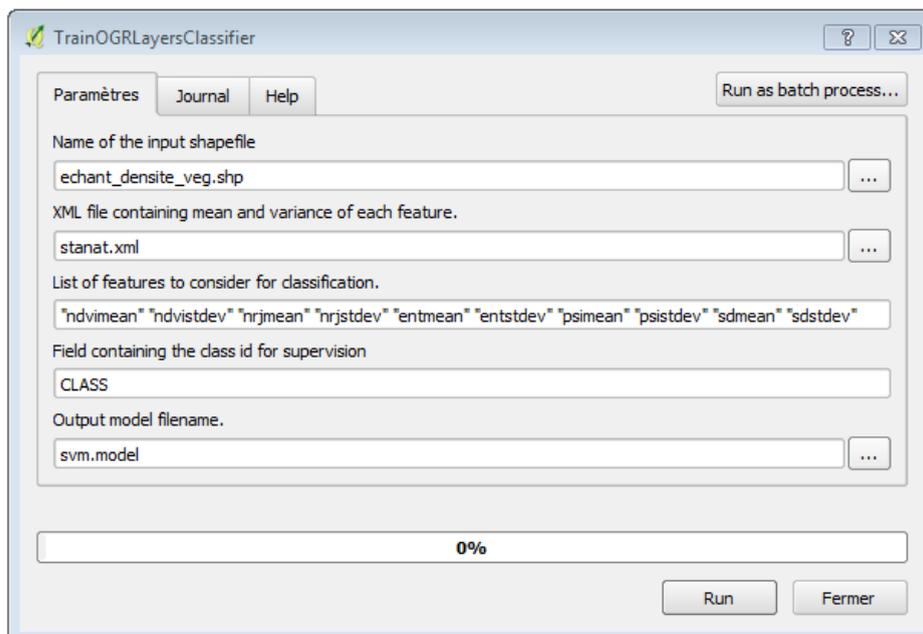
Les noms des variables prédictives (features) listés doivent être écrits entre guillemets.

4.3. Création du modèle de classification

- Entraînement du classifieur SVM sur la couche vectorielle des sites d'entraînement

Dans la **Boîte à outils de traitements**

- **barre à outils Orfeo : Segmentation > TrainOGRLayersClassifier**

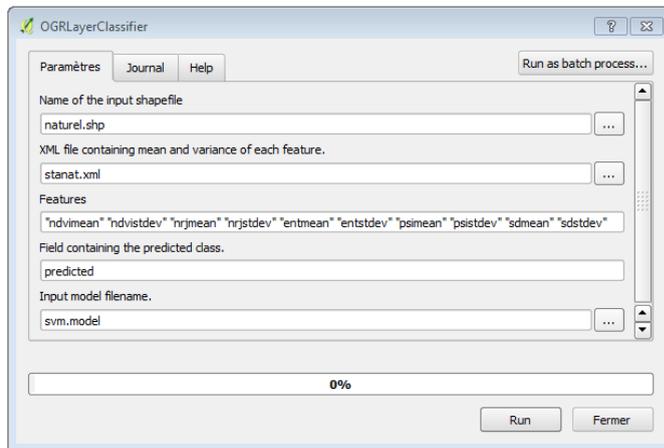


4.3. Classification des milieux naturels

L'étape précédente a permis la création d'un modèle prédictif. Ce modèle est maintenant appliqué aux segments à classifier.

Dans la **Boîte à outils de traitements**

- barre à outils Orfeo : Segmentation > OGRLayersClassifier



La table attributaire de la couche **naturel.shp** contient un nouveau champ appelé '*predicted*' correspondant à la valeur de la classe prédite.

Remarque :

- Si le fichier **naturel.shp** est déjà ouvert dans QGIS, le nouveau champ pourrait ne pas apparaître à l'ouverture de la table. Dans ce cas, fermer la couche et la réouvrir afin de rafraîchir l'affichage.

Etape 5 : Extraction du niveau de productivité des zones herbacées

Pour analyser la productivité maximale de biomasse végétale des zones herbacées, les segments identifiés comme appartenant à la classe « herbacée » subissent une analyse multi-temporelle. Puisqu'il s'agit de distinguer la productivité annuelle des zones herbacées, la classification se fait sur la base du NDVI maximum annuel (NDVImax) dérivé des images de Sentinel 2.

