



Licence Professionnelle

Cartographie, Topographie et Système d'Information Géographique

Promotion : 2021 – 2022

Cédric LEPICIER

RAPPORT DE STAGE

***Qualification d'un jeu de données générique
Elaboration d'une chaîne de contrôle qualité autonome***

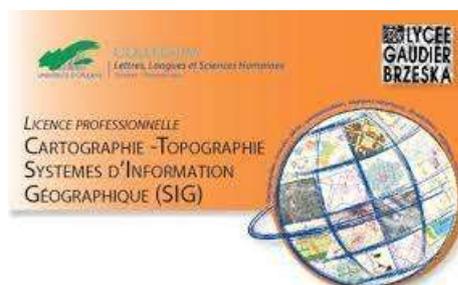


Sous la direction de :

Mr Stéphane ROLLE, Maître de stage : CRIGE – PACA

(Centre de ressources en Information Géographique – Provence - Alpes - Côte-d'Azur)

Mr Abdelkrim BENSALD, Tuteur de stage : Université d'Orléans



RESUME

Une partie des activités du CRIGE – PACA consiste à diffuser de la donnée géographique numérique. Dans ce cadre et pour répondre aux exigences de qualité qu’il s’est fixé au travers de son engagement et de sa participation dans diverses structures en lien direct avec la diffusion, la promotion et la production de l’open data en France , le CRIGE – PACA souhaite valoriser et qualifier ses données de sources hétérogènes et variées en se dotant d’une chaîne de qualification et de traitement automatisé capable de pouvoir traiter des fichiers génériques afin d’évaluer leur niveau de qualité avant leur diffusion sur leur plateforme ([.https://www.crige-paca.org/](https://www.crige-paca.org/)). J’ai donc travaillé à l’élaboration d’une chaîne de traitement autonome, au calcul des indicateurs de qualité et à leur diffusion.

Mots clés : SIG - Contrôle - Qualité – Qualification - traitement

ABSTRACT

Part of the activities of CRIGE – PACA consists in disseminating digital geographic data. In this context and to meet the quality requirements that it has set itself through its commitment and participation in various structures directly linked to the dissemination, promotion and production of open data in France, the CRIGE – PACA wishes to enhance and qualify its data from heterogeneous and varied sources by acquiring a qualification and automated processing chain capable of processing generic files in order to assess their level of quality before their distribution on their platform ([. https://www.crige-paca.org/](https://www.crige-paca.org/)). I therefore worked on the development of an autonomous processing chain, the calculation of quality indicators and their dissemination.

Key Word : GIS - Control - Quality – Qualification - processing



SOMMAIRE

RESUME	1
SOMMAIRE.....	2
REMERCIEMENTS.....	3
INTRODUCTION	4
LE CRIGE – PACA, PORTRAIT D’UNE STRUCTURE ATYPIQUE.....	5
Présentation du CRIGE – PACA	5
Situation géographique du CRIGE – PACA.....	6
Le CRIGE PACA au service de l’interopérabilité des données et de ses utilisateurs.....	7
La qualité, une problématique portée par le CRIGE	8
Les différents aspects de la qualité	8
La norme ISO 19157, norme de référence de la qualité des données géographiques.....	8
La qualité interne.....	12
La qualité externe.....	13
La qualité, un élément des métadonnées parmi d’autres.....	13
Objectif du stage et définition du périmètre du sujet.....	13
Une réponse aux enjeux de la qualité	13
Un périmètre adapté	14
MISE EN ŒUVRE D’UNE CHAÎNE DE CONTROLE QUALITE AUTONOME	16
Définition des principes directeurs mis en œuvre dans la chaîne de traitement	16
La logique adoptée pour le processus de traitement.....	17
Principe de fonctionnement du logiciel «FME»	18
Description et fonctionnement du processus de traitement de la donnée.....	20
Les données d’entrée	20
Les trois chaînes de traitement principales.....	22
La chaîne de traitement des emprises spatiales de références	23
La chaîne de traitement de la structure du «SHAPE»	24
La chaîne de traitement de la structure interne du «SHAPE».....	27
Principe de calcul de la notation globale.....	27
Principe de calcul du critère de cohérence logique.....	27
Définition du sous-critère de cohérence de format.....	27
Définition du sous-critère de cohérence de topologique.....	27
Notation de l’intégrité géométrique	27
Notation des doublons	27
Notation des micro-polygones	27
Principe de calcul du critère d’exhaustivité.....	28
QUALIFICATION DES DONNEES : RESULTATS ET AXES D’AMELIORATION	28
Application de la chaîne de traitement, des résultats cohérents et pertinents	28
Préparation des répertoires de travail pour une nouvelle analyse.....	28
Analyse et validation ou rejet de la projection utilisée.....	29
Analyse des informations générales du «SHAPE»	31
Dénombrement des géométries en entrée.....	32
Validation ou rejet des géométries présentes	32
Principe de notation des éléments constitutifs des sous-critères et critères	37
Traitement des géométries valides en fonction des seuils de calcul	38
Traitement des géométries en doublon.....	41
Calcul du sous-critère de cohérence topologique.....	43
Calcul du critère d’exhaustivité	44
Traitement des géométries valides en fonction des seuils de calcul et de l’emprise spatiale de référence	45



Calcul de la notation du sous critère de cohérence topologique.....	47
Calcul de la notation globale.....	50
Méthode de confection des rapports finaux.....	51
Envoi automatique des mails.....	53
Un outil opérationnel mais perfectible	55
Difficultés spécifiques à l'élaboration du script	55
Les axes de développement du qualificateur.....	56
Axes d'amélioration possibles en début de traitement.....	57
Axes d'amélioration possibles en cours de traitement	57
Axes d'amélioration possibles en fin de traitement.....	57
FME un logiciel propriétaire sous licence	57
CONCLUSION	58
TABLES DES FIGURES.....	58
TABLES DES ANNEXES	60
WEBOGRAPHIE.....	60
BIBLIOGRAPHIE	60
TABLES DES ACRONYMES.....	61
ANNEXES.....	64

REMERCIEMENTS

Je tiens à remercier toutes les personnes qui ont contribué à rendre possible cette reprise d'étude après une période d'activité professionnelle conséquente.

Tous d'abord, mes remerciements s'adressent à tous les professeurs et intervenants de la licence professionnelle en Système d'Information Géographique Cartographie topographie qui tout au long de la formation nous ont transmis une partie de leurs savoir-faire et de leur expérience. La valeur cardinale que j'ai pu constater au sein de cette formation est la constante bienveillance et un certain niveau d'exigence. M. BENSAID et M. NEDJAI incarnent ces valeurs et ont toujours transmis leurs connaissances en restant disponible et à l'écoute de leurs étudiants.

Je tiens également à remercier toute l'équipe du Centre de Ressource en Information Géographique Provence - Alpes - Côte d'Azur.

Je leur suis très reconnaissant de m'avoir permis de réaliser mon stage de fin d'étude au sein de leur structure. Et tout particulièrement, Merci à mon maître de stage : M. STEPHANE ROLLE pour la qualité de son écoute et de nos échanges, il m'a permis d'appréhender les différentes dimensions du sujet proposé. son expertise, sa hauteur de vue sur le sujet et son support technique m'ont été très utiles pour apporter les réponses les plus pertinentes possibles à la problématique posée.

Enfin je remercie tous mes proches, famille et amis qui ont grandement contribué à créer les conditions propices à la réussite de cette année de reprise d'étude.



INTRODUCTION

Le CRIGE – PACA par l’entremise de M. STEPHANE ROLLE est fortement impliqué dans le traitement des problématiques liées à la qualité. A ce titre et en lien avec le groupe de travail QuaDoGéo du CNIG, le CRIGE souhaite proposer une solution assez autonome et automatique pour qualifier des fichiers de données géographiques.

Le choix du CRIGE s’est porté sur un «ETL» comme support d’une chaîne de traitement pour calculer et mettre en évidence un certain nombre d’indicateurs pertinents afin d’établir le niveau de qualité de la donnée avant diffusion sur sa plateforme. Cependant, la problématique de la qualité étant très vaste, il n’était pas possible de traiter l’ensemble des aspects de cette problématique dans le temps imparti par notre formation au stage de fin d’étude. Le parti pris de traiter une partie seulement est donc volontaire. Le résultat de ce stage de fin d’étude sera présenté et proposé au groupe de travail QuaDoGéo lors de leur prochaine réunion.

Ce travail inabouti sur certains aspects de la problématique de la qualité pourra toutefois être repris, poursuivi et complété par d’autres personnes qui pourront s’appuyer sur les éléments et la méthodologie détaillée dans ce rapport.

LE CRIGE – PACA, PORTRAIT D’UNE STRUCTURE ATYPIQUE

Présentation du CRIGE – PACA

Toutes les régions de France disposent d’un CRIGE cependant, les noms, les organisations, les statuts juridiques sont différents mais leurs missions sont identiques et de ce point de vue le CRIGE – PACA ne déroge pas à la règle.



Fig. 1 : Carte des CRIGE en France

(Source AFIGEO 2022)

Situation géographique du CRIGE – PACA

Les locaux du CRIGE-PACA se situent sur la commune d'AIX EN PROVENCE au cœur du «TECHNOPOLE DE L'ENVIRONNEMENT ARBOIS MEDITERRANE» sur le plateau de l'Arbois à l'est de la zone d'activité d'AIX EN PROVENCE dite des milles près du quartier de LA DURANNE.



Fig. 2 : Plan de situation du CRIGE (petite échelle)

(Source : Plan IGN géoportail)

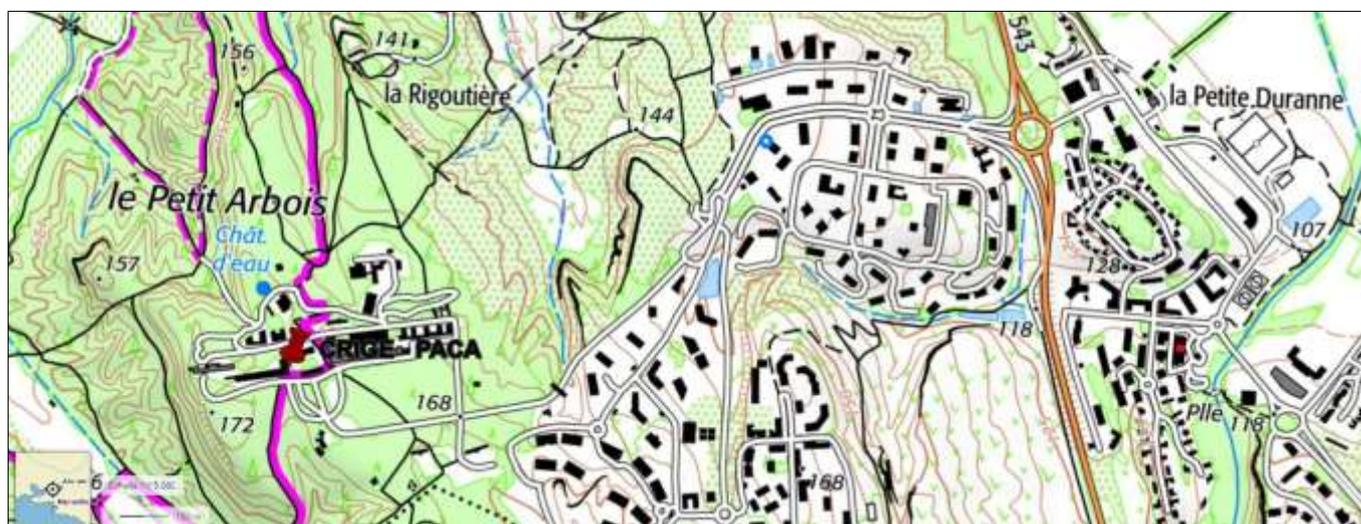


Fig. 3 : Plan de situation du CRIGE (grande échelle)

(Source : carte topographique IGN géoportail)

Le CRIGE PACA au service de l'interopérabilité des données et de ses utilisateurs

Le CRIGE – PACA structure atypique par son statut juridique dans le secteur d'activité de l'information géographique travail depuis le début des années 2000 à la promotion et à la diffusion de l'open data. Son modèle économique est tout aussi atypique car il repose sur les subventions et sur les adhésions de divers acteurs du secteur départements, communes, EPCI entreprises.

Remarquable par son indépendance politique, cet association loi 1901 est fortement impliquée dans les organisations qui en France organisent gèrent et promeuvent la donnée géographique sous toutes ses formes.

En effet, le CRIGE - PACA participe au sein de l'AFIGEO ou du CNIG à plusieurs groupes de travail.

D'ailleurs Mme ARCHIAS Directrice du CRIGE – PACA est vice-présidente de la commission de l'AFIGEO mais aussi responsable de la commission des standards au sein du CNIG. Elle est donc fortement impliquée dans le mouvement de l'open data en France. Cette implication est d'ailleurs une marque caractéristique des membres de cette petite structure de quatre personnes qui a récemment subi une restructuration significative.

D'autres membres du personnel du CRIGE PACA sont également membres actifs de différents groupes de travail au sein du CNIG ou de l'AFIGEO et notamment le groupe de travail QUADOGEO qui s'occupe de la thématique de la qualité des données géographiques.

Les activités du CRIGE PACA sont variées :

- il produit de la donnée qualitative conforme aux normes et standards en vigueur
- il simplifie la gestion et l'intégration des données dans les systèmes internes des organisations adhérentes,
- il facilite l'ouverture, la diffusion et la promotion des données géographiques,
- il développe des outils collaboratifs avec LIZMAP par exemple ou les utilisateurs peuvent enrichir la base de données
- Au travers de ces activités techniques le CRIGE travaille également beaucoup sur l'aspect relationnel et la communication, il favorise la synergie entre les différentes strates et compétences territoriales.

Pour atteindre ses objectifs le CRIGE – PACA s'est fixé plusieurs missions :

- favoriser, porter, valoriser les projets en créant du lien entre les acteurs du secteur.
- fournir une expertise technique et juridique
- proposer des applications en lignes
- concevoir des standards de données localement
- favoriser le développement des bons usages
- simplifier le travail quotidien de ses adhérents
- mutualisation des couts

Le CRIGE est également attaché à certaines valeurs telles que :

- le partage et la mutualisation des données, des méthodes, des outils et des expériences
- la neutralité et la subsidiarité dans le positionnement stratégique et les choix techniques
- la souplesse et un esprit ouvert sur les financements, la conduite des projets, et le dialogue avec les différents partenaires
- la proximité et la réactivité par l'adaptation aux besoins dans la gestion des priorités et le respect des objectifs des projets engagés
- la motivation et l'engagement grâce à une équipe d'experts motivés et mus par une certaine idée du service aux publics et du développement territorial



Le CRIGE PACA propose donc plusieurs activités et services telles que :

- la fourniture de services géographiques en ligne en proposant des données raffinées prêt à l'emploi structurées à façon selon le besoin de l'adhérent
- montage et coordination de démarches et projets en proposant des solutions techniques puis en les déployant au profit d'autres organisations
- assistance technique, information et formation en proposant une assistance en ligne et en produisant des documents tutoriaux
- coordination, accompagnement, animation du réseau en dialoguant avec les autres régions ou avec le niveau national
- en faisant de la communication auprès des professionnels du secteur et du grand public

Le CRIGE -PACA est donc un lieu d'échange qui favorise le dialogue entre les différents acteurs du secteur qui permet aux différentes strates administratives et à leurs responsables de dialoguer en neutralisant la part d'égoïsme personnels et l'inertie des structures importantes.

La qualité, une problématique portée par le CRIGE

Les différents aspects de la qualité

La norme ISO 19157, norme de référence de la qualité des données géographiques

« L'ISO 19157:2013 établit les principes de description de la qualité des données géographiques. Elle

- définit des composants destinés à décrire la qualité de données;
- spécifie des composants et la structure du contenu d'un registre de mesures de qualité des données;
- décrit des procédures générales d'évaluation de la qualité des données géographiques;
- pose les principes de la description de la qualité des données dans des rapports.

L'ISO 19157:2013 définit également un ensemble de mesures de qualité des données destinées à l'évaluation et à la mise en place de rapports sur la qualité de données. Elle s'applique aux producteurs de données fournissant des informations de qualité pour décrire et évaluer la façon dont un jeu de données répond à sa spécification de produit et aux utilisateurs cherchant à déterminer si des données géographiques spécifiques sont ou non de qualité suffisante pour leur application particulière.

L'ISO 19157:2013 ne cherche pas à définir des niveaux minimums acceptables de qualité en matière de données géographiques.»

(Source : <https://www.iso.org/obp/ui/fr/#iso:std:32575:fr>)

Avec le développement des technologies numériques, pour les simples citoyens comme pour les professionnels du secteur, l'abondance et la facilité d'accès à de gros volumes de données géolocalisées est maintenant une réalité. L'effet pervers de ce foisonnement de données libres et facilement accessibles est qu'il n'est pas évident de trouver la donnée facilement utilisable adaptée et structurée selon le besoin exprimé.

Dans nos activités de géomaticiens, la donnée est la matière première de nos productions. Cela pose donc la question suivante comment évaluer, selon quels critères avec quels outils la qualité des données que nous incorporons dans nos productions ?



La communauté de géomaticiens réfléchit depuis longtemps sur cette problématique mais force est de constater qu'il n'existe pas de réponse facile et rapide même si le secteur tend à définir des recommandations des normes et des standards.

Il est vrai que la notion de qualité est complexe, abstraite et la première question qui vient à l'esprit de l'utilisateur n'est pas de savoir si la donnée que l'on s'apprête à utiliser est-elle de bonne qualité ou pas ? Néanmoins, l'importance de la réponse à cette question est cruciale car elle conditionne en partie la qualité des productions en aval réalisées sur l'exploitation de cette donnée que nous allons qualifier.

Même si la qualité peut se définir selon de multiples critères, elle se divise en deux grandes catégories :

LA QUALITE INTERNE et **LA QUALITE EXTERNE**.

Aussi pour «vulgariser» et rendre plus facilement compréhensible les principes édictés par la norme ISO 19157 : 2013, le CEREMA a publié en 2017 une série de onze fiches techniques qui reprennent tous les aspects de la qualité. D'ailleurs certains critères dont les calculs sont traités dans le script s'inspirent du contenu de ces fiches.

Ces fiches sont accessibles et téléchargeable sur le site du CRIGE – PACA dans sa partie dédiée à la qualité, Elles sont également disponible sur le site du CEREMA (<https://www.cerema.fr/fr/centre-ressources/boutique/qualifier-donnees-geographiques>)

Fiches Qualité du CEREMA



Qualifier les données géographiques
la démarche de la norme ISO 19157

Connaître la qualité d'une donnée géographique fiabilise son utilisation

Le contenu de ces fiches est accessible en format PDF et en format ePub. Elles sont disponibles gratuitement sur le site du CEREMA.

Les fiches sont composées de deux séries : les fiches "**méthodes**" qui précisent les concepts et les modes opératoires généraux, et les fiches "**critères**", qui décrivent et sélectionnent les critères et les mesures définis par la **norme ISO 19157** et retenus ici pour l'évaluation de la qualité des données géographiques. Elles proposent également des **méthodes d'évaluation** et de rapportage de la qualité selon différents critères.

- **Fiche n°01** : Connaître la qualité d'une donnée géographique fiabilise son utilisation
- **Fiche n°02** : Généralités sur la qualité des données géographiques
- **Fiche n°03** : Éléments de contexte pour le contrôle qualité
- **Fiche n°04** : Éléments statistiques + 2 ANNEXES disponibles dans le résumé
- **Fiche n°05** : Méthodes d'échantillonnage
- **Fiche n°06** : Modes de représentation
- **Fiche n°07** : Critère de cohérence logique
- **Fiche n°08** : Critère d'exhaustivité
- **Fiche n°09** : Critère de précision thématique (màj 12/2021)
- **Fiche n°10** : Critère de précision de position (màj 12/2021)
- **Fiche n°11** : Critère de qualité temporelle

Organisation	Publié le	Format	Poids	Téléchargé	
Cerema	16/02/2018	PDF	2 Mo	55 Fois	TÉLÉCHARGER

Fig. 4 : Espace de téléchargement des fiches du CEREMA sur le site du CRIGE - PACA

(Source : <https://www.crige-paca.org/projets/quadogeo/>)



Qualifier les données géographiques
Un décryptage de la norme ISO 19157

Critère de cohérence logique

La connaissance de la qualité des données, en sécurisant l'utilisateur, incite davantage à leur réutilisation.

Ce décryptage de la norme ISO 19157 a pour vocation de donner un cadre méthodologique pour qualifier les données lors de leur diffusion.

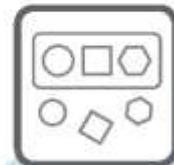
L'essor des données ouvertes et géolocalisées et la profusion d'usages existant et à venir nous rend tous progressivement producteur et utilisateur de données géographiques.

Les activités régaliennes ou les politiques publiques s'appuient sur de l'information maîtrisée où la qualité des données produites ou utilisées devient un entrant indispensable. Pour autant, tout le monde ne dispose pas des moyens des producteurs institutionnels de données et il paraît utile de fournir des recommandations et des méthodes plus adaptées au contexte de chacun, pour qualifier les données géographiques, communiquer sur les résultats obtenus voire savoir les interpréter. C'est l'objectif que s'est fixé le Cerema en proposant cette collection de fiches, à l'interface des productions et des usages.

Cette fiche propose des méthodes pour exprimer la qualité d'un jeu de données relatives à leur cohérence logique. Mesurer la cohérence logique n'a de sens que s'il existe des spécifications qui décrivent le schéma conceptuel et la structure physique attendus. La vérification se fait alors par rapport aux règles décrites.

Pour autant, même en l'absence de spécifications, certaines règles « de bon sens » peuvent être vérifiées comme la projection cartographique utilisée et l'unicité des identifiants d'objets.

Fiche n° 07 - décembre 2017



Collection | Connaissances

Fig. 5 : Extrait de la Fiche N°7 du CEREMA sur le critère de cohérence logique page 1/12

(Source <https://www.cerema.fr/>)



2.4 Cohérence topologique

Ce sous-critère regroupe plusieurs mesures caractérisant la qualité topologique de construction des objets géométriques (auto-intersections, micro-surfaces, connexions aux extrémités, etc.)

Il ne traite pas des relations topologiques entre deux classes d'objets (adjacence, inclusions, intersections) que l'on retrouve dans le § 2.1 sur la cohérence conceptuelle.

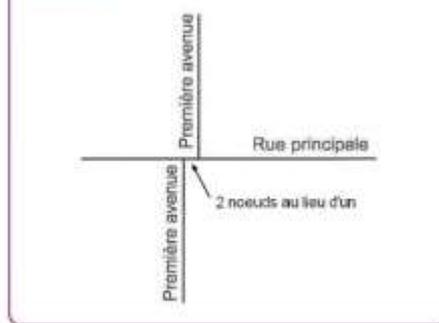
Il peut se contrôler avec des fonctionnalités dédiées des logiciels SIG.

■ Nombre de connexions arc-noeud erronées

Définition • Comptage des connexions arc-noeud erronées.

Type de valeur • Nombre entier.

Exemple : carrefour en « baionnette » injustifié



■ Taux de connexions arc-noeud erronées

Définition • Nombre de connexions arc-noeud erronées par rapport au nombre de connexions supposées.

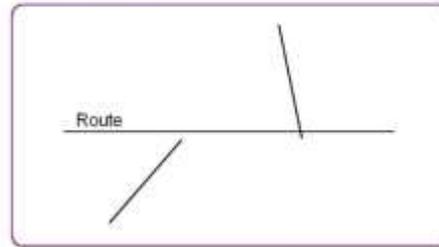
Type de valeur • Nombre réel (pourcentage).

■ Nombre de connexions manquantes en raison de ligne trop courte

Définition • Comptage des éléments du jeu de données disjoints en raison de ligne trop courte par rapport à une tolérance donnée.

Paramètre • Tolérance de recherche à partir de l'extrémité d'une ligne pendante.

Type de valeur • Nombre entier.



■ Nombre de connexions manquantes en raison de ligne trop longue

Définition • Comptage des éléments du jeu de données disjoints en raison de lignes trop longues.

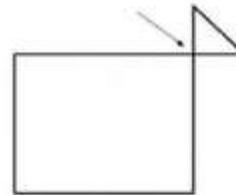
Paramètre • Tolérance de recherche sur la longueur minimale admise dans le jeu de données.

Type de valeur • Nombre entier.

■ Nombre de micro-surfaces non valides

Définition • Comptage de toutes les micro-surfaces non valides.

Description • Une micro-surface est une zone non souhaitée, apparaissant lorsque des surfaces adjacentes ne sont pas numérisées correctement. Les bords de surfaces adjacentes peuvent se chevaucher ou s'écarter de manière involontaire, en faible distance, et provoquer cette erreur topologique.



Paramètre 1 • Surface maximale d'une micro-surface (les éléments de surface supérieures ne seront pas considérés comme des micro-surfaces).

Fig. 6 : Extrait de la Fiche N°7 du CEREMA sur le critère de cohérence logique page 5/12

(Source <https://www.cerema.fr/>)

La qualité interne

Elle peut s'évaluer en comparant la donnée avec un cahier des charges une norme ou des valeurs cibles. Les différents validateurs existants travaillent sur cet aspect en comparant si la donnée correspond à la valeur cible ou à la plage de valeur cible, le résultat est alors soit conforme soit non conforme.



Fig. 7 : Exemple du validateur de référence IBAN

(Source : <https://www.iban.com/>)

Exemple ci-dessus, le validateur de numéro IBAN ne fait que vérifier si le Numéro est cohérent et existe et renvoi un document validant



Fig. 8 : Extrait du document de retour du validateur

(Source : <https://www.iban.com/>)

Le Numéro de compte soumis au validateur est donc bien valide

Autre exemple de validateur : le validateur de la *Base adresse locale*



Fig. 9 : Exemple du validateur de la Base Adresse Locale

(Source <https://adresse.data.gouv.fr/bases-locales/validateur>)

La qualité externe

Elle s'évalue plutôt sur des critères de pertinence et d'adéquation entre le besoin exprimé et la nature de la donnée, cette notion est plus ardue à définir et plus difficile à évaluer.

La qualité des données géographique dans leur globalité est encadrée par la norme ISO 19157 qui fait appel à des critères et des méthodes complexes peu accessibles et compréhensibles pour la majeure partie de la communauté des utilisateurs de la donnée géolocalisée.

La qualité, un élément des métadonnées parmi d'autres

Les métadonnées, données sur la donnée renseignent sur les différentes caractéristiques de la donnée contenues dans un fichier.

C'est le fichier d'extension .xml qui contient ces métadonnées. Cependant, deux constats s'imposent à nous, tout d'abord le fichier d'extension .xml n'est pas toujours présent, loin de là, parfois il peut être vide partiellement vide ou renseigné de manière inadaptée, ensuite il n'y a pas de norme pour renseigner cette section.

Les données de référence, elle-même renseignent peu sur cet aspect particulier.

Informations sur la qualité de la donnée géographique de la ressource	
Source de la donnée Direction Datas	Les données géométriques sont compatibles avec le COG de l'IGN. Les données de publication sont à jour en ce qui concerne les informations portant sur le contenu de l'IGN (fusion, création de communes, changement de nom des communes, etc.). Sur la version 3.0, aucune méthode de généralisation n'est mise en œuvre, ce qui ne permet pas de superposer les données de la version 2.0 avec celles des versions antérieures. La version 2.0 contient de nouvelles classes d'objets et reprend le nom de pseudo-couleur.
	voir message voir message

Fig. 10 : Extrait de la métadonnée de «ADMIN-EXPRESS_3-1» - section dédiée à la qualité

(Source : IGNF_ADE_3-1_SHP_WGS84G_FRA.html)

En ce sens la production d'indicateurs relatifs à la donnée constitue une création de métadonnées qui pourrait alors alimenter les sections dédiées à la qualité.

Objectif du stage et définition du périmètre du sujet

Une réponse aux enjeux de la qualité

Le sujet proposé par le CRIGE-PACA pour ce stage de fin d'étude est très vaste et revêt de multiples aspects dont certains sont complexes à appréhender. Aussi très rapidement, le constat a été clairement établi que la totalité des critères ne pourraient pas être traités sous tous leurs aspects. Une grande liberté et une grande autonomie m'a été laissée pour traiter le sujet. Cependant, seulement deux contraintes m'ont été définies le processus d'analyse et de contrôle des données devait être réalisé avec l'ETL «FME» et les données d'entrées à analyser et qualifier sont exclusivement au format «SHAPE».

L'objectif de ce script est donc de traiter de manière générique n'importe quel fichier au format «SHAPE» pour tenter d'en quantifier la qualité. La démarche se distingue en cela des validateurs qui établissent seulement de manière binaire .si la donnée est conforme ou non conforme. Concrètement, il s'agit de produire et de communiquer de manière automatique et autonome au travers d'un rapport synthétique les indicateurs de qualité au producteur de la donnée avant diffusion sur la plate-forme du CRIGE.

En complément du script «FME», deux autres documents seront rédigés, il s'agit de deux notices d'utilisations destinées à deux publics différents. Une première notice explique comment utiliser le script pour un utilisateur



basique de l'outil, une seconde pour un public plus aguerri, décrit les différents paramètres et la façon de les modifier pour faciliter une éventuelle reprise ou évolution du script.

Un périmètre adapté

Après nos échanges avec mon maître de stage où les principes de traitement de la donnée et les attentes ont été définis et une rapide formation sur le logiciel «FME», il est ressorti une première visualisation de la tâche à accomplir.

