



Détection automatique des cours d'eau sur le SCAN 25 Historique

DREAL PACA > SCADE > UIC > Pôle Géomatique > Mathieu Rajerison, le 13 Janvier 2016

La note "Appui Onema et IGN à l'inventaire des cours d'eau police de l'eau" propose d'utiliser le SCAN 25 Historique¹ afin de compléter les cours d'eau de l'édition BDTOP0 151 transmise aux services.

Le ciblage et l'enregistrement des cours d'eau issus de ce SCAN 25 historique constitue un travail fastidieux et chronophage pour les services.

Ce document présente le programme conçu à la DREAL PACA afin de détecter les cours d'eau sur le SCAN 25 Historique.

Il décrit les données produites munies de leurs précautions d'utilisation.

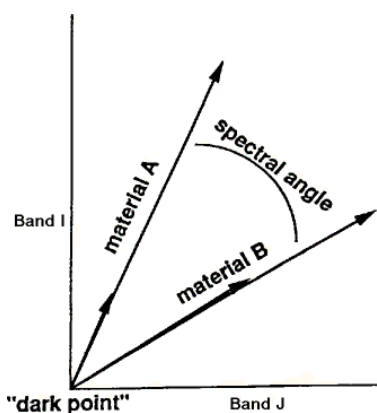
Table des matières

Méthode de constitution des couches.....	2
<i>Détection des cours d'eau bleus par Spectral Angle Mapping.....</i>	<i>2</i>
<i>Connexion automatique par snapping.....</i>	<i>2</i>
Lancement du programme d'identification.....	3
<i>Pré-requis.....</i>	<i>3</i>
<i>Données en entrée du script</i>	<i>3</i>
<i>Configuration.....</i>	<i>4</i>
Environnement de travail.....	4
Couleurs.....	5
Configuration avancée.....	5
<i>Lancement.....</i>	<i>6</i>
<i>Résultats.....</i>	<i>6</i>
Les données livrées.....	7
<i>Leur structure.....</i>	<i>7</i>
<i>Précautions d'utilisation.....</i>	<i>7</i>
Omission.....	7
Commission (surplus détecté).....	8
Annexes.....	9
<i>Visualiser la donnée sous QGIS.....</i>	<i>9</i>
<i>Filter la couche en fonction de données exogènes.....</i>	<i>9</i>
<i>Remanier la couche de lignes.....</i>	<i>10</i>
Fusionner la couche, puis la découper aux intersections.....	10
Obtenir une couche des tronçons.....	10
<i>Nettoyer la couche en se basant sur les propriétés géométriques</i>	<i>10</i>

1 Dans cette note, en PJ, intitulée Appui_Onema_inventaire_des_CE-volet IG_V9.doc, vous trouverez une description du SCAN 25 Historique

Méthode de constitution des couches

Détection des cours d'eau bleus par Spectral Angle Mapping



La méthode de classification utilisée afin d'extraire les cours d'eau bleus s'appelle le Spectral Angle Mapping. Manuellement, on fournit à l'algorithme une série de couleurs de référence, assorties de seuils de tolérance.

La couleur d'un pixel peut être représentée dans un espace à trois dimensions défini par les axes Rouge, Vert et Bleu.

Le Spectral Angle Mapping compare l'angle de la couleur de pixels d'une image à classifier à celui correspondant à la couleur de pixel de référence. Si l'angle du pixel est assez proche, alors ce dernier est considéré de la bonne classe.

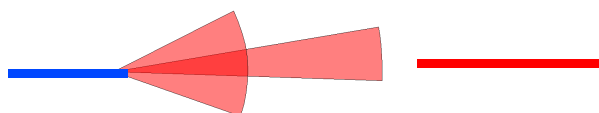
Une fois que tous les pixels sont calculés, on obtient une image binaire avec ce qui est considéré comme cours d'eau et ce qui ne l'est pas.

Connexion automatique par snapping

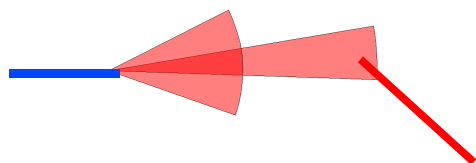
Sur le SCAN 25, certains cours d'eau sont figurés par des tirets. Aussi, des éléments du terrain, tels que des routes, peuvent interrompre le tracé.

