



Production partagée d'une base de données « Occupation des Sols » à grande échelle en Région Grand Est (OCS GE2)

CONTRÔLE QUALITÉ DPT 51

Version 3

David LOY – Jean-Pierre GERARD – Sharif LEWIS-LAMBERT

TTI Production



Earth Observation Consulting Services

janvier 2024

Table des matières

Introduction.....	4
Types de contrôles.....	5
Contrôles sur la forme.....	7
LE NOMMAGE	7
FICHIERS SHAPE FILE d’ArcGis (Contrôle automatisé)	7
ENCODAGE	7
PROJECTION	7
DIMENSION	7
SUPERPOSITION COUCHES	7
DELIMITATION DU DEPARTEMENT	7
STRUCTURE DE LA TABLE ATTRIBUTAIRE	7
Types et noms de champs.....	7
Relation entre polygones et attributs	11
Contrôles topologiques	12
TOPOLOGIE DE NIVEAU 1	12
Présence de micro-intersections de précision	12
Absence d’intersection et de superposition	13
Absence de vides entre les polygones	14
Absence d’arc pendant ou de polygones ouverts	15
Contrôles géométriques.....	15
CONTRAINTES DE PRODUCTION	15
Respect des surfaces minimales.....	15
Contrôle des polygones voisins identiques.....	20
Contrôles sémantiques	20
REPLISSAGE DE LA TABLE ATTRIBUTAIRE	20
Contrôle des ID	20
Contrôle des intitulés dans les tables attributaires.....	20
Contrôle des surfaces de polygones.....	21
L’IMPERMÉABILISATION AU NIVEAU 5	21
Taux d’erreurs sur le Niveau 5 de la nomenclature	22
Contrôles thématiques	23
ÉCHANTILLONNAGE	23
Methodologie.....	23
Methodologie de contrôle	25
RÉSULTATS.....	28
Taux d’erreurs sur le Niveau 4 de la nomenclature	29

Contrôle qualité de la base de données « Occupation des Sols » à grande échelle en Région Grand Est – Millésime 2021