Ceux-ci sont résumés dans le visuel suivant qui reprend les critères et sous-critères constitutifs de la norme ISO 19157.

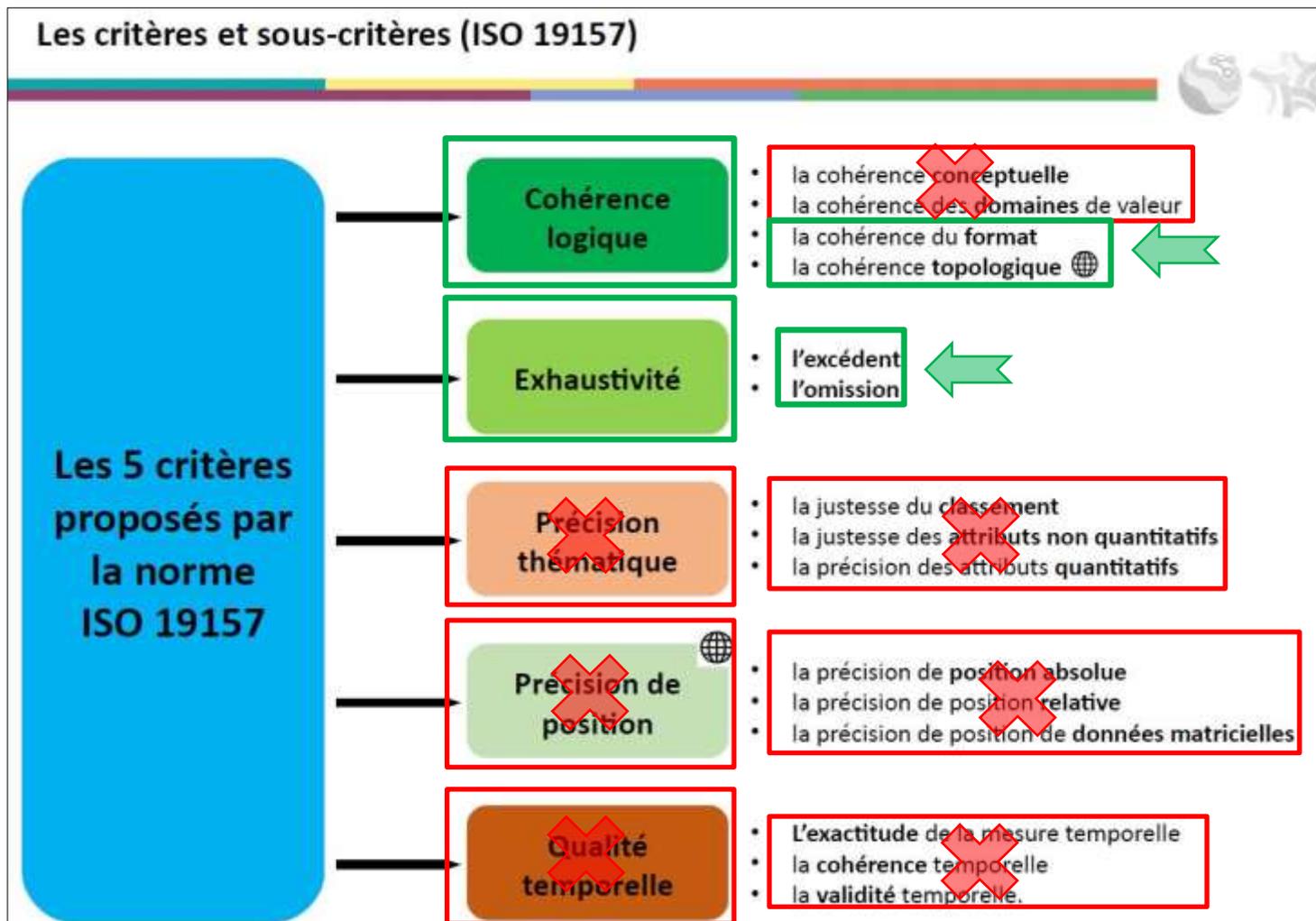


Fig. 11 : Les critères et sous critères de la norme ISO 19157 retenus pour l'élaboration du script

(Sources CEREMA 2019)



Critère ou sous critère volontairement écarté de la chaîne de qualification des données



Critère ou sous critère traité dans la chaîne de qualification des données



Notre réflexion s'est donc appuyée sur une base déjà structurée pour ses principes mais aussi sur une analyse moins formelle qui se présente ci-dessous dans sa forme originelle.

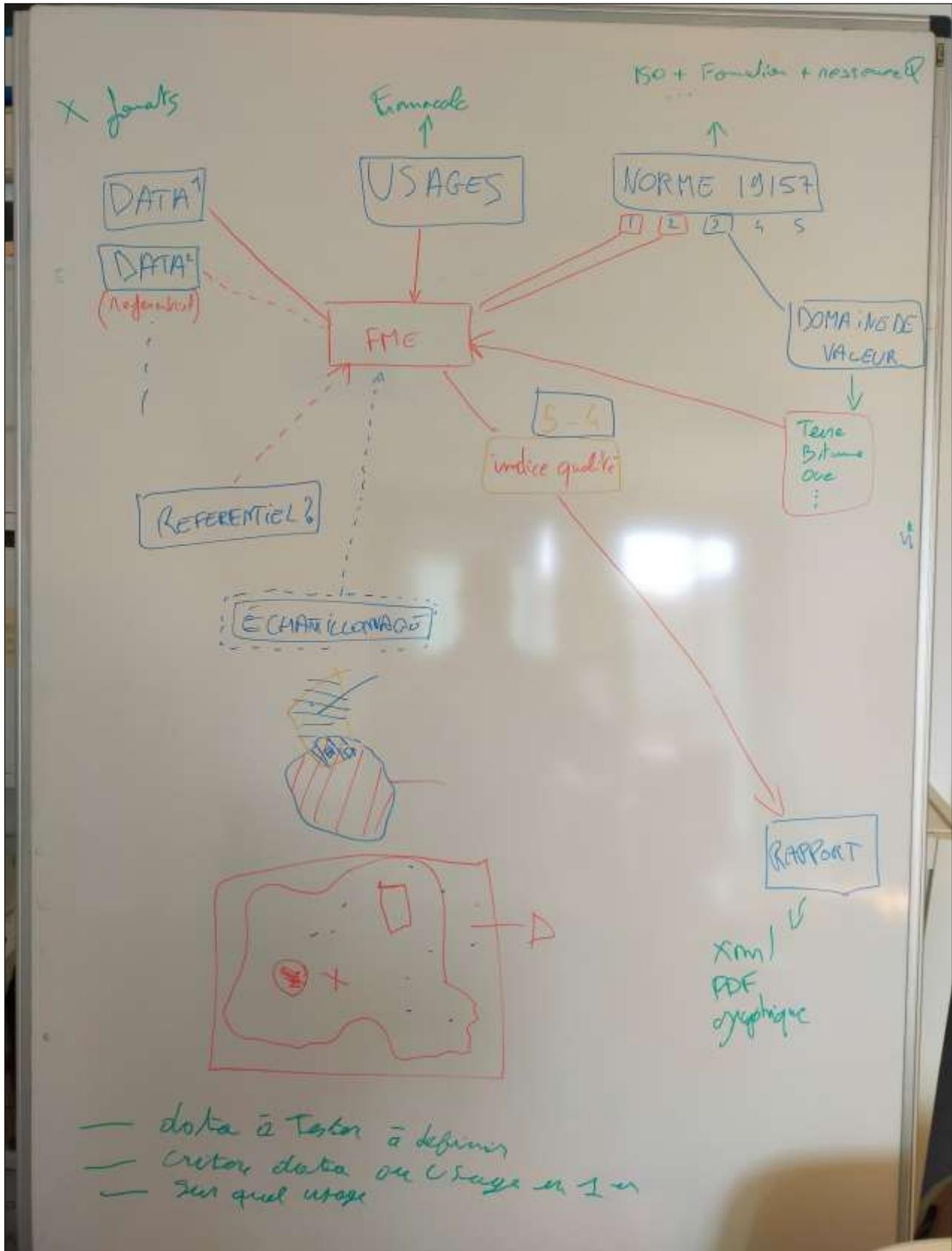


Fig. 12 : Base de la réflexion pour la conception du script

L'intérêt concret et immédiat pour le CRIGE et d'avoir une évaluation rapide de la qualité de la donnée basée sur des critères objectifs et quantifiables car dans la mesure où leur responsabilité sinon à minima leur réputation pourrait être engagée en cas de diffusion de données erronées ou inadaptées, le CRIGE a donc tout intérêt à vérifier en amont de leur mise à disposition aux publics privés et institutionnels les données qui lui sont transmises pour diffusion.

La forme de principe retenue pour visualiser les indices de qualité est un rapport qui s'inspire des rapports utilisés dans la grande distribution pour évaluer les biens de consommation courants tels qu'un téléphone portable ou un appareil photo. Les notes finales du rapport se situent donc dans une fourchette comprise entre 0 et 5.



Fig. 13 : Exemple de rapport de notation avec note globale et sous-critères d'évaluation.

MISE EN ŒUVRE D'UNE CHAÎNE DE CONTRÔLE QUALITÉ AUTONOME

Définition des principes directeurs mis en œuvre dans la chaîne de traitement

Pour rester fidèle aux valeurs du CRIGE-PACA et aux principes sous-jacents de l'open data la réalisation du script a été pensée et anticipée dès le début pour faciliter l'appropriation du programme et favoriser une reprise ou une poursuite de son écriture par une tiers personne. Par conséquent certaines instructions (transformers) sont désactivées ou ne sont pas indispensables à son fonctionnement dans son état actuel mais permettent une poursuite du travail facilitée.

Par ailleurs, le qualificateur de données ne se substitue pas au producteur de la donnée, il l'informe et évalue le fichier «SHAPE» proposé au qualificateur mais en aucun cas il ne corrige les éléments en erreurs que le fichier pourrait contenir. En cours ou en fin de processus selon le cas, le producteur reçoit un mail généré automatiquement par «FME» accompagné d'un rapport synthétique au format HTML et éventuellement des fichiers d'extension .shz mettant en évidence les éléments en défauts révélés par le processus de contrôle.

La méthode d'évaluation et certains critères relèvent du parti pris. Cependant, ce parti pris s'appuie toujours sur une réflexion et une analyse des données, c'est pourquoi toutes les notations et notamment les pondérations intermédiaires qui participent au calcul des notes finales sont identifiées accessibles et facilement modifiables dans le script.

La logique adoptée pour le processus de traitement

Pour traiter au mieux les deux critères principaux et leurs sous-critères associés :

- D'une part **LA COHERENCE LOGIQUE** composée des sous-critères

De *cohérence du format* et de *la cohérence topologique*

- Et d'autre part **l'EXHAUSTIVITE** caractérisée par l'*excédent* ou l'*omission* d'éléments,

le traitement s'articule au départ autour de trois axes d'analyses, ces trois «processus» principaux analysent d'une part la structure, la composition et la complétude du fichier «SHAPE» soumis au qualificateur et d'autre part la structure interne soit le contenu du fichier «SHAPE» en comparaison avec les emprises spatiales de références cohérentes avec le territoire de la région Provence – Alpes - Côte d'Azur. Les périmètres administratifs retenus pour cette comparaison peuvent être

- les communes
- les EPCI
- les départements
- la région PACA

Leurs définitions géométriques sont issues des couches "COMMUNE", "EPCI", "DEPARTEMENT" et "REGION" de la BD TOPO® (IGN).

Les contours au format «SHAPE» de ces échelons administratifs sont compilés dans le fichier «EMPRISES_SPATIALES_REF_ASS» et «FME» vient récupérer le contour sélectionné par l'utilisateur qui soumet son fichier au qualificateur.

Le processus de traitement travaille aussi sur le tri et l'analyse des géométries de bases qui peuvent constituer un fichier au format «SHAPE» soit :

- des polygones
- des lignes
- des points

Les trois processus de traitement, un pour chaque nature de géométrie, aboutissent à une évaluation sous la forme d'une note sur cinq qui est la résultante de critères et sous critères eux aussi évalués sur cinq.

Principe de fonctionnement du logiciel «FME»

L ETL «FME» édité par «SAFE SOFTWARE» est distribué en France par «VEREMES», il comprend plusieurs modules distincts.

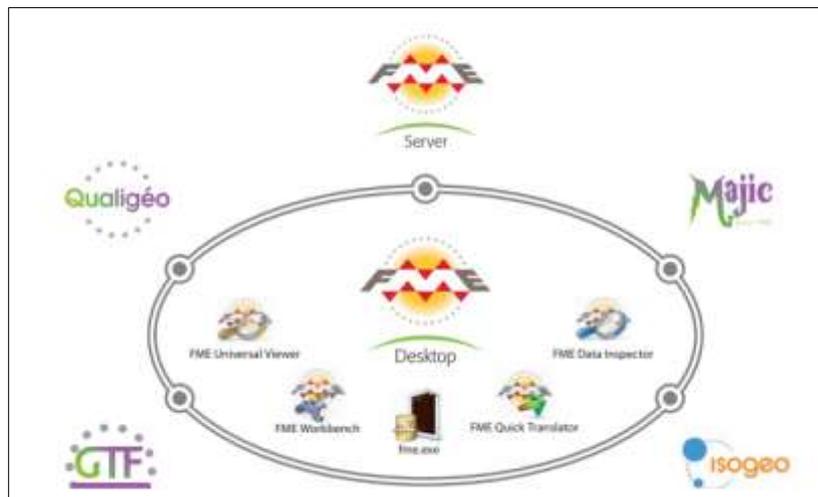


Fig. 14 : Illustration des différents modules de «FME»

(Sources livre blanc VEREMES)

Ce logiciel est compatible avec plusieurs centaines de formats de fichiers et fonctionne à partir d'éléments de bases appelés «transformers», il en existe plus de 500, personnellement, dans cette chaîne de traitement, je n'ai utilisé que le module «FME Workbench» et pour réaliser ce projet, je n'ai utilisé que trente «transformers» différents soit 6%. Chacun de ces «transformers» doit être paramétré pour que les données d'amont et d'aval soient compatibles et composent une chaîne de traitement cohérente dans sa globalité.

«FME» permet également d'exécuter différents codes dans un autre langage comme «PYTHON» ou langage system.

Aggregator	FeatureWriter
AreaCalculator	GeometryFilter
AttributeCreator	GeometryValidator
Attribute Keeper	HTMLReportGenerator
AttributeManager	LengthCalculator
Attribute Remover	ListBuilder
Attribute Renamer	ListConcatenator
Attribute Rounder	ListExploder
Clipper	ListSearcher
CoordinateSystemDescriptionConverter	Matcher
CoordinateSystemExtractor	StatisticsCalculator
Counter	SystemCaller
CRCCalculator	Terminator
Deaggregator	Tester
Emailer	TestFilter
ExpressioEvaluator	

Fig. 15 : Liste des «transformers» utilisés

Ces «transformers» capables de traiter des données vectorielles, des points, des nuages de points et encore bien d'autres natures de données héritent des données transmises par les autres «transformers» situés en amont dans le processus de traitement réalise une action sur la donnée et la transmettent à leur tour aux autres «transformers» situés en aval de la chaîne.

Cet outil puissant et très personnalisable permet d'automatiser un grand nombre de tâches.

Ainsi un script répondant au même objectif réalisé par deux personnes différentes pourra remplir sa mission de manière aussi performante mais la façon dont est écrite le script, les étapes et enchainements de «transformers» pourront être très différents selon le concepteur du script.

Un transformer peut calculer, créer, transformer, supprimer, fusionner, tester, filtrer, renommer, et réaliser encore bien d'autres action sur la donnée.

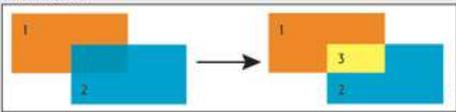
AreaOnAreaOverlay <i>Analyse spatiale Filtres et jointures</i>	Calcule l'intersection généralisée d'entités polygonales. L'intersection entre toutes les entités surfaciques entrant dans le Transformer est calculée et toutes les entités résultantes sont créées et envoyées vers le port Output. Les polygones générés héritent des attributs des entités d'origine dans lesquelles ils sont inclus. 
AttributeCompressor <i>Attributs Chaînes de caractères</i>	Comprime et éventuellement crypte les valeurs des attributs spécifiés.
AttributeCopier <i>Attributs</i>	Crée de nouveaux attributs en leur affectant la valeur d'attributs existants. Les attributs existants demeurent et un nouvel attribut est créé avec un nom différent et la même valeur.
AttributeCreator <i>Attributs</i>	Ajoute de nouveaux attributs à l'entité et leur affecte des valeurs constantes, des valeurs d'attributs ou des expressions. Les valeurs peuvent se référer à des entités adjacentes.
AttributeDecompressor <i>Attributs</i>	Décompresse et décrypte les valeurs d'attributs spécifiés ayant été préalablement compressés par AttributeCompressor.
AttributeDereferencer <i>Attributs</i>	Copie la valeur de l'attribut dont le nom se trouve dans l'attribut source vers un attribut nouvellement créé.
AttributeEncoder <i>Chaînes de caractères Web</i>	Encode les valeurs d'attributs définis en paramètres.
AttributeExploder <i>Attributs Chaînes de caractères</i>	Crée une nouvelle paire d'attribut (nom d'attribut / valeur d'attribut) à partir de chaque attribut de l'entité entrante. Ils sont soit orientés vers une nouvelle entité, soit ajoutés en tant qu'élément de liste à l'entité originale. Dans les deux cas, il est possible de conserver ou de supprimer les attributs et la géométrie de l'entité d'origine.
AttributeExposer <i>Attributs</i>	Expose des attributs cachés dans Workbench de manière à ce qu'ils puissent être utilisés par les Transformers et Writers suivants.
AttributeFileReader <i>Attributs Exécution Spécifique à un format</i>	Lit le contenu d'un fichier et stocke le contenu dans un attribut dont le nom est défini en paramètre.
AttributeFileWriter <i>Attributs Exécution Spécifique à un format</i>	Ecrit le contenu de l'attribut spécifié dans un fichier.
AttributeFilter <i>Filtres et jointures</i>	Dirige les entités vers différents ports de sortie en fonction de la valeur d'un attribut.
AttributeKeeper <i>Attributs</i>	Supprime tous les attributs et attributs de liste à l'exception de ceux spécifiés par l'utilisateur qui sont alors conservés.
AttributeManager <i>Attributs</i>	Modifie de multiples attributs en les renommant, copiant, supprimant, triant ou ajoutant. Définit des valeurs pour les attributs existants, nouveaux et modifiés selon n'importe quelle combinaison de constantes, valeurs d'attributs, conditions, expressions et paramètres. Les valeurs peuvent se référer à des entités adjacentes.
AttributePivoter <i>Chaînes de caractères Filtres et jointures Valeurs calculées</i>	Restructure et regroupe les entités en entrée selon un attribut de regroupement et calcule des statistiques pour former un tableau croisé dynamique en sortie.

Fig. 16 : Extrait du guide des transformers FME 2020- illustration des multiples possibilités de traitement de la donnée

(Sources guide des transformers 2020)

Plusieurs «Transformers» peuvent également être rassemblés dans un signet afin de créer une chaîne de traitement thématique ou un morceau de la chaîne de traitement globale ayant une fonction, une cohérence, un objet commun (exemple : établir la projection d'un fichier).

Le signet même s'il n'est pas indispensable au fonctionnement du script permet de rassembler les «transformers» et de créer des ensembles ou des parties de chaînes de traitement cohérentes, enfin et surtout l'utilisation des signets permet au créateur du script de se retrouver dans l'enchaînement des traitements en créant une architecture graphique dans laquelle il est plus facile de naviguer par rapport à l'enchaînement des traitements intermédiaires.

«FME» propose également plusieurs fenêtres de travail avec une vue simultanée des données attributaires et de leur représentation graphique.



Fig. 17: Illustration des différentes fenêtres de travail de «FME»

Description et fonctionnement du processus de traitement de la donnée.

Les données d'entrée

Tout d'abord l'utilisateur du qualificateur renseigne les éléments suivants :

- Nom du demandeur (*entrée sans contraintes particulière*)
- L'adresse Email du demandeur sert à envoyer les mails à l'adresse désirée (*l'adresse doit être valide et fonctionnelle*).
- Le répertoire contenant le «SHAPE» à qualifier sert pour l'analyse de la structure (*les noms de fichiers et chemin supérieurs à 256 caractères longs sont à proscrire*).
- Le fichier «SHAPE» à qualifier sert pour l'analyse du contenu. (*les noms de fichiers et chemin supérieurs à 256 caractères longs sont à proscrire*)
- L'emprise spatiale de référence parmi la liste proposée sert à définir l'emprise de travail cohérente avec le fichier à qualifier (*choix unique parmi dans une liste déroulante*).

Fig. 18 : Visualisation du masque de saisie des informations pour lancer le traitement.

En complément du script, deux notices d'utilisation ont été rédigées :

- une première pour guider l'utilisateur inexpérimenté dans l'utilisation de l'outil d'analyse de la qualité
- une seconde pour les futurs administrateurs du script, le but de cette notice est de faciliter l'adaptation, la reprise ou l'enrichissement du script existant.

Condition préalable à l'exécution du script :

Le répertoire qui contient le fichier à qualifier ne doit contenir que celui-ci

```

D:\F1 (D) > test_qualifier

Nom
  test_qualifier.cpp
  test_qualifier.dxf
  test_qualifier.prj
  test_qualifier.qsd
  test_qualifier.shp
  test_qualifier.xls
  
```

- 1 Renseignez l'identité du demandeur directement dans les champs dédié
- 2 Entrez l'adresse mail à laquelle vous voulez que soit transmis le rapport et les pièces jointes. Directement dans le champs dédié
- 3 Choisissez le répertoire dans lequel se trouve le fichier à qualifier en cliquant sur
- 4 Choisissez le fichier .shp à analyser en cliquant sur
- 5 Vous accédez à la liste de choix des emprises spatiales en cliquant sur
Puis choisissez dans cette boîte de dialogue l'emprise qui vous intéresse.
Vous avez le choix entre les 946 Communes, les 51 EPCL, le 6 départements ou la région PACA

CRIGE PACA - Domaine du Petit Arbois - Bât. Lannec 13100-Aix-en-Provence - Tel : 04 42 90 71 22 - Mail : contact@crige-paca.org

Page 2 sur 3

Fig. 19 : Extrait de la notice d'utilisation du script pour un utilisateur basique.

Les trois chaînes de traitement principales

Les trois chaînes de traitement répondent respectivement à une fonction particulière soit :

- l'analyse du contenu du «SHAPE»
- l'analyse de la structure du «SHAPE»
- l'analyse du fichier compilant les emprises spatiales de référence

leurs traitements s'effectuent en parallèle et concomitamment sous «FME». En réalité les traitements semblent s'opérer simultanément mais en fait dans le détail «FME» travaille branches par branches. Cet étalement du traitement dans le temps ne peut se constater qu'avec des fichiers volumineux.

Ces trois chaînes principales font appel à des données d'entrées spécifiques

Analyse des emprises spatiales de références

Le point d'entrée pour cette analyse est le fichier d'extension .shp correspondant aux différentes emprises administratives compilées dans un fichier unique «EMPRISES_SPATIALES_REF_ASS».

Analyse de la structure du «SHAPE»

La structure du «SHAPE» est analysée directement dans le répertoire de dépôt des données en relevant la présence ou l'absence des différents formats de fichiers potentiellement constitutifs d'un «SHAPE». Le répertoire de dépôt est défini par l'utilisateur.

Analyse de la structure interne du «SHAPE»

La structure interne est analysée à partir du fichier d'extension .shp, «FME» récupère ensuite dans les autres fichiers constitutifs du «SHAPE» les autres informations associées.

Le point d'entrée pour cette analyse est le fichier d'extension .shp.

La chaîne de traitement des emprises spatiales de références

L'utilisateur choisit une emprise unique dont le contour géométrique est défini dans le fichier des emprises «EMPRISES_SPATIALES_REF_ASS» qui compile toute les emprises possibles cohérentes avec le territoire de la région PACA.

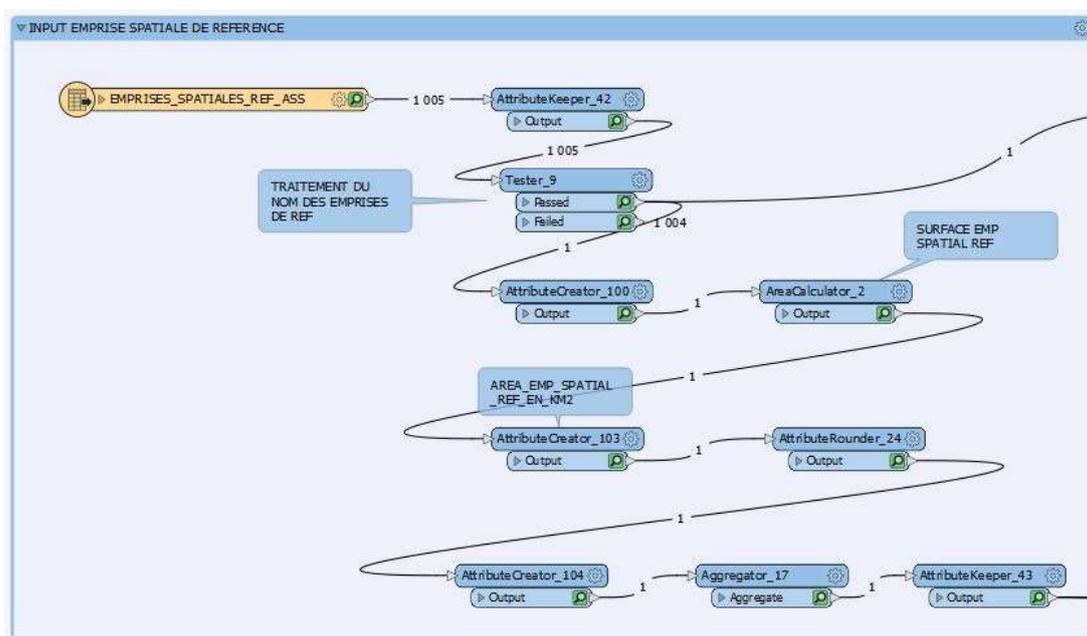


Fig. 20 : Visualisation du signet traitant l'emprise spatiale de référence

Un fichier d'extension .shz est créé à partir de l'emprise unique choisie par l'utilisateur parmi les 946 communes, les 51 EPCI, les 6 départements ou la région.

Ce fichier d'emprise permet à l'utilisateur après sa réception de corriger les éventuelles erreurs ou défauts qui lui seront transmis par mail.

Les détails de ce processus seront développés plus loin dans ce document.

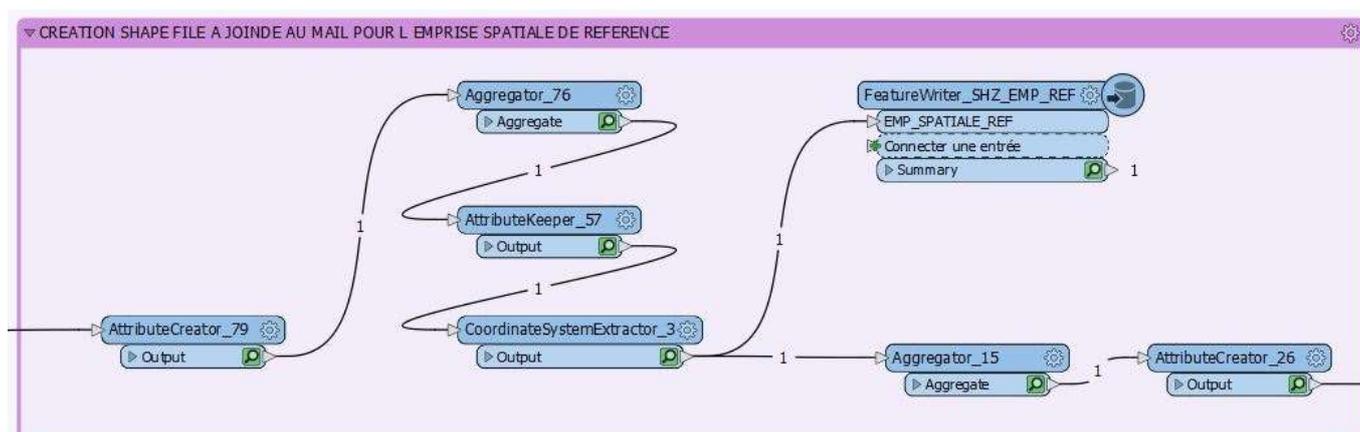


Fig. 21 : Visualisation du signet de création du fichier .shz de l'emprise spatiale de référence joint aux mails

La chaîne de traitement de la structure du «SHAPE»

«FME» fait un état des fichiers présents dans le répertoire de dépôt du «SHAPE» défini par l'utilisateur. Il calcule également le poids des données transmises et établit un diagnostic sur la structure du «SHAPE» pour définir un statut de validité.

La structure du «SHAPE» peut être :

- complète et valide
- incomplète et valide
- incomplète et invalide (de fait dans la version actuelle du script, ce statut n'apparaîtra jamais dans le rapport de synthèse car un statut invalide stoppe le traitement avant son terme et déclenche automatiquement l'envoi d'un mail d'erreur)

Ce diagnostic s'appuie sur la somme des notes individuelles attribuées pour chaque type d'extensions présent. En fonction du résultat de la somme obtenue le texte associé au score est alors validé pour la suite du traitement.

