



HAL
open science

Utilisation des modules SAGA GIS dans QGIS

Paul Passy, Sylvain Theyry

► **To cite this version:**

Paul Passy, Sylvain Theyry. Utilisation des modules SAGA GIS dans QGIS. QGIS et outils génériques, 1, ISTE Editions, pp.119-163, 2018, Utilisation de QGIS en télédétection, 9781784053383. hal-03102527

HAL Id: hal-03102527

<https://hal.science/hal-03102527>

Submitted on 7 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Utilisation des modules SAGA GIS dans QGIS

4.1. SAGA GIS dans QGIS

4.1.1. Développement de SAGA GIS

SAGA GIS (*System for Automated Geoscientific Analyses*) est un logiciel SIG libre et open source, développé depuis le début des années 2000. SAGA GIS est principalement maintenu et développé par une équipe basée au Département de géographie physique de l'Université de Hambourg (Allemagne). C'est un logiciel multi-plateformes utilisable sous Windows et Linux et plus difficilement sous Mac OSX.

SAGA GIS est divisé en modules permettant de faire des traitements raster et vecteur. Un très large éventail d'applications est possible. Différents modules sont dédiés aux traitements d'images multi-spectrales, à la géomorphologie, aux géostatistiques, à l'hydrologie, à l'écologie du paysage, etc. Même si SAGA GIS est réputé essentiellement pour ses fonctionnalités raster, il existe de nombreux modules dédiés aux traitements de données vecteurs (analyse spatiale, zones tampon, polygones de Thiessen, etc.). Le logiciel utilise les bibliothèques GDAL et OGR pour gérer respectivement les formats raster et vecteur. De ce fait, de très nombreux formats de données sont supportés.

SAGA GIS propose des fonctionnalités très puissantes, issues de recherches dans des laboratoires de science de l'environnement du monde entier qui s'avèrent très utiles. Malheureusement, ce logiciel pêche par son manque de documentation détaillée.

Depuis quelques versions de QGIS, les modules de SAGA GIS peuvent être utilisés directement au sein de QGIS. QGIS peut donc faire office d'interface à SAGA GIS. La souplesse et l'ergonomie de QGIS se retrouvent ainsi couplées à la puissance de SAGA GIS. Il est cependant nécessaire de bien noter que quelques modules ne sont pas accessibles *via* QGIS.

4.1.2. Interfaçage de SAGA GIS dans QGIS

4.1.2.1. Installation

Sur Windows, le module SAGA GIS est directement installé lors de l'installation de QGIS. D'une manière générale il est intéressant d'installer QGIS en passant par l'interface pour « utilisateurs avancés » nommé « OSGeo Network Installer ». *Via* cette interface il est possible de configurer finement les modules et les différentes versions que l'on souhaite installer, notamment le module SAGA (figure 4.1).

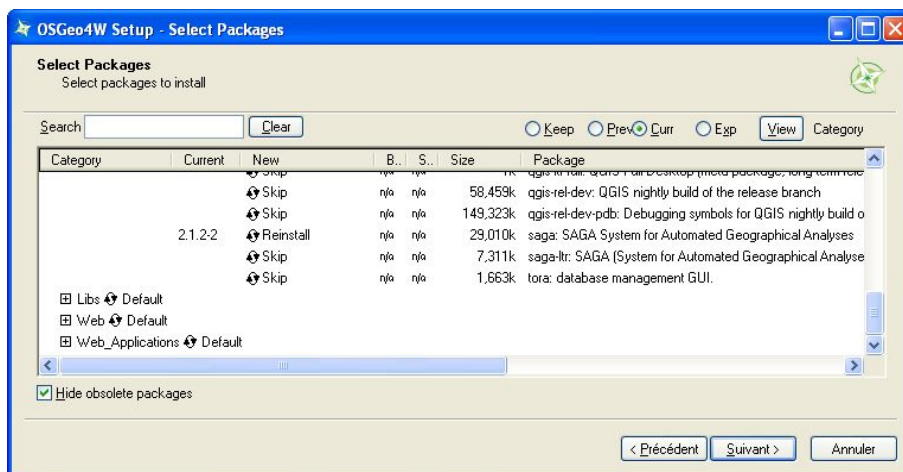


Figure 4.1. Installation de SAGA GIS via l'interface OSGeo

L'installation sur Linux varie en fonction de la distribution utilisée (Debian, Ubuntu, Linux Mint, etc.). Le plus simple est de suivre les explications sur la page dédiée à l'installation sur le site officiel de QGIS (<https://qgis.org>).

4.1.2.2. Appel des modules SAGA GIS

L'appel des modules SAGA GIS dans QGIS se fait très simplement. Tout d'abord, il est nécessaire de s'assurer que la « Boîte à outils de traitements » est bien installée et disponible. Pour cela, allez dans le menu « Extensions » de QGIS, cherchez l'extension nommée « Processing », installez-la si nécessaire et cochez la case correspondante pour rendre cette boîte à outils disponible dans l'interface de QGIS (figure 4.2).

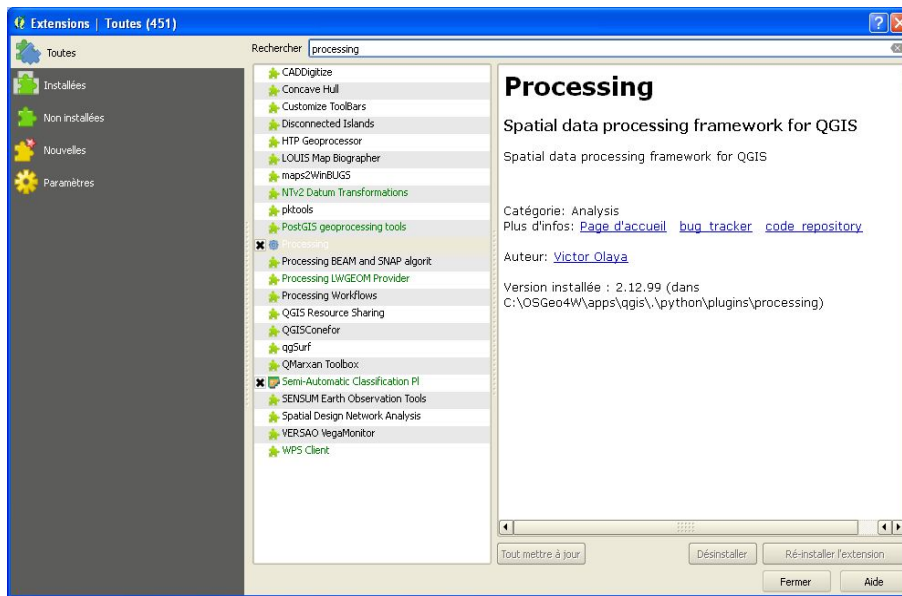


Figure 4.2. Installation de la Boîte à outils de traitements

Pour accéder à cette boîte à outils, cliquez sur le menu « Traitement » et sélectionnez « Boîte à Outils ». Un nouveau panneau apparaît à droite de la fenêtre QGIS, intitulé « Boîte à outils de traitements » (figure 4.3).

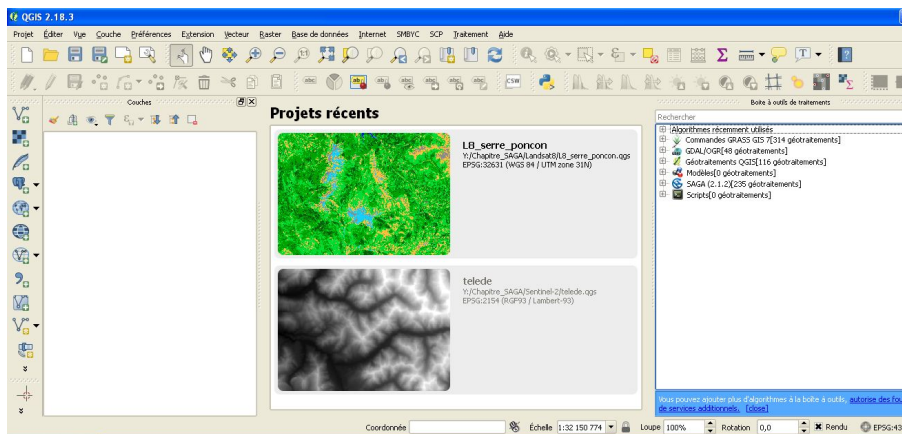


Figure 4.3. Le panneau « Boîte à outils de traitement » sur la droite de la fenêtre de QGIS

Si jamais les modules SAGA n'apparaissent pas, au bas de cette boîte à outils, cliquez sur « autorise des fournisseurs de services additionnels ». Une nouvelle fenêtre « Options de traitements » apparaît. Déroulez le menu « Fournisseurs de traitements » et « SAGA », et cochez la ligne « Activer » (figure 4.4).

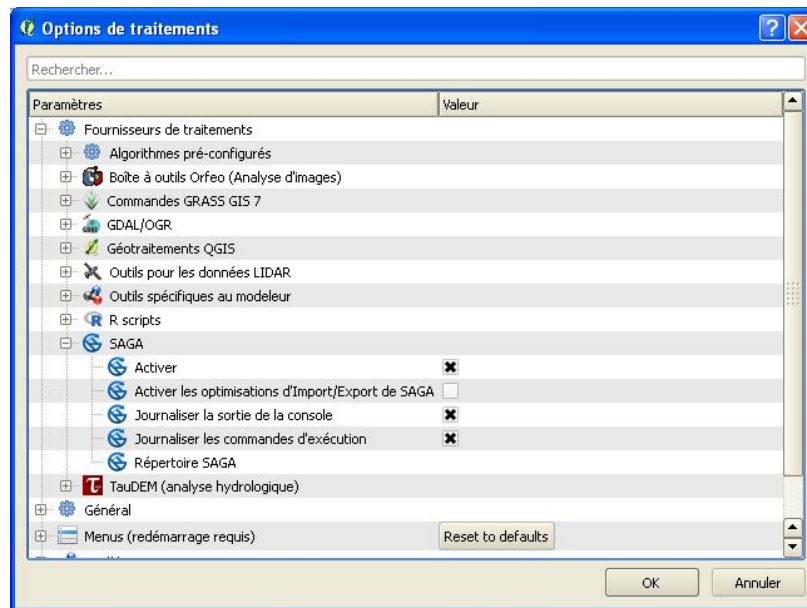


Figure 4.4. Activation des modules SAGA dans les « Options de traitements »

Nous pouvons noter que dans ce panneau « Boîte à outils de traitements », outre les modules SAGA, nous avons accès aux modules GDAL, GRASS GIS ou aux géotraitements QGIS.

Une fois cette installation terminée, les modules SAGA GIS se retrouvent en déroulant la ligne « SAGA » de la Boîte à outils de traitements. 235 géotraitements SAGA sont disponibles. Notons toutefois que la version de SAGA GIS disponible dans QGIS n'est pas la plus récente. Les modules SAGA sont rangés par grande catégorie : « Geostatistics », « Raster tools », « Vector general tools », etc. Lorsque nous déroulons un groupe, nous accédons aux différents modules. Par exemple, en déroulant le groupe « Raster tools », nous trouvons les modules « Agrégation », « Close gaps »... Pour utiliser un module, il suffit de double cliquer sur son nom. Si nous connaissons le nom du module, nous pouvons également y accéder en tapant son nom dans la barre de recherche se trouvant en haut du panneau « Boîte à outils de traitements ». L'ensemble des modules SAGA GIS est décrit à ce lien : www.saga-gis.org/saga_tool_doc/2.1.4/a2z.html. Nous pouvons noter que, malheureusement, la documentation n'est pas le point fort de SAGA GIS.

Ce chapitre est divisé en cinq parties. Dans la première, nous verrons comment utiliser SAGA pour analyser des images satellites multi-spectrales. Dans la deuxième, nous verrons comment extraire un réseau hydrographique à partir d'un modèle numérique de terrain. La troisième partie sera consacrée à l'analyse géomorphologique. La quatrième partie sera une synthèse des trois parties précédentes où nous croiserons différentes couches d'informations spatiales pour créer de nouvelles données. Enfin, la cinquième partie sera consacrée à l'interpolation.

4.2. SAGA GIS pour le traitement d'images satellites multi-spectrales

4.2.1. Méthodologie

Dans cette section, nous allons présenter comment utiliser les modules SAGA GIS et profiter de leur intégration au sein de QGIS pour construire une carte d'occupation du sol à partir d'une image satellite haute résolution Landsat 8. Cet exercice est centré sur les alentours du lac de Serre-Ponçon (figure 4.5) situé sur la Durance, dans les Alpes françaises (44°31'47" Nord, 6° 23'35" Est).



Figure 4.5. Localisation de Serre-Ponçon dans les Alpes françaises

La couche d'occupation du sol de la région sera dérivée à partir de l'image satellite en utilisant la technique de l'arbre de décision. Cette technique permet à l'utilisateur de définir clairement, pour chaque pixel, les différents critères d'appartenance à telle ou telle classe, de manière hiérarchique. Cet arbre de décision sera fondé sur la définition de seuils sur les bandes Landsat brutes, mais aussi sur des indices dérivés de ces bandes comme le SAVI (*Soil-adjusted Vegetation Index*) et l'AWEI (*Automated Water Extraction Index*). Les différentes étapes de l'analyse sont résumées ci-dessous :

- acquisition des données ;
- correction radiométrique des images satellites ;
- découpage des images satellites selon l'étendue de la zone étudiée ;
- construction de compositions colorées et exploration de la zone ;
- calculs des indices SAVI et AWEI ;
- définition de l'arbre de décision ;
- mise en pratique de l'arbre de décision ;
- application d'un filtre majoritaire sur la classification obtenue ;
- simplification de l'occupation du sol par reclassification des pixels.

4.2.2. Acquisition et présentation des données utilisées

Dans cet exemple, nous utilisons une image satellite multi-spectrale datée du 23 août 2016, issue du capteur Landsat 8. Ce capteur est le dernier représentant du programme américain Landsat qui a débuté au début des années 1970. Landsat 8 a été lancé en février 2013 et survole une même portion de la surface terrestre tous les 16 jours en moyenne. C'est un capteur multi-spectral. Il est en effet capable d'enregistrer l'énergie réfléchie et émise par la surface dans 11 bandes spectrales différentes (tableau 4.1). La résolution spatiale des images dépend de la bande spectrale considérée et s'étend de 15 m pour la plus précise et à 100 m pour la plus grossière. Les images Landsat sont mises à disposition du public gratuitement et sont librement téléchargeables sur différents sites. Un des portails de téléchargement les plus utilisés est celui maintenu par l'USGS : EarthExplorer (<https://earthexplorer.usgs.gov/>). Il suffit de s'inscrire gratuitement pour accéder à toutes les archives Landsat ainsi qu'à d'autres produits fournis par la NASA ou même l'ESA (European Space Agency). Notons que sur ce portail il est également possible de télécharger des images Landsat déjà corrigées des effets de l'atmosphère.