Nous ne disposons pas d'échantillon permettant de calibrer l'indice temporel NDVI max avec des données d'observation de terrain. Par conséquent, le NDVImax est décliné en 3 classes au regard de l'histogramme de distribution des valeurs. Cette approche se base sur l'hypothèse forte que la productivité maximale de biomasse est directement corrélée au NDVImax annuel. L'indicateur est décliné en 3 niveaux de productivité : faible, moyenne et forte. Le seuil est calculé en fonction de la distribution statistique des valeurs de l'indice de production végétale. En théorie, dans une distribution qui suit une loi normale (ou distribution en cloche), les valeurs de l'intervalle comprises entre $[\mu - \sigma ; \mu + \sigma]$ (où μ est la moyenne et σ est l'écart-type) représentent 68% des valeurs de la distribution.

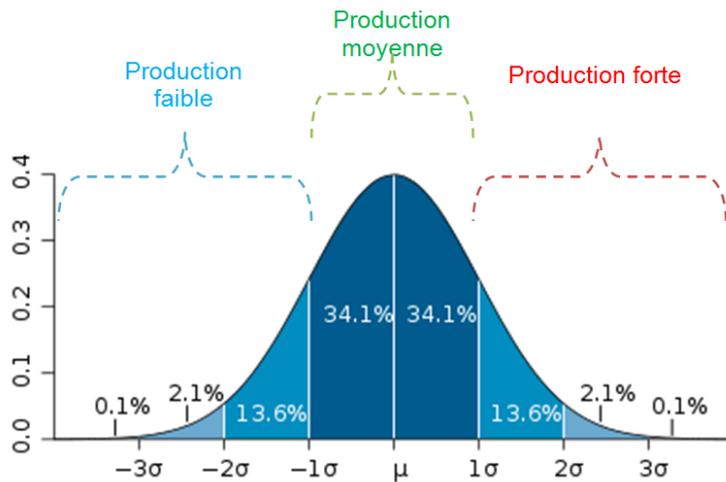


Figure 4 : Seuil théorique des 3 classes de production végétale en fonction d'une distribution normale.

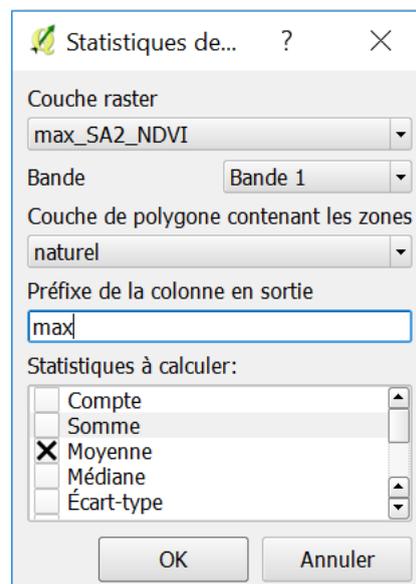
5.1. Statistique de zone sur le maximum du NDVI de la série temporelle sentinelle

A ce niveau, il s'agit d'extraire les valeurs moyennes du maximum de NDVI annuel, d'autre part pour chaque polygone classé en milieu « naturel » N.

Dans la **barre de menu**:

- **Raster > Statistique de zones > Statistique de zones** 

Choisir le fichier **max_SA2_NDVI.tif** comme couche raster et le fichier **naturel.shp** comme couche de polygone contenant les zones. Saisir un préfixe de colonne significatif comme **max** et cocher la statistique désirée c'est-à-dire **moyenne**.



5.2. Extraction du niveau de productivité des prairies

Les différents niveaux de productivité sont classés en fonction de la moyenne et de l'écart-type du NDVI_{max} sur l'ensemble des polygones du fichier *naturel.shp* correspondant aux herbacées (ou prairies naturelles).

Pour extraire ces informations :

- Dans la attributaire, sélectionner les polygones correspondants à la classe 3 (H) pour le champ « *predicted* » par une requête sélection 

- Onglet **Vue>Résumé statistique** ou 

Une fenêtre nommée « Statistiques » s'ouvrira. Sélectionner la couche vecteur, le champ *maxmean* et cocher Entité(s) sélectionnée(s) uniquement .