Toutes les extensions n'ont pas la même importance, certains fichiers sont plus indispensables que d'autres.

aih	fbx	sbn
ain	ixs	sbx
atx	mxs	shp
cpg	prj	shp.xml
dbf	qix	shx
fbn	qmd	

Fig. 22 : Liste des extensions testées pour l'analyse de la structure du «SHAPE»

Les extension en rouge sont indispensables et l'absence d'une seule d'entre elles déclenche l'envoi automatique d'un mail au producteur de la donnée pour l'informer du caractère invalide du «SHAPE».

Les extension en vert sont importantes mais non indispensables leur absence conditionne le statut incomplet du «SHAPE». Les autres extensions de fichier sont plus rares, elle sont donc recensées, évaluées et prise en compte pour le calcul du **sous-critère de cohérence du format** mais n'interviennent pas pour statuer sur le caractère incomplet du «SHAPE».

L'administrateur du script peut affecter des pondérations à la notation individuelle de chaque fichier cela permet de personnaliser l'évaluation des fichiers en fonction des besoins, des thématiques et des problématiques rencontrées.

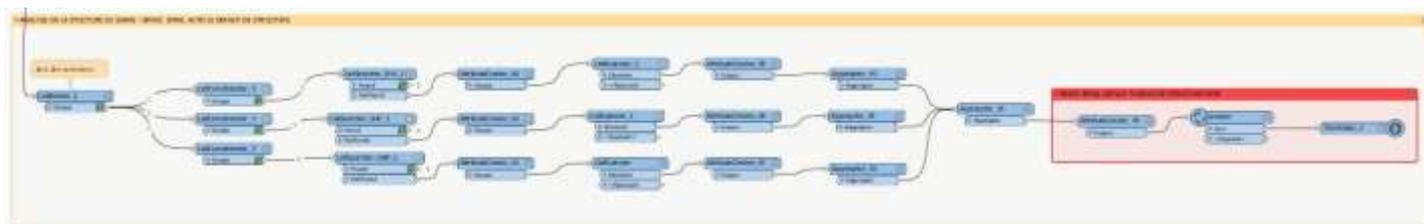


Fig. 23 : Visualisation du signet testant la validité de la structure du «SHAPE»

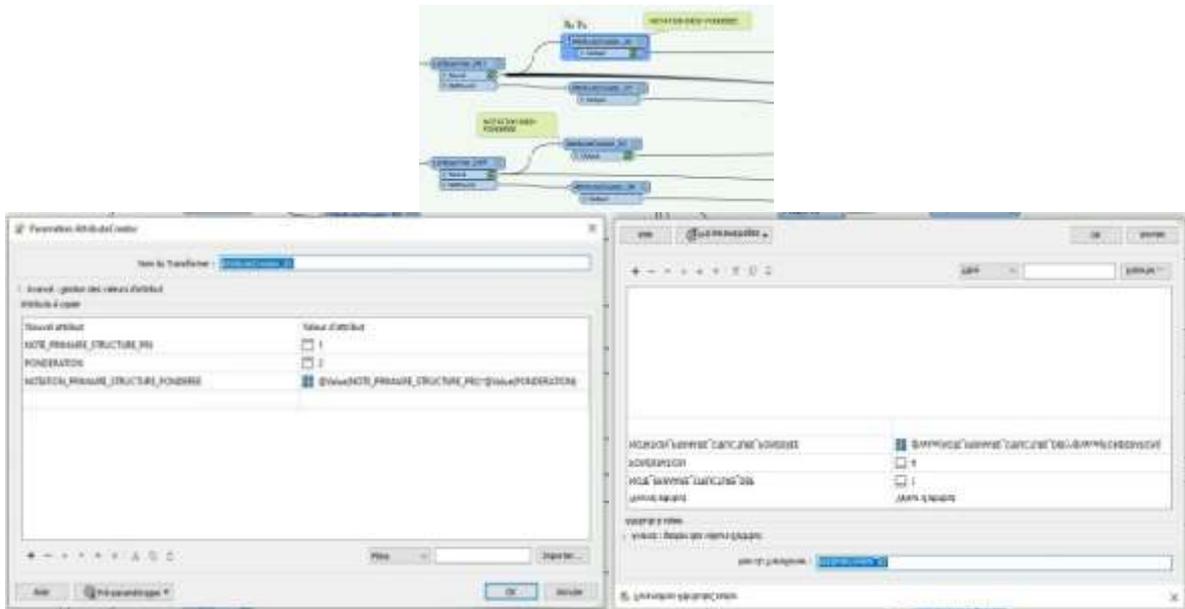


Fig. 24 : Visualisation des «transformers» et de leur paramétrage pour la notation individuelle des fichiers constitutifs du «SHAPE»

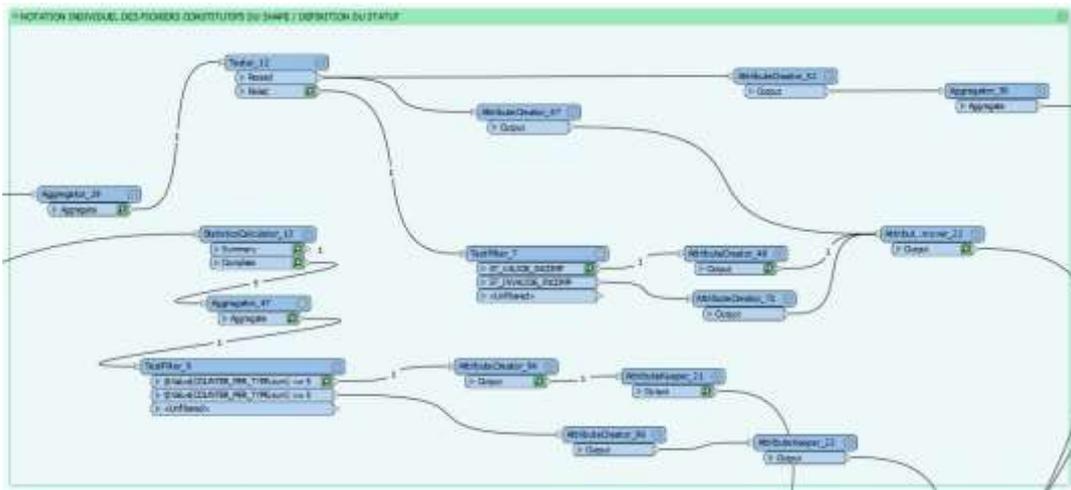


Fig. 25 : Visualisation du signet statuant définissant le statut du «SHAPE»



Fig. 26 : Visualisation du message affiché par «FME» en cas d'absence de fichier indispensable au «SHAPE»

Le poids total des données du fichier «SHAPE» à qualifier est également établi par addition des poids individuels des fichiers.



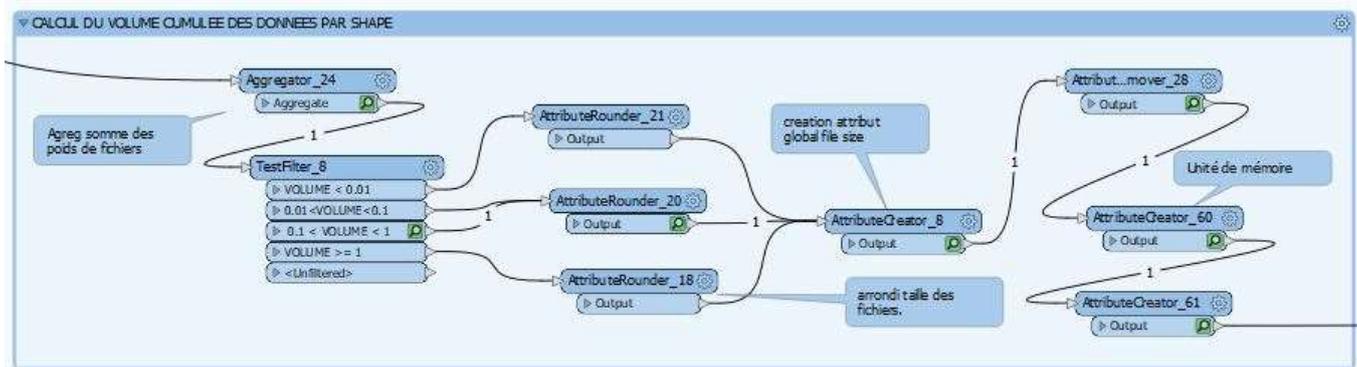


Fig. 27 : Visualisation du signet calculant le poids des données transmises

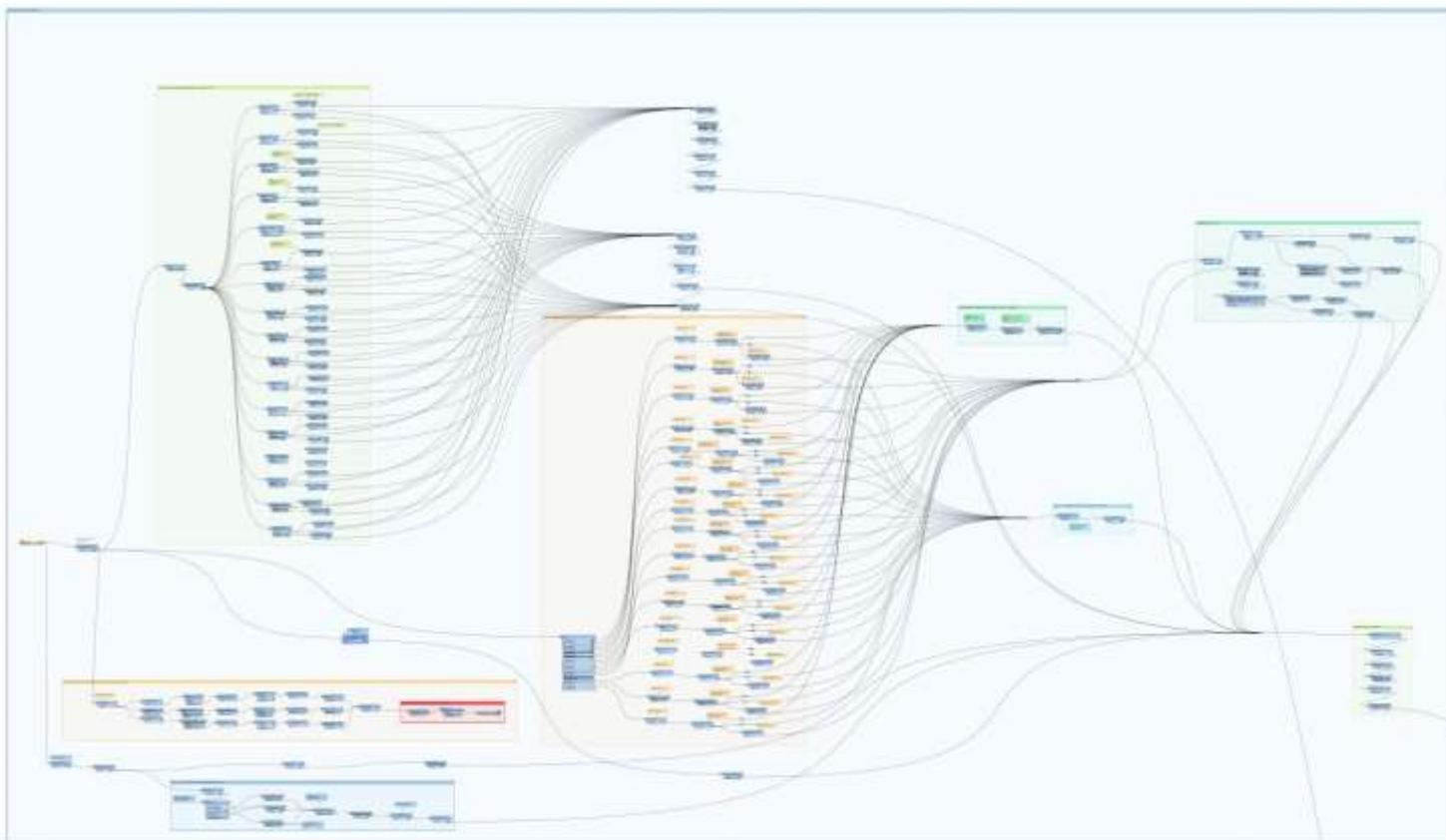


Fig. 28 : Visualisation globale du signet traitant la structure du «SHAPE»

SOUS CRITÈRE DE COHÉRENCE DU FORMAT = 5 sur 5		
Nature du contrôle	Modalité du contrôle	Résultat
Poids des fichiers	Calcul du poids des données transmises	0.53 Mo
Nature du contrôle	Modalité du contrôle	Résultat
Liste des fichiers et type d'extension	contrôle sur le type d'extension	test_sourçage_EPSR_2154.dwg test_sourçage_EPSR_2154.gxd test_sourçage_EPSR_2154.qxd test_sourçage_EPSR_2154.shp test_sourçage_EPSR_2154.shx
Nature du contrôle	Modalité du contrôle	Résultat
Structure de fichier	Vérification du format de fichier	Les fichiers d'extension ".cpg" ".shp" ".shx" ".dbf" ".gdb" ".gdbx" ".sbn" ".sbnx" ".mxd" ".mxdx" ".mxd" sont absents
	La présence des fichiers obligatoires ".shp", ".shx" et ".dbf" est testée et l'absence des fichiers ".gdb" ou ".cpg" est relevée pour définir la complétude et la validité du "SHAPE"	La structure du "SHAPE" est valide mais incomplète

Fig. 29 : Extrait du rapport du rapport de synthèse final - partie dédiée au sous critère de cohérence du format

La chaîne de traitement de la structure interne du «SHAPE»

Cette troisième chaîne concentre le gros du traitement en effet, elle est la plus importante et la plus complète car outre le fait qu'elle hérite du flux de donnée et des indicateurs des deux chaînes précédentes pour les incorporer dans le calcul des sous-critères et critères finaux, elle traite toute la données de la structure interne du fichier «SHAPE».

Une stratégie de traitement successif à été mise en application pour écarter les données indésirables ou en défauts au regard des paramètres définis dans le script. Tant que la donnée testée répond aux exigences que l'on s'est fixé le traitement continu jusqu'au calcul de la notation final. En fonction de l'importance de l'erreur ou du défaut de la donnée, le processus de traitement peut être interrompu avant le terme du traitement. Le producteur en est alors informé par mail.

Le principe de calcul des notations consiste en un calcul de moyennes à partir d'autres moyennes.

Principe de calcul de la notation globale

La notation globale est la moyenne arithmétique sans pondération du critère d'exhaustivité et du critère de cohérence logique. Elle est exprimée sous la forme d'une note globale sur 5.

Principe de calcul du critère de cohérence logique

La notation du critère de cohérence logique est le résultat de la moyenne arithmétique sans pondération entres le sous-critère de cohérence de format et le sous-critère de cohérence topologique. Cette notation intermédiaire est exprimée sous la forme d'une note sur 5.

Définition du sous-critère de cohérence de format.

Il est calculé selon un score qui est le résultat de la somme des notes individuelles affectées à chaque fichiers constitutifs du «SHAPE» à qualifier. Ce score est ensuite rapporté à une note sur cinq.

Définition du sous-critère de cohérence de topologique

Il est défini comme la moyenne arithmétique sans pondération des notations de la vérification de l'intégrité des géométries, de la recherche des géométries en doublons et de la notation des micro-géométries. Il s'exprime sous la forme d'une note sur cinq.

Notation de l'intégrité géométrique

Cette notation sur cinq correspond à la conversion du ratio exprimé en pourcentage des géométries valides du fichier à qualifier

Notation des doublons

Cette notation sur cinq correspond à la conversion du ratio exprimé en pourcentage des géométries en doublons du fichier à qualifier.

Notation des micro-polygones

Cette notation sur cinq correspond à la conversion du ratio exprimé en pourcentage des micro-géométries définies comme telles en fonction des valeurs de seuils définies.



Principe de calcul du critère d'exhaustivité

Le **critère d'exhaustivité** est le résultat unique de la notation de la relation spatiale entre le fichier et l'emprise spatiale de référence choisie. Il est exprimé sous la forme d'une note intermédiaire sur 5.

Tous les détails et formules de calcul sont explicités et développés dans la partie suivante du rapport traitant les résultats leur mise en forme et leur communication.

QUALIFICATION DES DONNEES : RESULTATS ET AXES D'AMELIORATION

Application de la chaîne de traitement, des résultats cohérents et pertinents

Préparation des répertoires de travail pour une nouvelle analyse

«FME» commence son traitement, par l'exécution d'une commande système pour supprimer les fichiers contenus dans certains répertoires définis par l'administrateur dans les paramètres privés. Cette opération garantit que les fichiers envoyés par les mails automatiques correspondent au fichier à qualifier.

En effet, en fonction des types de géométries présents dans le «SHAPE» et des erreurs rencontrées lors des traitements précédents sur d'autres fichiers, des fichiers résiduels non cohérents avec le fichier à qualifier pourraient alors être indument envoyés. Cette commande permet une remise à zéro des fichiers créés par «FME» à l'occasion de chaque traitement.

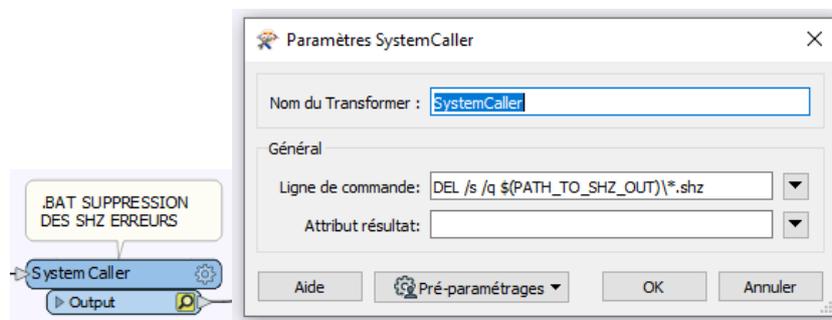


Fig. 30 : Visualisation du transformeur «SystemCaller» et de son paramétrage



Fig. 31 : Visualisation du répertoire de travail avant exécution du «SystemCaller»



Fig. 32 : Visualisation du répertoire de travail après exécution du «SystemCaller»

Le répertoire a été purgé et il est prêt à recevoir les fichiers d'erreurs du traitement à venir.

Analyse et validation ou rejet de la projection utilisée

Le processus analyse alors si la projection du fichier à qualifier est cohérente avec le territoire de la région PACA. Les seules projections autorisées sont :

- RGF 93 / Lambert 93 / Code EPSG = 2154
- RGF 93 / Conique Conforme 43 (Zone 2) / Code EPSG = 3943
- RGF 93 / Conique Conforme 44 (Zone 3) / Code EPSG = 3944
- RGF 93 / Conique Conforme 45 (Zone 4) / Code EPSG = 3945

Si la projection utilisée correspond à l'un de ces choix, le traitement se poursuit sinon le processus de traitement s'interrompt et envoi un mail automatique.

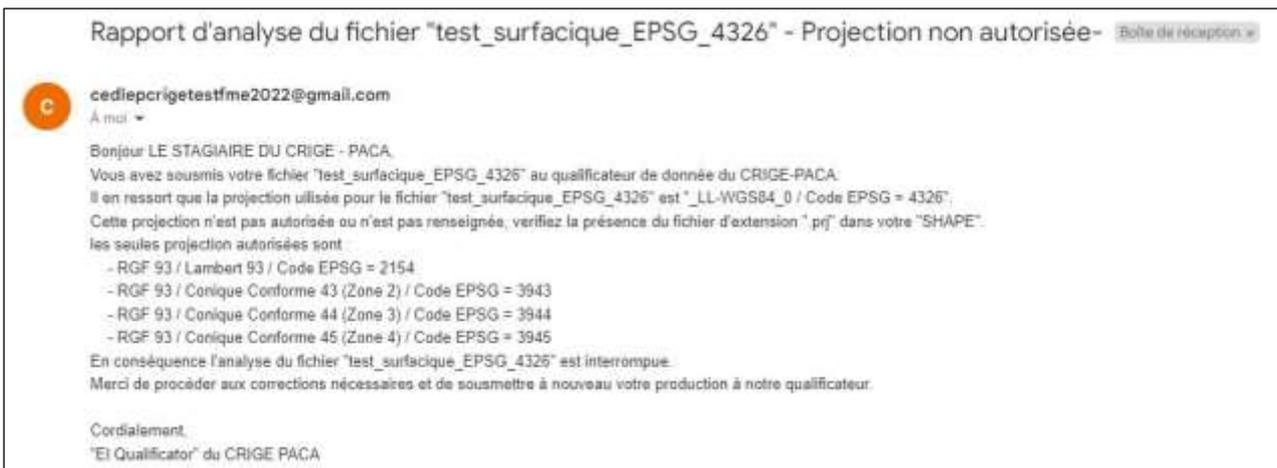


Fig. 33 : Visualisation du mail envoyé automatiquement en cas de mauvaise projection

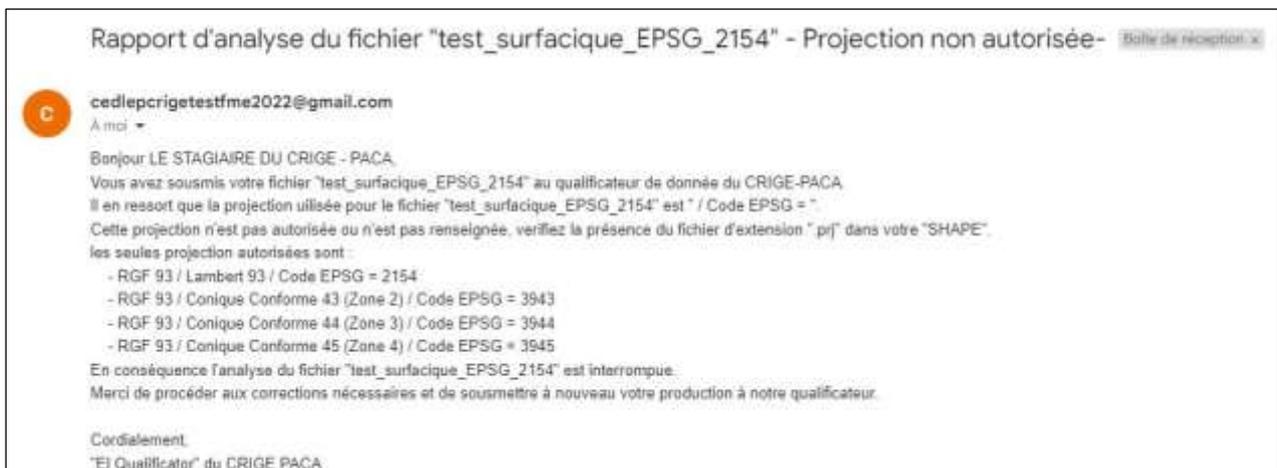


Fig. 34 : Visualisation du mail envoyé automatiquement en cas d'absence de projection

Tout au long du traitement, il est possible de suivre l'évolution du traitement et l'avancement des tâches programmées dans la fenêtre des logs de traitement. En cas de projection inappropriée, «FME» affiche les messages suivant:

```

Log du traitement
12 Erreurs 4 Alertes Information
1679 Terminator_3 (TestFactory): Terminator_3: Termination Message: 'TRAITEMENT INTERROMPU POUR CAUSE DE MAUVAISE PROJECTION UTILISEE.
1680 LA PROJECTION UTILISEE EST_LL-WGS84_0.
1681 LES SEULES PROJECTIONS AUTORISEES SONT L93 / CC43 / CC44 / CC45.
1682 UN MAIL D INFORMATION AU PRODUCTEUR DU FICHIER A ETE ENVOYE A cedlepcrigetestfme2022@gmail.com.'
1683 A fatal error has occurred. Check the logfile above for details
1684 PythonFactory output feature was rejected downstream. Check the logfile above for details
1685 A fatal error has occurred. Check the logfile above for details
  
```

Fig. 35 : Visualisation du message affiché par «FME» en cas de projection non autorisée

```

Log du traitement
12 Erreurs 4 Alertes Information
1635 Terminator_3 (TestFactory): Terminator_3: Termination Message: 'TRAITEMENT INTERROMPU POUR CAUSE DE MAUVAISE PROJECTION UTILISEE.
1636 LA PROJECTION UTILISEE EST .
1637 LES SEULES PROJECTIONS AUTORISEES SONT L93 / CC43 / CC44 / CC45.
1638 UN MAIL D INFORMATION AU PRODUCTEUR DU FICHIER A ETE ENVOYE A cedlepcrigetestfme2022@gmail.com.'
1639 A fatal error has occurred. Check the logfile above for details
  
```

Fig. 36 : Visualisation du message affiché par «FME» en cas d'absence de projection- absence de fichier .prj

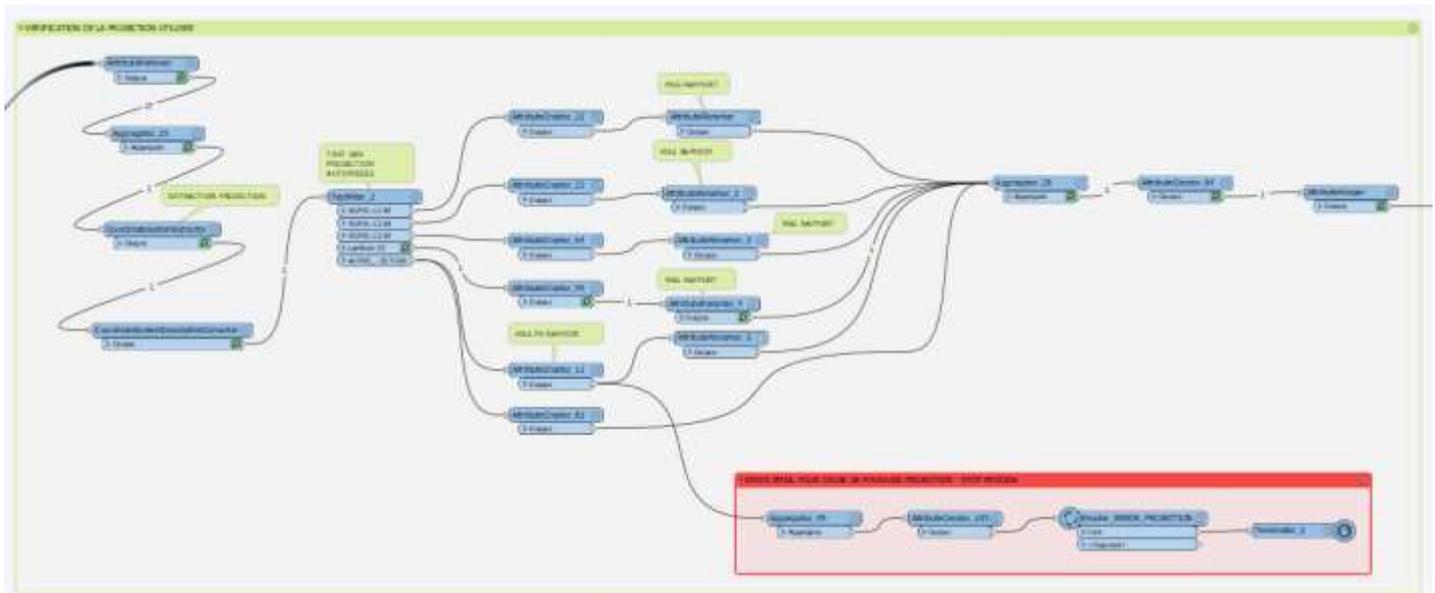


Fig. 37 : Visualisation du signet statuant sur la validité de la projection utilisée dans le «SHAPE»

Analyse des informations générales du «SHAPE»

Le nom du fichier ainsi que l'horodatage du traitement sont extraits du «SHAPE» et sont communiqués dans le rapport de synthèse final joint au mail d'information transmis à l'utilisateur du qualificateur.

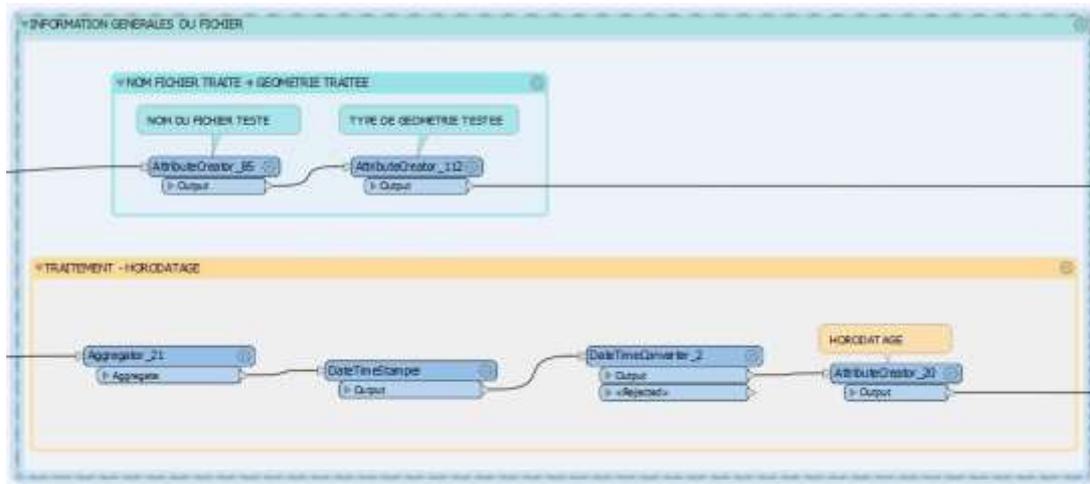


Fig. 38 : Visualisation du signet traitant les informations générales du «SHAPE»



Crige
Centre de Ressources
en Information Géographique
Provence-Alpes-Côte d'Azur

RAPPORT D'ANALYSE DU FICHIER "test_surfacique_EPSG_2154"

INFORMATIONS GÉNÉRALES SUR LE FICHIER

Demandeur :	LE STABIANW DU CRIGE - PACA	Mail :	cedd@unigetext@red22@gmail.com
Fichier analysé :	test_surfacique_EPSG_2154		
Date / Heure du traitement :	Le 21/06/2022 à 14h07		
Seuils de calcul :	Longueur = 3 en (kilomètre(s))	Surface = 3 en (mètre(s) ²)	
Emprise spatiale de référence choisie (1) :	"EPSG 00 de l'Office de Saint-Tropez"	Surface = 425,3 km ²	
Projection du fichier :	REP27 / Lambert 93 / Code EPSG = 2154		
Type de géométrie analysé :	POLYGOONE		
Nombre de POLYGOONE(s) analysé(s) :	6		

(1) Les emprises spatiales de référence sont issues des couches "COMMUNE", "EPCI", "DEPARTEMENT" et "REGION" de la BD TOPO (IGN). Par conséquent selon le source des emprises spatiales de référence utilisé pour votre fichier, des erreurs peuvent se présenter dues à une différence de précision entre vos données et l'emprise spatiale de référence choisie.