Bande spectrale	Longueur d'onde (μm)	Domaine	Résolution (m)
-----------------	-----------------------------------	---------	----------------

1	0,433–0,453	Bleu profond	30
2	0,450–0,515	Bleu	30
3	0,525–0,600	Vert	30
4	0,630–0,680	Rouge	30
5	0,845–0,885	Proche IR	30
6	1,560–1,660	Moyen IR 1	30
7	2,100–2,300	Moyen IR 2	30
8	0,500–0,680	Panchromatique	15
9	1,360–1,390	IR lointain	30
10	10,6–11,2	IR thermique 1	100
11	11,5–12,5	IR thermique 2	100

Tableau 4.1. Caractéristiques des bandes Landsat 8

En plus des images Landsat 8, pour cet exemple, un *shapefile* délimitant la zone d'étude est également fourni. Il contient un seul polygone de forme rectangulaire qui permet de découper les images Landsat afin de ne traiter que notre zone d'intérêt.

4.2.3. Correction des images satellites

Avant de pouvoir correctement traiter et analyser les images satellites, certaines corrections sont nécessaires. La première correction est d'ordre géométrique. Il s'agit de géoréférencer l'image dans un système de projection et de l'orthorectifier en utilisant un modèle numérique de terrain et des points de contrôle. Heureusement, la plupart des images satellites distribuées auprès du grand public ont déjà subi ces corrections. Dans notre cas, nos images sont projetées dans le système UTM zone 31N dont le code EPSG est 32631. Nous conserverons ce système de projection tout au long de notre analyse.

Le deuxième niveau de correction permet de passer des valeurs numériques telles qu'enregistrées par le satellite aux valeurs de réflectance à la surface de l'atmosphère terrestre. Après cette correction, nous obtenons des images dites « Top of Atmosphere (TOA) reflectance » [OSE 16]. Cette correction permet de s'affranchir du « bruit » s'ajoutant aux réflectances sur le chemin allant de l'atmosphère au capteur satellite.

Il existe un troisième niveau de correction, dit de correction atmosphérique. Il s'agit ici d'enlever le bruit dû à l'atmosphère. Ce bruit peut être dû aux aérosols, aux gouttelettes d'eau, à la poussière, etc. Il existe différentes méthodes de correction atmosphériques plus ou moins délicates à mettre en place. Cette correction n'est pas toujours nécessaire, surtout comme ici, où nous étudions une seule date sur une zone assez petite [SON 01]. Il existe en effet un risque de bruite encore plus l'image en appliquant cette correction. Ainsi, dans notre exemple, nous n'appliquerons pas de corrections atmosphériques.

Nous allons maintenant voir comment concrètement mettre en place la seule correction qui nous intéresse ici, à savoir la transformation des données satellites « brutes » en données « Top of Atmosphere reflectance » (TOA). Il est possible de faire cette correction manuellement ou automatiquement. Dans ce chapitre, nous appliquerons cette correction de façon automatique en utilisant un module supplémentaire intégré à QGIS.

Le module que nous allons utiliser s'appelle « Semi-Automatic Classification Plugin ». Ce module est avant tout dédié à la classification supervisée d'images satellites, mais permet également de prétraiter les images. Pour installer ce module,

allez dans « Extensions → Installer/Gérer les extensions ». Recherchez l'extension en tapant son nom dans la barre de recherche, et cliquez sur « Installer » (figure 4.6).

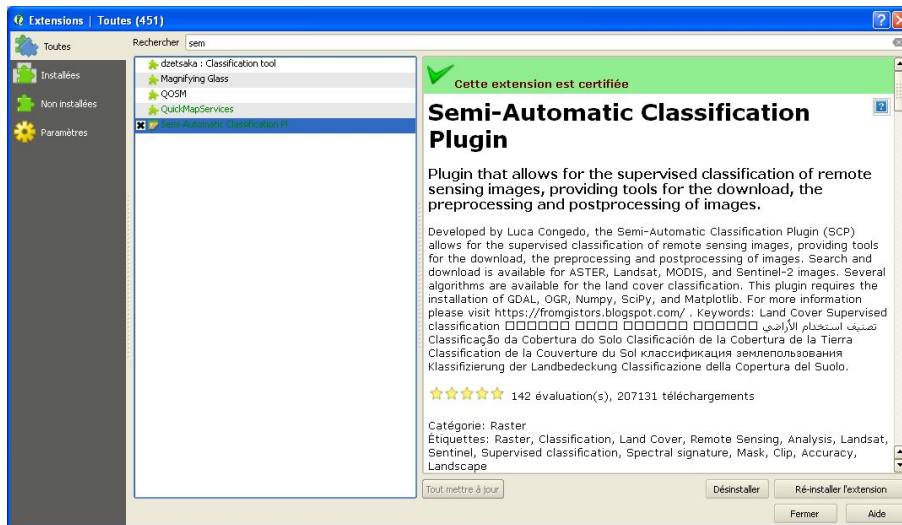


Figure 4.6. Installation du module « Semi-Automatic Classification Plugin : SCP »

Une fois le module installé, un nouveau menu intitulé « SCP » apparaît. Nous allons maintenant corriger nos images Landsat. Pour cela, allez dans le menu « SCP → Preprocessing → Landsat ». Dans la nouvelle fenêtre qui s'affiche, à la ligne « Directory containing Landsat bands », pointez vers le répertoire contenant les bandes, c'est-à-dire les images Landsat. Ce répertoire doit également contenir le fichier de métadonnées (MTL) associées. Ce fichier est indispensable pour appliquer les corrections. Une fois le chemin renseigné, le module détecte par lui-même (grâce au fichier de métadonnées) les informations associées aux images (capteur, date, bandes spectrales... Voir figure 4.7).

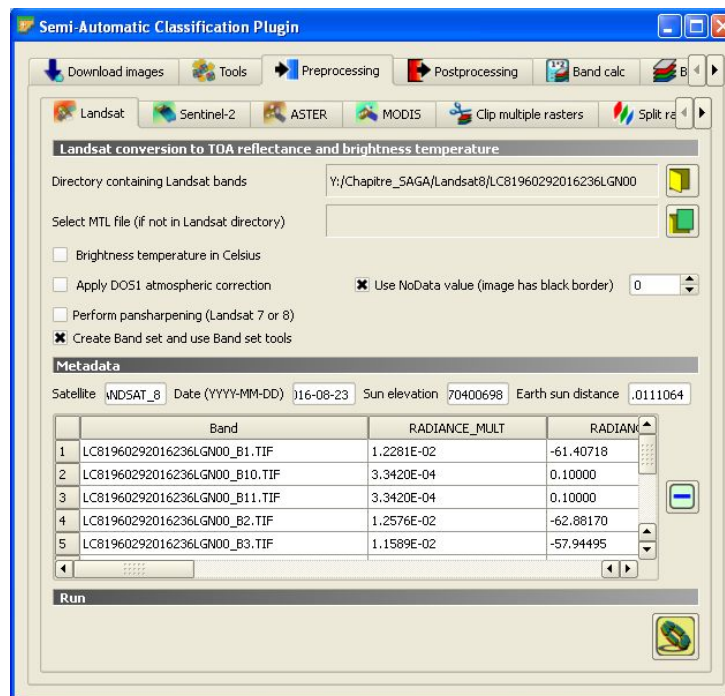


Figure 4.7. Initialisation du module SCP

Bien que nous ne l'utiliserons pas ici, à cette étape, il est possible de convertir les bandes de l'infrarouge thermique en degrés Celsius, d'appliquer une correction atmosphérique via la méthode DOS1, et d'appliquer un rééchantillonnage des pixels des différentes bandes en fonction de la bande possédant la meilleure résolution (*pan-sharpening*). Nous pouvons

laisser les options par défaut à ce niveau. Pour lancer le processus de correction, cliquez sur le bouton jaune et vert en bas à droite de la fenêtre. Il est maintenant nécessaire de spécifier le répertoire dans lequel nous allons stocker les images corrigées. Appelons-le « L8_2016-08-23_TOA » par exemple. Cette correction peut prendre quelques minutes.

Une fois la correction réalisée, les images corrigées apparaissent dans QGIS. Il est possible de vérifier la valeur de certains pixels à l'aide de l'outil « Identifier les entités ». Sur les images corrigées, les valeurs des pixels sont comprises entre 0 et 1. Nous sommes donc bien en présence de valeurs de réflectance. Pour rappel, un pixel avec une valeur de 1 dans une bande spectrale donnée signifie que ce pixel réfléchit la totalité du rayonnement incident dans cette bande spectrale. Ce genre de pixel apparaîtra en blanc. À l'inverse, un pixel présentant une valeur de 0 signifie que ce pixel ne réfléchit rien dans cette bande spectrale (ce pixel apparaîtra en noir).

4.2.4. Découpage des images satellites selon la zone étudiée

Nous allons maintenant découper nos images selon le *shapefile* fourni (« zone_etude_32631.shp ») afin de nous consacrer sur notre zone d'étude. Il faut tout d'abord charger cette couche dans QGIS. Celle-ci est déjà projetée dans le même système de projection que celui de nos images Landsat (UTM 31N – EPSG 32631). Réorganisez vos couches afin d'avoir la zone d'étude tout en haut et les bandes Landsat de la première à la dernière (bande 1 en haut et bande 11 en bas), comme indiqué sur la figure 4.8.

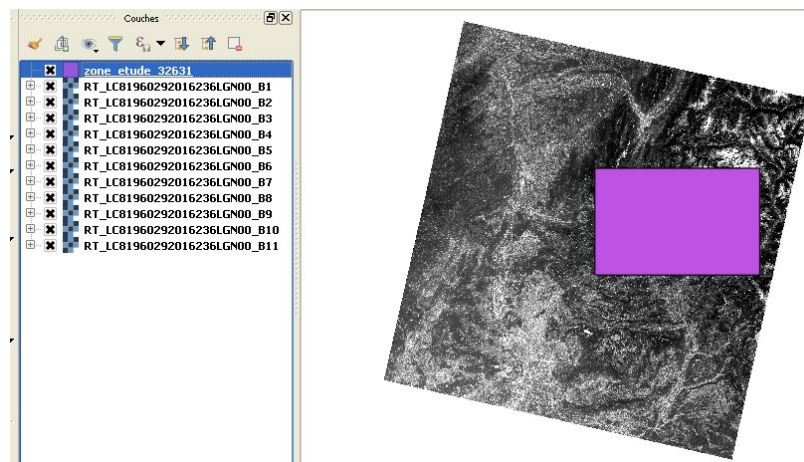


Figure 4.8. Préparation des bandes pour le découpage selon la zone d'étude

Le découpage de nos images raster va se faire en utilisant le module SAGA intitulé « Clip raster with polygones ». Ce module se trouve dans la « Boîte à outils de traitements », dans le menu « SAGA » et « Vector <-> Raster ». Un autre moyen pratique est de le chercher *via* la barre de recherche de cette même Boîte à outils de traitements. Une nouvelle fenêtre « Clip raster with polygon » apparaît. Comme nous allons découper plusieurs images en fonction du même *shapefile*, nous allons utiliser l'outil « Exécuter comme processus de lot ». Dans la nouvelle fenêtre, nous allons renseigner toutes les images à découper et le fichier de découpe à utiliser. Commençons par cliquer sur l'icône « ... » de la colonne « Input », et choisissons « Sélectionner depuis les couches ouvertes ». Dans la fenêtre « Sélection multiple » qui s'affiche, nous sélectionnons les images que nous voulons découper. Dans notre cas, nous souhaitons découper les sept premières bandes. Nous ne travaillerons pas avec les autres, nous ne les découpons donc pas. Dans la colonne « Polygones », selon le même principe, nous sélectionnons le *shapefile* que nous souhaitons utiliser pour découper. Comme nous utilisons le même polygone pour toutes les découpes, nous renseignons le même fichier à chaque ligne. Enfin, dans la colonne « Clipped », nous devons renseigner le chemin où nous souhaitons stocker les images découpées. Nous les enregistrons dans le répertoire que nous avons créé précédemment (L8_2016-08-23_TOA), et nous les nommons « BX_L8_2016-08-23_TOA_sp », en remplaçant le X par le numéro de la bande considérée. « Sp » signifie « Serre-Ponçon ». Entre chaque renseignement de ligne de la colonne « Clipped », sélectionnez « Ne pas remplir automatiquement ». Nous obtenons une fenêtre similaire à la figure 4.9. Nous pouvons charger directement les images ainsi créées dans QGIS. Cliquez sur « Run » pour lancer le processus de découpage par lot. Une fois le processus terminé, une fenêtre vous l'indique.

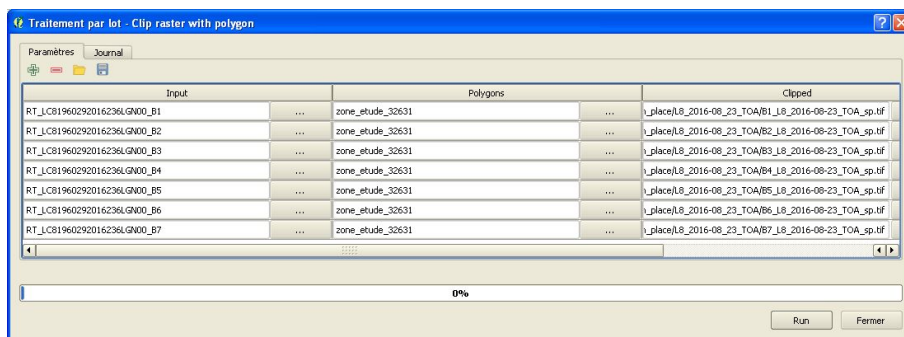


Figure 4.9. Découper des images raster par lot via le module SAGA « Clip raster with polygon »

Les couches ainsi créées apparaissent dans QGIS. Afin de mieux nous organiser, nous fermons toutes les couches, et nous ne réimportons que les couches découpées, que nous trions par ordre croissant de bande (bande 1 sur le dessus). Nous disposons maintenant des sept premières bandes spectrales de notre image Landsat, exprimées en réflectance à la surface de l'atmosphère et découpées selon notre zone d'intérêt (figure 4.10).

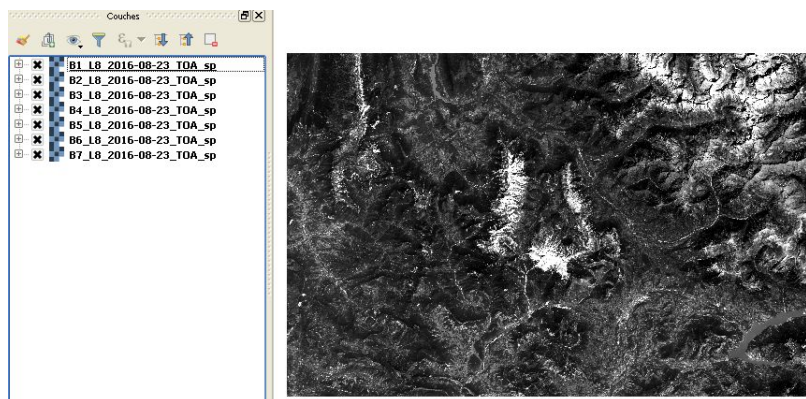


Figure 4.10. Les images Landsat découpées selon notre zone d'étude (44°37' Nord, 5° 57" Est)

4.2.5. Construction de compositions colorées et exploration de la région

Afin de construire des compositions colorées, nous allons regrouper nos bandes spectrales Landsat dans un raster virtuel à l'aide d'un outil GDAL. Pour cela, nous allons dans le menu « Raster → Divers → Construire un Raster Virtuel (Catalogue VRT) ». Dans la nouvelle fenêtre qui s'affiche, nous cochons la case « Utiliser les rasters visible comme entrée ». Nous spécifions un chemin où stocker ce raster virtuel à la ligne « Fichier en sortie » et nous l'appelons « L8_2016-08-23_TOA_stack ». Il est nécessaire de cocher la case « Séparer » (figure 4.11). Cela nous permettra de gérer ce raster virtuel bande par bande.