Une fonction a été conçue afin de relier des tronçons proches qui suivent la même trajectoire.

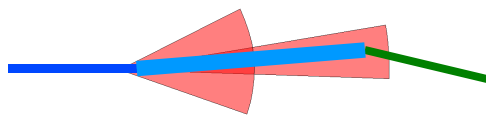
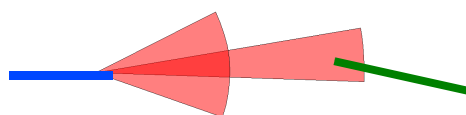
Dans le produit final, il est possible de dissocier les segments détectés de ces liaisons.



Le segment n'est **pas assez proche** pour être ciblé






Le segment est compris dans la zone de la tête chercheuse mais les **angles** des deux segments sont trop différents pour qu'il puisse être ciblé.



Le segment est dans la zone et la différence d'angle entre les deux segments est tolérable, donc on remplit l'espace entre les deux.

Lancement du programme d'identification

Pré-requis

L'outil s'appuie sur 3 logiciels libres : le logiciel de statistique  , l'outil  et la librairie  . R est embarqué au sein du programme et ne nécessitera pas d'être installé. Par contre, vous devrez procéder à l'installation des deux autres².

Le programme doit être lancé sur une machine de travail de performances suffisantes pour des traitements raster.

Données en entrée du script

Le script a besoin :

1. de l'image du **SCAN 25 historique** au format ECW ou tout autre format supporté par GDAL
2. d'une **couche de couverture**.

Le programme fonctionne au fil de l'eau, selon les objets qu'il trouvera dans la couche de couverture. Par exemple, si vous choisissez une couche de bassins versants et que vous lancez le script, les couches d'extraction correspondant aux objets bassins versants seront produites successivement. L'image ci-dessous en est une illustration.

Il est déconseillé de choisir une couche de départements en entrée car le processus, qui est gourmand en ressources, risque de bloquer votre machine.