Note : En cas d'erreurs sur les géométries, les fichiers d'erreurs d'extension ".shp" sont joints au mail.

Fig. 39 : Extrait du rapport de synthèse final - partie dédiée aux informations générales

Les seuils de calcul des surfaces et des longueurs sont gérés et accessibles uniquement par l'administrateur du script, ils peuvent être ajustés pour correspondre aux différents fichiers proposés et surtout pour mettre en évidence des réalités, des thématiques et des problématiques qui peuvent être très différentes.

Dénombrement des géométries en entrée

Le nombre de géométries en entrée est recensé, cette information est également communiquée dans le rapport de synthèse, Elle participe au calcul des statistiques qui seront présentées plus loin dans ce rapport.

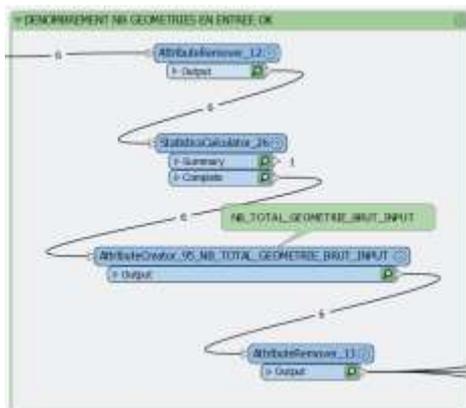


Fig. 40 : Visualisation du signet dénombrant les géométries en entrée

Validation ou rejet des géométries présentes

La vocation des données à être publiée implique que celles-ci répondent a minima à certains critères de validité, dans cet esprit, la validité des géométries que ce soient des polygones, des lignes ou des points doit être établie. Bien que le transformer propose une réparation, les géométries défectueuses ne sont pas corrigées. Le producteur de la donnée est et doit être responsable de son fichier et le CRIGE – PACA n’as pas vocation à se substituer aux producteurs.

Cette étape est très importante dans la mesure où elle peut conditionner l’arrêt du traitement en cas de géométries en défaut. Pour ma part, j’ai opté pour une poursuite du traitement car toutes les géométries peuvent ne pas être en défaut. Le «transformer» «GeometryValidator» fait le tri entre les géométries valides et les géométries en défaut.

Les géométries valides poursuivent le processus de traitement et les géométries en défauts sont dirigées vers le «transformer» «FeatureWriter».

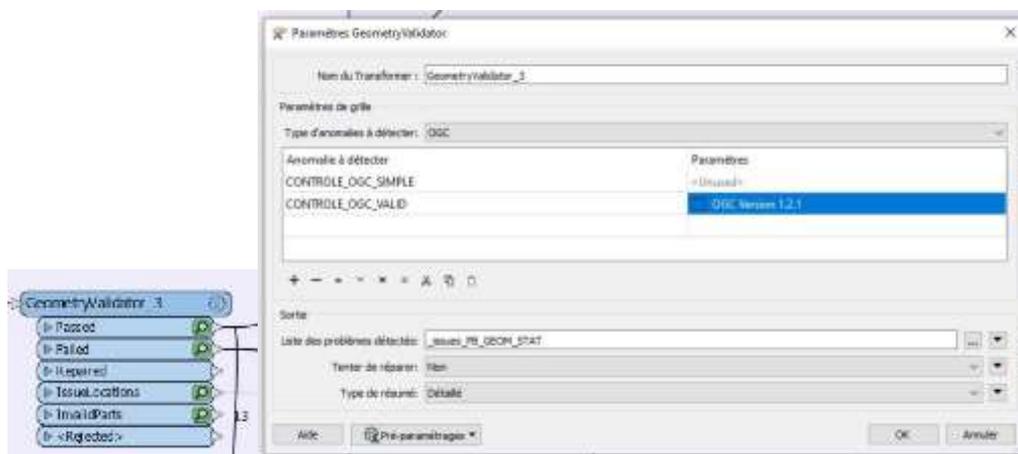


Fig. 41 : Visualisation du «transformer» «GeometryValidator» et de son paramétrage

Le contrôle de la géométrie réalisé par le logiciel s'appuie sur la norme OGC Version 1.2.1

▼ Conforme à l'OGC

Une géométrie est "conforme à l'OGC" si elle peut être représentée sans perte dans le modèle géométrique défini par l'OGC.

Voir <http://www.opengeospatial.org/standards/sfa> pour plus de détails sur les géométries qui peuvent être représentées dans le modèle géométrique Simple Feature Access de l'OGC.

Le paramètre *OGC Version* spécifie la version de la norme à utiliser.

Fig. 42 : Extrait de l'aide en ligne de «FME» sur la norme OGC

Géométries dégénérées ou corrompues	pas de NaN ou de valeurs infinies en entrée
Auto-intersections en 2D	pas de NaN ou de valeurs infinies, de données dégénérées ou corrompues en entrée
Astuce : pour éviter un comportement indéfini, sélectionner Géométries dégénérées ou corrompues et Contient des NaN ou infini avant de sélectionner Auto-Intersection en 2D.	
Surfaces non planes	pas de NaN ou de valeurs infinies en entrée
Bordures de solide invalides ou Vides dans solide invalides	pas de NaN ou de valeurs infinies, de doublons, de données dégénérées, corrompues ou d'auto-intersections dans les faces individuelles en entrée.
Vides dans solides invalides	pas de bordures de solide invalides
Dépendances de sortie des anomalies, en supposant que les dépendances d'entrée sont respectées	
Réparer les anomalies suivantes peut produire d'autres anomalies :	
Contient des -0, NaN ou des valeurs infinies	peut produire en sortie une donnée dégénérée
Géométries dégénérées ou corrompues	peut produire en sortie une donnée auto-intersectée
Contient des parties géométriques nulles	peut produire en sortie une donnée dégénérée
Points consécutifs dupliqués (doublons)	peut produire en sortie une donnée dégénérée
Auto-intersections en 2D	peut produire en sortie une donnée dégénérée
Bordures de solide invalides	peut produire des faces dégénérées en bordure du solide

Fig. 43 : Extrait de l'aide en ligne de «FME» sur les géométries corrompues détectées

La topologie doit être conforme et respecter les principes du document ci-dessous.

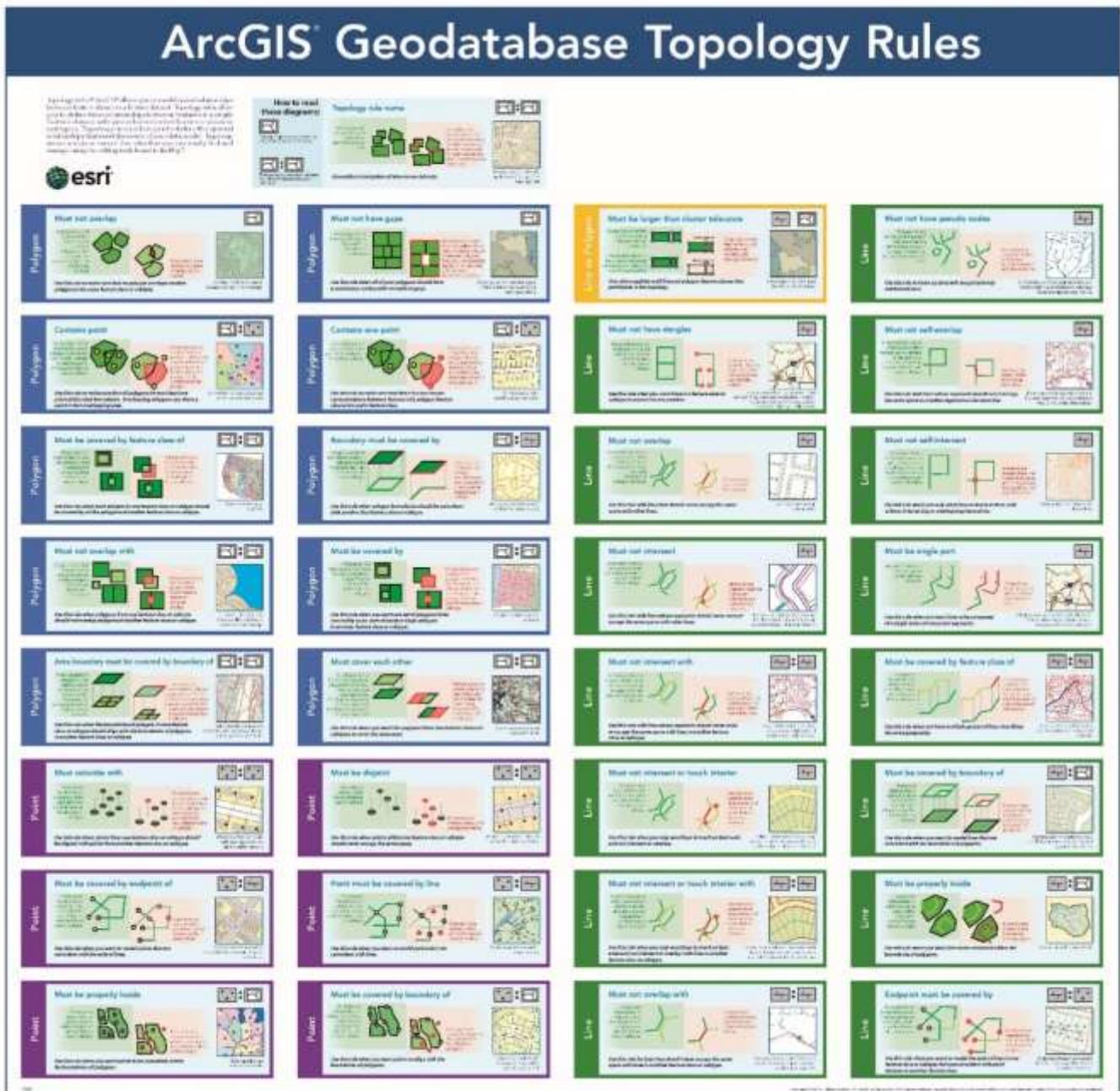


Fig. 44 : Synthèse des erreurs topologiques recherchées et mises en évidence

(Source <https://pro.arcgis.com>)

Le transformateur «FeatureWriter» crée ensuite un fichier au format désiré à partir des entités qui lui arrivent. Dans ce cas, le format d'export défini est le «SHAPE».

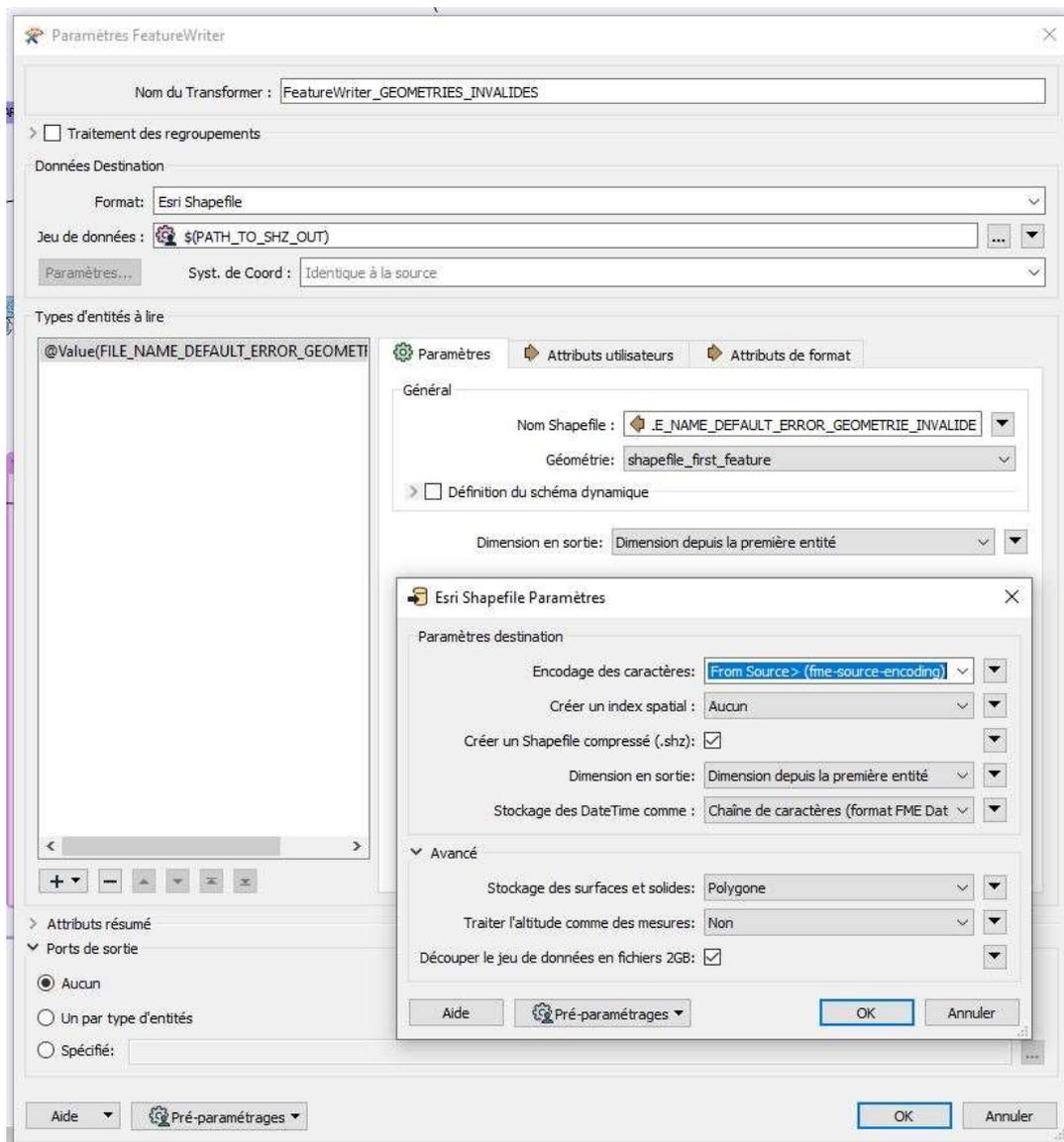


Fig. 45 : Visualisation du «transformer» «FeatureWriter» et de son paramétrage

Le fichier «SHAPE» généré par le «FeatureWriter» indique à l'utilisateur les géométries en défaut. Ce fichier d'extension .shz est transmis en pièce jointe en accompagnement du mail final.

En ouvrant le fichier transmis par mail sous QGIS, la table attributaire laisse apparaître l'état invalide des géométries (2ème colonne). Ce paramétrage s'effectue en amont Avec «FME». Il permet de définir les informations que l'on veut voir apparaître sous QGIS et sous quelle forme.

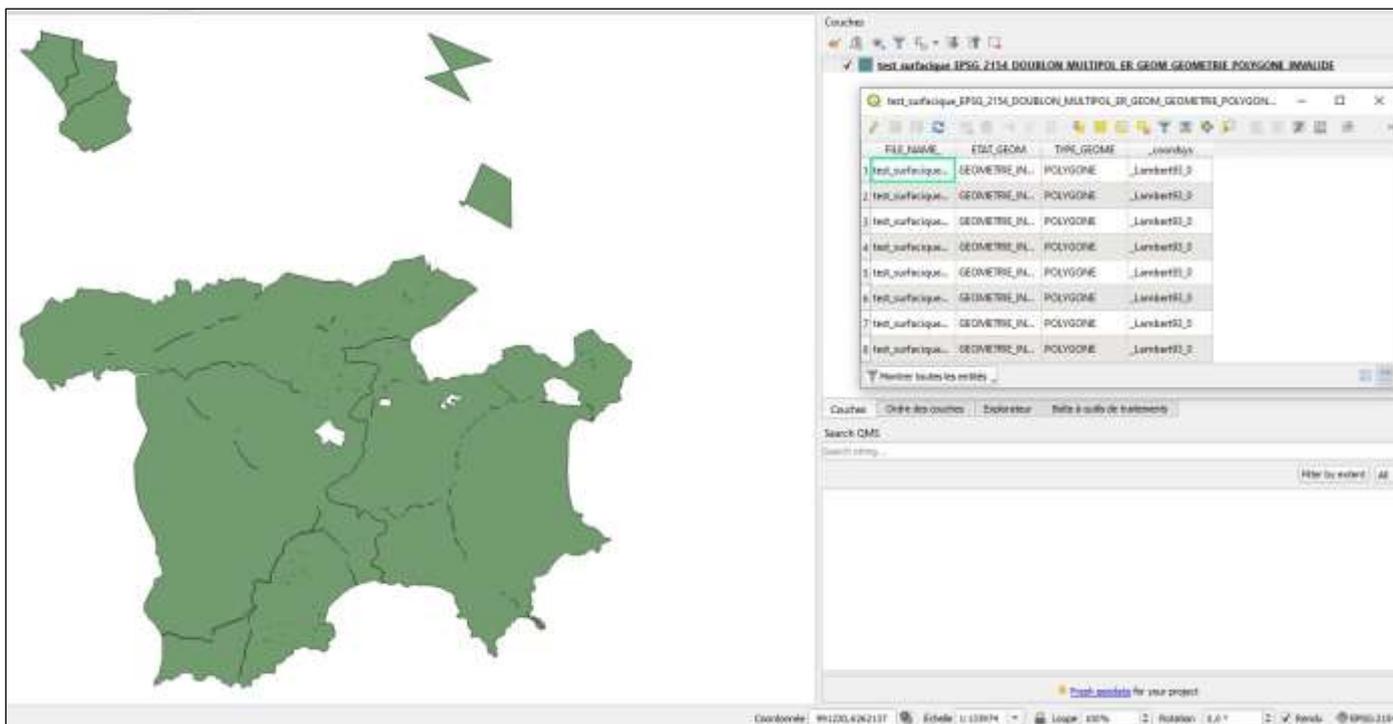


Fig. 46 : Visualisation du fichier «SHAPE» des polygones en défaut sous QGIS

Les géométries valides continuent dans la chaîne de traitement et font l'objet d'une notation qui participe au calcul du sous-critère de cohérence topologique.

Les géométries en défaut quelle que soit leur nature sont écartées de la suite du traitement et ne seront plus prises en compte.

Les ratios et les notations de ratios sont toujours calculés selon la même méthode.

Pour les géométries valides, le ratio est égal à

$$((\text{nombre de géométrie valides}) / (\text{nombre total de géométrie en entrée})) \times 100.$$

La formule de calcul pour la notation des ratios est égale à

$$(\text{Ratio des géométries valides} / 100) \times 5.$$

Cette formule de conversion de ratio vers notation de ration assure une cohérence entre le ratio et sa notation.

La même méthode sera utilisée que ce soit pour les polygones, les lignes ou les points.

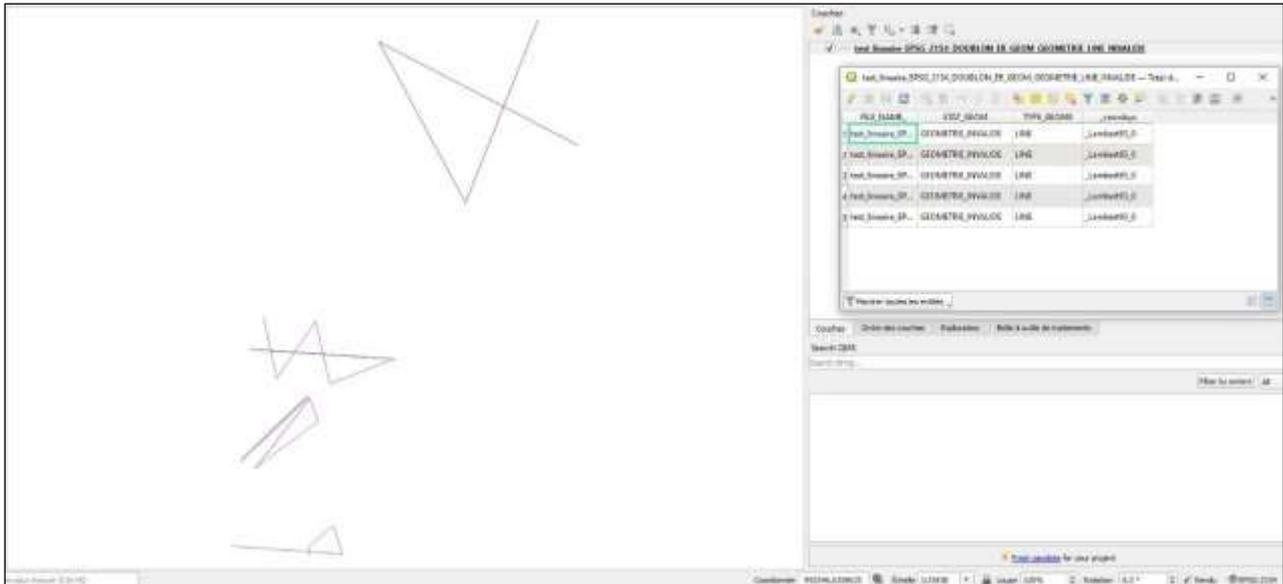


Fig. 47 : Visualisation du fichier «SHAPE» des lignes en défaut sous QGIS

La seconde colonne de la table attributaire précise l'état de la géométrie en l'occurrence ici elles sont invalides.

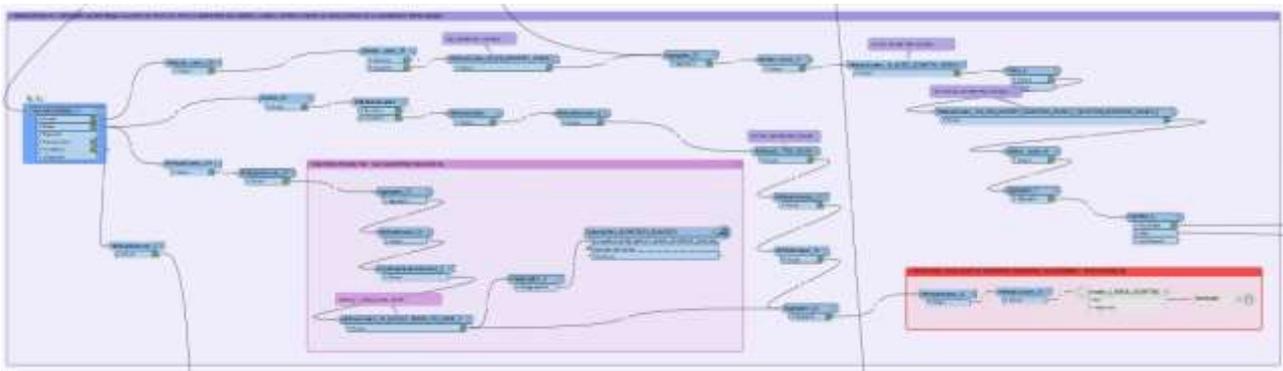


Fig. 48 : Visualisation du signet traitant la validité des géométries

Principe de notation des éléments constitutifs des sous-critères et critères

La notation s'applique toujours sur les éléments positifs car il est compliqué de noter en négatif. Les notations sont également toujours complémentaires, si 25% des géométries sont corrompues alors implicitement 75% ne le sont pas. Les modalités de contrôle définies dans le rapport de synthèse décrivent ce principe en explicitant la méthode d'analyse qui recherche les éléments en défaut pour exprimer les éléments valides de manière complémentaires.

SOUS CRITERE DE COHERENCE TOPOLOGIQUE = 3.9 sur 5

Nature du contrôle	Modalité du contrôle	Résultat	Notation
Vérification de l'intégrité géométrique des LINES	Recherche des LINES corrompus. Les objets classés en défaut à cet étape ne sont plus pris en compte pour la suite du traitement.	12 LINE(S) valide(s) sur 17 soit 70,6 %	3,5 sur 5

Fig. 49 : Extrait du rapport de synthèse illustrant la complémentarité entre les modalités de contrôle et la notation de cet élément



Traitement des géométries valides en fonction des seuils de calcul

Contrairement à l'analyse de la validité de la géométrie qui fournit un résultat binaire valide ou non valide. L'étape suivante fait intervenir en partie les seuils de calculs appliqués aux surfaces de polygones et aux longueurs de lignes.

Les géométries sont d'abord filtrées puis dispatchées en fonction de leur nature.

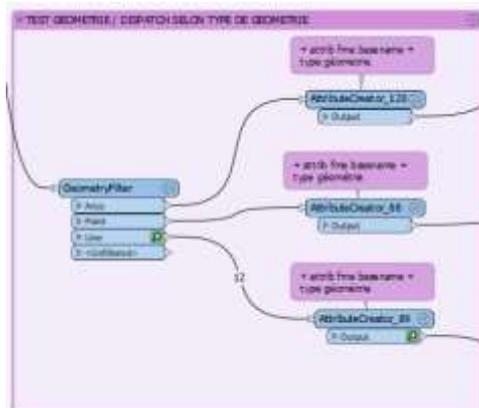


Fig. 50 : Visualisation du signet filtrant les géométries selon leur nature

Après le passage du flux de données par le «transformer» «Deaggregator» qui sépare les géométries multiples en géométries simples et individuelles, les surfaces de chaque polygones ainsi que les longueurs de chaque ligne sont alors comparées à la valeur des seuils de surface et de longueur défini dans les paramètres privés.

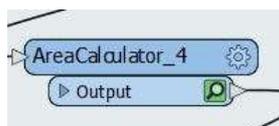
La méthode retenue pour mettre en évidence les géométries trop petites au regard des seuils définis par l'administrateur du script ne s'applique évidemment qu'aux polygones et lignes.

Il convient de s'affranchir de ces erreurs géométriques pour plusieurs raisons.

Tout d'abord elles sont le résultat d'une mauvaise application des règles de topologie et dans sa logique de diffusion de données de qualité et des bonnes pratiques, le CRIGE – PACA pour rester cohérent avec lui-même ne souhaite pas diffuser de telles données. Au-delà de cet aspect, un trop grand nombre de ces micro-géométries peut faire échouer certains traitements et requêtes. Il convient donc de s'en affranchir avant diffusion et exploitation des données.

Les surfaces et longueurs sont donc calculées à l'aide de «transformers» spécifiques, les longueurs et surfaces sont alors comparées à la valeur du seuil

«FME» permet d'avoir une visualisation interactive et en temps réel d'une ou plusieurs entités et de ses attributs, dans ce cas la dernière colonne correspond à la surface des polygones calculée par le logiciel.



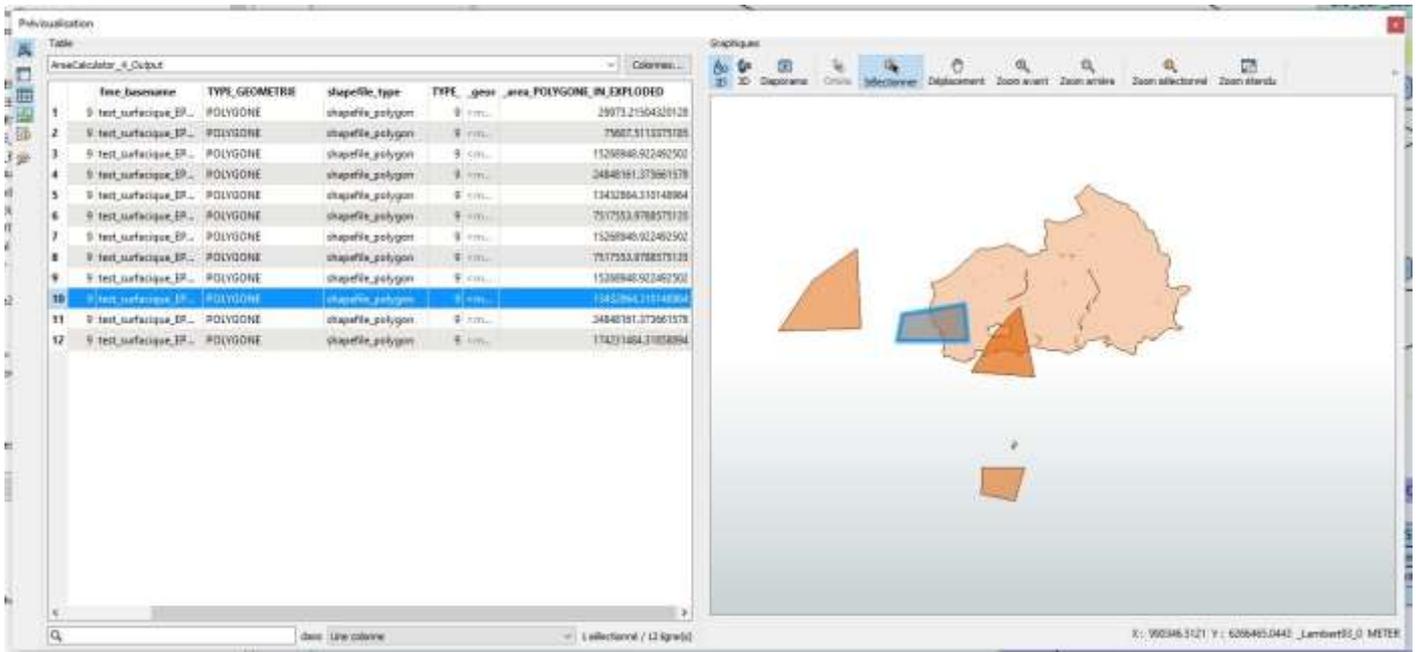


Fig. 51 : Visualisation du «transformer» «Area Calculator» et résultat de son action

La dernière colonne du tableau affiche les surfaces calculées par «FME».