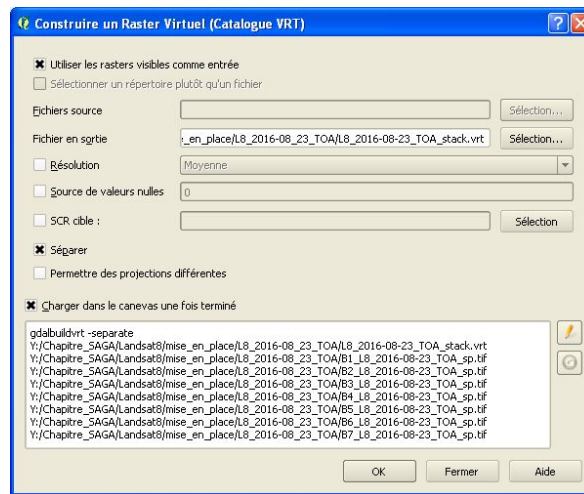


Figure 4.11. Construction du raster virtuel à l'aide de GDAL. La syntaxe GDAL utilisée est visible (et éditable) dans le panneau du bas

Une fois le processus terminé, le raster virtuel apparaît directement dans QGIS. Il est possible d'ouvrir ce raster virtuel avec un éditeur de texte. Il s'agit en fait simplement d'un fichier XML pointant vers les rasters d'origine. Dans notre cas, nous nous servons de cette propriété pour créer des compositions colorées. Nous allons construire une composition colorée en fausses couleurs afin de faire ressortir les zones végétalisées. Nous allons donc colorer la bande verte en bleu, la rouge en vert et la proche infrarouge (dans laquelle la végétation réfléchit le plus) en rouge. Faites un clic droit sur votre raster virtuel et allez dans « Propriétés » et « Style ». Pour le type de rendu, nous sélectionnons « Couleur à bandes multiples ». Pour la bande rouge, nous choisissons la bande 5 (le proche infrarouge), pour la bande verte nous prenons la bande 4 (le rouge) et pour la bande bleue, nous sélectionnons la bande 3 (le vert). Nous améliorons les contrastes en choisissant l'option « Étirer jusqu'au Min/Max », et nous cliquons sur « Charger » dans le panneau « Charger les valeurs Min/Max » afin de récupérer les valeurs extrêmes de chaque bande (figure 4.12).

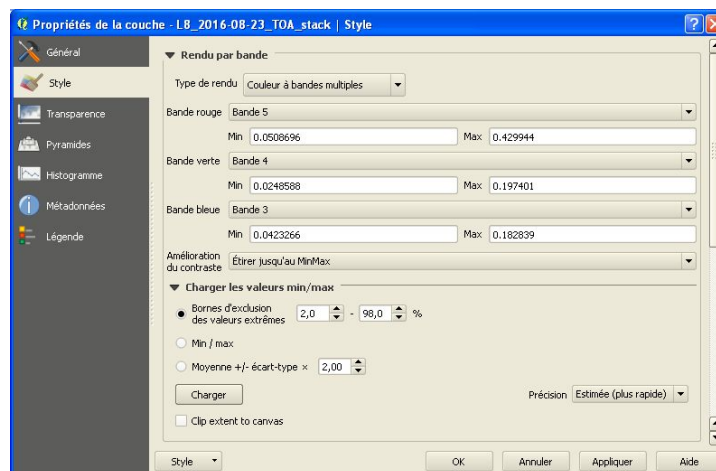


Figure 4.12. Construction d'une composition colorée à partir d'un raster virtuel

Nous obtenons la composition colorée suivante (figure 4.13). Avec cette composition, les zones végétalisées apparaissent en rouge plus ou moins vif selon l'état de la végétation. En effet, une végétation en « bonne santé » réfléchit fortement dans l'infrarouge. Sur cette composition, l'infrarouge étant codé en rouge, une végétation apparaîtra d'autant plus rouge qu'elle est en bonne santé.



Figure 4.13. Composition colorée en fausses couleurs avec le proche infrarouge coloré en rouge (44°37' Nord, 5° 57" Est)

Cette région est montagneuse et les zones sombres sur le flanc nord-ouest des reliefs correspondent à des zones d'ombre. Les réflectances enregistrées dans les différentes bandes sont affectées par ce phénomène. Ainsi, les classifications d'occupation du sol sur ce type de terrain sont plus compliquées que celles effectuées sur terrain plat. Cependant, nous verrons que dès que nous calculons des indices combinant des rapports entre bandes spectrales, ces artefacts diminuent grandement.

4.2.6. Calculs d'indices afin d'extraire la végétation et les surfaces en eau

Dans cette sous-section, nous allons calculer deux indices. Nous commencerons par calculer le SAVI (*Soil-adjusted Vegetation Index*) qui nous permettra d'extraire différentes classes de végétation. Le second sera l'AWEI (*Automated Water Extraction Index*), permettant d'extraire les surfaces en eau.

4.2.6.1. Extraction de la végétation avec le SAVI

Un des indices de végétation les plus connus et les plus utilisés est le NDVI (*Normalized Difference Vegetation Index*). Néanmoins, de nombreux autres indices de végétation existent, dont le SAVI. Cet indice a été développé en ajoutant un facteur d'ajustement prenant en compte la réflectance du sol plus ou moins visible à travers la canopée. La formule de cet indice est la suivante :

$$SAVI = \frac{(1+L)(PIR-R)}{PIR+R+L} \quad [4.1]$$

Avec PIR, la bande du proche infrarouge, R, la bande du rouge, et L, le facteur d'ajustement du sol. Une valeur de 0,5 pour le facteur L est considérée comme optimale dans la plupart des cas [JEN 13].

Le calcul du SAVI peut se faire directement au sein de SAGA via le module « Vegetation index (slope based) », qui se trouve dans le sous-menu « Image analysis » des commandes SAGA. Lorsque nous exécutons ce module, une nouvelle fenêtre « Vegetation index (slope based) » apparaît. En premier lieu, il s'agit de renseigner les bandes spectrales correspondant au proche infrarouge et au rouge. Dans notre cas (Landsat 8), il s'agit des bandes 5 et 4 respectivement (figure 4.14). Ensuite, nous devons sélectionner les indices que nous souhaitons calculer et préciser où nous les stockons. Ici, nous ne calculerons que le SAVI, le dernier indice proposé, et nous le stockons dans notre répertoire « L8_2016-08-23_TOA » et le nommons « savi_L8_2016-08-23 ».

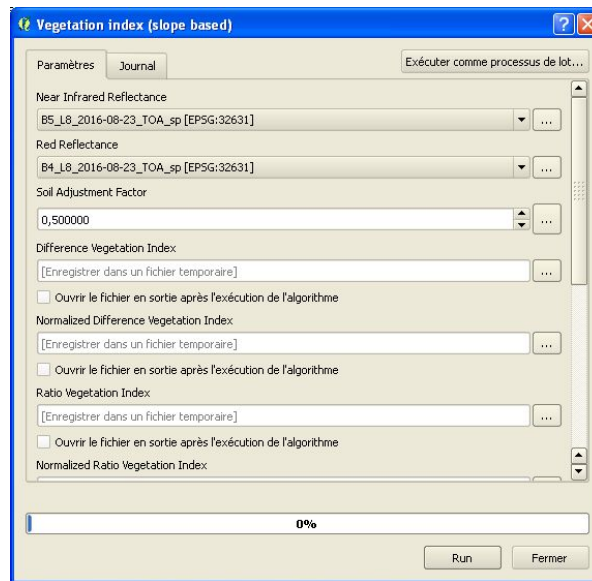


Figure 4.14. Calcul du SAVI via le module SAGA « Vegetation index (slope based) »

Une nouvelle couche correspondant au SAVI apparaît alors dans QGIS. Il est possible de changer son apparence en allant dans les propriétés de cette couche et dans l'onglet « Style ». Dans le « Type de rendu », nous sélectionnons « Pseudo-couleur à bande unique », et pour la « Couleur », nous choisissons la palette « Greens » (figure 4.15).

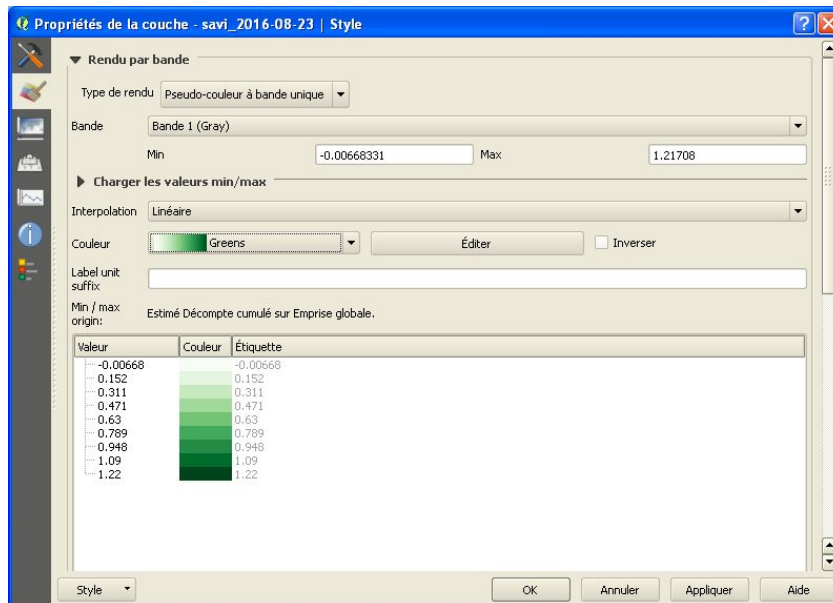


Figure 4.15. Modification du style du raster de SAVI

Il suffit ensuite de cliquer sur « Classer » et « Appliquer », et nous obtenons une représentation du SAVI en teintes de verts (figure 4.16).

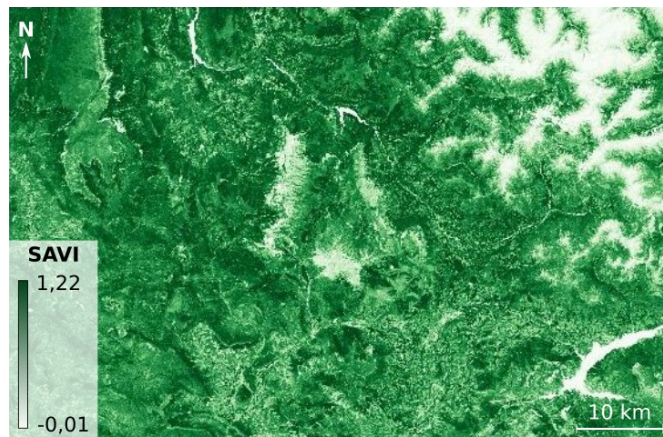


Figure 4.16. Calcul du SAVI via le module SAGA « Vegetation index (slope based) »

Sur cette image (figure 4.16), les pixels apparaissant en vert profond sont les pixels dont les valeurs sont proches de 1 (valeurs maximales), et les pixels tirant sur le blanc sont les pixels avec des valeurs de SAVI très faibles (inférieures à 0). Or, les pixels à fort SAVI sont les pixels densément végétalisés alors que les pixels à faible SAVI sont les pixels dépourvus de toute couverture végétale. Finalement, les différentes teintes de vert, donc les différentes valeurs de SAVI s'étalant entre 0 et 1, traduisent des densités de végétation différentes.

4.2.6.2. Extraction des surfaces en eau avec le AWEI

Il existe de nombreux indices permettant d'extraire les surfaces en eau à partir d'une image satellite multi-spectrale. Un indice robuste et pratique est l'AWEI [FEY 14]. Il s'agit d'un indice empirique fondé sur les bandes du domaine du bleu, de vert, du proche infrarouge et du moyen infrarouge. L'AWEI a l'avantage de rendre la discrimination entre les surfaces en eau et les surfaces terrestres très facile. En effet, du fait de sa construction, les pixels en eau apparaissent avec une valeur positive et les pixels terrestres avec une valeur négative. Cet indice est décliné en deux variantes. La première est à utiliser sur les images dépourvues de zones d'ombres et la seconde sur les images où des ombres sont présentes. Ainsi, dans notre cas, nous utiliserons la seconde variante de l'indice, dont la formule est donnée ci-dessous [FEY 14] :

$$AWEI_{sh} = B + 2,5 \times V - 1,5 \times (PIR + MIR1) - 0,25 \times MIR2 \quad [4.2]$$

AWEI_{sh} est l'indice AWEI pour les images présentant des zones d'ombres (*shadow areas*). B correspond au canal du domaine du bleu (B2 de Landsat 8), V à celui du Vert (B3), PIR à celui du proche infrarouge (B5) et MIR1 et MIR2 à ceux des deux canaux du moyen infrarouge (B6 et B7).

Contrairement à l'indice SAVI, il n'est pas possible de calculer directement l'AWEI au sein de SAGA GIS. Il est nécessaire d'utiliser la calculatrice raster de QGIS, qui se trouve dans le menu « Raster → Calculatrice Raster ». La formule de l'AWEI doit être entrée manuellement en prenant les bandes adéquates en entrée. Nous sauvegardons la couche résultante dans notre répertoire de travail et la nommons « awei_L8_2016-08-23 » (figure 4.17).