Conclusion30
Retours du CQC31
Matrice de confusion.....33

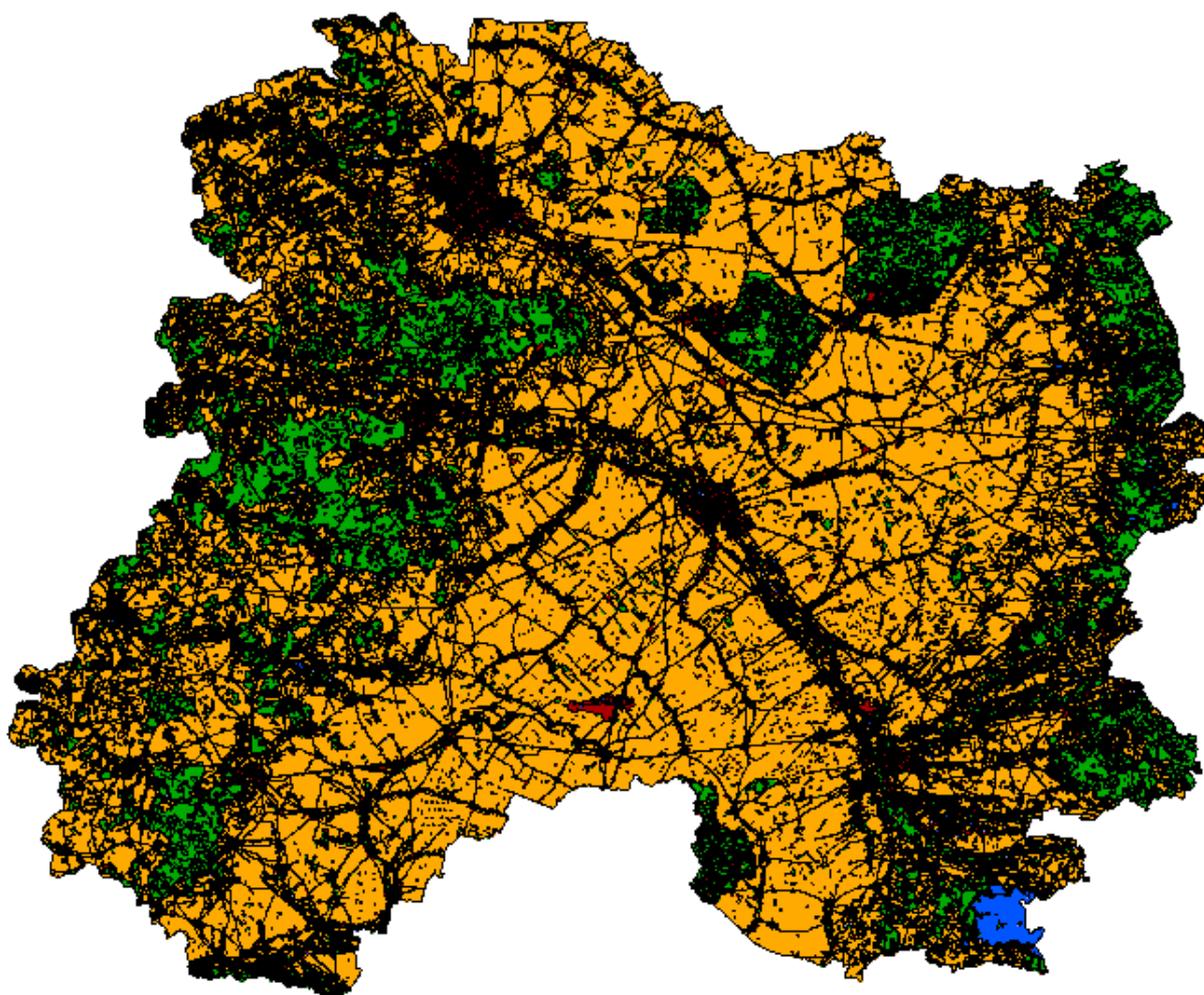
Introduction

Le phasage de production du département 51 est le suivant :

1. Production du socle (squelette + bâtis) sur le département.
2. Production de l'occupation du sol 2021 et mise à jour 2010 et 2019 (*)
3. Contrôle qualité
 - a. CQI – Contrôle Interne du producteur (SIRS)
 - b. CQE – Contrôle Externe du lot 3 (TTI Production)
 - c. CQC – Contrôle groupe projet (COPIL & COTECH)

(*) Pour simplifier la compréhension, les millésimes seront nommés 2021, 2019 et 2010. Néanmoins, il conviendra de noter que selon les départements, les dates diffèrent sensiblement.

Ce contrôle concerne l'intégralité du département 51.



Contrôle qualité de la base de données « Occupation des Sols » à grande échelle en Région Grand Est – Millésime 2021

Fichiers originaux de l'OS :

ocsge2_GE_dpt51_2010.shp

Département 51 (8301.65 km²), Millésime 2010 > 331103 polygones

ocsge2_GE_dpt51_2019.shp

Département 51 (8301.65 km²), Millésime 2019 > 345481 polygones

ocsge2_GE_dpt51_2021.shp

Département 51 (8301.65 km²), Millésime 2021 > 351818 polygones

Fichiers multi-dates et d'évolution au niveau 5 de l'OS :

ocsge2_GE_dpt51_2010_2019_2021_v2.shp

Département 51 (8301.65 km²), Millésimes 2010, 2019 & 2021 > 406310

polygones

ocsge2_GE_dpt51_mutations_2010_2019.shp

Département 51 (499.97 km²), Millésimes 2010 & 2019 > 80151 polygones

ocsge2_GE_dpt51_mutations_2019_2021.shp

Département 51 (105.07 km²), Millésimes 2019 & 2021 > 22392 polygones

ocsge2_GE_dpt51_mutations_2010_2019_2021.shp

Département 51 (568.25 km²), Millésimes 2010, 2019 & 2021 > 93537 polygones

NOTE : la multodate, les simplifications sur le niveau 4, le lcr et le cs ne sont pas des livrables prévus au CCTP. La multodate a néanmoins été fournie par le producteur.

Types de contrôles

Le contrôle qualité porte sur de nombreux points. Certains contrôles sont réalisés sur les couches vectorielles entières et d'autres sur un échantillonnage. De même, une partie des contrôles est automatisée (ou semi-automatisée lorsqu'un contrôle visuel vient vérifier les erreurs) ou intégralement visuelle.