Grâce à un nouveau «transformer», une nouvelle variable nommée «ANALYSE_SURF_POLYgone_INSIDE_CF_SEUIL» est définie. Celle-ci est le résultat de la formule surface calculée du polygone – valeur de seuil des surfaces.

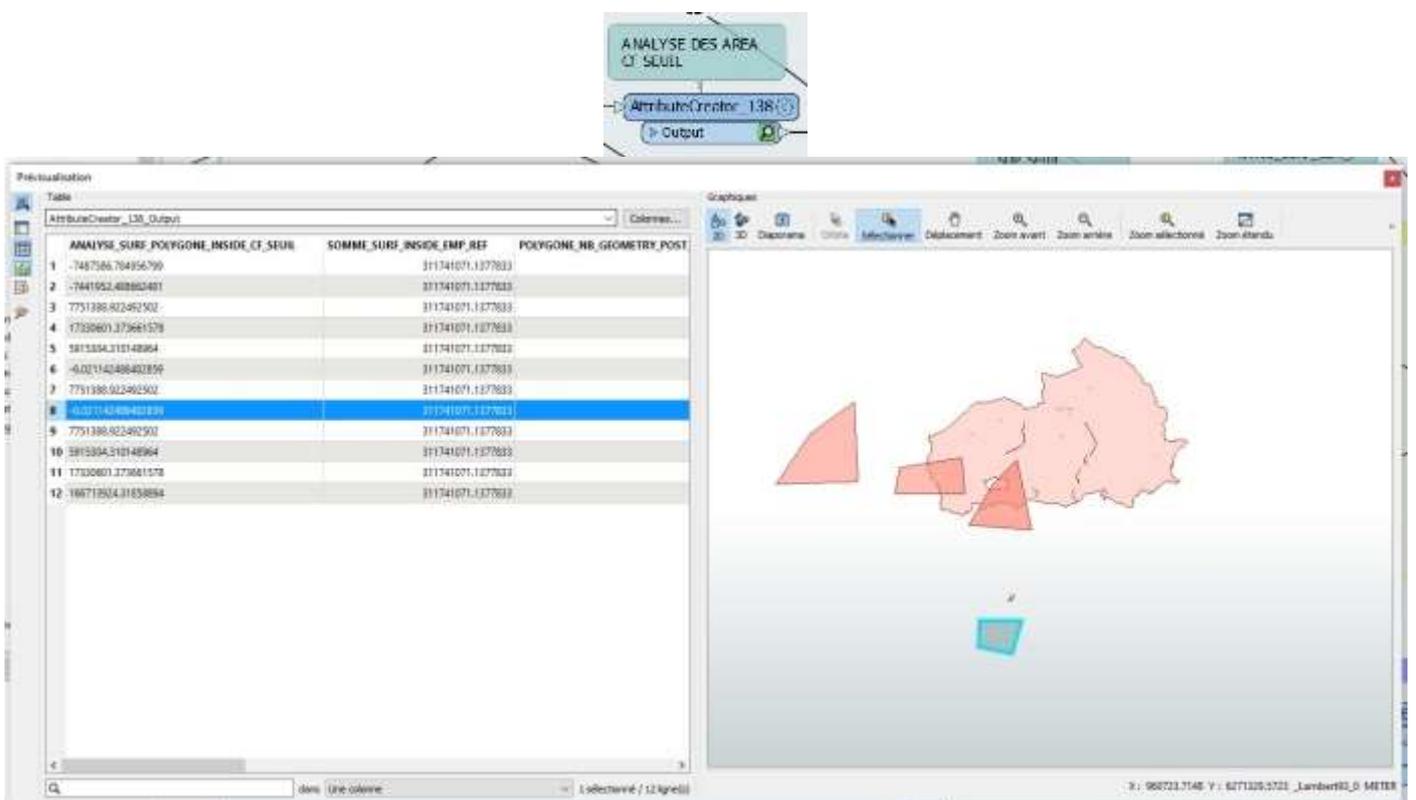


Fig. 52 : Visualisation du «transformer» «Attribut Creator» qui génère la variable «ANALYSE_SURF_POLYgone_INSIDE_CF_SEUIL» et résultat de son action

Ici la première colonne du tableau laisse apparaître des valeurs négatives.



La variable «ANALYSE_SURF_POLYGONE_INSIDE_CF_SEUIL» est alors soumise à un test, via le «transformer» «Tester» en l'occurrence il s'agit de savoir si la valeur de cette variable est positive ou négative.

En fonction de la réussite ou de l'échec à ce test, les entités sont distribuées sur deux ports de sorties différents «PASSED» ou «FAILED». Dans notre cas, le fichier comporte des polygones dont la surface est supérieure au seuil mais aussi des polygones dont la surface est inférieure au seuil.

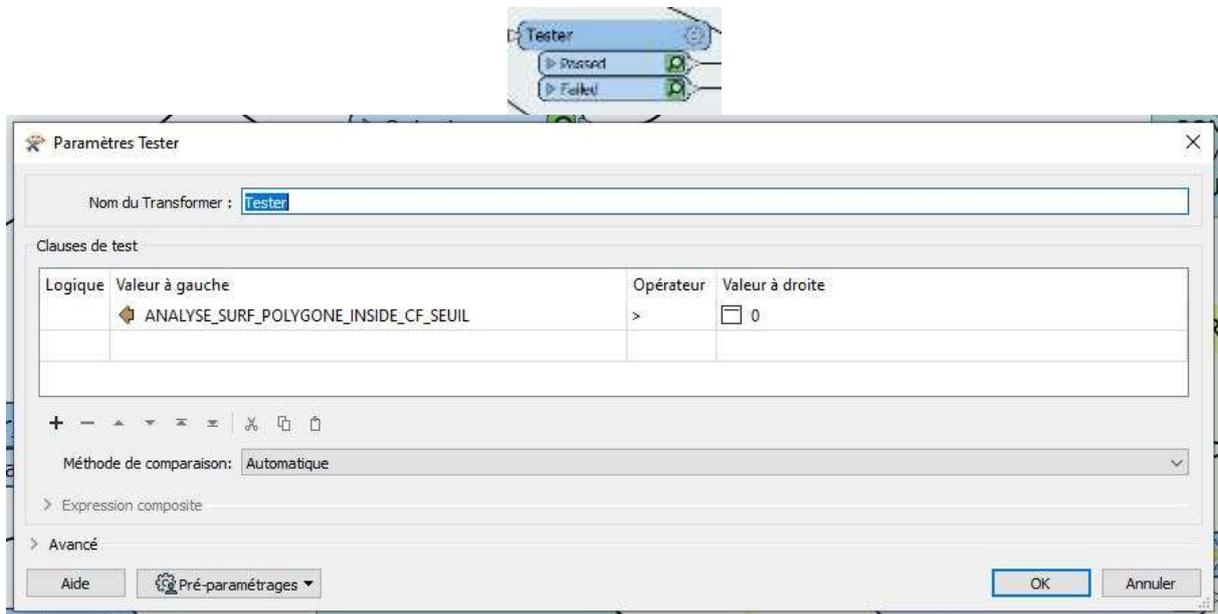


Fig. 53 : Visualisation du «transformer» «Tester» et de son paramétrage

Les entités qui passent le test avec succès poursuivent le processus d'analyse.

Les surfaces individuelles supérieures au seuil sont additionnées pour être comparées à la surface totale des polygones en entrée et ainsi générer la notation correspondant à la recherche des micro-polygone qui se fait sur le même principe développé précédemment à savoir le calcul d'un ratio.

Les entités en échec sur ce test sont dirigées vers une chaîne de traitement secondaire qui génère un fichier d'erreur .shz contenant les polygones identifiés en fonction de leur surfaces inférieures au seuil.

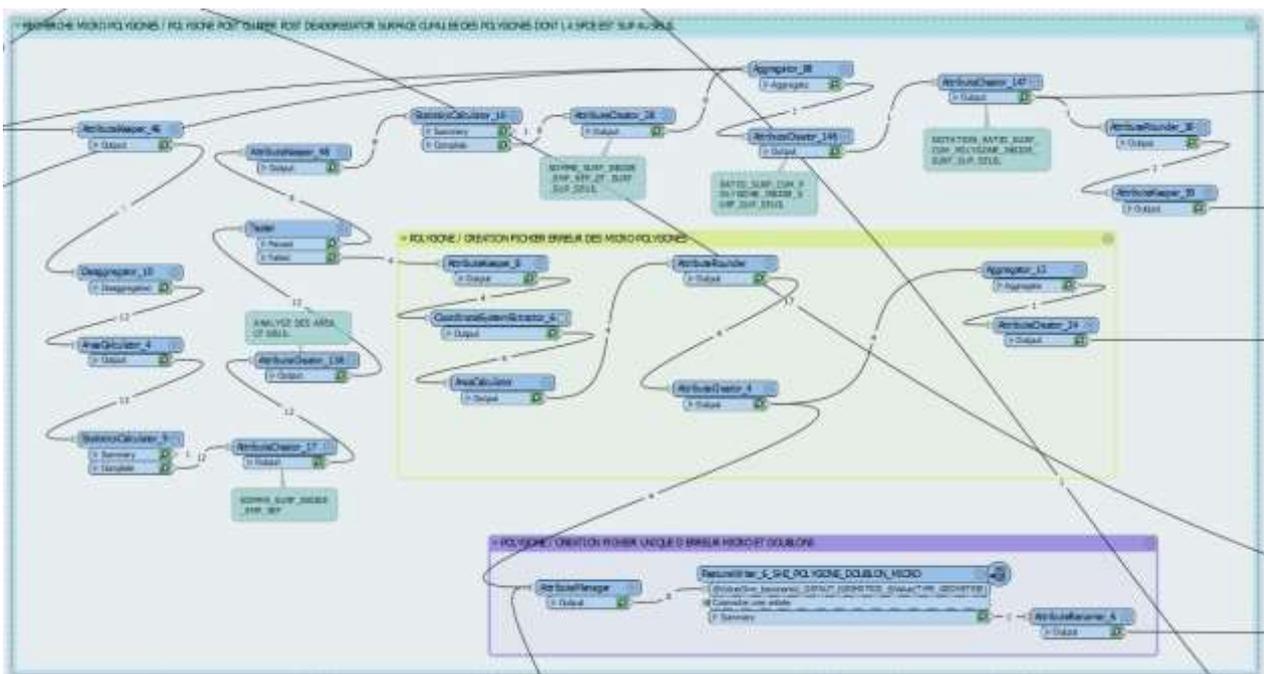


Fig. 54 : Visualisation du signet traitant les micro-polygones

Nature du contrôle	Modalité du contrôle	Résultat	Notation
Recherche des "micro"-POLYGONES dont la surface individuelle est inférieure au seuil	Recherche des POLYGONES dont la surface est inférieure au seuil (S = 7317360 m ²)	Les surfaces cumulées des POLYGONES dont la surface individuelle est supérieure au seuil représentent 256608382.3 m ² sur 311741673.1 soit 95.1 %	4.0 sur 3

Fig. 55 : Extrait du rapport de synthèse final - partie dédiée à la notation des «micro-polygones»

Pour illustrer, le principe de traitement des micro-polygones, et pour s'adapter au fichier en présence, le seuil de surface a été volontairement paramétré avec une valeur très grande.

Cependant, même si les valeurs de seuil sont à l'unique discrétion de l'administrateur du script, en conditions opérationnelles, les valeurs des seuils des surfaces seront fixées de façon beaucoup plus cohérente.

Traitement des géométries en doublon

Les géométries en doublons n'ont aucun intérêt, en effet la même information répétée de multiples fois n'apporte aucune valeur ajoutée au fichier. De plus, le doublonnage des données dans le même fichier est pénalisant car il alourdit le poids des données. Cette notion de poids des données est très importante car l'administration, la gestion, le stockage et la transmission des données ont un cout qui augmente avec le poids croissant des données.

Il convient préalablement, à la recherche des géométries en doublons de définir exactement les caractéristiques du doublonnage :

- un objet est considéré en doublon si et seulement si la géométrie et tous les attributs sont strictement identiques. Le transformer «Matcher» est prévu à cet effet

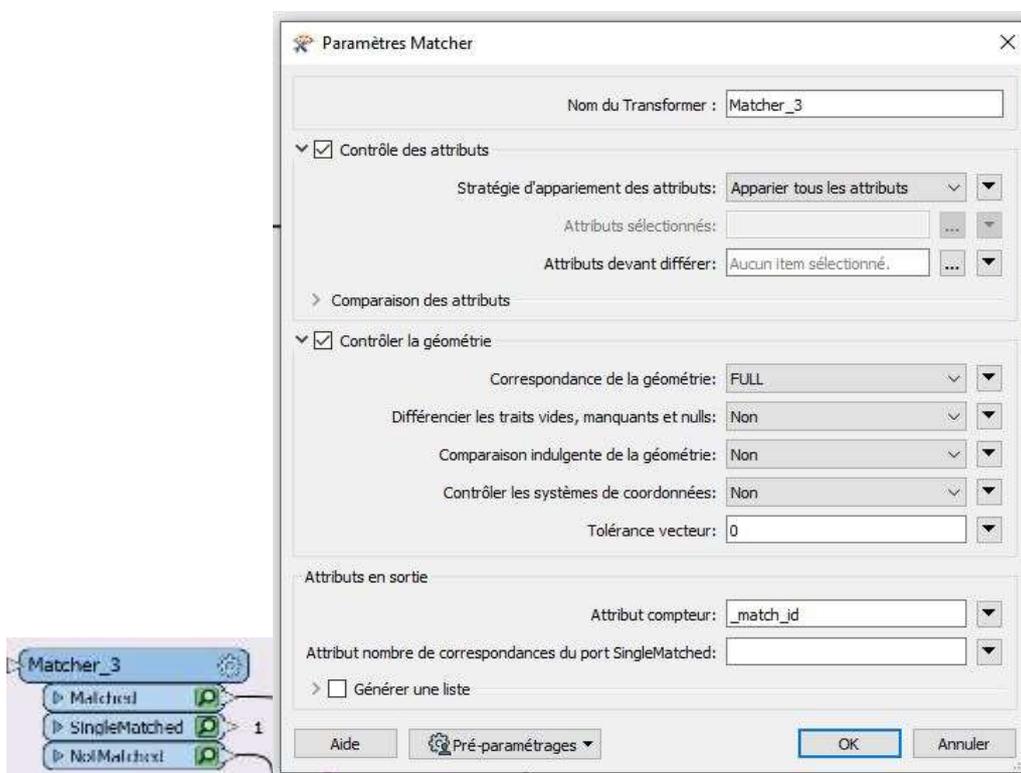


Fig. 56 : Visualisation du «transformer» «Matcher» et de son paramétrage

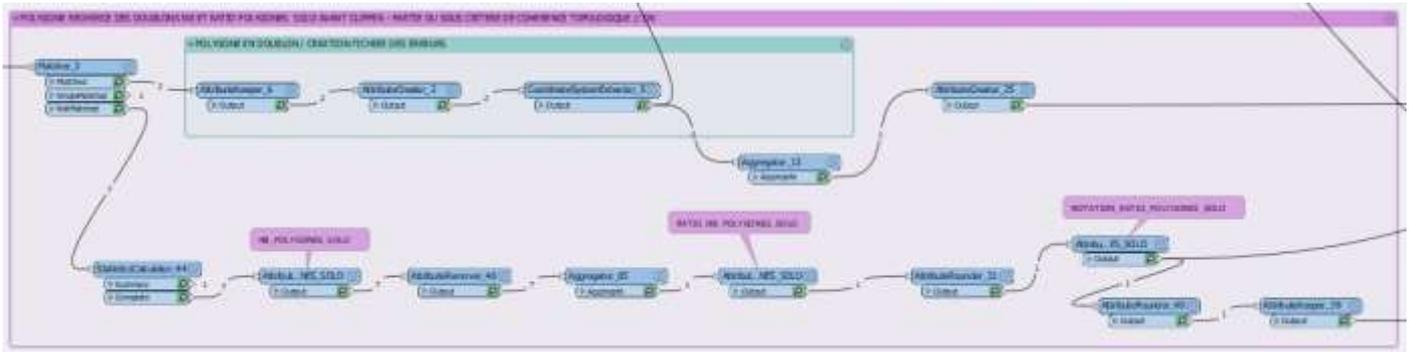


Fig. 57 : Visualisation du signet traitant les polygones en doublons

Les polygones identifiés en doublons sont dirigés vers un port dédié qui rejoint la chaîne de traitement des micro-polygones ou un «transformer» «FeatureWriter» génère un unique fichier d'erreur .shz rassemblant les micro-polygones et les polygones en doublons.

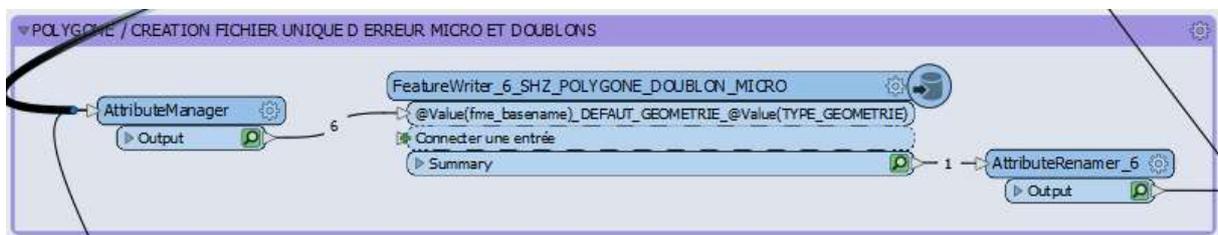


Fig. 58 : Visualisation du signet générant le fichier des erreurs .shz des micro-polygones et des doublons

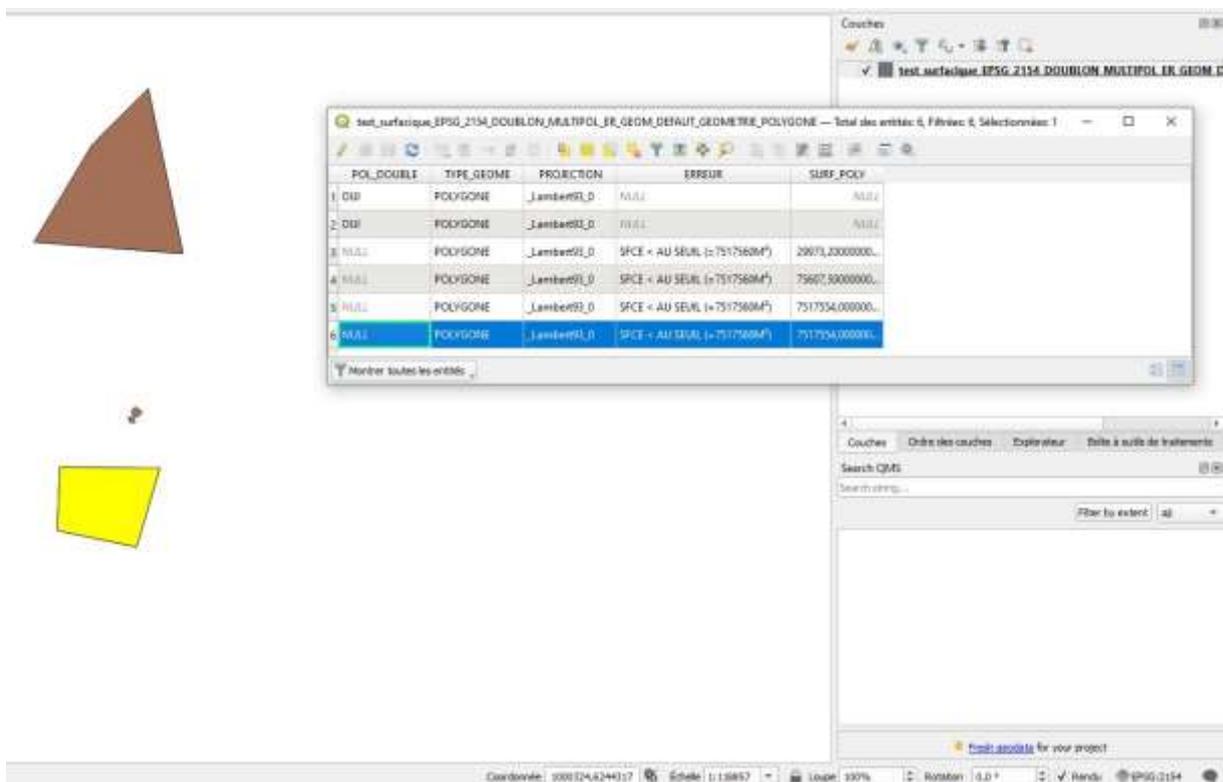


Fig. 59 : Visualisation du fichier «SHAPE» des micro-polygones et polygones en doublons sous QGIS

L'utilisateur du qualificateur recevra en pièce jointe du mail le fichier .shz des erreurs ou la table attributaire est renseignée du type d'erreur pour chaque polygones.

Le même principe de traitement est appliqué pour les micro-lignes et lignes en doublons, c'est pourquoi, celui-ci ne sera pas développé en détail.

Calcul du sous-critère de cohérence topologique

Les notations individuelles pour les polygones sur les éléments suivants

- vérification de l'intégrité géométrique des polygones
- Recherche des doublons
- Recherche des «micro-polygones»

Sont maintenant calculées.

Le **sous-critère de cohérence topologique** qui est la résultante des trois composantes ci-dessus énoncées est alors calculé comme une simple moyenne arithmétique sans pondération.

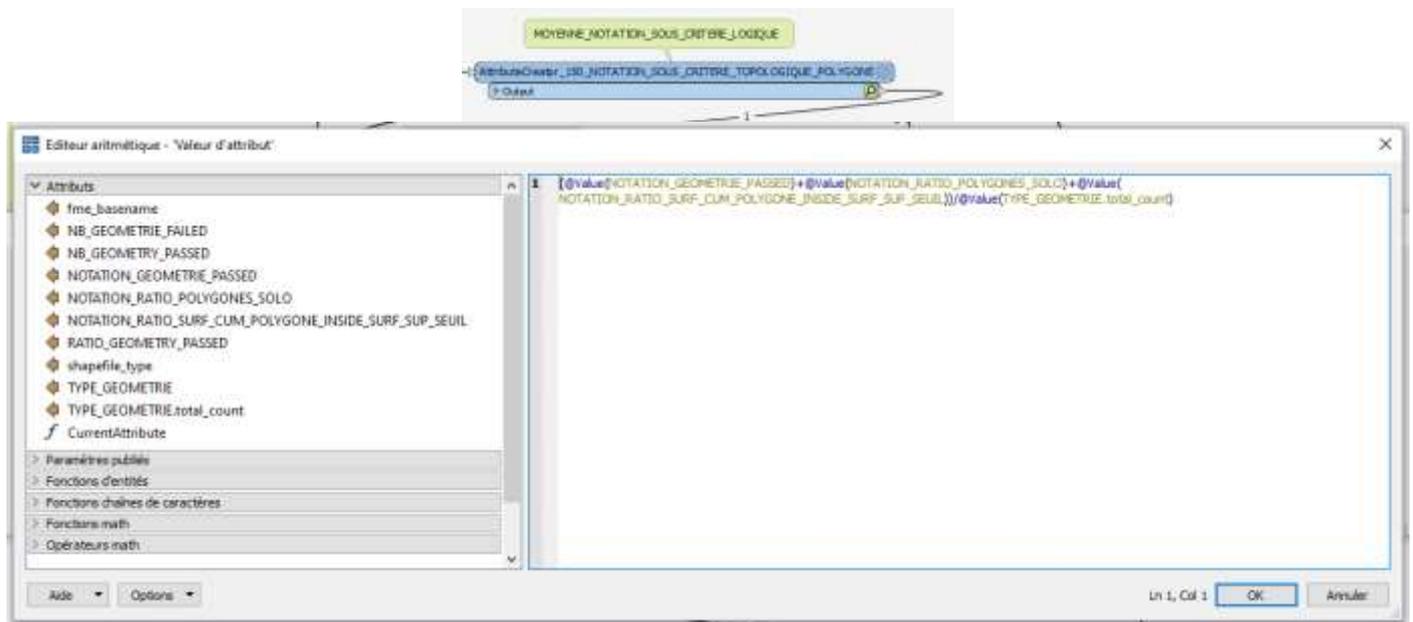


Fig. 60 : Visualisation du «transformer» gérant le calcul du sous-critère de cohérence topologique des polygones et son paramétrage

L'introduction de coefficients est facilement paramétrable pour chaque composante, je n'ai pas opté pour cette solution estimant qu'aucune des composantes n'était prépondérante sur les autres. Cependant pour rester dans l'objectif d'anticipation des évolutions futures de ce script, le nombre d'entrées dans le «transformer» est automatiquement calculé et la moyenne s'ajuste donc de manière autonome.

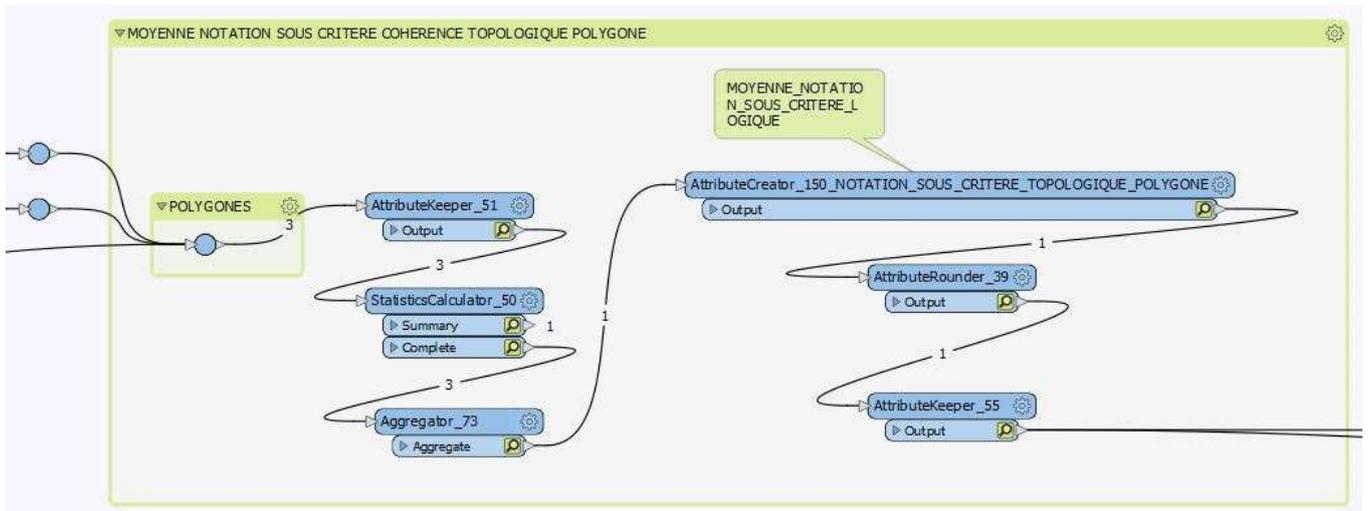


Fig. 61 : Visualisation du signet de calcul du sous-critère de cohérence topologique des polygones

Calcul du critère d'exhaustivité

Le calcul de ce critère ne fait intervenir qu'un seul indicateur pour les trois géométries (polygones, lignes, points). En substance, il quantifie la proportion de géométrie, longueurs cumulées pour les lignes, Surfaces cumulées pour les polygones ou nombre de points à l'intérieur de l'emprise spatiale de référence par rapport aux cumul des surfaces, cumul des longueurs ou nombre total de points après validation des géométries.

Dans notre cas de figure, le critère d'exhaustivité est égale au paramètre «NOTATION_RATIO_LONGUEUR_LINE_INSIDE» car c'est le seul qui est calculé.