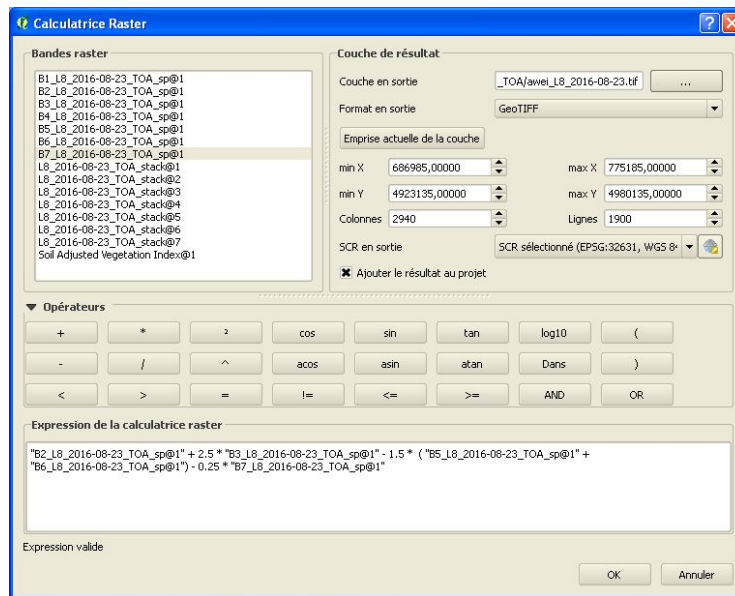


Figure 4.17. Calcul de l'indice AWEI via la calculatrice raster de QGIS

Le résultat apparaîtra directement dans QGIS une fois le calcul terminé. En suivant le même principe que dans le cas du SAVI, nous pouvons changer la coloration de cette couche. Cette fois-ci, nous ne prenons pas la palette « greens », mais la palette « blues ». Nous obtenons la figure 4.18.



Figure 4.18. Indice AWEI calculé pour la région de Serre-Ponçon à partir de l'image Landsat 8 du 23 août 2016

Sur cette image, les pixels en bleu profond sont les pixels présentant une valeur d'AWEI positive. Ce sont donc les pixels en eau. Les pixels tirant vers le blanc sont les pixels avec une valeur négative. Ce sont donc les pixels terrestres. Néanmoins, du fait de la topographie de notre zone d'étude, les zones les plus à l'ombre ainsi que les zones enneigées ont tendance à apparaître comme les surfaces en eau. Nous atteignons là une des limites de cet indice, que nous surmonterons en partie dans la section suivante dédiée à la classification de l'occupation du sol en appliquant un arbre de décision.

4.2.6.3. Classification de l'occupation du sol

4.2.6.3.1. Définition de l'arbre de décision

Nous allons maintenant classifier l'occupation du sol de notre zone d'intérêt en utilisant notre image Landsat 8, les indices dérivés et la technique de l'arbre de décision. Comme dit précédemment, cette dernière permet à l'utilisateur de guider exactement le processus de classification en définissant lui-même, pour chaque pixel, les critères d'appartenance à chaque classe d'occupation du sol. Cette méthode peut demander du temps et de l'expertise thématique de la part de l'utilisateur pour la mettre en place. L'avantage est qu'une fois les critères mis en place, ils peuvent être appliqués à d'autres dates pour la même région, en faisant l'hypothèse d'une très bonne calibration de ces images.

Ainsi, nous souhaitons décrire l'occupation du sol de la région de Serre-Ponçon selon 8 classes d'occupation du sol : neige sèche, neige fondue, surfaces en eau, sol nu/artificialisé, végétation rase, prairie/culture, formation arbustive et forêt dense.

Commençons par caractériser la neige sèche à partir des images dont nous disposons. La première image sur laquelle la neige sèche se dégage bien est celle correspondant à l'AWEI. Sur cette couche, nous observons que les régions enneigées présentent un AWEI supérieur à 0,12. Néanmoins, les surfaces en eau présentent également des valeurs supérieures à 0,12. Il faut donc trouver un critère permettant de distinguer l'eau libre de la neige sèche. Après observation de la réponse de ces deux classes sur les autres images d'indices, nous nous apercevons que le SAVI est utilisable pour différencier ces deux classes (neige sèche et surfaces en eau). En effet, les zones recouvertes de neige sèche présentent en général un SAVI inférieur à -0,25 alors que les surfaces en eau présentent un SAVI supérieur à cette valeur. Nous avons donc nos critères pour différencier nos deux classes. Si un pixel a une valeur d'AWEI supérieur à 0,12 et une valeur de SAVI inférieur à -0,25, alors ce pixel est recouvert de neige sèche. En revanche, si un autre pixel a lui aussi une AWEI supérieure à 0,12, mais un SAVI supérieur à -0,25, alors ce pixel correspond à de l'eau libre. Nous remarquons que normalement, le seuil de l'AWEI pour détecter de l'eau est 0, alors qu'ici nous avons choisi un seuil de 0,12. Ce seuil nous permet d'éliminer quelques pixels terrestres se trouvant dans des zones d'ombre ou présentant une surface particulièrement sombre.

En utilisant une approche similaire pour toutes les classes désirées, nous obtenons l'arbre de décision présenté en figure 4.19. Par exemple, un pixel est couvert de végétation rase si sa valeur d'AWEI est supérieure à 0,12, sa réflectance dans la première bande (bleu profond) est supérieure à 0,7 et son SAVI compris entre 0,3 et 0,7.

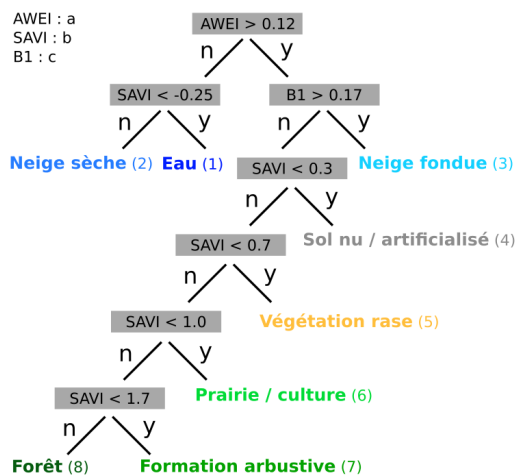


Figure 4.19. Arbre de décision pour la classification de l'occupation du sol de la région de Serre-Ponçon en 8 classes

4.2.6.3.2. Mise en pratique de l'arbre de décision

Une fois l'arbre de décision défini, nous allons concrètement l'appliquer *via* le module « Raster calculator » de SAGA GIS. Nous pouvons localiser ce module en le cherchant simplement dans la barre de recherche de la « Boîte à outils de traitements ». Dans ce module, il faut tout d'abord spécifier les couches que nous allons utiliser. Dans la rubrique « Main input layer », nous sélectionnons la première couche à utiliser. Par exemple, nous sélectionnons « awei_L8_2016-08-23 » qui correspond à l'indice AWEI précédemment calculé. Cette couche sera appelée par la lettre « a » dans la formule que

nous construirons dans un second temps. Ensuite, dans le sous-menu « Additional layers », accessible en cliquant sur l'icône « ... » en bout de ligne, nous spécifierons les couches additionnelles que nous allons utiliser. Ici, il s'agit du SAVI (« savi_L8_2016-08-23_TOA_sp ») et de la première bande de notre image Landsat (« B1_L8_2016-08-23 »). Nous ferons référence à ces deux couches respectivement par « b » et « c », conformément à leur ordre de sélection.

Le point délicat est la traduction de notre arbre de décision en une syntaxe intelligible par SAGA GIS. Nous allons utiliser l'instruction de test logique « ifelse ». La syntaxe est la suivante :

```
ifelse(test à effectuer, valeur si vraie, valeur si faux)
```

Par exemple, si nous prenons le tout premier niveau de notre arbre de décision portant sur l'AWEI, nous aurons quelque chose de cette forme :

```
ifelse(l'AWEI est-il supérieur à 0.12, si oui alors je teste si le SAVI est inférieur à -0.25, si non alors je teste si la bande 1 est supérieure à 0.17)
```

Par un jeu d'emboîtement de « ifelse », dans lequel il faut faire attention aux différentes parenthèses ouvrantes et fermantes, nous obtenons l'arbre de décision suivant, que nous inscrivons dans la rubrique « Formula » :

```
ifelse(a>0.12,ifelse(b<(-0.25),1,2),ifelse(c>0.17,3,ifelse(b<0.3,4,ifelse(b<0.7,5,ifelse(b<1,6,ifelse(b<1.17,7,8))))))
```

Attention, lorsque nous testons une valeur négative, cette valeur doit apparaître entre parenthèses. Ainsi, nous testons « b<(-0.25) » et non « b<-0.25 ». La fenêtre du module « raster calculator » ressemble finalement à la figure 4.20.

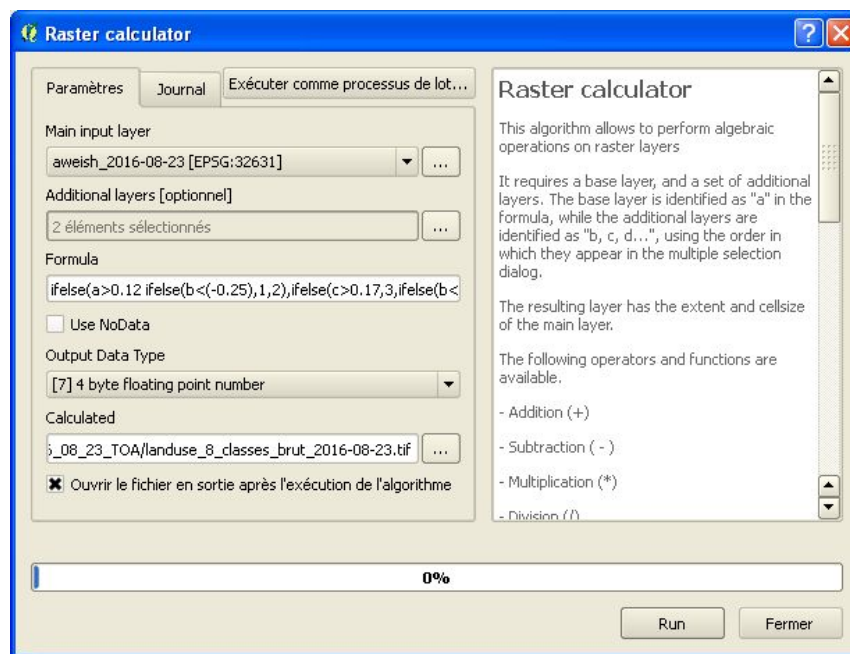


Figure 4.20. Application de l'arbre de décision dans le module « Raster calculator » de SAGA GIS

Nous spécifions le nom et le lieu de stockage de la classification résultante en cliquant sur l'icône « ... » de la rubrique « Calculated ». Nous stockons le tout dans notre répertoire de travail et le nommons « usol_L8_2016-08-23 ». La classification apparaît directement dans QGIS une fois le calcul terminé. Il est possible de changer son nom en allant dans les « Propriétés » de la couche et dans l'onglet « Général ». Nous avons maintenant une occupation du sol de la région de Serre-Ponçon en 8 classes (tableau 4.2).


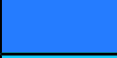






Valeur du pixel	Occupation du sol	Code hexa	Teinte
1	Surface en eau	#001dff	
2	Neige sèche	#257cff	
3	Neige fondue	#13d0ff	
4	Sol nu/artificiel	#8c8c8c	
5	Végétation rase	#ffbe3b	
6	Prairie/culture	#00db36	
7	Formation arbustive	#02a200	
8	Forêt	#055c0f	

Tableau 4.2. Les 8 classes d'occupation du sol de la région de Serre-Ponçon

Il est possible de changer le style de cette classification en allant dans l'onglet « Style » des « Propriétés » de la couche. Le style peut être défini manuellement en choisissant un « Type de rendu » par « Pseudo couleur à bande unique » ou alors il peut être chargé à partir d'un fichier fourni. Dans notre cas, vous pouvez charger le fichier de style « L8_usol_8_classes.qml » en cliquant sur « Style → Changer le style... ». Nous obtenons la classification suivante (figure 4.21).

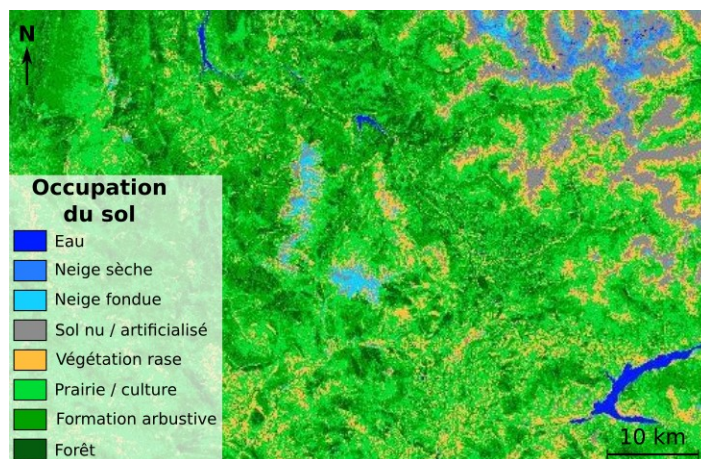


Figure 4.21. Occupation du sol de la région de Serre-Ponçon en huit classes, dérivé de l'image Landsat 8 du 23 août 2016

4.2.6.3.3. Simplification de l'occupation du sol via un filtre majoritaire

La méthode employée (arbre de décision) pour créer l'occupation du sol sur notre site d'étude est une méthode dite « pixellaire ». En effet, chaque pixel de l'image est classé dans une catégorie d'occupation du sol en fonction de ces caractéristiques dans certaines couches. Un désavantage de ce genre de méthode est de produire une classification avec un effet « poivre et sel », où de nombreux pixels isolés sont mal étiquetés et ont tendance à bruyter le résultat. Il est difficile de faire la part entre bruit et résultat cohérent, surtout avec des images à haute résolution spatiale, mais un pixel isolé au milieu d'un groupe homogène de pixels peut souvent être considéré comme du bruit. Une méthode classique pour s'affranchir de cet artefact est d'appliquer un filtre majoritaire.

Un filtre majoritaire est en fait une fenêtre glissante, dont la taille et la forme sont définies par l'utilisateur. L'algorithme cherche quelle est la valeur de pixel majoritaire au sein de cette fenêtre et remplace la valeur du pixel central par cette valeur si elle est initialement différente. SAGA GIS propose un module permettant d'appliquer ce genre de filtre facilement. Nous allons utiliser le module nommé « Majority filter ». Dans la fenêtre qui apparaît, il faut tout d'abord spécifier la couche que nous souhaitons filtrer. Dans notre cas, nous sélectionnons notre occupation du sol « usol_L8_2016-08-23 ». Ensuite, dans la rubrique « Search mode », nous spécifions la forme de la fenêtre (carrée ou circulaire) que nous voulons utiliser. Ce choix est à la discrétion de l'utilisateur. Ici, nous prenons une fenêtre carrée, c'est-à-dire « Square ». Ensuite, le « Radius » correspond à la taille de la fenêtre. Nous pouvons définir une taille de 5 par exemple. La fenêtre glissante aura alors une taille de 5 pixels par 5 pixels. Enfin, dans la partie « Filtered grid », nous définissons le chemin où nous souhaitons sauvegarder notre résultat et son nom. Nommons-le « usol_L8_2016-08-23_filtre5 » et stockons-le dans notre espace de travail (figure 4.22).