Relever la valeur moyenne et l'écart-type.

Calculer les intervalles

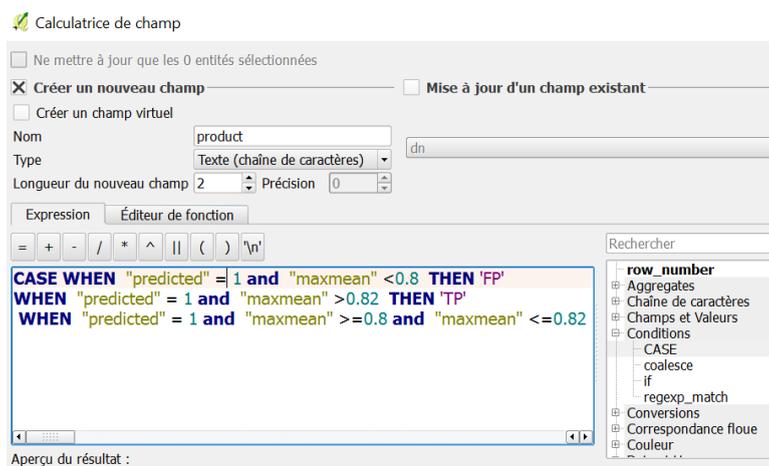
- De faible productivité (FP) $< \mu - \sigma$
- De moyenne productivité (MP) $[\mu - \sigma ; \mu + \sigma]$
- De très forte productivité (TF) $> \mu + \sigma$

Ouvrir la calculatrice de champ  afin de créer un champ de texte « *production* » qui recevra les valeurs FP, MP ou TF en fonction de la valeur *maxmean*. Ceci est réalisé par une expression de la forme suivante :

```
CASE WHEN "predicted" = 1 and "maxmean" <0.8 THEN 'FP'
```

```
WHEN "predicted" = 1 and "maxmean" >0.82 THEN 'TP'
```

```
WHEN "predicted" = 1 and "maxmean" >=0.8 and "maxmean" <=0.82 THEN 'MP' END
```



La couche a été automatiquement mise en mode édition, par conséquent quitter ce mode après exécution du calcul 

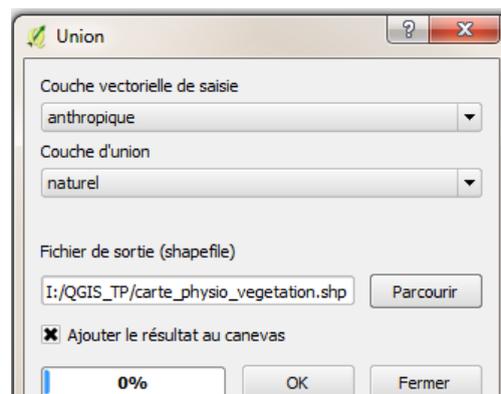
Etape 6. Cartographique finale

6.1. Union des cartes

Le fond final correspond à l'ensemble des couches vectorielles extraits depuis le début de ce tp soit, *anthropique.shp* et *naturel.shp*.

Pour fusionner ces 2 couches et en faire une seule et unique carte nommée ***carte_physio_vegetation.shp*** :

- Ouvrir toutes les couches dans QGIS
- Dans la barre de menu **Vecteur>Outils de géotraitement> Union**



6.2. Nettoyage de la table attributaire

Vous pouvez maintenant supprimer les champs inutiles de la nouvelle table attributaire de ***carte_physio_vegetation.shp***: activer le mode édition  puis à travers le bouton « supprimer le champ » .

Ne conserver que les champs « *predicted* » et « *production* ». Puis sauver et sortir du mode édition.