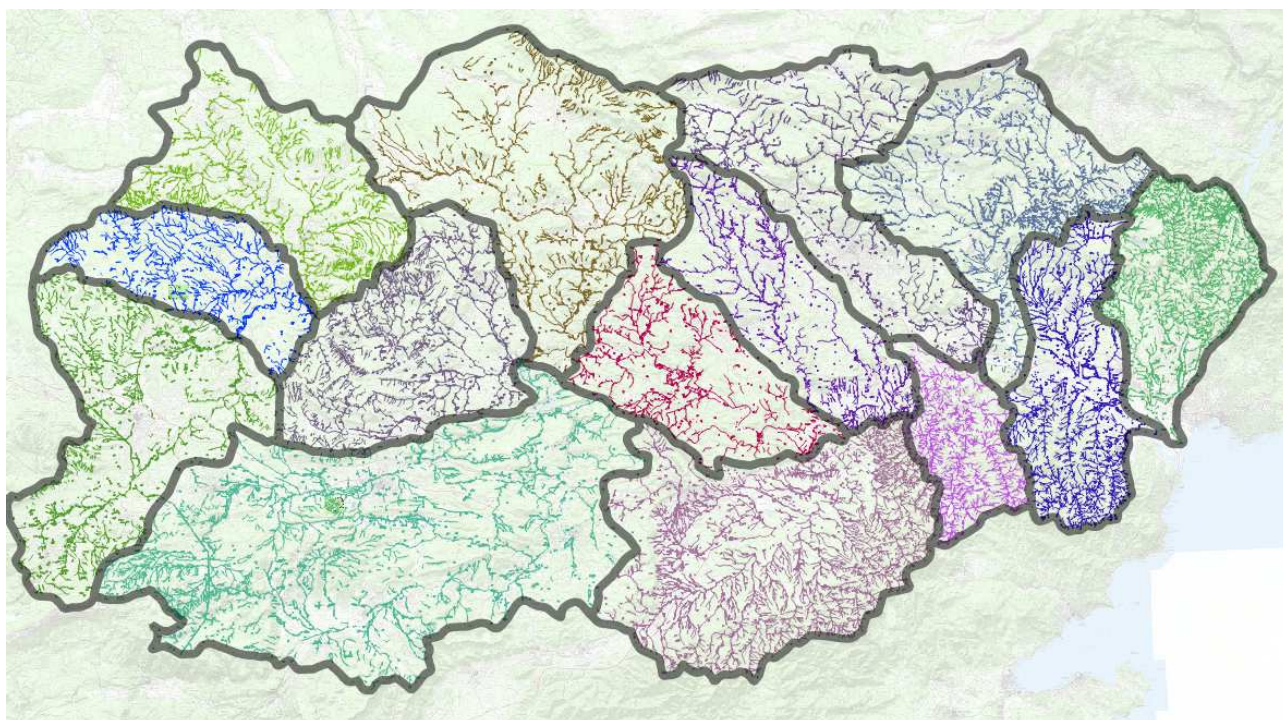


Illustration 1: Extractions réalisées sur le bassin versant de l'Argens

² <https://grass.osgeo.org/grass64/binary/mswindows/native/WinGRASS-6.4.4-1-Setup.exe>
<http://home.gdal.org/fwtools/FWTools247.exe>

Configuration

La configuration se réalise à partir de **deux fichiers .ini + un facultatif**. Ce dernier gère des comportements particuliers de l'outil. Tous ces fichiers ont été pré-remplis mais vous devrez intégrer vos valeurs pour ce qui est des chemins vers les dossiers et programmes.

```
## INPUTS
scan25=in/scan25_d83.ecw
cover=in/bassins_versants83_IGN.shp

## OUTPUTS
tmpFolder=tmp
outputFolder=out/83_IGN

## GDAL
FWTools=C:/FWTools2.4.7/bin

## GRASS
gisBase=C:/GRASS-6.4.4
gisDbase=D:/GRASSDB
location=paca
mapset=mapset

## R
memoryLimit=500000
```

```
R,G,B,seuil
85,134,154,0.1
119,173,239,0.15
119,239,229,0.1
0,37,105,0.3
7,92,217,0.3
```

```
## BIG PATCHES
bigPatch=2000

## TOLERANCE POUR LES PIXELS NOIRS
blackThresh=50

## LISSAGE
ncells=9

## THINNING
thinningIterations=200

## SUPPRESSION DES SCORIES (valeur en mètres)
dangleThresh=5

## DIRECTIONAL THRESH (distance : angle en radians)
angTol=0.8
radarHead=100:0.5,200:0.1
```

Environnement de travail

Dans un premier temps, on configure les chemins d'accès vers les logiciels et les répertoires grâce au fichier **config.ini**. La plupart des champs peuvent être laissés tels quels.

	paramètre	description	Remarques
INPUTS	scan25	Le chemin vers le fichier scan25, par rapport à l'emplacement du programme	
	cover	Le chemin vers la couche de couverture, par rapport à l'emplacement du programme	
OUTPUTS	outputFolder	Chemin vers le dossier de sortie où seront produites les extractions au fil de l'eau	Le chemin renseigné peut être laissé tel quel
	tmpFolder	Dossier temporaire de cache	
GDAL	FWTools	Chemin vers FWTools	
GRASS	gisBase	Chemin vers l'exécutable GRASS	L'outil comprend déjà une base GRASS pré-configurée
	gisDbase	Chemin vers la base de données GRASS	
	location	Nom du secteur	
	mapset	Nom du jeu de données	
R	memoryLimit	Limite mémoire que l'on affecte à l'exécution de R	La valeur pré-renseignée de 500 Mo est raisonnable. Elle peut être laissée tel quel, voire augmentée selon les capacités de la machine.

Couleurs

Dans un second temps, on liste l'ensemble des couleurs de référence qui serviront à la classification supervisée dans le fichier **couleurs.ini**. La dernière colonne correspond à la tolérance affectée à la couleur. Si la tolérance est élevée, les couleurs qui l'avoisinent dans un plus grand rayon angulaire seront intégrées dans la classification.