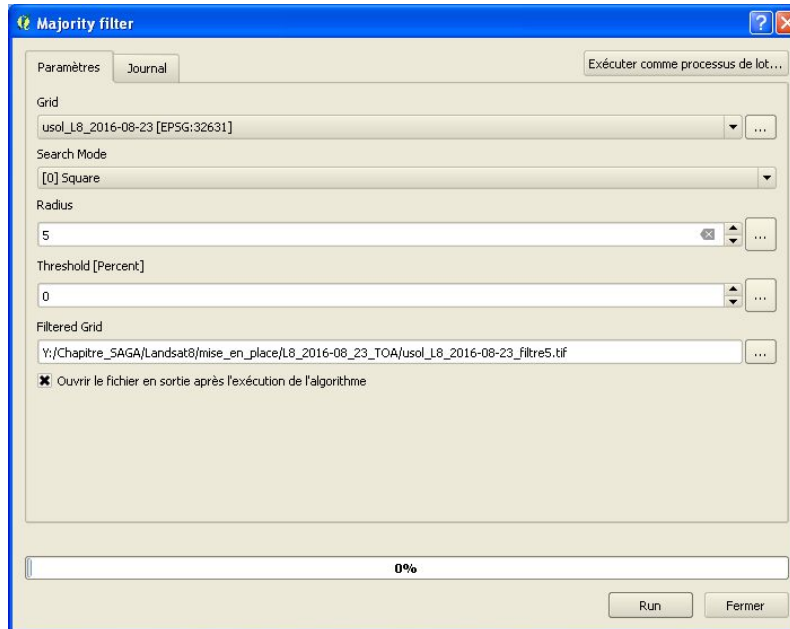


Figure 4.22. Application d'un filtre majoritaire à l'aide du module « Majority filter » de SAGA GIS

Une fois le traitement réalisé, il est possible de changer le nom de la couche produite et son style par un simple copier-coller à partir du style de la classification précédente. Nous obtenons une classification beaucoup moins bruitée et aux contours plus adoucis (figure 4.23). Attention de ne pas trop simplifier les classifications, il faut savoir trouver le juste milieu entre simplification et cohérence de la classification.

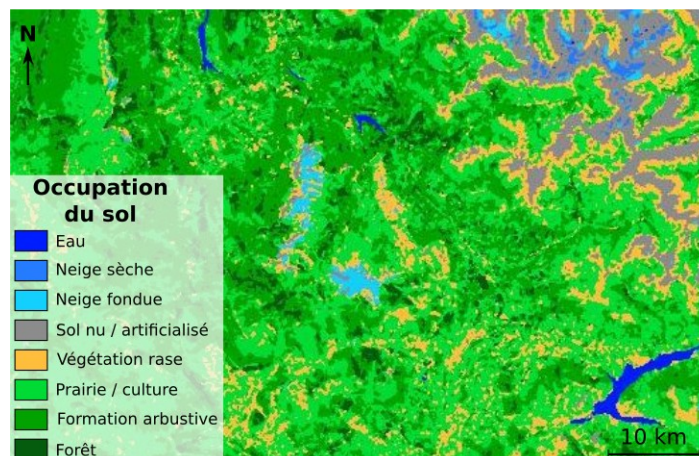


Figure 4.23. Occupation du sol après application d'un filtre majoritaire à partir d'une fenêtre carrée de taille 5

A ce stade, si nous disposons d'une référence terrain, il serait nécessaire de construire une matrice de confusion afin de quantifier la validité de notre classification. Un exemple de matrice de confusion est donné dans le chapitre consacré à la détection des coupes rases [OSE 16].

4.2.6.3.4. Réduction du nombre des classes d'occupation du sol

Dans certaines applications, il peut s'avérer utile et nécessaire de simplifier l'occupation du sol en grandes classes. Par exemple, ici, nous allons simplifier notre occupation du sol en 8 classes en une occupation du sol en 4 classes. Les classes que nous désirons sont les suivantes : eau, neige, végétation, sol nu. Nous pouvons résumer la simplification que nous désirons par le tableau 4.3.

Classe originale	Valeur originale	Classe finale	Valeur finale
Surface en eau	1	Surface en eau	1
Neige sèche	2	Neige	2
Neige fondue	3	Neige	2
Sol nu/artificiel	4	Sol nu	3
Végétation rase	5	Végétation	4
Prairie/culture	6	Végétation	4
Formation arbustive	7	Végétation	4
Forêt	8	Végétation	4

Tableau 4.3. Reclassification de l'occupation du sol à 8 classes en une occupation du sol à 4 classes

Ce processus de reclassification va se faire à l'aide du module SAGA GIS « Reclassify values (simple) ». À l'ouverture de ce module, il faut tout d'abord spécifier le raster que nous désirons reclassifier. Ici, nous choisissons notre classification préalablement filtrée « usol_L8_2016-08-23_filtre5 ». À la ligne « Lookup Table », nous allons spécifier quelles valeurs nous voulons changer et quelles nouvelles valeurs nous voulons attribuer. Nous cliquons sur l'icône « ... » de la ligne « Lookup Table » (figure 4.24).

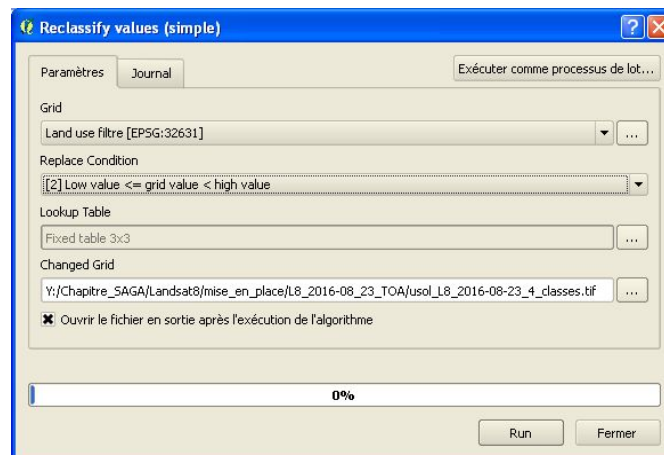


Figure 4.24. Reclassification de l'occupation du sol

Nous remplissons le tableau de la nouvelle fenêtre en s'inspirant du tableau 4.3 (figure 4.25).

	minimum	maximum	new
1	1	1	1
2	2	3	2
3	4	4	3
4	5	8	4

Figure 4.25. Table de correspondance de reclassification de l'occupation du sol en 8 classes vers une occupation du sol en 4 classes

En cliquant sur « OK », nous revenons à la fenêtre précédente. Il faut maintenant choisir l'opérateur « Operator » de reclassification. Par construction de notre tableau, nous souhaitons reclasser à la fois les valeurs minimums et les valeurs maximums. Par exemple, la nouvelle classe 4 « végétation » contiendra les anciens pixels 5, 6, 7 et 8 ; 5 et 8 étant compris dans les valeurs à reclasser. Par conséquent, nous choisissons l'opérateur « Low value <= grid value <= high value », qui est le choix numéro 2. Enfin, nous pouvons décocher les options « reclass no data values », « replace other values » et spécifier l'endroit où nous stockons le résultat. Appelons cette nouvelle couche « usol_L8_2016-08-23_4_classes » (figure 4.24). Nous obtenons la couche présentée en figure 4.26.

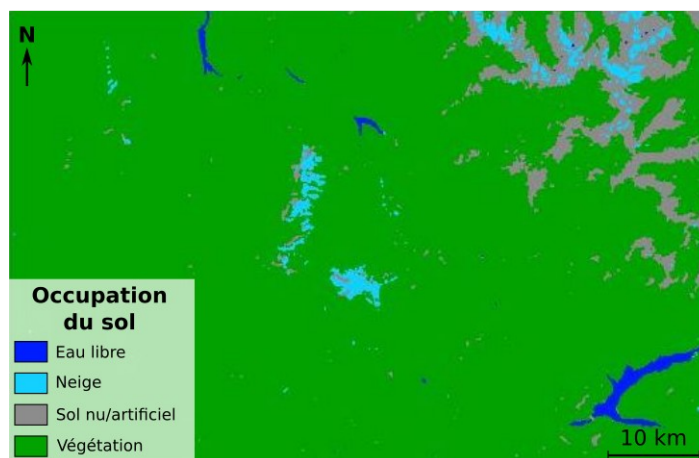


Figure 4.26. Occupation du sol reclassée en 4 classes principales. Vert : végétation, cyan : neige, gris : sol nu, bleu : eau libre

4.3. Extraction d'un réseau hydrographique avec SAGA GIS dans QGIS

4.3.1. Intérêts

Afin d'étudier le fonctionnement hydrologique, biogéochimique ou autre d'une rivière et de son bassin versant, il est souvent nécessaire d'extraire le réseau hydrographique associé. Sur ce réseau, il sera alors possible de calculer des indices morphologiques (longueur, densité de drainage, indice de méandrage, etc.). Ces réseaux extraits sont une représentation simplifiée de l'hydrosystème sous forme vectorielle ou raster. Sous forme vectorielle, les rivières sont représentées sous forme de polygones orientés et les bassins versants sont représentés par des polygones.

Cette extraction de réseau hydrographique se fait la plupart du temps à partir de modèles numériques de terrain (MNT). Certains sont disponibles à l'échelle mondiale (GTOPO30, SRTM) et d'autres aux échelles nationales (BD Alti IGN en France) ou encore plus locales. Ces MNT sont caractérisés par une résolution spatiale donnée (dimension de chaque pixel) et par une précision altimétrique, qui déterminent la qualité et la précision de la représentation que l'on fera de l'hydro-système.

À travers un exemple centré sur la même région de Serre-Ponçon, nous allons voir comment utiliser QGIS et son module SAGA pour générer de tels jeux de données. SAGA implémente les méthodes classiques, basées sur la définition d'un seuil unique de surface à partir duquel on considère que débute un écoulement de surface de l'eau [OCA 84]. Il implémente également les méthodes basées sur le *stream burning* [HUT 89]. Ici, nous utiliserons une méthode propre à SAGA, basée sur l'ordination de Strahler [STR 57].

4.3.2. Préparation du MNT – mosaïquage et découpage

Nous pouvons télécharger un extrait des données SRTM (MNT quasi mondial –de la latitude 56°S à 60° N – créé par la NASA) [FAR 07] depuis le portail EarthExplorer déjà évoqué. Une fois identifié, il est possible de télécharger au format geoTIFF les 2 dalles SRTM englobant notre région d'étude. Nous choisissons la dernière version du SRTM disponible avec une résolution d'une seconde d'arc géoréférencée en WGS84, équivalent à une résolution d'environ 30 m de côté à l'équateur [JAR 08]. Dans l'interface « EarthExplorer », nous choisissons ainsi le produit « SRTM 1 Arc-Second Global » se trouvant dans le panneau « Data Sets → Digital Elevation → SRTM ». Les dalles correspondant à la zone que nous souhaitons sont les suivantes : n44_e005_1arc_v3.tif et n44_e006_1arc_v3.tif.

Afin d'optimiser nos traitements, nous allons mosaïquer ces dalles SRTM et les reprojeter en UTM zone 31 N (EPSG 32631) comme dans la partie dédiée à l'analyse de l'occupation du sol. Cette reprojection est nécessaire, car SAGA GIS gère mieux les données projetées dans un système métrique. C'est d'ailleurs plus facile pour l'utilisateur de raisonner en mètres carrés qu'en degrés carrés.

Nous allons ici mosaïquer nos dalles SRTM afin de ne manipuler qu'un seul fichier plutôt que deux fichiers. Nous utilisons l'outil SAGA « Mosaic raster layers » du menu « Raster tools ». Dans la fenêtre qui apparaît, nous renseignons les deux dalles à assembler à la ligne « Input grids », nous laissons les autres éléments par défaut et nous sauvons notre mosaïque sous le nom « mos_tmp.tif » (figure 4.27). Notre résultat apparaît à l'écran.

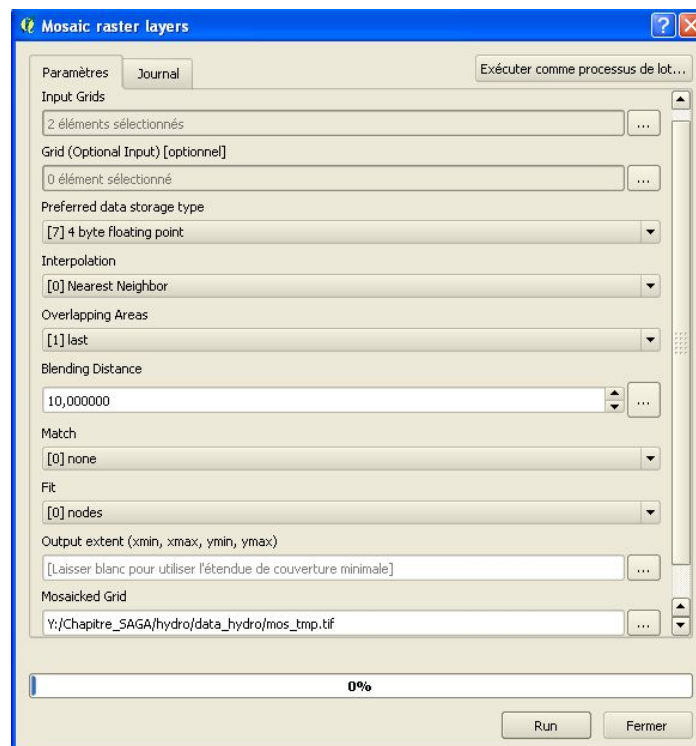


Figure 4.27. Création de la mosaïque des deux dalles SRTM de notre zone d'étude

Il est maintenant nécessaire de projeter cette mosaïque en UTM Zone 31 N afin de travailler dans une projection à coordonnées métriques et pertinente. Pour cela, nous cliquons droit sur la couche, nous sélectionnons « Enregistrer sous... ». A la ligne « Enregistrer sous », nous spécifions l'endroit où nous stockons la nouvelle couche et nous la nommons « mos_tmp2.tif », et enfin, nous spécifions que nous voulons reprojeter cette couche en UTM Zone 31 N à la ligne « SCR ». Nous sélectionnons cette projection dont le code EPSG est 32631, comme indiqué sur la figure 4.28.

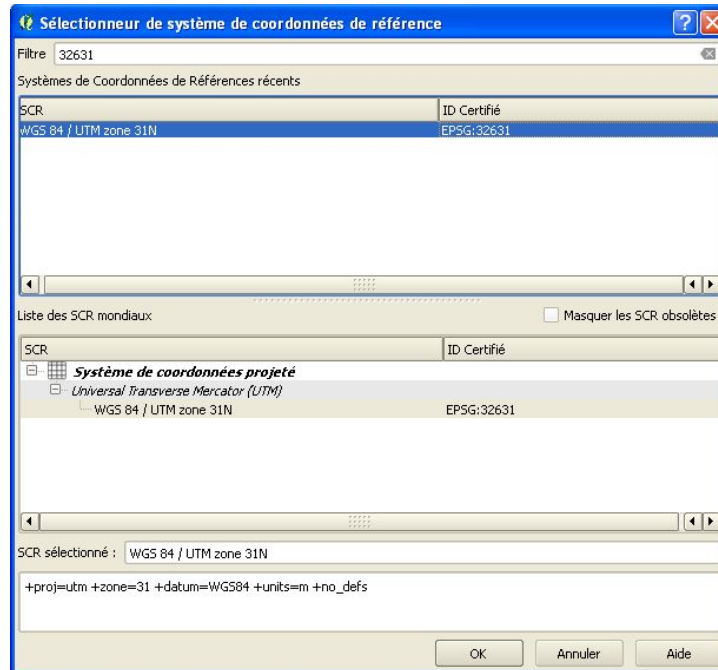


Figure 4.28. Reprojection en UTM zone 31 N de la mosaïque de dalles SRTM

Finalement, afin d'optimiser les temps de calcul, nous allons découper notre mosaïque selon l'emprise de notre zone d'étude telle que définie par le fichier « zone_etude_32631.shp ». Nous réglons tout d'abord la projection du projet en UTM 31N. Pour cela, nous cliquons sur l'icône en bas à droite de l'écran QGIS contenant le mot « EPSG ». Dans la nouvelle fenêtre qui apparaît, dans l'onglet « SCR », nous configurons le projet en UTM 31N via son code EPSG (figure 4.29).