Création d'un champ pour la nomenclature finale

Un nouveau champ appelé « *classe_veg* » est créé afin de recevoir une nomenclature finale et compréhensible par tous :

- Ouvrir la calculatrice de champ  et saisir l'expression suivante :

```
CASE WHEN "predicted" = 3 and "production" = 'FP' THEN 'herbacées de faible productivité'
```

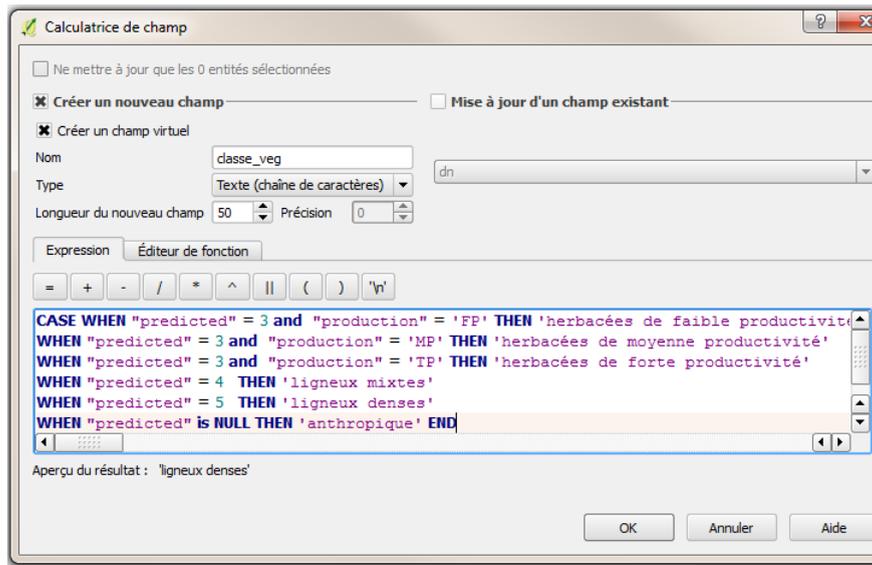
```
WHEN "predicted" = 3 and "production" = 'MP' THEN 'herbacées de moyenne productivité'
```

```
WHEN "predicted" = 3 and "production" = 'TP' THEN 'herbacées de forte productivité'
```

```
WHEN "predicted" = 4 THEN 'ligneux mixtes'
```

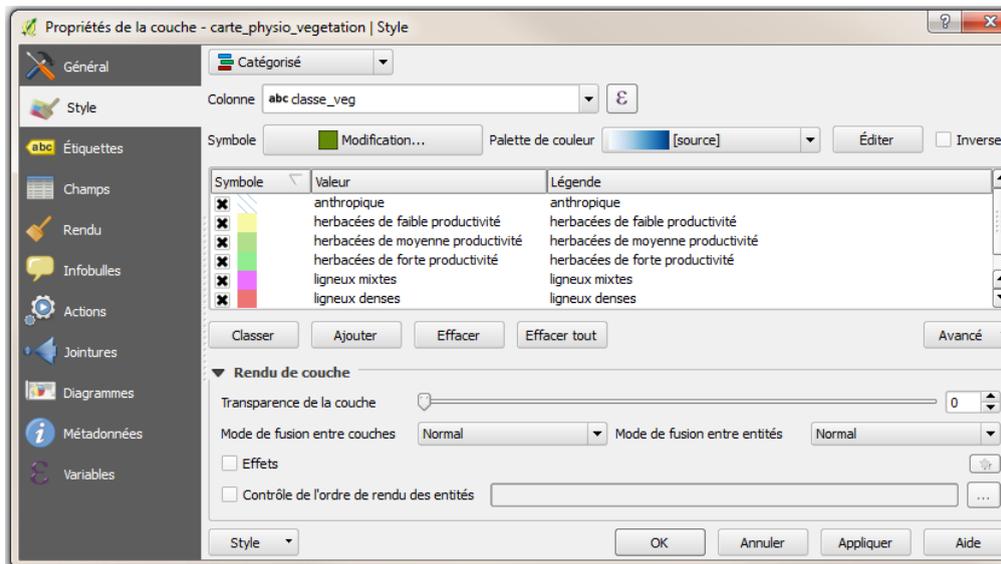
```
WHEN "predicted" = 5 THEN 'ligneux denses'
```

```
WHEN "predicted" is NULL THEN 'anthropique' END
```