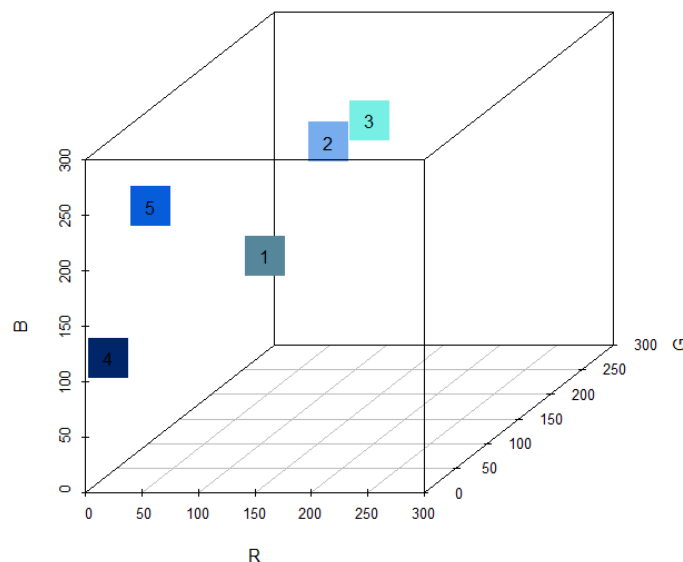
Une série de tests visant une extraction maximale de bonne qualité a permis de définir les couleurs qui sont dans le fichier couleurs.ini. Néanmoins, il se peut que vous souhaitiez intégrer vos paramètres.

A noter que si vous utilisez beaucoup de couleurs de référence, le programme sera plus lent. 3 teintes de référence est un minimum, 5, un chiffre raisonnable.

Voici le fichier de configuration tel que vous le trouverez, avec 5 couleurs de référence :

```
R,G,B,seuil
85,134,154,0.1
119,173,239,0.15
119,239,229,0.1
0,37,105,0.3
7,92,217,0.3
```

Et la représentation de ces 5 teintes dans un espace RGB :

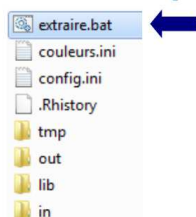


Configuration avancée

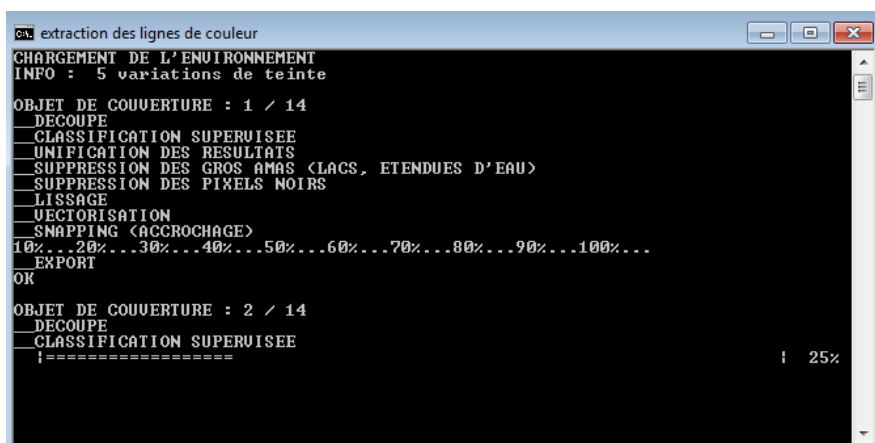
Un fichier appelé **advanced.ini** permet de modifier le comportement de fonctions particulières de l'application : surface maximale en nombre de pixels des patches à retenir (pour la suppression des étendues d'eau, etc...), intensité du lissage, longueur maximale des scories, tolérance lors de la simplification, paramètres de la tête chercheuse et angle de tolérance pour l'accrochage des segments,...

Lancement

Pour lancer le script, il suffit de double-cliquer sur le fichier **extraire.bat**