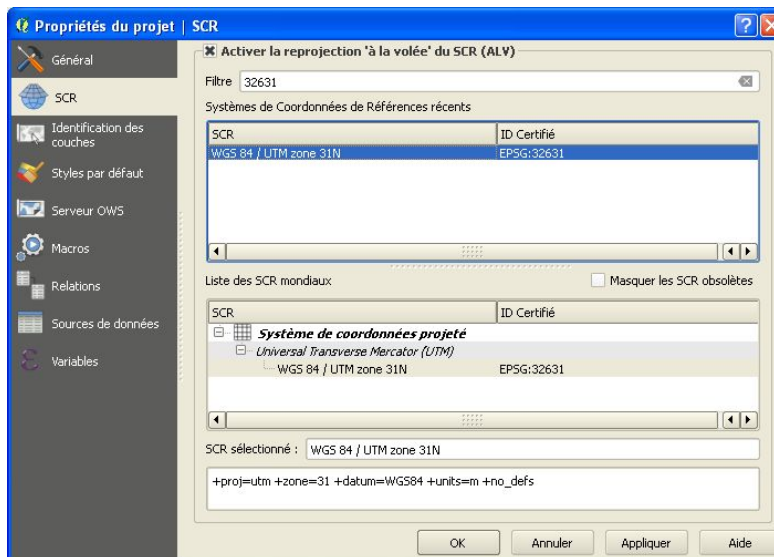


Figure 4.29. Configuration du projet en UTM zone 31N

Nous chargeons la couche « zone_etude_32631.shp » dans QGIS, puis à l'aide du module « Découper... » se trouvant dans le menu Raster → Extraction, nous allons extraire le MNT pour notre zone d'étude. Dans la fenêtre qui s'affiche, pour le fichier source, nous sélectionnons notre mosaïque « mos_tmp2.tif », nous nommons notre fichier résultat « srtm_durance_utm31n.tif », nous configurons les valeurs nulles à « -99999 » (valeurs nulles par défaut de SAGA), nous fixons le mode de découpage à « couche de masquage » à l'aide de la couche « srtm_sp_utm31n.shp » et nous cochons l'option « Découper l'emprise du jeu de données ciblé selon l'emprise du trait de coupe » (figure 4.30).

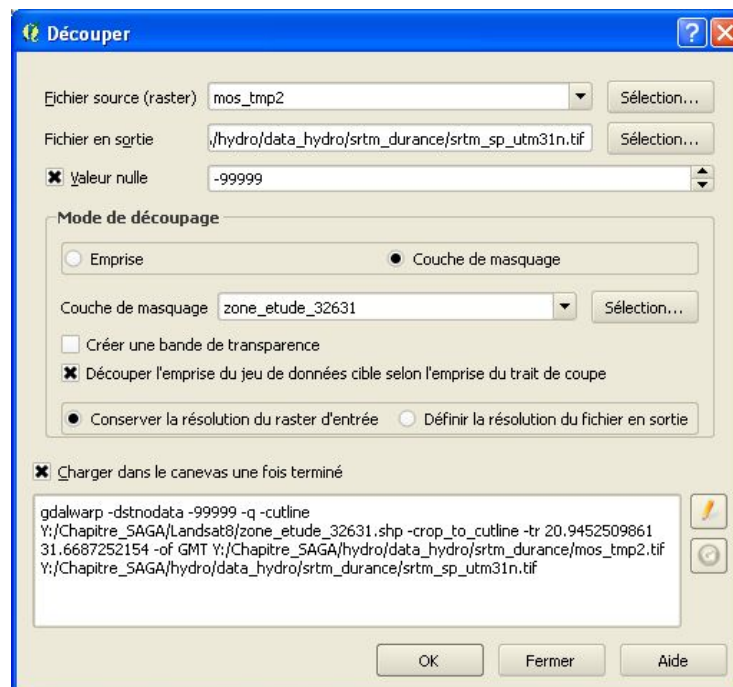


Figure 4.30. Découpage du SRTM en fonction de notre zone d'étude

Nous obtenons notre SRTM dimensionné à notre zone d'étude. Nous pouvons enlever tous les rasters intermédiaires et ne garder que le fichier final pour la suite.

4.3.3. Remplissage des puits

L'utilisation d'un MNT à des fins hydrologiques implique un premier traitement de comblement des puits qui peuvent y figurer, et qui empêcheraient l'écoulement théorique d'une rivière. Ceci survient typiquement lorsqu'une cellule d'un raster, qui n'est pas l'exutoire d'une rivière, est entourée d'autres cellules dont l'altitude est supérieure à la sienne, formant ainsi une sorte de cuvette. Il peut s'agir d'une réalité du terrain (avens et dolines en zones karstiques par exemple), mais aussi d'artefacts dus à la résolution du MNT ou à ses conditions d'acquisition. La création d'un réseau d'arcs orientés symbolisant un réseau hydrographique nécessite donc de combler ces cellules afin d'éviter des discontinuités dans le réseau. L'opération est nommée « Comblement des puits » (*fill sinks* en anglais) et consiste à interpoler les valeurs du bord du puits vers l'intérieur du puits. De plus, dans le module que nous utilisons par la suite, la pente générale est également reconstituée sur les valeurs interpolées.

Il existe plusieurs méthodes pour procéder à cette opération. SAGA implémente ainsi la méthode de [PLA 01] qui les comble en simulant ce qui resterait d'un passage d'une lame d'eau sur le modèle numérique de terrain. Pour l'exercice, nous utiliserons la méthode de [WAN 06], qui offre l'avantage de conserver un gradient de pente dans les cellules comblées. Celle-ci se trouve dans le menu « Terrain Analysis → Hydrology → Fill sinks [WAN 06] ». Dans la fenêtre qui s'affiche, nous spécifions d'abord le MNT (DEM en anglais) que nous souhaitons combler, à savoir « *srtm_sp_utm31n.tif* ». Nous laissons la valeur de « Minimum slope » par défaut. Nous spécifions où nous stockons le raster comblé (*Filled DEM*) que nous nommons « *srtm_sp_filled_utm31n.tif* ». Nous ne nous servons pas des rasters « Flow Directions » et « Watershed basins », nous pouvons par conséquent décocher leurs options « Ouvrir le fichier après exécution » (figure 4.31).

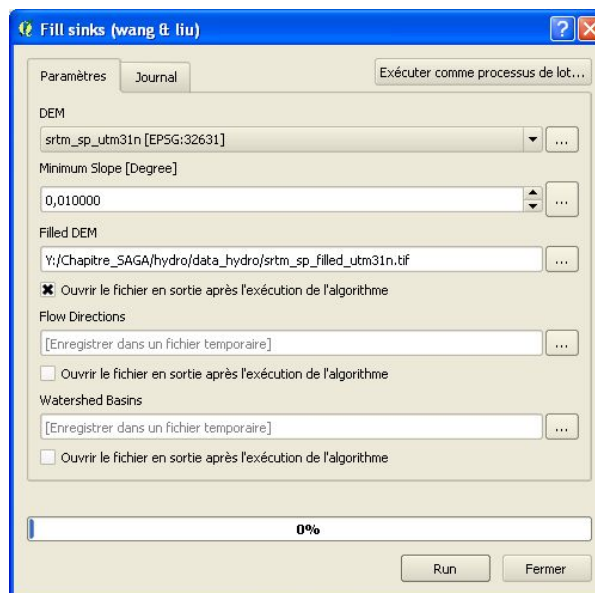


Figure 4.31. Remplissage des puits et définition du raster de direction des flux

À la fin du traitement, nous obtenons un nouveau MNT très semblable au premier, seules les valeurs des puits ont changé.

4.3.4. Extraction du réseau hydrographique

En plus des méthodes classiques basées sur un seuil de drainage, SAGA propose une méthode originale d'extraction du réseau hydrographique fondée sur l'ordination de Strahler de chacun des pixels du MNT. La logique est exactement la même que pour l'ordination des rivières. Les pixels qui n'ont aucun pixel amont sont des pixels d'ordre 1. Lorsqu'un pixel est en aval de deux pixels d'ordre 1, il devient un pixel d'ordre 2. Si ce pixel d'ordre 2 reçoit un pixel d'ordre 1, le pixel reste d'ordre 2. Il ne passera à l'ordre 3 que lorsqu'il recevra un autre pixel d'ordre 2, et ainsi de suite.

Avec cette méthode, nous devons configurer un seuil sur l'ordre de Strahler au-dessus duquel nous estimons qu'une rivière peut se former de façon pérenne. Si le terrain est très imperméable, ce seuil sera faible. Au contraire, lorsque le terrain est perméable, ce seuil sera élevé. Nous allons appliquer cette méthode sur notre MNT comblé à l'aide du module « Channel network and drainage basins » dans le menu « Terrain Analysis - Channels ». La fenêtre suivante présentée en figure 4.32 apparaît.

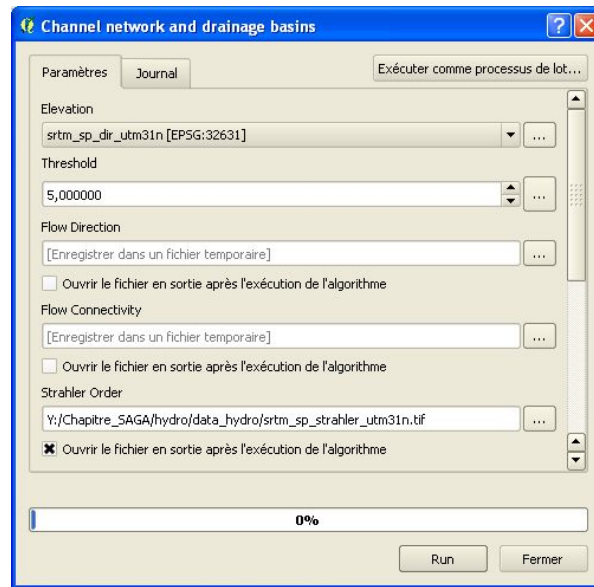


Figure 4.32. Extraction du réseau hydrographique par la méthode de l'ordination de Strahler

Nous précisons d'abord sur quel MNT nous souhaitons travailler. Dans notre cas, nous sélectionnons le MNT comblé « srtm_sp_fill_utm31n.tif ». Le seuil (*threshold*) est par défaut égal à 5. Ce seuil signifie que tout pixel présentant un ordre de Strahler supérieur à 5 sera vu comme une rivière pérenne. Nous pouvons laisser ce seuil dans un premier temps. Nous ne sauons que le raster « Strahler Order » que nous nommons « srtm_sp_strahler_utm31n.tif » et le réseau hydrographique « Channels » sous forme de polygones que nous nommons « channels5_sp_utm31n.shp ». Le résultat apparaît en figure 4.33.

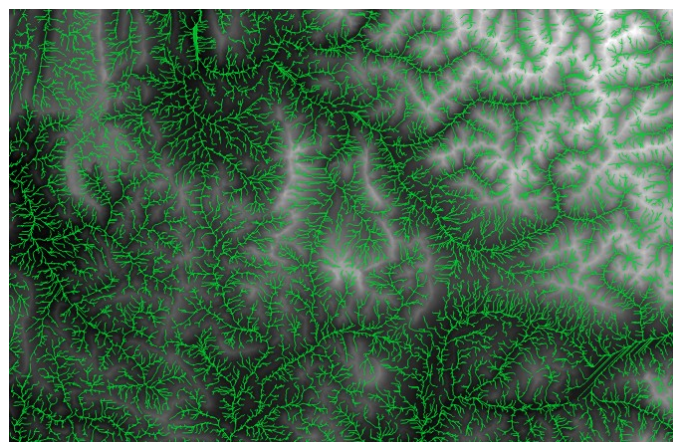


Figure 4.33. Le réseau hydrographique extrait pour la région de Serre-Ponçon avec un seuil de 5 sur l'ordre de Strahler. L'image de fond est le MNT (SRTM)

Afin de bien comprendre la logique utilisée par SAGA et la signification du seuil de Strahler, il est possible de zoomer sur un début de rivière, une source en quelque sorte (figure 4.34).



Figure 4.34. Zoom sur deux débuts de rivières (en rouge). Le raster de fond est le raster d'ordination de Strahler pixel par pixel « srtm_sp_strahler_utm31n.tif » (résolution du pixel : 30 mx30 m)

Sur la figure 4.34, nous repérons les deux débuts de rivière en rouge. Le raster sous-jacent est le raster d'ordination de Strahler pixel par pixel. Ainsi, chaque pixel a une valeur correspondant à son ordre de Strahler. Les pixels les plus noirs ont un ordre de -3 et les pixels les plus clairs, où les rivières débutent, ont un ordre de 1. Par définition, l'ordination de Strahler ne peut pas être négative, elle commence à 1. En fait, ici il faut se fier au rang et non à la valeur. Si nous reclassons les pixels afin de mettre les -3 à 1, nous obtenons le tableau x.4.

Strahler du raster	-3	-2	-1	0	1
Strahler « réel »	1	2	3	4	5

Tableau 4.4. Correspondance entre les ordres de Strahler du raster et les « vrais » ordres de Strahler

Les résultats montrent que les rivières débutent bien à un ordre de Strahler « réel » de 5. Nous sommes donc bien en accordance avec le seuil précédemment défini. Grâce à ce décalage, les premières rivières ont bien un ordre de Strahler, hydrologiquement pertinent, de 1.

Afin de sentir la sensibilité au seuil de Strahler, nous pouvons recommencer l'extraction du réseau hydrographique, mais en choisissant un seuil de 8. Nous nommons le résultat « channels8_sp_utm31n.shp ». Nous obtenons le réseau suivant présentée en figure 4.35.