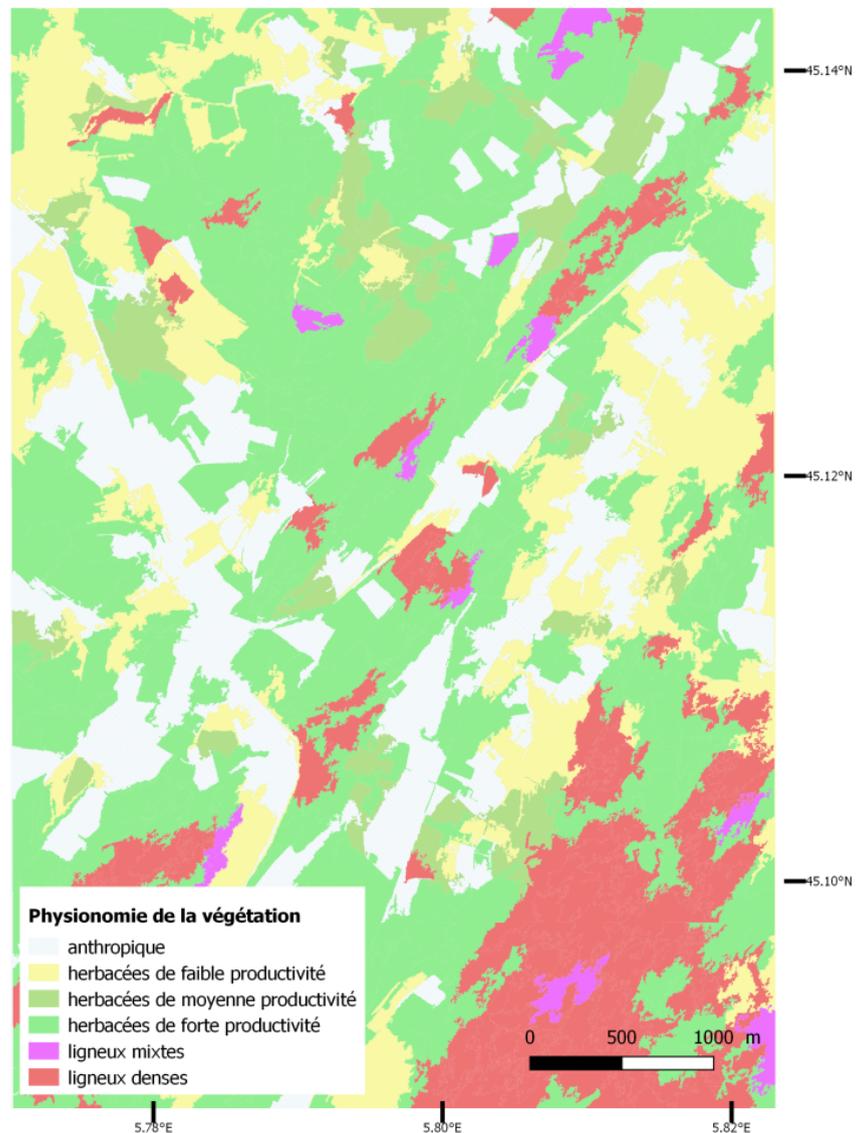
6.3. Symbologie

- **Propriété**
- Choisir l'onglet style
- A la rubrique colonne, choisir « classe_veg »
- Cliquer sur le bouton 
- Choisir une symbologie appropriée pour chaque classe et faire appliquer



6.4. Représentation finale

Voici un exemple de cartographie finale créée avec le composeur d'impression. Cette étape n'est pas détaillée dans ce chapitre.



Bibliographie

[CRE 17] Cresson R., Grizonnet M., Michel J., xxxxx

[HAR 79] Haralick, R.M. (1979). Statistical and structural approaches to texture. Proc. IEEE 67, 786–804.

[HUA 07] Huang X., Zhang L., Li P., Classification and Extraction of Spatial Features in Urban Areas Using High-Resolution Multispectral Imagery, IEEE Geoscience and Remote Sensing Letters, vol. 4, no. 2, pp. 260-264, 2007

[NUS 06] Nussbaum, S., Niemeyer, I., and Canty, M.J. (2006). SEaTH–A new tool for automated feature extraction in the context of object-based image analysis. Proc OBIA.

[OSE 16] Ose, K., Demagistri, L., and Corpetti, T. (2016). Traitement d'images satellitaires multispectrales. In Télédétection Pour l'Observation Des Surfaces Continentales, (ISTE Editions), pp. 73–133.