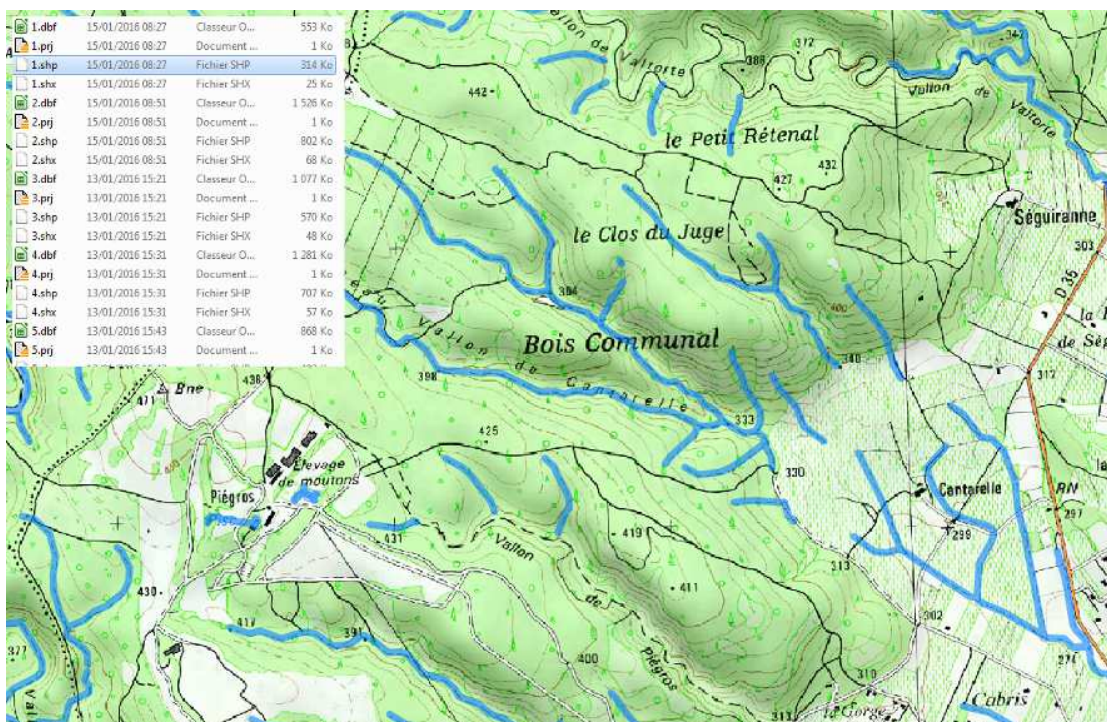


Une fenêtre s'ouvre et affiche l'avancement des opérations.



Résultats

Dans le dossier de sortie précisé par le fichier *config.ini*, les fichiers au format ESRI Shapefile sont créés au fil de l'eau au fur et à mesure de la progression du calcul, pour chacun des objets de la couche de couverture.



Les données livrées

Leur structure

Le nom de la couche indique l'identifiant de l'objet, soit son rang dans la couche de couverture. Avec un logiciel SIG standard, vous pourrez procéder à l'agrégation des données produites.

La structure est sommaire :

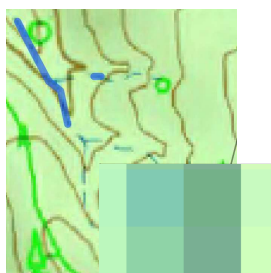
Nom de la colonne	description	Valeurs possibles
id	Il s'agit d'un identifiant unique, par catégorie d'objet (la catégorie pouvant être Détection ou Liaison, voir ci-après)	
type	Il s'agit de la nature de la ligne : ligne détectée sur le SCAN 25, ou segment calculé automatiquement par le programme pour connecter deux tronçons disjoints.	Détection Liaison
idCov	C'est l'identifiant de l'objet de couverture qui a fourni l'emprise d'extraction. Grâce à cet identifiant, vous pourrez, par jointure, récupérer les attributs de l'objet de couverture auquel il est associé.	
parameters	Ce sont les couleurs de référence ayant servi à la classification de l'image, assorties de leur seuil de tolérance. La valeur se présente sous la forme de liste de valeurs RGB séparées par un [R],[G],[B]:[seuil de tolérance en radians]	85,134,154:0.1 119,173,239:0.15 119,239,229:0.1 0,37,105:0.3 7,92,217:0.3

Précautions d'utilisation

Omission

Dans certains cas, les pixels appartenant aux cours d'eau sur le SCAN 25 Historique se fondent un peu avec le vert de l'environnement dans lequel ils se situent si bien qu'ils ne sont pas pris en compte.

Des lacunes peuvent également être causées par l'algorithme lors de la phase de nettoyage des scories.



Commission (surplus détecté)

Le programme ne fait pas abstraction des sources, réservoirs et toponymes car cela nécessiterait une étape additionnelle pour identifier ces objets. Ces derniers figurent alors sous forme de lignes dans la couche produite.

Cela dit, en s'appuyant sur des données exogènes telles que les réservoirs, les points d'eau de la BDTOPO, voire sur les propriétés géométriques des lignes (longueur, densité...), il est envisageable de les cibler pour les enlever par la suite.