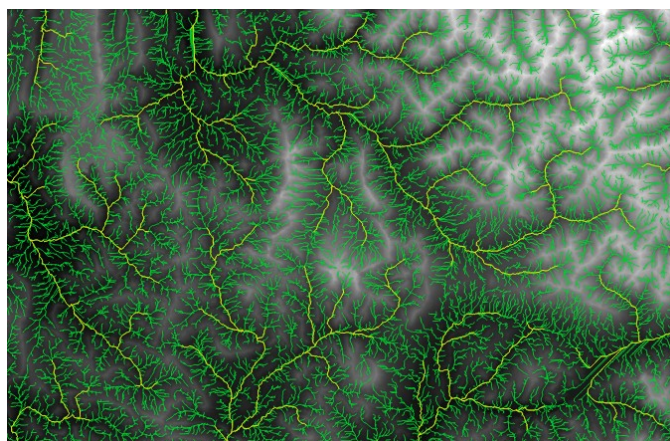


Figure 4.35. Comparaison des réseaux extraits avec un seuil de 5 (vert) et 8 (jaune)

Sans surprise, le réseau extrait avec un seuil de 8 est beaucoup moins dense que celui extrait avec un seuil de 5. La définition de ce seuil dépend de la région étudiée. Il peut être déterminé par comparaison du résultat avec un réseau hydrographique dont l'utilisateur est sûr (carte topographique, image satellite, BD Carthage IGN [IGN 14], etc.).

Plusieurs méthodes d'extraction de réseaux hydrographiques existent. Il serait possible de les comparer, à l'instar de [SCH 17] qui, de plus, propose une nouvelle approche non plus basée uniquement sur des données d'altitude, mais aussi sur des données de climatologie et de lithologie.

4.4. Interpolation avec SAGA GIS

4.4.1. Principe de l'interpolation

L'interpolation spatiale consiste à reconstruire les valeurs d'une variable spatialisée sur un domaine à partir d'échantillons connus en un nombre limité de points. De très nombreuses méthodes, plus ou moins complexes, existent. Dans l'exemple que nous utiliserons, nous verrons une méthode simple, facile à mettre en place et couramment utilisée. Concrètement, lorsque nous réalisons une interpolation spatiale, nous passons d'un fichier de points (par exemple, stocké en *shapefile*) à un raster. Nous ne discuterons pas ici des différences entre les différentes méthodes, ni de leurs avantages et de leurs inconvénients¹. L'interpolation peut en effet être un sujet de géostatistiques en elle-même.

4.4.2. Interpolation de mesures de qualités des eaux souterraines dans le bassin de la Seine

4.4.2.1. Présentation des données

Pour cet exemple, nous disposons de mesures de qualité des eaux souterraines dans le bassin de la Seine (figure 4.36), dans la partie nord de la France (48°N, 2°E).

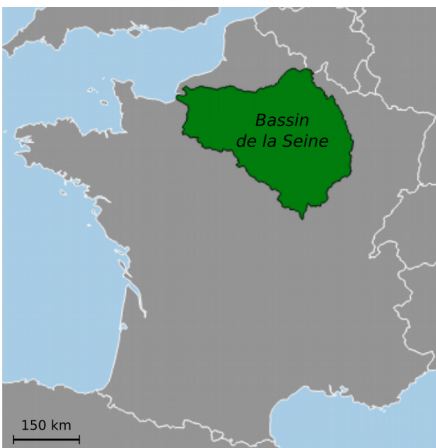


Figure 4.36. Localisation du bassin de la Seine dans la partie nord de la France

Ces données sont stockées dans le *shapefile* « *eaux_souterraines_oct_2006.shp* » et sont fournies par l'Agence de l'eau Seine Normandie. Il s'agit de données de bicarbonate (HCO_3 en mgC/l), de pH et de température ($^{\circ}\text{C}$). Toutes ces données ont été mesurées en octobre 2006. Nous allons maintenant voir comment passer de données ponctuelles en données continues. Concrètement, nous allons transformer notre information qui est actuellement stockée sous forme d'un fichier de points en un fichier raster dans lequel chaque pixel aura une valeur de HCO_3 , de pH ou de température. Dans ce cas, l'interpolation peut être utile dans le cas où l'on veut avoir une idée de la qualité de l'eau souterraine en un point précis qui ne serait pas un point de mesure.

¹ cg.ensmp.fr/bibliotheque/public/WACKERNAGEL_Publication_02091.pdf.

4.4.2.2. Préparation des données

Nous allons interpoler séparément les données relatives au bicarbonate, au pH et à la température. Nous souhaitons également pouvoir quantifier la qualité de notre interpolation, ce qui nous permettra de choisir la meilleure méthode d'interpolation. Notre échantillon présente 239 points. Nous allons le diviser en deux parties égales. La première nous servira à interpoler, et la seconde nous servira à qualifier notre interpolation.

Afin de sélectionner aléatoirement 50 % de nos points, nous utilisons le module « Sélection aléatoire » qui se trouve dans la « Boîte à outils de traitements → Géotraitements QGIS → Outils de sélection de vecteur → Sélection aléatoire ». Dans la fenêtre qui s'affiche alors (figure 4.37), nous sélectionnons notre couche d'entrée (« eaux_souterraines_oct_2006.shp ») et pour la méthode, nous choisissons « pourcentage d'entités sélectionnées », que nous fixons à 50 %.

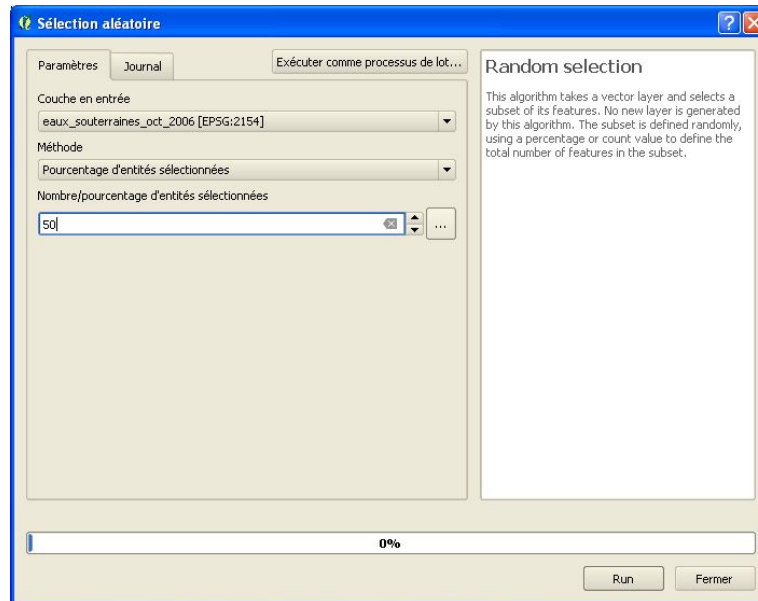


Figure 4.37. Sélection aléatoire de 50 % des points de mesures

Nous sauvons cette sélection qui nous servira pour la validation en faisant un clic droit sur la couche et en sélectionnant « Enregistrer sous... ». Il ne faut pas oublier de cocher la case « N'enregistrer que les entités sélectionnées ». Dans le champ « file name », nous spécifions l'endroit où nous sauvons ce fichier. Nous le nommons « eaux_souterraines_oct_2006_validation.shp ». La nouvelle couche de points doit normalement s'afficher.

Nous allons maintenant sauver le symétrique de ces points, qui nous servira à l'interpolation. Pour cela, nous ouvrons la table attributaire de notre couche d'origine (la sélection précédente doit toujours être active) par un clic droit « Ouvrir la table d'attributs ». Les lignes correspondant aux points sélectionnés apparaissent surlignées. Dans le menu du haut, nous cliquons sur « Inverser la sélection », et nous obtenons une sélection symétrique de la précédente. Comme pour la première moitié, nous sauvons ces points dans un nouveau fichier que nous nommons « eaux_souterraines_oct_2006_pour_interpolation.shp » (n'oubliez pas de cocher la case « N'enregistrer que les entités sélectionnées »). Nous avons maintenant deux jeux de données distincts et complémentaires.

4.4.2.3. Mise en place de l'interpolation

Nous allons maintenant interpoler nos mesures selon le champ « HCO3 » à l'aide de modules SAGA GIS. Il existe de nombreuses fonctions d'interpolation, nous allons en tester trois différentes qui sont couramment utilisées : « B-spline approximation », « Inverse Distance Weighted Interpolation » (IDW) et « Thin plate » (global).

Commençons avec la première technique. Nous allons dans le menu SAGA GIS, « Raster creation tools » et nous sélectionnons « B-spline approximation ». Dans la fenêtre qui apparaît, nous sélectionnons la couche de points que nous souhaitons interpoler (« eaux_souterraines_oct_2006_pour_interpolation.shp »), le champ sur lequel nous voulons interpoler (« HCO3 ») et changeons la résolution à 500. Plus notre zone d'étude est vaste, plus ce paramètre doit être élevé. À la ligne « Grid », nous indiquons où nous souhaitons sauver notre interpolation. Nommons-la « HCO3_interpol_B-spline.tif » (figure 4.38).

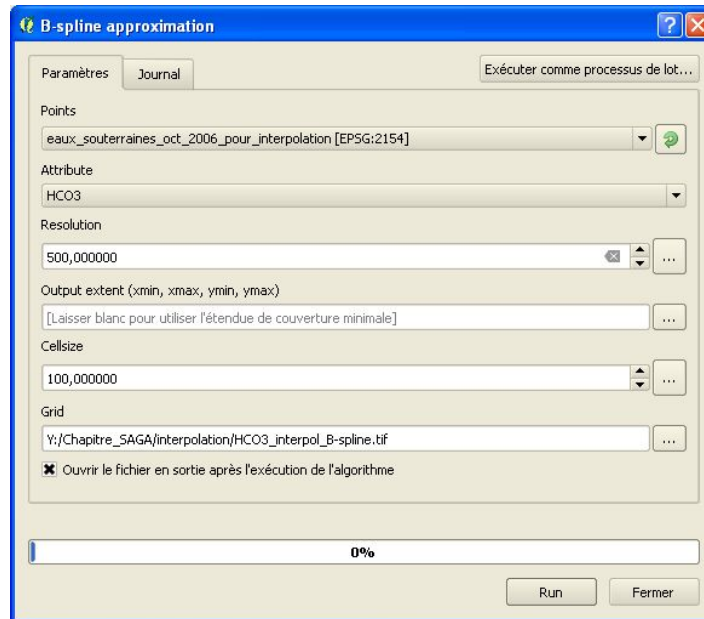


Figure 4.38. Paramétrage de l'interpolation par la méthode « B-spline »

Le résultat présenté en figure 4.39 s'affiche dans la fenêtre principale, après réglage du style dans les propriétés de la couche.

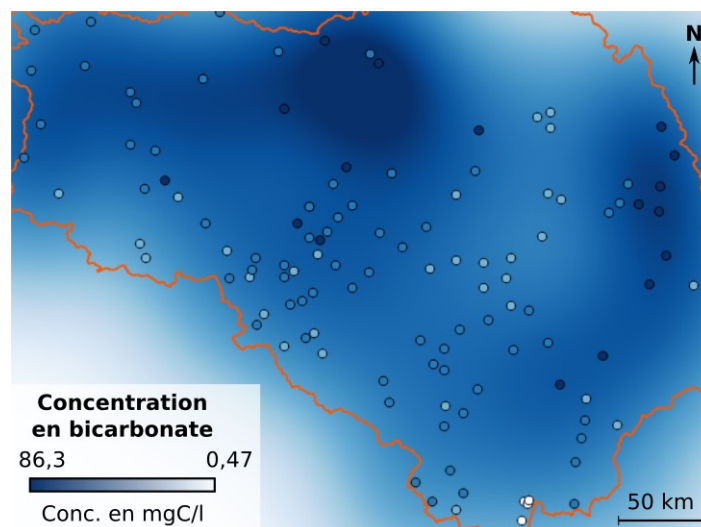


Figure 4.39. Interpolation des bicarbonates sur le bassin de la Seine (les limites du bassin sont en orange) par interpolation en utilisant la méthode « B-spline interpolation »

Nous réalisons maintenant la même interpolation, mais avec la méthode « Inverse Distance Weighted Interpolation ». Le module correspondant porte le même nom. Dans la fenêtre qui apparaît, nous sélectionnons notre couche d'entrée « eaux_souterraines_oct_2006_pour_interpolation.shp » et notre attribut à interpoler « HCO3 ». Nous laissons la méthode par défaut « Inverse distance to a power » et ses paramètres par défaut (2 et 1). Nous changeons le « search range » par « no search radius (global) », car notre zone d'étude est assez vaste. Nous pouvons laisser les autres options par défaut et nous nommons notre résultat « HCO3_interpol_idw.tif » (figure 4.40).

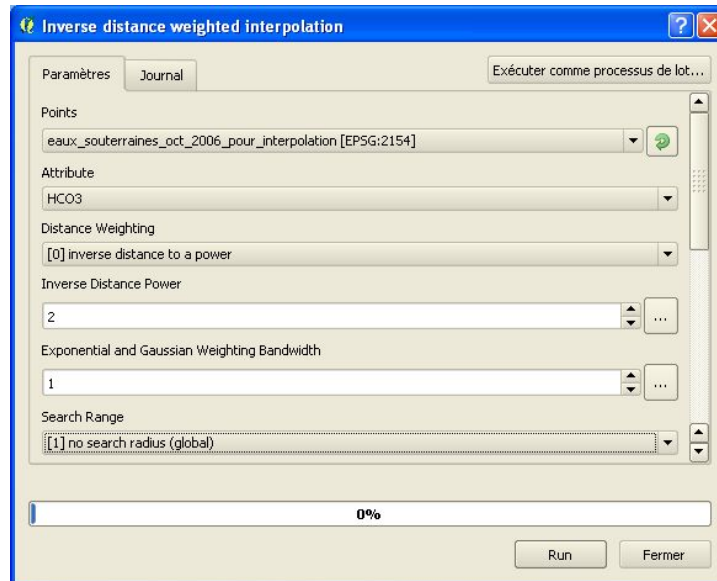


Figure 4.40. Paramétrage de l'interpolation par la méthode « IDW »

Nous obtenons le résultat illustré sur la figure 4.41.

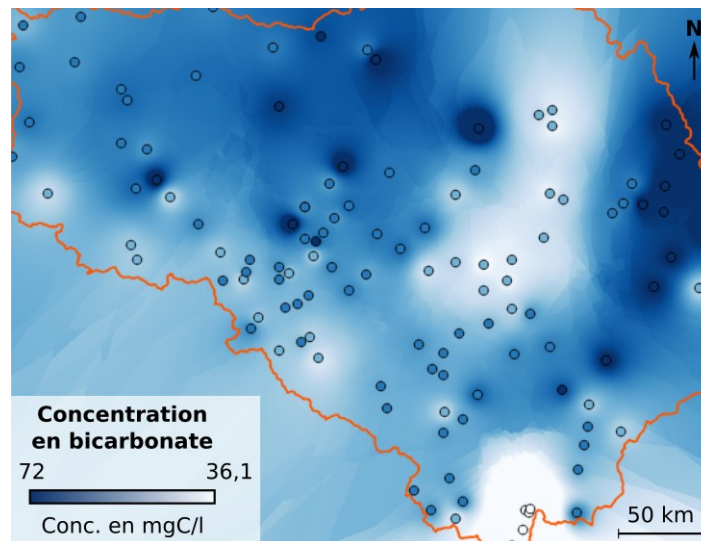


Figure 4.41. Interpolation des bicarbonates sur le bassin de la Seine par interpolation en utilisant la méthode « Inverse distance to a power »

Pour finir, nous faisons le même travail en utilisant la méthode « Thin plate spline (global) » via le module portant le même nom. Dans la fenêtre du module, nous spécifions toujours notre couche d'entrée, notre attribut à interpoler et le répertoire où nous sauons notre résultat « HCO3_interpol_thin_plate.shp » (figure 4.42).