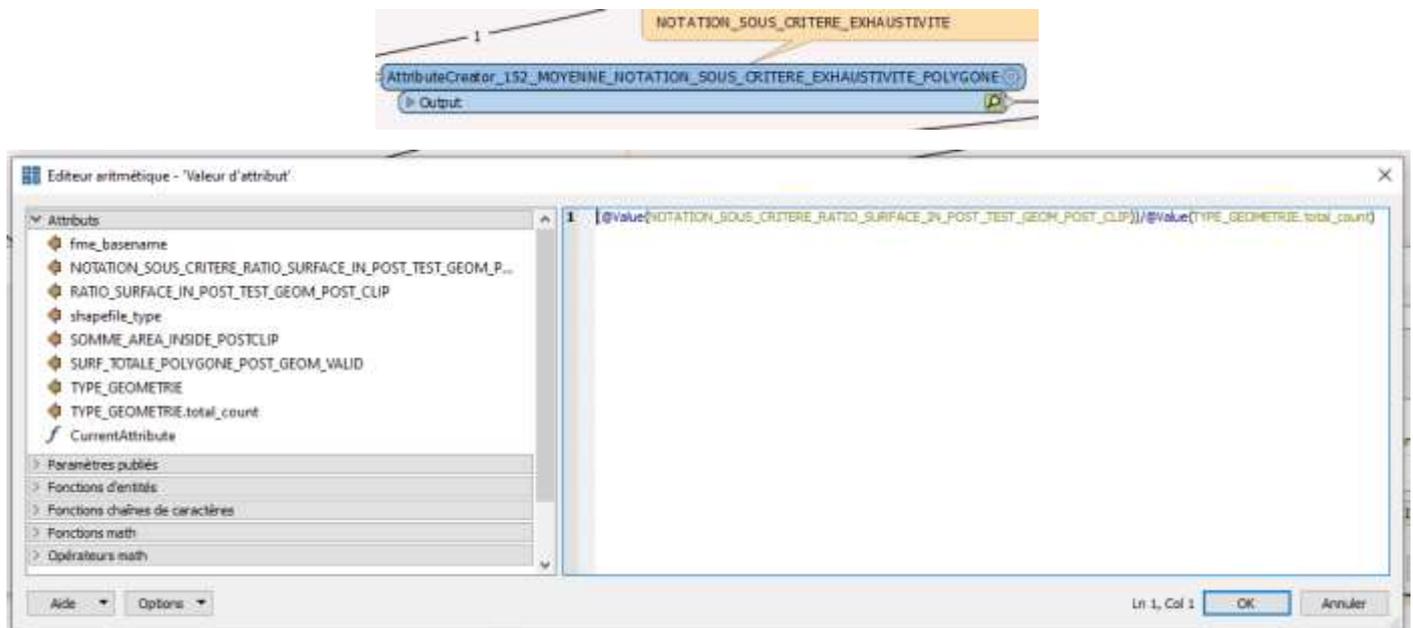


Fig. 62 : Visualisation du «transformer» de calcul du critère d'exhaustivité des polygones et son paramétrage

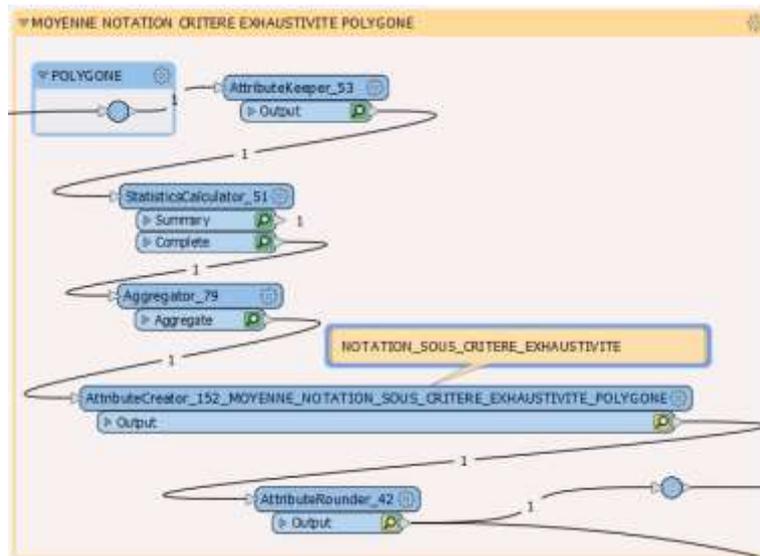


Fig. 63 : Visualisation du signet de calcul du critère d'exhaustivité des polygones

Traitement des géométries valides en fonction des seuils de calcul et de l'emprise spatiale de référence

Seule l'évaluation du sous-critère de cohérence topologique fait intervenir les deux paramètres seuil de calcul et emprise spatiale de référence.

Cette étape du traitement est stratégique car en fonction de l'emprise spatiale de référence c'est-à-dire l'échelon administratif défini par l'utilisateur en tout début du processus de traitement, le «transformer» «Clipper» agit comme un emporte-pièce et découpe toutes les autres entités entrantes en fonction du contour défini dans le fichier source des emprises.

L'emporte-pièce utilise alors l'emprise spatiale d'un seul territoire parmi les 946 Communes, les 51 EPCI, les 6 départements ou la région PACA.

Au-delà de la définition géométrique des polygones d'emprise, le choix du bon niveau administratif est primordial car un mauvais choix d'emprise peut engendrer une surévaluation ou une sous-évaluation de la notation. De fait le critère d'exhaustivité peut passer de 0 à 5 ou inversement. En effet, si toutes les géométries se trouvent exclusivement dans l'emprise du département du var mais que l'emprise de référence spatiale choisie est celle du Vaucluse la notation du critère d'exhaustivité sera égale à zéro. Par contre, si avec les mêmes données d'entrées, c'est l'emprise de la région PACA qui est choisie ce même critère sera égale à 5 sans pour autant qualifier objectivement ce critère. L'emprise spatiale de référence doit donc être judicieusement choisie.

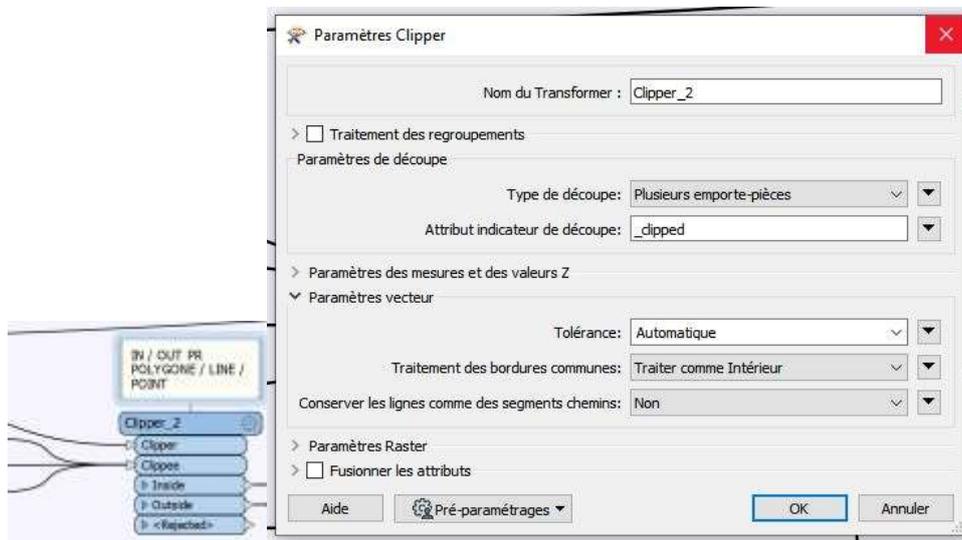


Fig. 64 : Visualisation du «transformer» «Clipper» et de son paramétrage

Les emprises spatiales de références sont issues des couches "COMMUNE", "EPCI", "DEPARTEMENT" et "REGION" de la BD TOPO® (IGN). Préciser leur source est important car selon l'emprise spatiale de référence utilisée par le producteur de la donnée et par la simple application du contour administratif choisi, des erreurs peuvent se générer.

En effet l'application de données représentant le même contour administratif issu de données de natures différentes réalisées avec une précision différente génère des erreurs dues aux échelles différentes de réalisation des données.

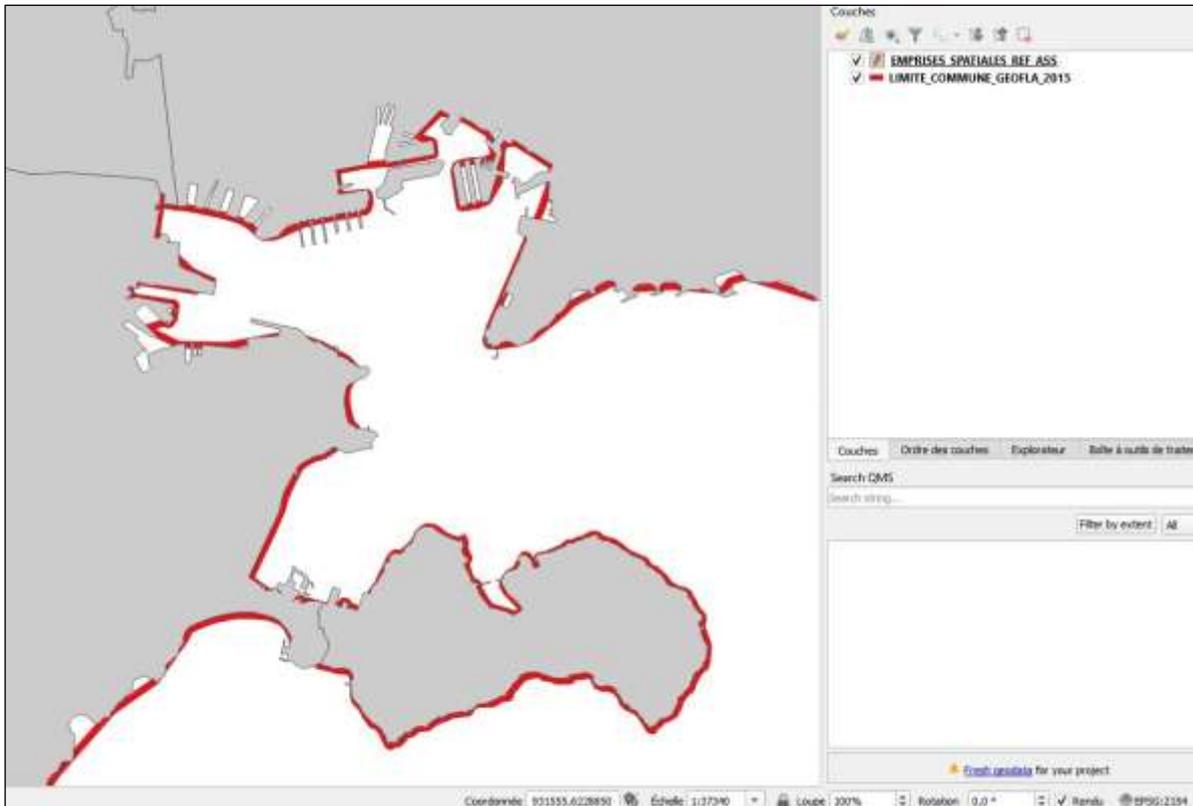


Fig. 65 : Illustration des erreurs générées par application de l'emprise spatiale de référence choisie sur la même emprise définie avec une autre précision.

Dans cet exemple qui correspond au secteur de la petite rade de TOULON dans le var, on constate que le tracé en rouge qui correspond aux limites de communes de la base GEOFLA® (IGN) de 2015 n'est pas du tout cohérent avec les limites du tracé en gris correspondant à la BDTOPO® (IGN) de 2022.

Cependant, ce résultat n'est guère étonnant dans la mesure où

- la résolution spatiale de la base GEOFLA® est égale à 1/ 100000
- la résolution spatiale de la base BDTOPO®GEOFLA® est égale à 1/ 10000

Cela illustre la nécessité d'utiliser des emprises spatiales de références cohérentes voire identiques d'où l'utilité et la justification de l'envoi de l'emprise spatiale de référence choisie.

Toutes les géométries sont impactées par le «transformer» «Clipper», les polygones et lignes sont découpés et les points se retrouvent alors de fait à l'intérieur ou à l'extérieur de l'emprise choisie. Suite à l'action de ce «transformer» Il n'est pas nécessaire de passer par le «Deaggregator» car le «Clipper» réalise de fait la même action avant de découper les géométries.

Calcul de la notation du sous critère de cohérence topologique

La formule de calcul du sous critère de cohérence topologique correspond à une simple moyenne arithmétique sans pondération à partir des notations sur cinq de la vérification de l'intégrité géométrique, de la recherche des géométries en doublons et des micro-géométries.

La même formule de calcul s'applique pour toutes les géométries polygones, lignes, points.

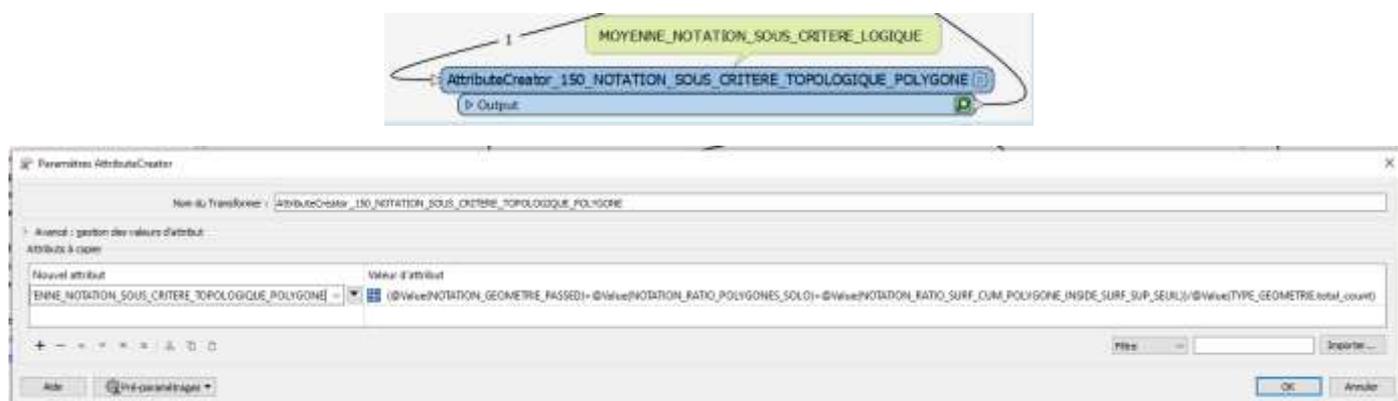


Fig. 66 : Visualisation du transformateur de calcul du sous-critère de cohérence topologique et son paramétrage

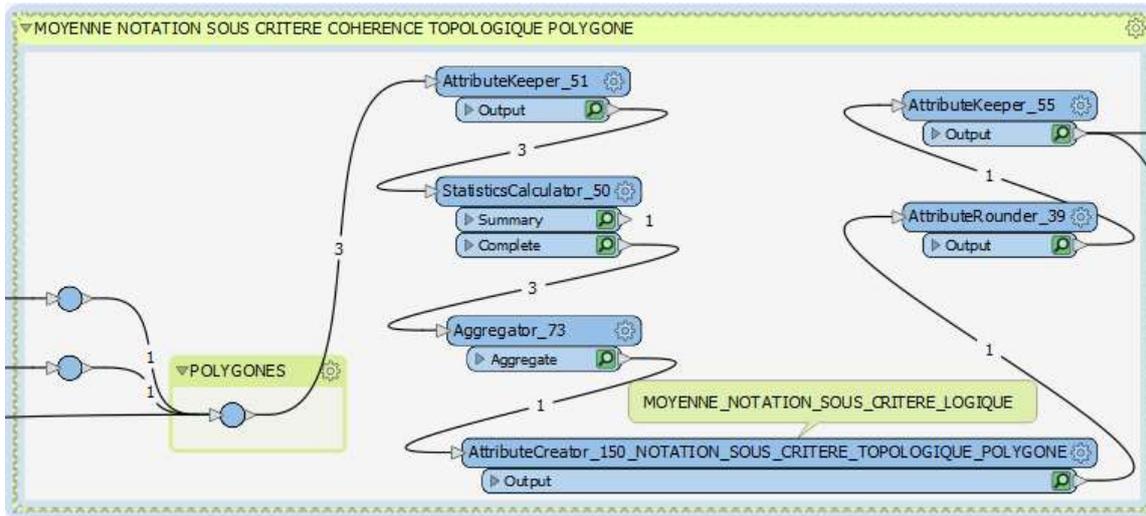


Fig. 67 : Visualisation du signet de calcul du sous-critère de cohérence topologique

SOUS CRITERE DE COHERENCE TOPOLOGIQUE = 3,7 sur 5

Nature du contrôle	Modalité du contrôle	Résultat	Notation
Vérification de l'intégrité géométrique des POLYONES	Recherche des POLYONES corrompus. Les objets créés en défaut à cet étape ne sont plus pris en compte pour la suite du traitement.	9 POLYONE(S) valide(s) sur 17 soit 52,9 %	2,8 sur 3
Recherche des doublons	Un POLYONE est considéré en double si et seulement si la géométrie et tous les attributs sont strictement identiques	7 POLYONE sans doublons sur 9 soit 77,8 %	2,9 sur 3
Recherche des "micre" POLYONES dont la surface individuelle est inférieure au seuil	Recherche des POLYONES dont la surface est inférieure au seuil (S = 7517500 m²). Les surfaces considérées des POLYONES dont la surface individuelle est supérieure ou sont représentées 296606382,5 m² sur 311741071,1 soit 95,1 %		4,2 sur 5

Fig. 68 : Extrait du rapport de rapport de synthèse final - partie dédiée au sous critère de cohérence topologique

L'objet du **critère d'exhaustivité** est de définir si la donnée soumise au qualificateur est suffisante en excédent ou manquante. Ce critère est défini et calculé en terme de relation spatiale entre les données situées à l'intérieur de l'emprise par rapport à l'ensemble des géométries en entrée après validation des géométries («transformer» «GeometryValidator»).

Les surfaces des polygones situées à l'intérieur sont additionnées puis comparées à la somme des surfaces des polygones géométriquement valides en entrée. Le résultat est un ratio lui-même converti en note sur cinq.

Le même principe de calcul s'applique pour les lignes et ne sera donc pas plus développé.

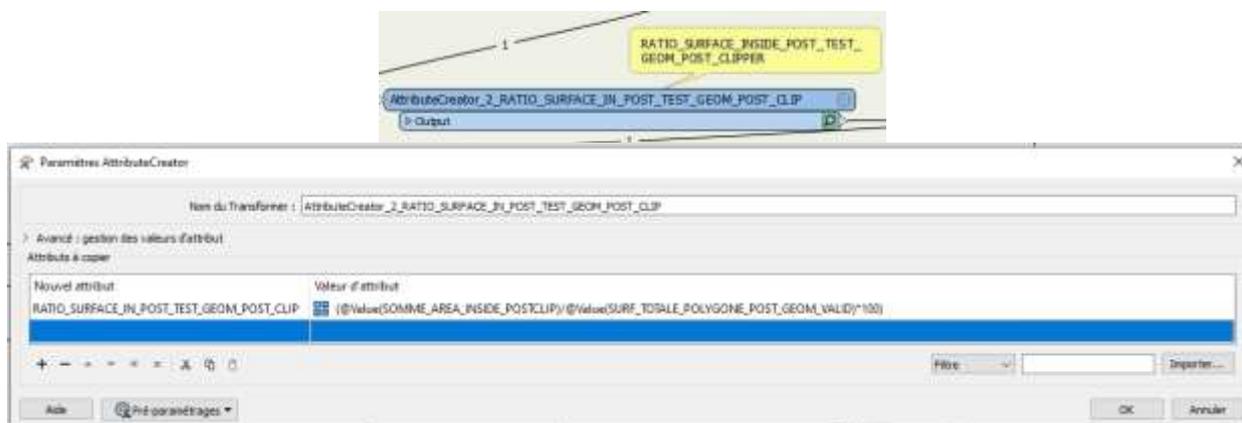


Fig. 69 : Visualisation du «transformer» «AttributeCreator» définissant le calcul du ratio et son paramétrage

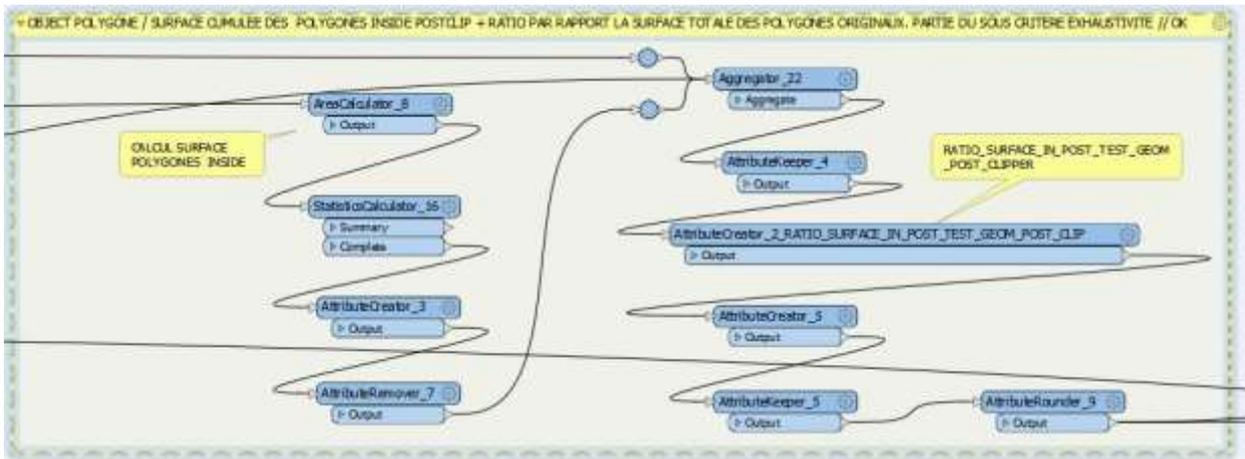


Fig. 70 : Visualisation du signet de calcul de la somme des surfaces de polygones à l'intérieur de l'emprise spatiale de référence

Le calcul de la notation du **critère d'exhaustivité** aurait pu être directement exploité ainsi. Cependant en gardant à l'esprit que l'écriture du script doit pouvoir se poursuivre le plus facilement possible, j'ai intégré un signet supplémentaire qui calcule une moyenne si d'autres variables venaient à être intégrées à ce calcul. Ce critère à l'image du **critère de cohérence logique** n'est pas pondéré. Cette anticipation s'applique à toutes les natures de géométries.

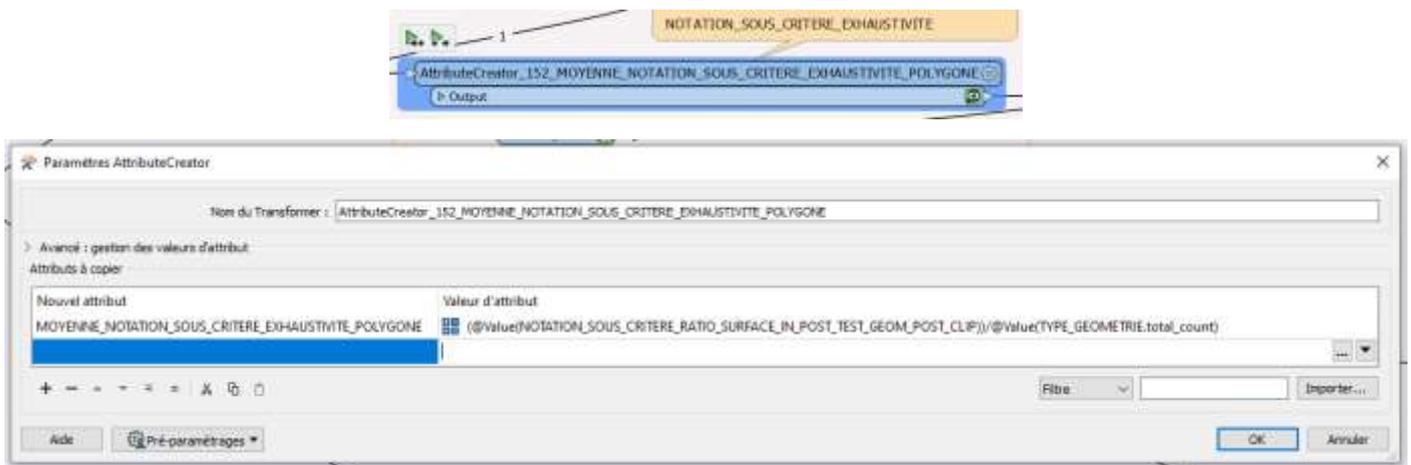


Fig. 71 : Visualisation du signet de calcul de la notation du critère d'exhaustivité des polygones et son paramétrage

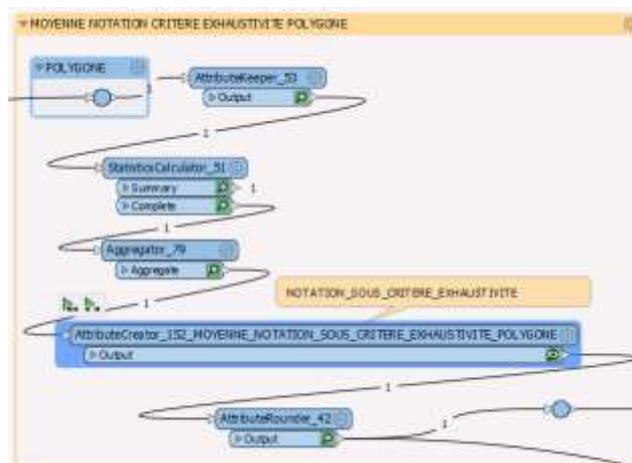


Fig. 72 : Visualisation du signet de calcul du critère d'exhaustivité des polygones

La notation de ce critère apparaît également dans le rapport de synthèse final.

CRITERE D' EXHAUSTIVITE = 4 sur 5			
Nature du controle	Modalité du controle	Résultat	Notation
Relation spatiale entre le fichier et l'emprise de référence choisie	Decoupe du fichier par l'emprise et comptage des surfaces	La somme des surfaces de POLYGOINES à l'intérieur de l'emprise représente 248396043 m ² sur 311741071,1 soit 79,7 %.	4 sur 5

Fig. 73 : Extrait du rapport de synthèse final - partie dédiée au critère d'exhaustivité

Calcul de la notation globale

Il s'agit du résultat final attendu par l'utilisateur, l'évaluation du fichier soumis au qualificateur fait intervenir dans son calcul les **critères d'exhaustivité** et de **cohérence logique** lui-même défini à partir des **sous critères de cohérence topologique** et de **cohérence de format**.

Ce calcul est en fait une simple moyenne arithmétique sans application de pondération.

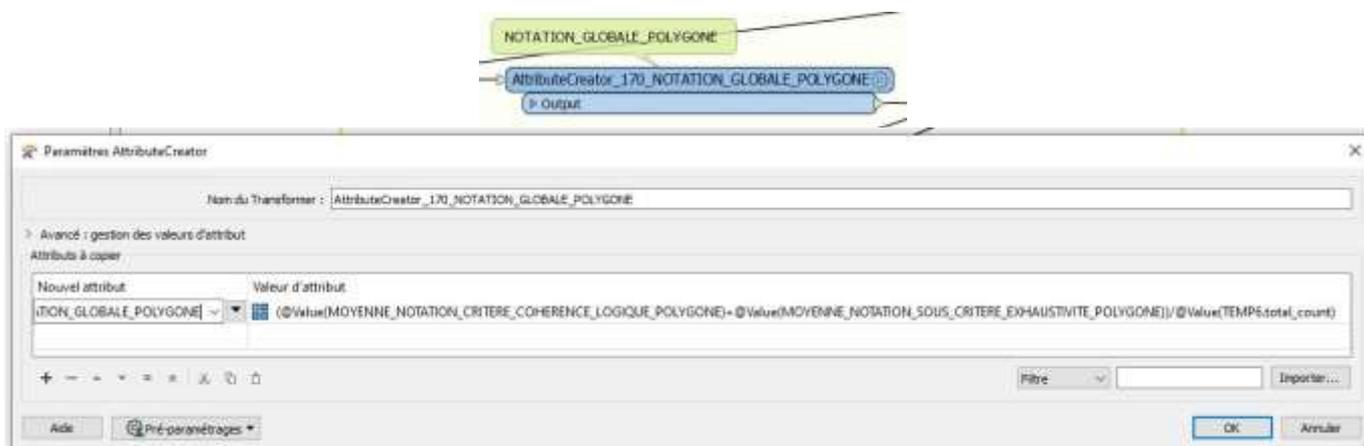


Fig. 74 : Visualisation du «transformer» «AttributeCreator» définissant le calcul de la note globale et de son paramétrage

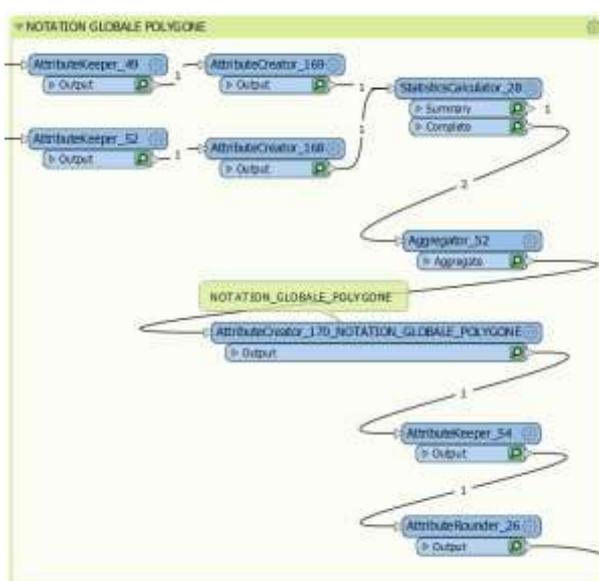


Fig. 75 : Visualisation du signet de calcul de la notation globale des polygones

NOTATION GLOBALE = 4.2 sur 5

Fig. 76 : Extrait du rapport de synthèse final – Notation globale du fichier

La même méthode avec les mêmes formules s'applique à toutes les natures de géométrie.