L'étape de lissage³ intégrée à l'algorithme peut entraîner des connexions intempestives, lorsque des lignes divergent tout en étant proches.

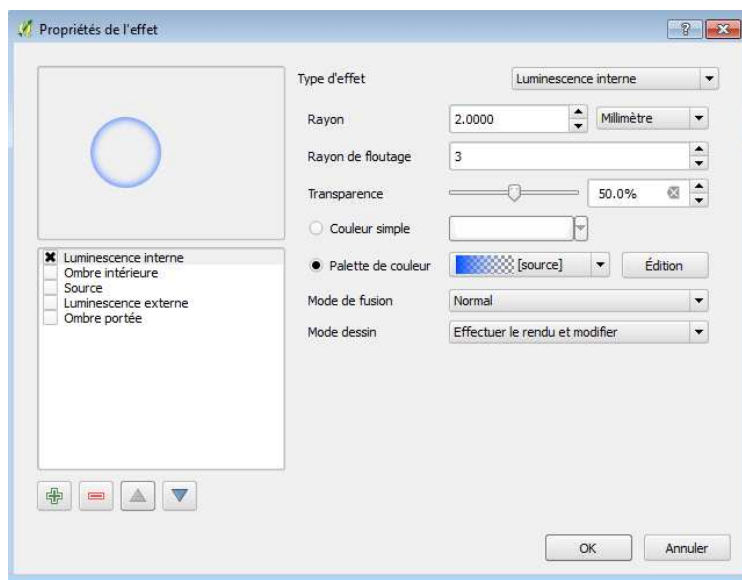


³ Voir la présentation au format odp

Annexes

Visualiser la donnée sous QGIS

Sous la version 2.12 de QGIS, appelée "Lyon", il est possible d'affecter un halo à des lignes. L'intérêt de ce style est que l'on peut représenter les lignes sous la forme d'un gradient qui va de l'opacité complète en bordure à une transparence complète au centre. Ainsi, il est possible de voir les cours d'eau du SCAN 25 Historique si ce dernier est positionné derrière.

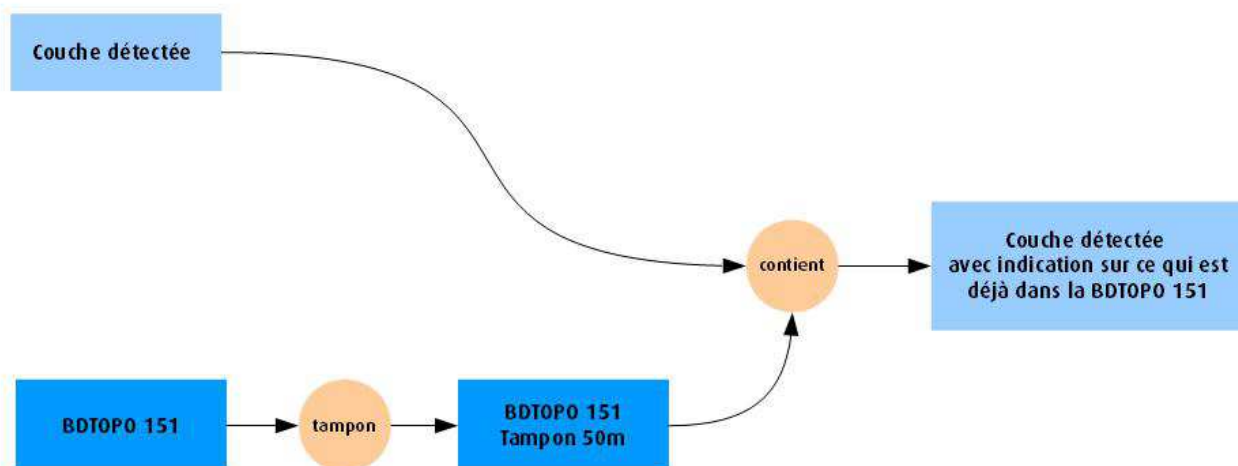


Filtrer la couche en fonction de données exogènes

En croisant la couche détectée avec les cours d'eau de la BDTOP0 151 ainsi que d'autres données exogènes telles que les points d'eau, réservoirs, toponymes, intégrés à cette même édition, il est possible de ne retenir que les tronçons qui apparaissent en plus des données existantes.