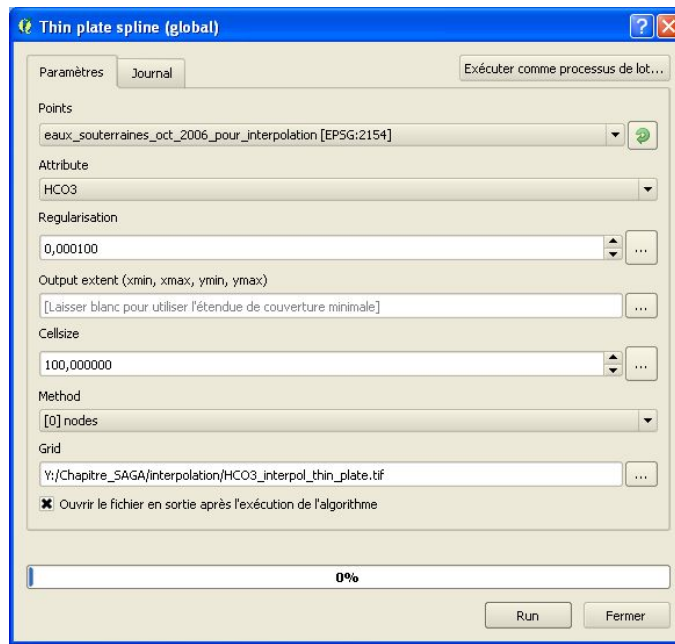


Figure 4.42. Paramétrage de l'interpolation par la méthode « thin plate »

Nous obtenons la figure 4.43.

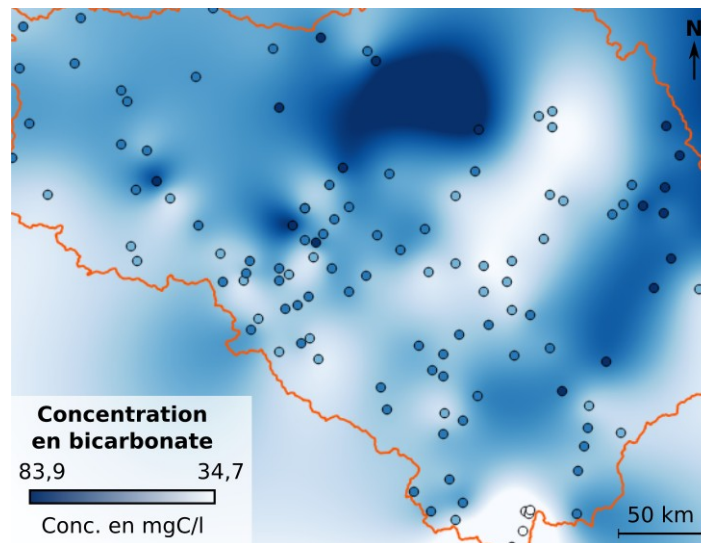


Figure 4.43. Interpolation des bicarbonates sur le bassin de la Seine par interpolation en utilisant la méthode « Thin plate spline (global) »

4.4.2.4. Qualification de l'interpolation

Nous avons maintenant trois interpolations que nous souhaitons caractériser afin de ne conserver que celle qui nous paraît être globalement la meilleure dans notre cas. Nous allons nous servir de la seconde moitié du jeu de données qui ne nous a pas servi pour l'interpolation « eaux_souterraines_oct_2006_validation.shp ». Le but est de calculer en chaque point de mesures la différence entre la valeur mesurée et la valeur interpolée sous-jacente.

Nous allons associer les valeurs de nos trois rasters d'interpolation à nos points de validation en utilisant le module SAGA GIS « Add raster values to points » dans le menu « Vector <-> raster ». Dans la fenêtre qui apparaît, nous

sélectionnons nos points de validation « eaux_souterraines_oct_2006_validation.shp », nous spécifions nos rasters contenant nos interpolations à la ligne « Grids » (leur ordre sera repris dans la table attributaire de la couche de points final), nous sélectionnons la méthode « Nearest Neighbor » afin de ne sélectionner que le pixel sous-jacent au point et nous sauvons notre résultat « interpolation_validation_temp.shp » (figure 4.44).

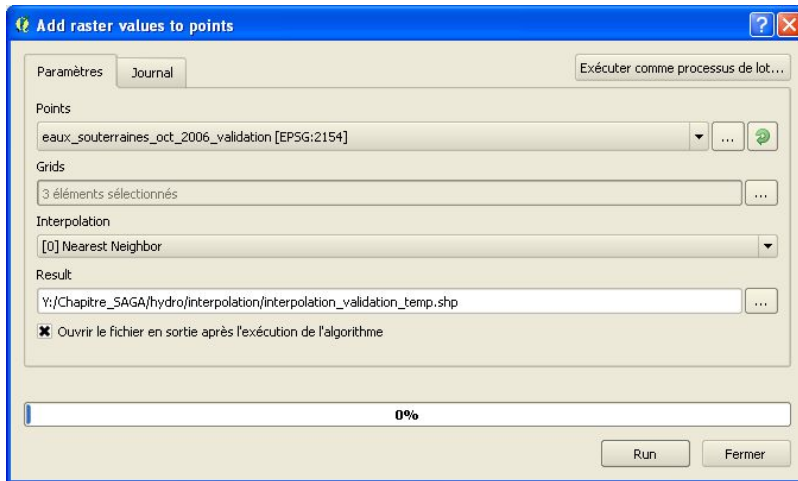


Figure 4.44. Ajout des valeurs interpolées à chaque point de validation

Nous pouvons voir dans la table d'attribut de la nouvelle couche de points trois nouveaux champs possédant les valeurs d'interpolation sous-jacentes à chaque point. L'ordre de ces champs est le même que celui dans lequel ils apparaissent dans le panneau de couches.

Nous pouvons renommer les champs afin de nous y retrouver plus facilement. Pour ce faire, nous utilisons le module « Refactoriser les champs » se trouvant dans les « Géotraitement QGIS ». Il suffit de renommer les champs souhaités en double cliquant dans la rubrique « Name » correspondante. Nous sauvons la couche avec les attributs renommés « interpolation_validation.shp ». Nous allons maintenant créer trois nouveaux champs dans la table attributaire de cette couche dans lesquels nous allons respectivement calculer les différences entre les mesures et chaque interpolation. Nous ouvrons la table d'attributs de la couche et nous ouvrons la « Calculatrice de champ », disponible dans le menu du haut. Nous choisissons « Créer un nouveau champ » et nous le nommons « d-thin ». Nous allons y calculer la différence entre les mesures et les valeurs d'interpolation « Thin plate » pour chaque point. Le champ sera donc de type « Nombre réel » et nous lui donnons une précision de 3 (valeur à la discrétion de l'utilisateur). Dans le panneau « Expression », nous entrons « HCO3 » – « thin », nous pouvons nous aider du panneau du milieu (figure 4.45).

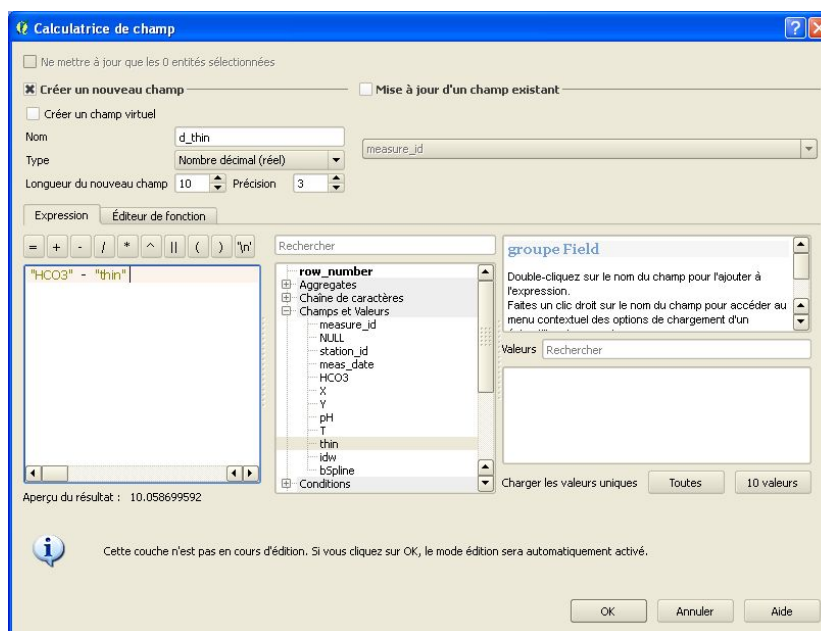


Figure 4.45. Calcul des différences entre mesures et valeurs interpolées via la calculatrice de champ

Après avoir cliqué sur « OK », le nouveau champ et les valeurs associées apparaissent dans la table d'attributs. Nous procédons de même pour les deux autres interpolations en nommant ces champs « d_idw » et « d_bspline ». Il faut sauvegarder la couche lorsque les calculs sont terminés. Nous avons maintenant trois nouveaux champs contenant les différences entre les mesures et les interpolations.

Nous allons caractériser la pertinence de nos interpolations simplement en regardant l'erreur moyenne pour chacune. Nous utilisons le module « Statistiques basiques pour les champs numériques » des « Géotraitement QGIS ». Nous spécifions la couche sur laquelle nous voulons travailler « interpolation_validation.shp », le champ dont nous voulons les statistiques « d_thin » pour commencer et nous spécifions où nous sauons le fichier contenant les statistiques « stats_thin » (figure 4.46).

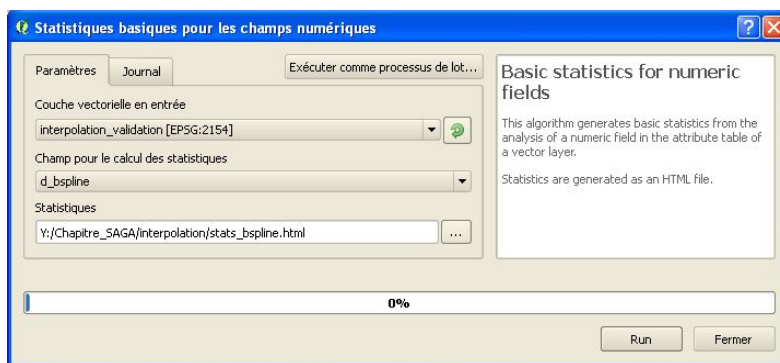


Figure 4.46. Caractérisation des différences entre les valeurs d'observation et les valeurs interpolées

Nous procédons de même pour les deux autres interpolations que nous nommons « stats_idw » et « stats_bspline ». Les résultats sont stockés dans un fichier html qu'il est possible d'ouvrir avec un navigateur web. Nous obtenons les erreurs moyennes suivantes : -0,66, -0,72 et -13,72 pour la méthode « Thin plate », « idw » et « B-spline » respectivement. La première méthode peut donc être considérée comme étant la meilleure dans notre cas. Notons que le résultat est dépendant de la sélection aléatoire effectuée dans le paragraphe 4.4.2.2. Comme les erreurs moyennes des méthodes « thin plate » et « idw » sont très proches, il se peut qu'avec une autre sélection aléatoire, la performance des deux méthodes soit inversée.

Pour finir, nous devons insister sur le fait que la méthode d'interpolation employée joue sur le résultat final. Il est nécessaire de bien poser son problème afin de choisir la méthode la plus pertinente accompagnée de ses paramètres les plus pertinents. Le point délicat est vraiment la définition de la méthode plutôt que sa mise en pratique.

4.5. Bibliographie

- [FAR 07] FARR T.G., ROSEN P.A., CARO E., CRIPPEN R., DUREN R. et al. (2007). «The Shuttle Radar Topography Mission», *Reviews of Geophysics*, 45, RG2004.
- [FEY 14] FEYISA G.L., MEILBY H., FENSHOLT R., PROUD S.R. (2014). « Automated Water Extraction Index : A new technique for surface water mapping using Landsat imagery », *Remote Sensing of Environment*, 140, 23–35.
- [HOR 45] HORTON R.E. (1945). « Erosional development of streams and their drainage basins : hydrophysical approach to quantitative morphology », *Bulletin of the Geological Society of America*, 56, 275–370, 1945.
- [HUT 89] HUTCHINSON M. (1989). « A new procedure for gridding elevation and stream line data with automatic removal of spurious pits », *J. Hydrol.*, 106(3–4), 211–232.
- [IGN 14] IGN (2014). « Base de Données sur la CARTographie Thématique des AGences de l'eau et du ministère chargé de l'environnement », *Institut national de l'information géographique et forestière*.
- [JAR 08] JARVIS A., REUTER H.I., NELSON A., GUEVARA E. (2008). «Hole-filled SRTM for the globe Version 4», *CGIAR-CSI SRTM 90m Database*, voir : srtm.csi.cgiar.org.
- [JEN 13] JENSEN J.R. (2013). *Remote Sensing of the Environment: Pearson New International Edition: An Earth Resource Perspective*.
- [OCA 84] O'CALLAHAN J.F., MARK D.M. (1984). « The extraction of drainage networks from digital elevation data », *Comput. Vis. Graph. Image Process.*, (28), 323–344.
- [OSE 16] OSE K., CORPETTI T., DEMAGISTRI L. (2016), « Traitement d'images satellitaires multi-spectrales », dans BAGHDADI N. et ZRIBI M. (dir.), *Observation des surfaces continentales par télédétection optique*, ISTE Press Ltd, London and Elsevier Ltd, Oxford.
- [PLA 02] PLANCHON O., DARBOUX, F. (2002). «A fast, simple and versatile algorithm to fill the depressions of digital elevation models », *CATENA*, vol. 46, Issues 2–3, 159-176.
- [SCH 17] SCHNEIDER A., JOST A., COULON C., SILVESTRE M., THÉRY S., DUCHARNE A. (2017). «Global-scale river network extraction based on high-resolution topography and constrained by lithology, climate, slope, and observed drainage density», *Geophys. Res. Lett.*, 44, 2773–2781.
- [SON 01] SONG C., WOODCOCK C.E., SETO K.C., LENNEY M.P., MACOMBER S.A. (2001). « Classification and Change Detection Using Landsat TM Data: When and How to Correct Atmospheric Effects? », *Remote Sensing of Environment*, 75:230–244.
- [STR 57] STRAHLER A.N. (1957). « Quantitative analysis of watershed geomorphology », *Trans. Am. Geophys. Union*, 38(6), 913–920.
- [WAN 06] WANG L., LIU H. (2006). «An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling», *International Journal of Geographical Information Science*, vol. 20, n° 2, 193-213.