C'est pourquoi le calcul de la note globale n'est pas détaillé pour les lignes et les points.

Méthode de confection des rapports finaux

Toutes les variables calculées et utiles à la confection des rapports finaux qu'elles soient numériques ou alphanumériques convergent vers un point d'entrée unique pour être incorporées dans l'édition du rapport. Ce rapport a pour vocation de communiquer au producteur de la donnée l'évaluation du fichier soumis au qualificateur. Mais pour communiquer efficacement le document doit être lisible et un minimum esthétique. Tout comme une belle carte favorise la transmission du message qu'elle veut faire passer, un rapport ne doit pas être rébarbatif et trop austère dans sa présentation. Par défaut «FME» propose une mise forme de rapport très rigide qui ne permet pas de mettre en valeur les critères calculés. Le rapport final a donc été réalisé en utilisant l'option «HTML personnalisé» qui permet de générer une présentation beaucoup plus attrayante pour le lecteur.

Le rapport dès sa conception a été pensé pour être clair, concis et précis. Il présente donc un nombre limité d'indicateurs mais ceux-ci doivent avoir du sens. C'est dans cet esprit que les rapports finaux ont été codés. Il paraissait également important de conserver l'identité graphique du CRIGE – PACA, c'est pourquoi les codes hexadécimaux des couleurs du logo ont été réutilisés dans le codage HTML



Fig. 77 : Logo du CRIGE utilisé pour récupérer les codes hexadécimaux des couleurs

Tous les rapports sont donc codés en langage HTML avec quelques instructions en CSS.

Un modèle de rapport spécifique est défini pour chaque géométrie polygones, lignes, points intégrant des variables communes mais aussi les variables particulières liées à chaque nature de géométrie.

Le «transformer» «HTMLReportGenerator» est prévu à cet effet.

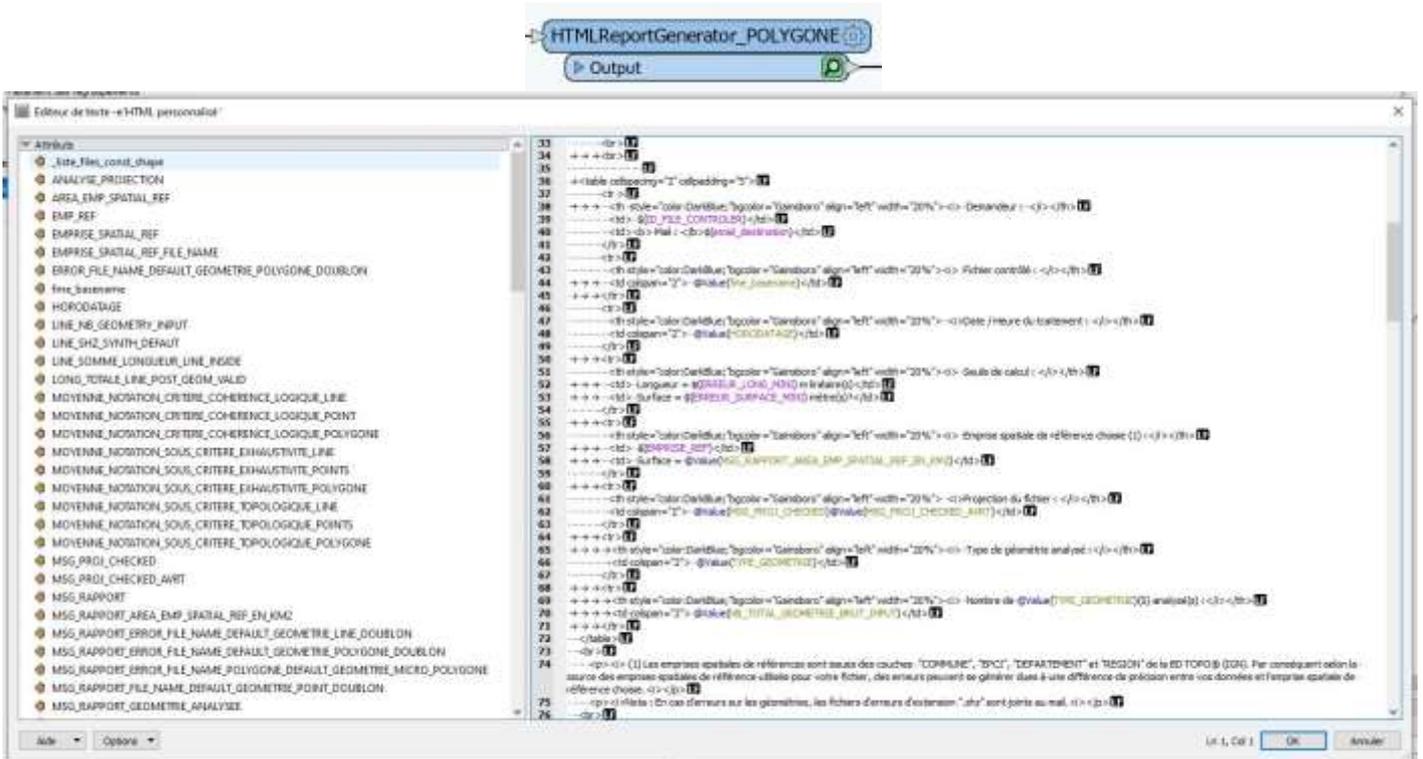


Fig. 78 : Visualisation du «transformer» «HTMLReportGenerator» et de son paramétrage

- les parties de codes en violet correspondent aux paramètres publiés
- les parties de code en marron clair correspondent aux variables numériques ou alphanumériques calculées ou générées par «FME»
- les parties de code en noir correspondent au code HTML

Sur la partie gauche se trouvent les variables disponibles et utilisables dans le rapport, et sur la partie droite, le code en lui-même qui reprend les définitions de variables.

```

<table cellpadding="2" cellspacing="5">
<tr>
<th style="font-size:30px;color:#40e0d0;background-color:#e0e0e0;align:center,width:80%;">INFORMATIONS GENERALES SUR LE FICHIER</th>
</tr>
</table>
<br>
<table cellpadding="2" cellspacing="5">
<tr>
<td style="color:DarkBlue;background-color:#d3d3d3" align="left" width="20%"><!-- Demandeur : --></td></tr>
<td style="color:DarkBlue;background-color:#d3d3d3" align="left" width="20%"><!-- Fichier contrôlé : --></td></tr>
<td colspan="2" style="color:DarkBlue;background-color:#d3d3d3" align="left" width="20%"><!-- Date / Heure de traitement : --></td></tr>
<tr>
<td colspan="2" style="color:DarkBlue;background-color:#d3d3d3" align="left" width="20%"><!-- Seuil de calcul : --></td></tr>
<tr>
<td colspan="2" style="color:DarkBlue;background-color:#d3d3d3" align="left" width="20%"><!-- Longueur = $(EGRET_LONG_MINI) m linéaire[ ] --></td></tr>
<tr>
<td colspan="2" style="color:DarkBlue;background-color:#d3d3d3" align="left" width="20%"><!-- Surface = $(ENSEMBL_SURFACE_MINI) mètre[ ] --></td></tr>
<tr>
<td colspan="2" style="color:DarkBlue;background-color:#d3d3d3" align="left" width="20%"><!-- Emprise spatiale de référence choisie [ ] --></td></tr>
<tr>
<td colspan="2" style="color:DarkBlue;background-color:#d3d3d3" align="left" width="20%"><!-- Projection du fichier : --></td></tr>
<tr>
<td colspan="2" style="color:DarkBlue;background-color:#d3d3d3" align="left" width="20%"><!-- Type de géométrie analysé : --></td></tr>
<tr>
<td colspan="2" style="color:DarkBlue;background-color:#d3d3d3" align="left" width="20%"><!-- Nombre de $(TYPE_GOMETRIE) [ ] analysé(s) --></td></tr>
</tr>
</table>

```

Fig. 79 : Extrait du code d'un rapport édité sous «notepad»



Pour développer le code des rapports et tester le résultat, un interpréteur libre en temps réel et en ligne a été utilisé. Le site <https://www.w3schools.com/> a permis de valider le rendu des rapports avant d'intégrer leur code HTML dans «FME».

The screenshot shows a web-based FME interpreter interface. On the left, there is a dark background with white text representing FME workspace code. On the right, the rendered HTML output is displayed. At the top right of the rendered page, the title is `"@Value(fme_basename)"`. Below the title, there is a section titled "INFORMATIONS GENERALES SUR LE FICHIER:" followed by a table with the following data:

Demander :	KID_FEE_CONTROLLER:	Mtd (Nom_du_basemap)
Fichier existant :	@Value(MOROLOGAGE)	
Date / Heure de traitement :	@Value(MOROLOGAGE)	
Année de calcul :	Longueur = KERRIER_LONG_MDS: m (mètres)	Surface = KERRIER_SURFACE_MDS: m (mètres)
Espace spatial de référence choisi (1) :	WGS84_REF	Surface = @Value(MO_RAPPORT_AREA_SVP_SPATIAL_REF_EN_RMG)
Projection de fichier :	@Value(MO_PROJ_CHECKED): @Value(MO_PROJ_CHECKED_AWT)	
Type de géométrie analysé :	@Value(TYPE_GEOMETRIE)	
Nombre de (1) analysés :	@Value(TOTAL_GEOMETRIE_BRUT_INPUT)	

Below the table, there is a note: "(1) Les espaces spatiaux de référence sont limités aux couches "COMMUNE", "ENCL", "REPARTEMENT" et "REGION" de la BD TOPOR (VGE). Par conséquent, selon la source des espaces spatiaux de référence utilisés pour votre fichier, des erreurs peuvent se glisser dues à une différence de précision entre vos données et l'espace spatial de référence choisi." Below this note is another note: "Note : En cas d'erreur sur les projections, les fichiers d'erreurs d'extension ".shp" sont joints au mail." At the bottom right of the rendered page, there is a section titled "NOTATION GLOBALE = @Value(NOTATION_GLOBALE_POLYgone) sur 5".

Fig. 80 : Capture d'écran de l'interpréteur en ligne

La partie gauche (fond d'écran noir) présente le code et partie droite présente sa représentation interprétée.

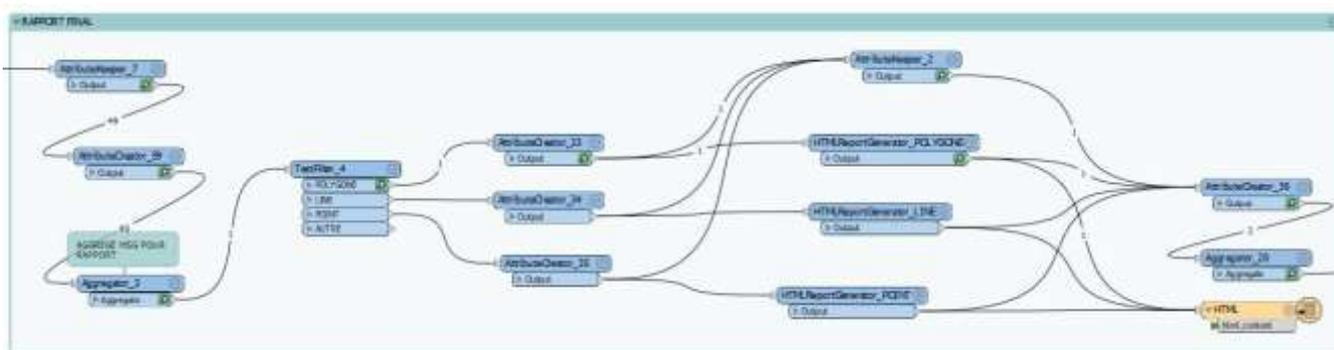


Fig. 81 : Visualisation du signet de création des rapports au format html

Un rapport spécifique est prévu pour chaque type de géométrie

Envoi automatique des mails

Le monde des SIG utilise maintenant massivement les consultations et téléchargements de données en ligne. Il convenait donc de proposer à l'utilisateur de récupérer à distance l'évaluation de son «SHAPE». C'est la solution du mail automatique qui a été alors retenue.

En avance de phase, une boîte mail sur le domaine «Gmail.com» a été spécifiquement créé pour tester et valider l'envoi et la réception des mails d'informations accompagnés de leurs pièces jointes.

Chaque contenu de mail a été défini que ce soit le corps de mail ou ses pièces jointes, le «transformer» «Testfilter» en fonction des caractéristiques d'erreurs transmises par le flux de données prévoit tous les cas de figure possibles. Les différents fichiers d'erreurs, ou leur absence déclenchent l'envoi du mail avec les pièces jointes.



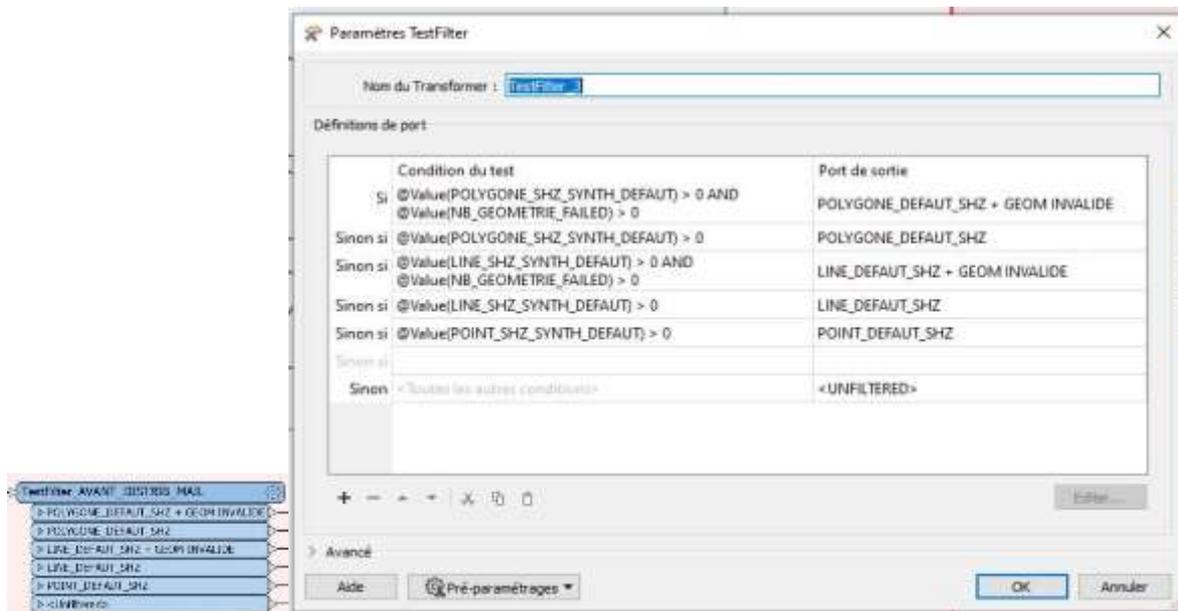


Fig. 82 : Visualisation du «transformer» «Testfilter» et de son paramétrage

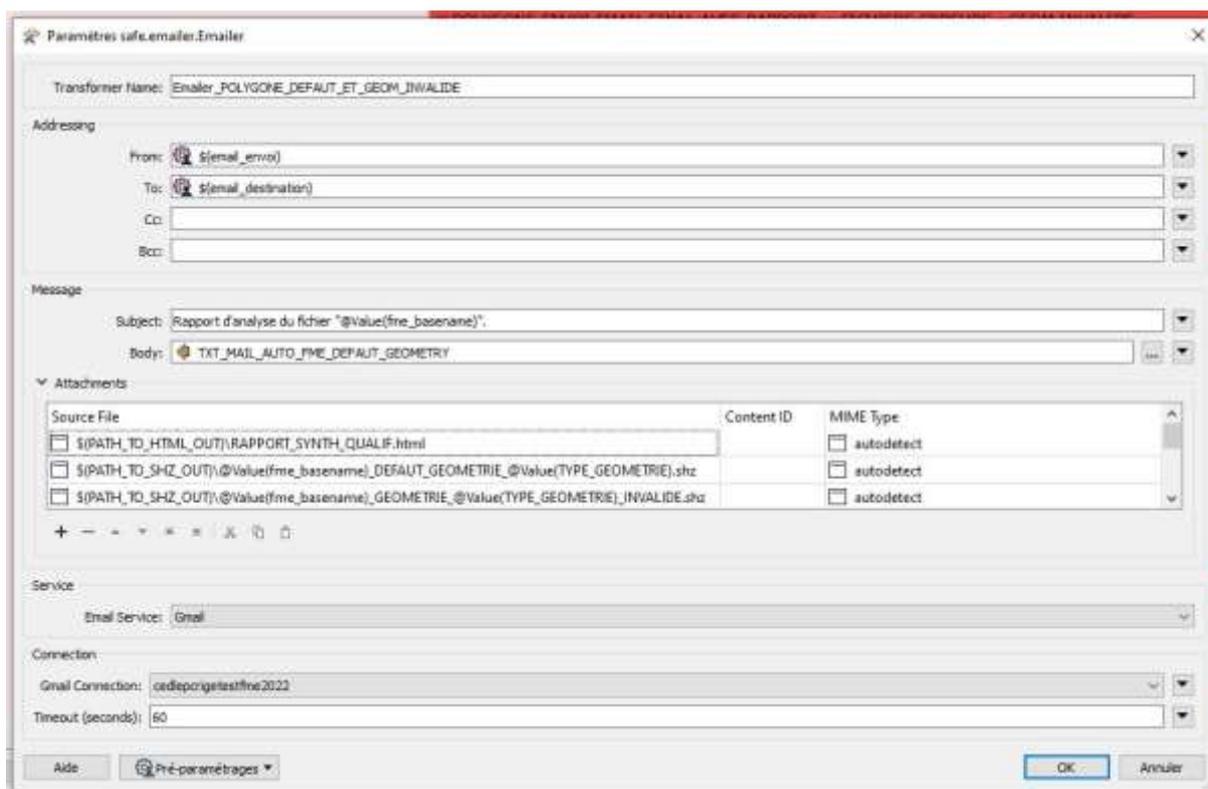
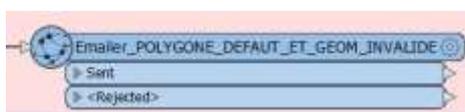


Fig. 83 : Visualisation du «transformer» «Emailer» et de son paramétrage

Une partie des noms de fichier est variable et correspond aux attributs définis au cours du traitement de «FME».

- Le fichier «EMP_SPATIALE_REF.shz» a un nom fixe et correspond au fichier de l'emprise spatiale de référence choisie.



- Le fichier d'extension .shz qui contient l'expression «_DEFAULT_GEOMETRIE_» correspond à l'export des géométries en doublons et / ou des micro-lignes ou micro-polygones en fonction de la géométrie contenu dans le «SHAPE»
- Le fichier d'extension .shz qui contient l'expression «_GEOMETRIE_...._INVALIDE.shz» correspond à l'export des géométries corrompues.
- Le fichier «RAPPORT_SYNTHT_QUALIF.html» à un nom fixe et correspond au rapport de synthèse.

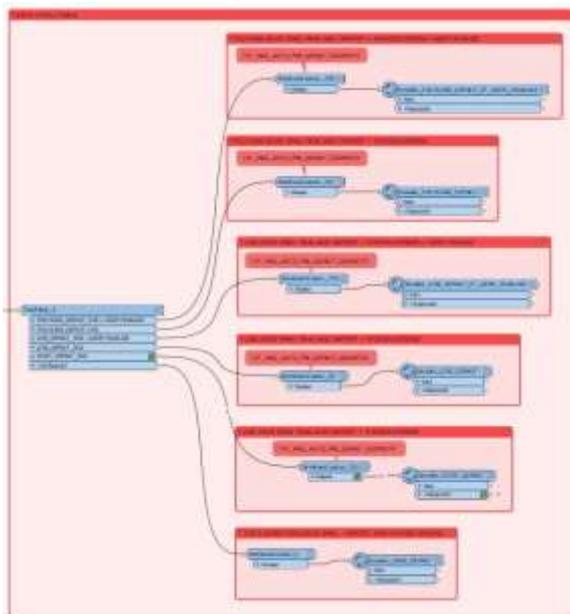


Fig. 84 : Visualisation du signet des «E-mailer» - fin du traitement

Cet ultime signet du script marque la fin du traitement.

Les rapport HTML, pièces jointes des mails automatiques, sont présentés en annexe 1 à 4.

Un outil opérationnel mais perfectible

Difficultés spécifiques à l'élaboration du script

La prise en main et l'apprentissage de «FME» ont nécessité un certain temps d'adaptation car il faut chercher dans le guide des transformers l'outil le plus adéquat pour le morceau de traitement que l'on veut réaliser. De plus il faut en permanence concilier les aspects macro et micro du script. Il faut donc régulièrement prendre du recul sur la chaîne de traitement pour lui conserver tous son sens. Le paramétrage précis et technique d'un transformer particulier doit trouver son sens et son utilité à l'intérieur du schéma général de traitement. Il s'agit de faire des allers-retours constants entre l'objectif général et le paramétrage particulier de chaque «transformer», l'enchaînement des «transformer» et de leurs paramétrages doit rester cohérent pour assurer une continuité cohérentes et pertinente entre l'amont et l'aval de chaque «transformer».

De plus, le script ne doit pas être pris en défaut et doit pouvoir s'adapter à tous les cas de figures il faut donc penser et prévoir toutes les variantes possibles.

Il a été aussi nécessaire de trouver une réponse au fonctionnement interne du script pour caractériser une absence, comment traiter et déclencher une action à partir d'une absence de donnée dans une branche de

traitement. C'est le cas notamment quand toutes les géométries sont à l'extérieur d'une emprise spatiale de référence.

«FME» est un outil puissant mais relativement rigide aussi a-t-il fallu renoncer à certains rendu pourtant envisagés à la conception du script.

Le «transformer» qui converti un HTML en PDF génère des erreurs, la source du problème est la conversion de l'encodage du fichier.

Le format HTML est supporté par «FME» mais le logiciel inscrit par défaut des instructions BOOTSTRAP en tête de fichier. Cela engendre des petites différences d'interprétation en fonction du navigateur web utilisé

Pour pallier à ces résultats de traitement qui ne donnaient pas satisfaction, il a donc fallu s'adapter à ces contraintes et trouver une solution à partir des outils à notre disposition.

Les axes de développement du qualificateur

Le script dans son état actuel est fonctionnel. La norme ISO 19157 décrit cinq critères principaux et ce projet de fin d'étude n'a permis en partie par manque de temps de ne traiter que partiellement deux de ces critères, c'est dire l'ampleur de la tâche pour traiter tous les critères et sous critères associées sous tous leurs aspects et notamment pouvoir déduire les usages possibles de la donnée à partir de l'évaluation sur cinq des critères objectifs.

La nomenclature des différents usages est encore en débat au sein du groupe de travail QuaDoGéo et l'état d'avancement du document ne permet pas encore son exploitation sur cet aspect qui s'avère complexe à mettre en place dans le script.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
	thème (sujet)	usages	info_3d	nivl_etap	l1/lms	exemple de données	sy visuels et / ou cartog: se aptier:	inventorier (recenser)	gener	localiser	analyse	planifier	suivre, observer	
1	agriculture	Visualiser et analyser rapidement les systèmes de cultures	visualisation / cartographie	cartographeur	visualiser	RPG	X							
2	agriculture	Simplifier la gestion de son exploitation agricole	Hors sujet	gener	gener	référentiels pédologiques			X					
3	agriculture	Interroger des données géolocalisées	visualisation / cartographie	localiser	localiser	(usage très, trop générique ?)								
4	Aménagement du territoire	Favoriser la dématérialisation de l'application du droit des sols	Hors sujet	gener	administrer				X					
5	Aménagement du territoire	Analyser et connaître les territoires	analyser, traiter les données		analyser	(usage très, trop générique ?)						X	X	
6	Aménagement du territoire	Lutter contre l'artificialisation des sols	visualisation / cartographie	planifier	planifier	OCSDL, MOS						X	X	
7	Biodiversité	Cartographier les habitats naturels	visualisation / cartographie	localiser	cartographeur		X							
8	Biodiversité	Lutter contre l'artificialisation des sols	visualisation / cartographie	planifier	planifier							X	X	
9	Biodiversité	Favoriser le développement de l'agrosylviculture	visualisation / cartographie	gener	planifier	Hors Sujet sur les usages...								
10	Chasse	Suivre les affectifs d'animaux	visualisation / cartographie	suivre/observer/evaluer	observer		X	X	X	X				
11	Chasse	Quantifier les dégâts agricoles	analyser, traiter les données	recenser	recenser			X						
12	Chasse	Cartographier les emprises chassées	visualisation / cartographie	cartographeur	cartographeur									
13	Climat	Surveiller le changement climatique	analyser, traiter les données	suivre/observer/evaluer	observer	trop métier	X	X	X	X				
14	Climat	Visualiser les modèles climatiques pour étudier les impacts	analyser, traiter les données	simuler	simuler	trop métier	X							
15	Climat	Fournir des données de référence sur la forêt	analyser, traiter les données	donnees pivot/ref	pivot/ref	bd forêt	X	X						
16	Climat	Prevenir les risques d'incendie	anticiper/modéliser	planifier ou gener	protéger	spécifique au métier								
17	Culture	Valeriser les lieux culturels et patrimoniaux	visualisation / cartographie	localiser	localiser		X	X		X				
18	Culture	Découvrir des lieux de culture grâce au jeu vidéo Minecraft	Hors Sujet	faciliter	???	Hors Sujet								
19	Culture	Contribuer à la description des territoires	visualisation / cartographie	analyser	décrire	trop vague								
20	Culture	Géolocaliser des événements culturels	visualisation / cartographie	localiser	localiser		X			X				
21	Défense	Produire des images satellitaires adaptées aux besoins militaires	visualisation / cartographie	donnees pivot/ref	pivot/ref	images satellitaires (données pivot)								

Fig. 85 : Visualisation du fichier de travail collaboratif en cours sur les usages de la donnée

(Source : https://lite.framacalc.org/9qfs-quadogeo_usages_ig)



Cependant, en se concentrant sur le script existant, il se dégage déjà des axes d'améliorations et d'optimisations, plus concret et immédiatement plus adaptable. Cela nécessite du temps de développement supplémentaire pour optimiser les performances et proposer la réponse la plus souple et la plus adaptée à tous les cas de figure.

Axes d'amélioration possibles en début de traitement

Le début du traitement pourrait être amélioré par les développements complémentaires suivants :

- une détection automatique de fichiers compressés
- un décompresseur de fichiers pour traiter les fichiers compressés
- la gestion d'autres formats de fichiers (.gpkg, geodatabase, geojson etc..)
- un traitement par lot des fichiers (batch)

Axes d'amélioration possibles en cours de traitement

Le cours du traitement pourrait également être optimisé par les développements complémentaires suivants :

- par la rationalisation et la diminution du nombre de «transformers» qui raccourcirait le temps de traitement (surtout pour les fichiers volumineux).
- l'insertion d'éléments graphiques dans les rapports dynamiserait et rendrait plus conviviale sa lecture
- l'intégration d'autres évaluations pour affiner le diagnostic
- le traitement de la complétude des tables attributaires.

Axes d'amélioration possibles en fin de traitement

La fin du traitement pourrait également être enrichie des développements complémentaires suivants :

- déduire l'usage potentiel de la donnée en fonction de l'évaluation du fichier
- améliorer l'esthétique des rapports par l'évolution du code HTML et l'introduction d'instructions en CSS
- intégrer un convertisseur de HTML vers PDF
- intégrer un compresseur de fichier

FME un logiciel propriétaire sous licence

L'utilisation de «FME» était l'une des contraintes à respecter, cependant il existe d'autres logiciels «ETL» disponibles sur le marché tels que GeoKeettle, Scriptelia, CloverDX, Apatar, Knowage, Talend, TIBCO Jaspersoft, Pentaho Data Integration (Source : <https://www.axysweb.com/a-la-decouverte-des-solutions-etl-open-source/>)

Ces logiciels ETL pourraient constituer une alternative à «FME» même si certaines versions sont également payantes et il convient au préalable d'en vérifier d'abord les caractéristiques, performances, compatibilités etc... .

Etablir un tel comparatif des performances et capacités de ces «ETL» open source constitue déjà en soit un travail important et cette tâche ne faisait pas partie de celles qui m'ont été confiées.

CONCLUSION

D'un point de vue factuel et opérationnel, Le script dans sa version actuelle, bien qu'incomplète au regard de l'ensemble des critères de la norme ISO 19157, fonctionne. Il a été testé sur des données de référence de l'IGN (BD.TOPO®) mais aussi sur des données issues du CRIGE - PACA de sources variées portant sur des thématiques diverses. Sur des données de référence, le script sur les données de référence ne renvoie pas d'erreur ce qui est logique.

Sur les fichiers créés ad hoc avec différents types d'erreurs caractéristiques, les rapports adéquats sont générés. Les sorties générées par le script sont donc pertinentes et cohérentes au regard du paramétrage du script.

Plus globalement, au-delà de l'apprentissage du logiciel «FME» qui constitue une nouvelle compétence technique intéressante à mobiliser et exploiter, cette immersion de fin d'étude dans le monde professionnel m'a permis de mieux m'imprégner de l'écosystème, des enjeux, des thématiques de la diversité des acteurs et pratiques du secteur d'activité.