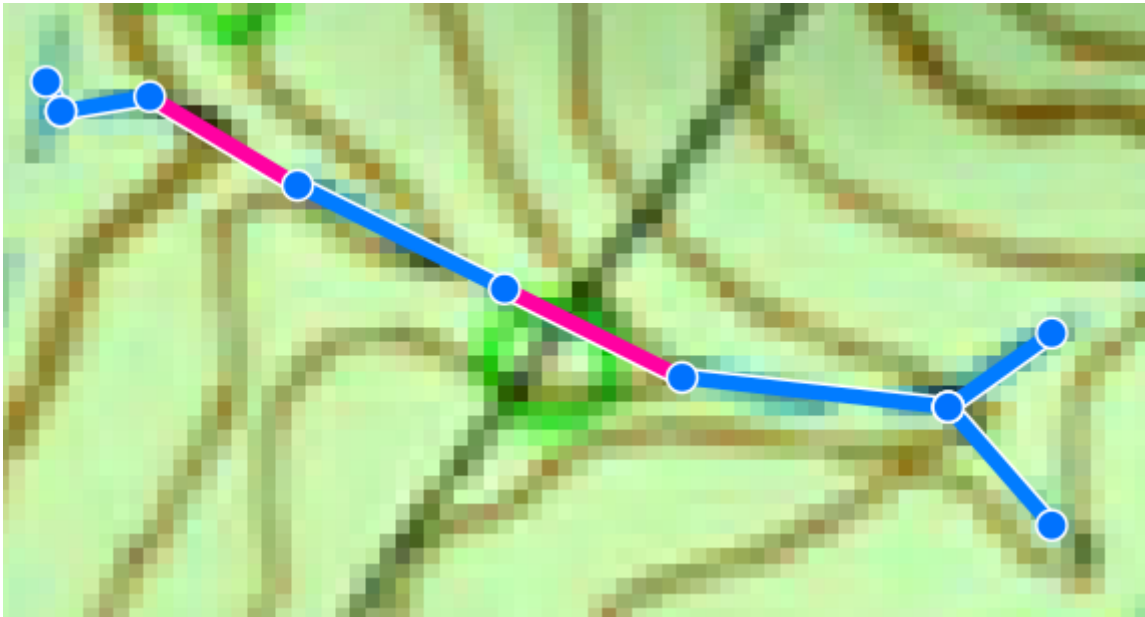
Dans le cas de la BDTOPAGE, cela permet de cibler les tronçons à rajouter à la livraison.

Voici un exemple de traitement basé sur la BDTOP0 151:



Remanier la couche de lignes

Actuellement, la couche est composée de polygones correspondant aux cours d'eau détectés. Ces **polygones** prennent fin à chaque intersection et elles sont parfois reliées à des **connexions** auto-calculées.



Fusionner la couche, puis la découper aux intersections

Imaginons que vous ayez fini de traiter une couche de données : vous avez validé les tronçons détectés et les connexions. Il se peut que vous désiriez fusionner le tout et obtenir de nouvelles lignes assemblées et découpées aux intersections.

Pour cela, il vous faudra d'abord fusionner vos lignes en une seule : sous QGIS, Éditer > Fusionner les entités sélectionnées. Enfin, intersecter la couche résultante avec son clone aboutira à une couche composée de lignes découpées aux intersections : Vecteur > Outils de géotraitement > Intersection.

Obtenir une couche des tronçons

Au contraire, pour obtenir, à partir de polygones, des lignes individuelles, il faut utiliser l'outil "Exploser des lignes" qui se situe dans la boîte à traitements de QGIS : Traitement > Boîte à outils > Exploder des lignes

Nettoyer la couche en se basant sur les propriétés géométriques

La longueur des segments peut servir d'indicateur afin de repérer les toponymes, et autres éléments bleus compris dans la classification et n'appartenant pas à la classe cours d'eau.

Dans le cas de toponymes, on remarque un réseau de lignes dense. La densité, propriété utilisée dans la théorie des graphes, indique si un réseau est très inter-connecté ou pas. Les cours d'eau, eux, forment un réseau peu dense. Obtenir cet indicateur nécessite d'utiliser des outils tels [pgRouting](#), la librairie pyqgis [network analysis](#) ou la librairie R [igraph](#).