Enfin, l'analyse fine des données, la construction d'un processus de diagnostic relativement autonome et sa réalisation au travers d'un script opérationnel m'ont permis de prendre le recul nécessaire sur le sujet pour proposer à mon niveau une réponse cohérente et la plus adaptée possible aux attentes de la communauté des SIG. Sur ces aspects techniques et en particulier sur les attentes des utilisateurs de la donnée géolocalisée, les échanges avec mon maître stage ont nourri et ont inspiré ma réflexion, ils m'ont conduit à me poser les bonnes questions pour y apporter la réponse la plus satisfaisante possible. Ce travail sera présenté et proposé au Groupe de Travail QuaDoGéo lors de leur prochaine réunion du 22 Octobre 2022 pour une éventuelle mise en production.

TABLES DES FIGURES

Fig. 1 : Carte des CRIGE en France.....	5
Fig. 2 : Plan de situation du CRIGE (petite échelle)	6
Fig. 3 : Plan de situation du CRIGE (grande échelle)	6
Fig. 4 : Espace de téléchargement des fiches du CEREMA sur le site du CRIGE - PACA.....	9
Fig. 5 : Extrait de la Fiche N°7 du CEREMA sur le critère de cohérence logique page 1/12.....	10
Fig. 6 : Extrait de la Fiche N°7 du CEREMA sur le critère de cohérence logique page 5/12.....	11
Fig. 7 : Exemple du validateur de référence IBAN	12
Fig. 8 : Extrait du document de retour du validateur	12
Fig. 9 : Exemple du validateur de la Base Adresse Locale	12
Fig. 10 : Extrait de la métadonnée de «ADMIN-EXPRESS_3-1» - section dédiée à la qualité.....	13
Fig. 11 : Les critères et sous critères de la norme ISO 19157 retenus pour l'élaboration du script.....	14
Fig. 12 : Base de la réflexion pour la conception du script	15
Fig. 13 : Exemple de rapport de notation avec note globale et sous-critères d'évaluation.	16
Fig. 14 : Illustration des différents modules de «FME»	18
Fig. 15 : Liste des «transformers» utilisés	18
Fig. 16 : Extrait du guide des transformers FME 2020- illustration des multiples possibilités de traitement de la donnée.....	19
Fig. 17 : Illustration des différentes fenêtres de travail de «FME»	20
Fig. 18 : Visualisation du masque de saisie des informations pour lancer le traitement.....	21
Fig. 19 : Extrait de la notice d'utilisation du script pour un utilisateur basique.	21
Fig. 20 : Visualisation du signet traitant l'emprise spatiale de référence	23
Fig. 21 : Visualisation du signet de création du fichier .shz de l'emprise spatiale de référence joint aux mails	23
Fig. 22 : Liste des extensions testées pour l'analyse de la structure du «SHAPE»	24
Fig. 23 : Visualisation du signet testant la validité de la structure du «SHAPE».....	24
Fig. 24 : Visualisation des «transformers» et de leur paramétrage pour la notation individuelle des fichiers constitutifs du «SHAPE»	25



Fig. 25 : Visualisation du signet statuant définissant le statut du «SHAPE».....	25
Fig. 26 : Visualisation du message affiché par «FME» en cas d'absence de fichier indispensable au «SHAPE».....	25
Fig. 27 : Visualisation du signet calculant le poids des données transmises.....	26
Fig. 28 : Visualisation globale du signet traitant la structure du «SHAPE».....	26
Fig. 29 : Extrait du rapport du rapport de synthèse final - partie dédiée au sous critère de cohérence du format.....	26
Fig. 30 : Visualisation du transformer «SystemCaller» et de son paramétrage.....	28
Fig. 31 : Visualisation du répertoire de travail avant exécution du «SystemCaller».....	28
Fig. 32 : Visualisation du répertoire de travail après exécution du «SystemCaller».....	29
Fig. 33 : Visualisation du mail envoyé automatiquement en cas de mauvaise projection.....	29
Fig. 34 : Visualisation du mail envoyé automatiquement en cas d'absence de projection.....	29
Fig. 35 : Visualisation du message affiché par «FME» en cas de projection non autorisée.....	30
Fig. 36 : Visualisation du message affiché par «FME» en cas d'absence de projection- absence de fichier .prj.....	30
Fig. 37 : Visualisation du signet statuant sur la validité de la projection utilisée dans le «SHAPE».....	30
Fig. 38 : Visualisation du signet traitant les informations générales du «SHAPE».....	31
Fig. 39 : Extrait du rapport de synthèse final - partie dédiée aux informations générales.....	31
Fig. 40 : Visualisation du signet dénombrant les géométries en entrée.....	32
Fig. 41 : Visualisation du «transformer» «GeometryValidator» et de son paramétrage.....	32
Fig. 42 : Extrait de l'aide en ligne de «FME» sur la norme OGC.....	33
Fig. 43 : Extrait de l'aide en ligne de «FME» sur les géométries corrompues détectées.....	33
Fig. 44 : Synthèse des erreurs topologiques recherchées et mises en évidence.....	34
Fig. 45 : Visualisation du «transformer» «FeatureWriter» et de son paramétrage.....	35
Fig. 46 : Visualisation du fichier «SHAPE» des polygones en défaut sous QGIS.....	36
Fig. 47 : Visualisation du fichier «SHAPE» des lignes en défaut sous QGIS.....	37
Fig. 48 : Visualisation du signet traitant la validité des géométries.....	37
Fig. 49 : Extrait du rapport de synthèse illustrant la complémentarité entre les modalités de contrôle et la notation de cet élément.....	37
Fig. 50 : Visualisation du signet filtrant les géométries selon leur nature.....	38
Fig. 51 : Visualisation du «transformer» «AreaCalculator» et résultat de son action.....	39
Fig. 52 : Visualisation du «transformer» «Attribut Creator» qui génère la variable «ANALYSE_SURF_POLYgone_INSIDE_CF_SEUIL» et résultat de son action.....	39
Fig. 53 : Visualisation du «transformer» «Tester» et de son paramétrage.....	40
Fig. 54 : Visualisation du signet traitant les micro-polygones.....	40
Fig. 55 : Extrait du rapport de synthèse final - partie dédiée à la notation des «micro-polygones».....	41
Fig. 56 : Visualisation du «transformer» «Matcher» et de son paramétrage.....	41
Fig. 57 : Visualisation du signet traitant les polygones en doublons.....	42
Fig. 58 : Visualisation du signet générant le fichier des erreurs .shz des micro-polygones et des doublons.....	42
Fig. 59 : Visualisation du fichier «SHAPE» des micro-polygones et polygones en doublons sous QGIS.....	42
Fig. 60 : Visualisation du «transformer» gérant le calcul du sous-critère de cohérence topologique des polygones et son paramétrage.....	43
Fig. 61 : Visualisation du signet de calcul du sous-critère de cohérence topologique des polygones.....	44
Fig. 62 : Visualisation du «transformer» de calcul du critère d'exhaustivité des polygones et son paramétrage.....	44
Fig. 63 : Visualisation du signet de calcul du critère d'exhaustivité des polygones.....	45
Fig. 64 : Visualisation du «transformer» «Clipper» et de son paramétrage.....	46
Fig. 65 : Illustration des erreurs générées par application de l'emprise spatiale de référence choisie sur la même emprise définie avec une autre précision.....	46
Fig. 66 : Visualisation du transformer de calcul du sous-critère de cohérence topologique et son paramétrage.....	47
Fig. 67 : Visualisation du signet de calcul du sous-critère de cohérence topologique.....	48
Fig. 68 : Extrait du rapport du rapport de synthèse final - partie dédiée au sous critère de cohérence topologique.....	48
Fig. 69 : Visualisation du «transformer» «AttributeCreator» définissant le calcul du ratio et son paramétrage.....	48
Fig. 70 : Visualisation du signet de calcul de la somme des surfaces de polygones à l'intérieur de l'emprise spatiale de référence.....	49
Fig. 71 : Visualisation du signet de calcul de la notation du critère d'exhaustivité des polygones et son paramétrage.....	49
Fig. 72 : Visualisation du signet de calcul du critère d'exhaustivité des polygones.....	49
Fig. 73 : Extrait du rapport de synthèse final - partie dédiée au critère d'exhaustivité.....	50
Fig. 74 : Visualisation du «transformer» «AttributeCreator» définissant le calcul de la note globale et de son paramétrage.....	50
Fig. 75 : Visualisation du signet de calcul de la notation globale des polygones.....	50
Fig. 76 : Extrait du rapport de synthèse final – Notation globale du fichier.....	51
Fig. 77 : Logo du CRIGE utilisé pour récupérer les codes hexadécimaux des couleurs.....	51



Fig. 78 : Visualisation du «transformer» «HTMLReportGenerator» et de son paramétrage	52
Fig. 79 : Extrait du code d'un rapport édité sous «notepad»	52
Fig. 80 : Capture d'écran de l'interpréteur en ligne	53
Fig. 81 : Visualisation du signet de création des rapports au format html.....	53
Fig. 82 : Visualisation du «transformer» «Testfilter» et de son paramétrage	54
Fig. 83 : Visualisation du «transformer» «Emailer» et de son paramétrage.....	54
Fig. 84 : Visualisation du signet des «Emailer» - fin du traitement.....	55
Fig. 85 : Visualisation du fichier de travail collaboratif en cours sur les usages de la donnée	56

TABLES DES ANNEXES

Annexe 1 : Rapport d'analyse sur un fichier de polygones valides, corrompues et doublonnés	65
Annexe 2 : Rapport d'analyse sur un fichier ou toutes les lignes sont en dehors de l'emprise spatiale de référence et certaines sont corrompues et / ou doublonnées.....	67
Annexe 3 : Rapport d'analyse sur un fichier dont la structure «SHAPE» est incomplète	69

WEBOGRAPHIE

De la donnée au SIG : collecte, structuration et valorisation. (2019, 24 juillet). LE MAGAZINE. Consulté le 24 août 2022, à l'adresse <https://www.gismartware.com/blog/donnees-geographiques-valorisation-sig/>

La qualité des données géographiques état des lieux pour un débat (ISSN : 1263–2570 ISRN : Certu/RE--10-14--FR Certu) (2010, novembre). CERTU.

ISO 19157. (2013). Online Browsing Platform (OBP). Consulté le 24 août 2022, à l'adresse <https://www.iso.org/obp/ui/#iso:std:iso:19157:ed-1:v1:fr>

Qualifier les données géographiques. (2017, décembre). CEREMA. Consulté le 29 août 2022, à l'adresse <https://www.cerema.fr/fr/centre-ressources/boutique/qualifier-donnees-geographiques>

Métadonnées relatives à la qualité des données géographiques. (2021, 24 février). data.geocatalogue.fr. Consulté le 24 août 2022, à l'adresse [//data.geocatalogue.fr/ncl/mesuresQuaDoGeo](https://data.geocatalogue.fr/ncl/mesuresQuaDoGeo)

BIBLIOGRAPHIE

Benoît David, Pascal Fasquel. Qualité d'une base de données géographique : concepts et terminologie. [Rapport de recherche] Institut géographique national. 1997, 53 p. hal-02372984



Rodolphe Devillers, R. D., & Robert Jeansoulin, R. J. (2005). *Qualité de l'information géographique*. Lavoisier.

TABLES DES ACRONYMES

AFIGEO

Association **F**rançaise pour l'**I**nformation **G**éographique

«Association loi 1901 créée en 1986, l'Afigéo poursuit trois principales missions :

Animer la communauté française des acteurs et des réseaux de l'information géographique : service de l'État et collectivités locales, institutions publiques et grands comptes, industries et entreprises de toutes tailles, universités et centres de formation et de recherche, associations et médias...

Promouvoir le secteur de l'information géographique en France et à l'international

Représenter la filière et ses différents acteurs et réseaux auprès des instances nationales»

(Source :

<https://www.afigeo.asso.fr/afigeo/association/#:~:text=Promouvoir%20le%20secteur%20de%20l,r%C3%A9seaux%20aupr%C3%A8s%20des%20instances%20nationales>)

BOOTSTRAP

«Bootstrap est une collection d'outils utiles à la création du design (graphisme, animation et interactions avec la page dans le navigateur, etc.) de sites et d'applications web. C'est un ensemble qui contient des codes HTML et CSS, des formulaires, boutons, outils de navigation et autres éléments interactifs, ainsi que des extensions JavaScript en option.»

(Source : [https://fr.wikipedia.org/wiki/Bootstrap_\(framework\)](https://fr.wikipedia.org/wiki/Bootstrap_(framework)))

CEREMA

Centre d'**E**tudes et d'expertise sur les **R**isques, l'**E**nvironnement, la **M**obilité et l'**A**maénagement

«Véritable centre de ressources et d'expertises scientifiques et techniques interdisciplinaires, le Cerema assure la diffusion et la promotion des travaux et des études liés à ses activités par le biais de publications d'ouvrages, de formations et par l'organisation de journées techniques.»

(Source:

<https://www.cerema.fr/fr/cerema/missions#:~:text=V%C3%A9ritable%20centre%20de%20ressources%20et%20l'organisation%20de%20journ%C3%A9es%20techniques.>)

CNIG

Conseil **N**ational de l'**I**nformation **G**éolocalisée

«Le Conseil national de l'information géographique a pour mission d'éclairer le Gouvernement dans le domaine de l'information géographique, notamment pour ce qui concerne la coordination des contributions des acteurs concernés et l'amélioration des interfaces entre ces derniers. Il prend en compte les besoins exprimés et en particulier les besoins des utilisateurs. Il peut formuler des avis sur toute question relative à l'information géographique.»

(Source:

http://cnig.gouv.fr/?page_id=843#:~:text=Le%20Conseil%20national%20de%20l,et%20l'am%C3%A9lioration%20des%20interfaces)



CSS

Cascading Style Sheets (Feuille de style en Cascade)

«Le CSS correspond à un langage informatique permettant de mettre en forme des pages web (HTML ou XML).

Ce langage est donc composé des fameuses « feuilles de style en cascade » également appelées fichiers CSS et contient des éléments de codage.»

(source : <https://www.atinternet.com/glossaire/css/>)

EPCI

Etablissement Public de Coopération Intercommunale

«Les établissements publics de coopération intercommunale (EPCI) sont des regroupements de communes ayant pour objet l'élaboration de « projets communs de développement au sein de périmètres de solidarité ». Ils sont soumis à des règles communes, homogènes et comparables à celles de collectivités locales. Les communautés urbaines, communautés d'agglomération, communautés de communes, syndicats d'agglomération nouvelle, syndicats de communes et les syndicats mixtes sont des EPCI.»

(source : <https://www.insee.fr/fr/metadonnees/definition/c1160>)

ETL

Extract Transform Load (Extraction, transformation, chargement)

FME

Feature Manipulation Engine

GT QuaDoGéo

Groupe de Travail sur la Qualité des Données Géographiques.

«Au-delà de la seule « ouverture » de la donnée, et dans un but de fiabilisation de la réutilisation, il est important d'informer l'utilisateur potentiel sur la qualité d'une donnée, pour qu'il puisse savoir si elle correspond à ses besoins pour palier au risque de mauvaise utilisation et d'incompréhension. Il devient ainsi indispensable de mettre en place des méthodes et des outils, accessibles à tous, permettant d'estimer et de communiquer sur la qualité des données diffusées.»

Mandat signé du Groupe de Travail « QuaDoGéo » (Source : http://cnig.gouv.fr/?page_id=18183)

HTML

Hyper Text Markup Language

«L'HyperText Markup Language, HTML, désigne un type de langage informatique descriptif. Il s'agit plus précisément d'un format de données utilisé dans l'univers d'Internet pour la mise en forme des pages Web. Il permet, entre autres, d'écrire de l'hypertexte, mais aussi d'introduire des ressources multimédias dans un contenu.»

(source : <https://www.journaldunet.fr/web-tech/dictionnaire-du-webmastering/1203255-html-hypertext-markup-langage-definition-traduction/>)

QGIS

Quantum Geographic Information System



IBAN

International **B**ank **A**ccount **N**umber

«l'IBAN est une norme internationale ISO qui a été mise en place pour sécuriser et faciliter les échanges et transferts bancaires internationaux sur tous les continents.»

(Source :<https://www.pricebank.fr/questions-reponses/questions-frequentes/qu-est-ce-que-le-code-iban.html>)

LIZMAP

Lizmap est un logiciel open-source, conçu par 3Liz. Il fonctionne avec QGIS et permet de réaliser de la cartographie par le WEB.

NORME ISO 19157 : 2013

Norme de référence sur la qualité des données géographiques

OGC

« L'**O**pen **G**eospatial **C**onsortium est un consortium international pour développer et promouvoir des standards ouverts, les spécifications Open GIS, afin de garantir l'interopérabilité des contenus, des services et des échanges dans les domaines de la géomatique et de l'information géographique.»

(Source :https://fr.wikipedia.org/wiki/Open_Geospatial_Consortium)



ANNEXES



Centre de Ressources
en Information Géographique
Provence-Alpes-Côte d'Azur

RAPPORT D'ANALYSE DU FICHIER "POLYGONE_TTE_ERREUR"

INFORMATIONS GENERALES SUR LE FICHIER

Demandeur :	LE STAGIAIRE DU CRIGE - PACA	Mail : cedleprigetestfme2022@gmail.com
Fichier contrôlé :	POLYGONE_TTE_ERREUR	
Date / Heure du traitement :	Le 25/08/2022 à 12h13	
Seuils de calcul :	Longueur = 5 m linéaire(s)	Surface = 7517560 mètre(s)²
Emprise spatiale de référence choisie (1) :	"EPCI CC du Golfe de Saint-Tropez"	Surface = 433.3 Km²
Projection du fichier :	RGF93 / Lambert 93 / Code EPSG = 2154	
Type de géométrie analysé :	POLYGONE	
Nombre de POLYGONE(S) analysé(s) :	17	

(1) Les emprises spatiales de références sont issues des couches "COMMUNE", "EPCI", "DEPARTEMENT" et "REGION" de la BD TOPO® (IGN). Par conséquent selon la source des emprises spatiales de référence utilisée pour votre fichier, des erreurs peuvent se générer dues à une différence de précision entre vos données et l'emprise spatiale de référence choisie.

Nota : En cas d'erreurs sur les géométries, les fichiers d'erreurs d'extension ".shx" sont joints au mail.

NOTATION GLOBALE = 4.2 sur 5

CRITERE DE COHERENCE LOGIQUE = 4.4 sur 5

SOUS CRITERE DE COHERENCE DU FORMAT = 5 sur 5

Nature du contrôle	Modalité du contrôle	Résultat
Poids des fichiers	Calcul du poids des données transmises	0,53 Mo

Nature du contrôle	Modalité du contrôle	Résultat

Liste des fichiers et type d'extension	contrôle sur le type d'extension	POLYGONE_TTE_ERREUR.cpg POLYGONE_TTE_ERREUR.dbf POLYGONE_TTE_ERREUR.prj POLYGONE_TTE_ERREUR.shp POLYGONE_TTE_ERREUR.shx
--	----------------------------------	---

Nature du contrôle	Modalité du contrôle	Résultat
Structure du fichier	Vérification du format de fichier	Les fichiers d'extension ".gmd" ".glt" ".shn" ".shx" ".fhn" ".fbx" ".ain" ".nh" ".atx" ".lxa" ".mxx" sont absents
	La présence des fichiers obligatoires ".shp", ".dbf," et ".shx" est testée et l'absence des fichiers ".prj" ou ".cpg" est relevée pour définir la complétude et la validité du "SHAPE"	La structure du "SHAPE" est valide et complète

SOUS CRITERE DE COHERENCE TOPOLOGIQUE = 3.7 sur 5

Nature du contrôle	Modalité du contrôle	Résultat	Notation
Vérification de l'intégrité géométrique des POLYGONES	Recherche des POLYGONES corrompus. Les objets classés en défaut à cet étape ne sont plus pris en compte pour la suite du traitement.	9 POLYGONE(S) valide(s) sur 17 soit 52.9 %.	2.6 sur 5

Nature du contrôle	Modalité du contrôle	Résultat	Notation
Recherche des doublons	Un POLYGONE est considéré en doublon si et seulement si la géométrie et tous les attributs sont strictement identiques	7 POLYGONE sans doublons sur 9 soit 77.8 %	3.9 sur 5

Nature du contrôle	Modalité du contrôle	Résultat	Notation
Recherche des "micro" - POLYGONES dont la surface individuelle est inférieure au seuil	Recherche des POLYGONES dont la surface est inférieure au seuil ($S = 7517560 \text{ m}^2$)	Les surfaces cumulées des POLYGONES dont la surface individuelle est supérieure au seuil représentent 296600382.5 m^2 sur 311741071.1 soit 95.1 %	4.8 sur 5

CRITERE D' EXHAUSTIVITE = 4 sur 5

Nature du contrôle	Modalité du contrôle	Résultat	Notation
Relation spatiale entre le fichier et l'emprise de référence choisie	Decoupe du fichier par l'emprise et comptage des surfaces	La somme des surfaces des POLYGONES à l'intérieur de l'emprise représente 248396043 m^2 sur 311741071.1 soit 79.7 %.	4 sur 5

CRIGE PACA - Domaine du Petit Arbois – Bât. Loarnec 13100 Aix-en-Provence - Tel : 04 42 90 71 22 - Mail : contact@crige-paca.org



RAPPORT D'ANALYSE DU FICHIER "LINE_TTE_ERREUR_ALL_OUTSIDE"

INFORMATIONS GENERALES SUR LE FICHIER

Demandeur :	LE STAGIAIRE DU CRIGE - PACA	Mail : cedleprigetstfme2022@gmail.com
Fichier contrôlé :	LINE_TTE_ERREUR_ALL_OUTSIDE	
Date / Heure du traitement :	Le 25/08/2022 à 12h16	
Seuils de calcul :	Longueur = 5 m linéaire(s)	Surface = 7517560 mètre(s)²
Emprise spatiale de référence choisie (1) :	"DEPARTEMENT Alpes-de-Haute-Provence"	Surface = 6994.2 Km²
Projection du fichier :	RGF93 / Lambert 93 / Code EPSG = 2154	
Type de géométrie analysé :	LINE	
Nombre de LINE(S) analysé(s) :	17	

(1) Les emprises spatiales de références sont issues des couches "COMMUNE", "EPCI", "DEPARTEMENT" et "REGION" de la BD TOPO® (IGN). Par conséquent selon la source des emprises spatiales de référence utilisée pour votre fichier, des erreurs peuvent se générer dues à une différence de précision entre vos données et l'emprise spatiale de référence choisie.

Nota : En cas d'erreurs sur les géométries, les fichiers d'erreurs d'extension ".shx" sont joints au mail.

NOTATION GLOBALE = 2.3 sur 5

CRITERE DE COHERENCE LOGIQUE = 4.5 sur 5

SOUS CRITERE DE COHERENCE DU FORMAT = 5 sur 5

Nature du contrôle	Modalité du contrôle	Résultat
Poids des fichiers	Calcul du poids des données transmises	0.004 Mo

Nature du contrôle	Modalité du contrôle	Résultat

Liste des fichiers et type d'extension	contrôle sur le type d'extension	LINE_TTE_ERREUR_ALL_OUTSIDE.cpg LINE_TTE_ERREUR_ALL_OUTSIDE.dbf LINE_TTE_ERREUR_ALL_OUTSIDE.prj LINE_TTE_ERREUR_ALL_OUTSIDE.shp LINE_TTE_ERREUR_ALL_OUTSIDE.shx
---	---	--

Nature du contrôle	Modalité du contrôle	Résultat
	Vérification du format de fichier	Les fichiers d'extension : ".qmd" ".qlx" ".sbn" ".shx" ".fbs" ".fbx" ".obn" ".obh" ".otx" ".lxs" ".mxx" sont absents
Structure du fichier	La présence des fichiers obligatoires ".shp", ".dbf," et ".shx" est testée et l'absence des fichiers ".prj" ou ".cpg" est relevée pour définir la complétude et la validité du "SHAPE"	La structure du "SHAPE" est valide et complète

SOUS CRITERE DE COHERENCE TOPOLOGIQUE = 3.9 sur 5

Nature du contrôle	Modalité du contrôle	Résultat	Notation
Vérification de l'intégrité géométrique des LINES	Recherche des LINES corrompues. Les objets classés en défaut à cet étape ne sont plus pris en compte pour la suite du traitement.	12 LINE(S) valide(s) sur 17 soit 70.6 %.	3.5 sur 5

Nature du contrôle	Modalité du contrôle	Résultat	Notation
Recherche des doublons	Une LINE est considérée en doublon si et seulement si la géométrie et tous les attributs sont strictement identiques	8 LINE(S) sans doublons sur 12 soit 66.7 %	3.3 sur 5

Nature du contrôle	Modalité du contrôle	Résultat	Notation
Recherche des "micro" - LINES dont la longueur individuelle est inférieure au seuil	Recherche des LINES dont la longueur est inférieure au seuil (S = 7517560 m)	Les longueurs cumulées des LINES dont la longueur individuelle est supérieure au seuil représentent 267276.2 m linéaire(s) sur 267276.2 soit 100 %	5 sur 5

CRITERE D' EXHAUSTIVITE = 0 sur 5

Nature du contrôle	Modalité du contrôle	Résultat	Notation
Relation spatiale entre le fichier et l'emprise de référence choisie	Decoupe du fichier par l'emprise et comptage des longueurs	La somme des longueurs des LINES à l'intérieur de l'emprise représente 0 m linéaire(s) sur 267276.2 soit 0 %.	0 sur 5

CRIGE PACA - Domaine du Petit Arbols – Bât. Loennec 13100 Aix-en-Provence - Tel : 04 42 90 71 22 - Mail : contact@crige-paca.org

Annexe 2 : Rapport d'analyse sur un fichier où toutes les lignes sont en dehors de l'emprise spatiale de référence et certaines sont corrompues et / ou doublonnées



RAPPORT D'ANALYSE DU FICHIER "STRUCT_SHAPE_PARTIEL"

INFORMATIONS GENERALES SUR LE FICHIER

Demandeur :	LE STAGIAIRE DU CRIGE - PACA	Mail : cedlepcrigitestfme2022@gmail.com
Fichier contrôlé :	STRUCT_SHAPE_PARTIEL	
Date / Heure du traitement :	Le 25/08/2022 à 12h40	
Seuils de calcul :	Longueurs = 5 m linéaire(s)	Surface = 7517560 mètre(s) ²
Emprise spatiale de référence choisie :	"EPCI CC du Golfe de Saint-Tropez"	Surface = 433.3 Km ²
Projection du fichier >	RGF93 / Lambert 93 / Code EPSG = 2154	
Type de géométrie analysé :	POINT	
Nombre de POINT(S) analysé(s) :	55	

(1) Les emprises spatiales de références sont issues des couches "COMMUNE", "EPCI", "DEPARTEMENT" et "REGION" de la BD TOPO® (IGN). Par conséquent selon la source des emprises spatiales de référence utilisée pour votre fichier, des erreurs peuvent se générer dues à une différence de précision entre vos données et l'emprise spatiale de référence choisie.

Nota : En cas d'erreurs sur les géométries, les fichiers d'erreurs d'extension ".shx" sont joints au mail.

NOTATION GLOBALE = 3.4 sur 5

CRITERE DE COHERENCE LOGIQUE = 3 sur 5

SOUS CRITERE DE COHERENCE DU FORMAT = 5 sur 5

Nature du contrôle	Modalité du contrôle	Résultat
Poids des fichiers	Calcul du poids des données transmises	0.02 Mo



Nature du contrôle	Modalité du contrôle	Résultat
Liste des fichiers et type d'extension	contrôle sur le type d'extension	STRUCT_SHAPE_PARTIEL.dbf STRUCT_SHAPE_PARTIEL.prj STRUCT_SHAPE_PARTIEL.shp STRUCT_SHAPE_PARTIEL.shx

Nature du contrôle	Modalité du contrôle	Résultat
Structure du fichier	Vérification du format de fichier	Les fichiers d'extension : ".cpg" ".qmd" ".qix" ".shx" ".shx" ".fbn" ".fbx" ".ain" ".aih" ".atx" ".lxs" ".mxx" sont absents
	La présence des fichiers obligatoires ".shp", ".dbf," et ".shx" est testée et l'absence des fichiers ".prj" ou ".cpg" est relevée pour définir la complétude et la validité du "SHAPE"	La structure du "SHAPE" est valide mais incomplète

SOUS CRITERE DE COHERENCE TOPOLOGIQUE = 1 sur 5

Nature du contrôle	Modalité du contrôle	Résultat	Notation
Recherche des doublons	Un POINT est considéré en doublon si et seulement si la géométrie et tous les attributs sont strictement identiques	11 POINT(S) sans doublon sur 55 soit 20 %	1 sur 5

CRITERE D EXHAUSTIVITE = 3.8 sur 5

Nature du contrôle	Modalité du contrôle	Résultat	Notation
Relation spatiale entre le fichier et l'emprise choisie	Découpe du fichier par l'emprise et comptage des objets	42 POINT(S) sur 55 sont à l'intérieur de l'emprise de référence choisie soit 76.4 %	3.8 sur 5

CRIGE PACA - Domaine du Petit Arbois - Bât. Laennec 13100 Aix-en-Provence - Tel : 04 42 90 71 22 - Mail : contact@crige-paca.org

Annexe 3 : Rapport d'analyse sur un fichier dont la structure «SHAPE» est incomplète