Contrôles sur la forme (Fichiers entiers)

- LE NOMMAGE (**visuel**)
- FICHIERS SHAPE FILE d'ArcGis (**automatisés**)
- ENCODAGE (**automatisés**)
- PROJECTION (**automatisés**)
- DIMENSION (**automatisés**)
- SUPERPOSITION COUCHES (**visuel**)
- STRUCTURE DE LA TABLE ATTRIBUTAIRE
 - Types et noms de champs (**automatisés**)
 - Relation entre polygones et attributs (**automatisés**)

Contrôles topologiques (Fichiers entiers)

- TOPOLOGIE DE NIVEAU
 - Absence d'auto-intersection et de micro polygones (**automatisés**)
 - Absence de vides entre les polygones (**automatisés**)
 - Absence d'arcs ou de points dupliqués (**semi-automatisés**)
 - Absence d'arc pendant ou de polygones ouverts (**automatisés**)
- TOPOLOGIE DE NIVEAU 2
 - Absence d'auto-intersection ou de polygones ouverts (**automatisés**)

Contrôles géométriques (Fichiers entiers)

- CONTRAINTES DE PRODUCTION
 - Respect des surfaces minimales (**automatisés**)

Contrôles sémantiques (Fichiers entiers)

- REMPLISSAGE DE LA TABLE ATTRIBUTAIRE
 - Contrôle des ID (**automatisés**)
 - Contrôle des intitulés dans les tables attributaires (**automatisés**)
 - Contrôle des surfaces de polygones (**automatisés**)

Contrôles complémentaires (Fichiers entiers)

- MÉTADONNÉES (**visuel**)
- STATISTIQUES (**automatisés**)

Contrôles thématiques (Échantillon de points) - (visuel)

- ÉCHANTILLONNAGE
 - Méthodologie
 - Classement des erreurs
 -
- RÉSULTATS SUR L'IMPERMÉABILISATION AU NIVEAU 5
 - Taux d'erreurs
 - Type d'erreurs
- RÉSULTATS SUR L'OS DE NIVEAU 4
 - Taux d'erreurs
 - Type d'erreurs
 - Matrice de confusion

La production s'est faite sur la base de la nomenclature validée lors de la première production 2010 - 2019.

Contrôles sur la forme

IMPORTANT : le contrôle sur la forme n'est réalisé que sur le niveau 5, les évolutions et le fichier multidates. En effet, le niveau 4, le cs et le lcr ne sont pas des livrables.

LE NOMMAGE

On constate que le nommage des fichiers est conforme à l'attente avec précision du département, du millésime et du type de produit.

FICHIERS SHAPE FILE d'ArcGis (Contrôle automatisé)

Les fichiers ne présentent aucune anomalie. Un fichier de style pour le SIG Q-Gis a été livré par le producteur.

ENCODAGE

L'encodage des fichiers est l'UTF-8, il convient donc d'importer ou d'utiliser les shape-files en respectant ce dernier.

PROJECTION

La projection des fichiers est leRGF93_Lambert_93 comme défini par le CCTP.

DIMENSION

Les fichiers sont bien en 2 D.

SUPERPOSITION COUCHES

La superposition des couches avec les BD-Ortho de référence est parfaite. On n'observe ni décalage ni distorsion.

DELIMITATION DU DEPARTEMENT

La limite externe des couches est compatible avec les limites départementales BD-Ortho 2023 de référence est parfaite.

STRUCTURE DE LA TABLE ATTRIBUTAIRE

Types et noms de champs

Un script a été réalisé afin de contrôler les noms de champs, le type (texte ou nombre) et la dimension (nombre de caractères ou nombre de chiffres après la virgule).

Contrôle qualité de la base de données « Occupation des Sols » à grande échelle en Région Grand Est – Millésime 2021

ocsge2_GE_dpt51_2010.dbf

Field Name	Type	Dec	Msize	Dsize	Moffs	Dofts	Dwd
gid	Double	0	8	10	0	1	10
cod_n1	Double	0	8	10	8	11	10
lib_n1	String	0	46	45	16	21	45
cod_n2	Double	0	8	10	64	66	10
lib_n2	String	0	61	60	72	76	60
cod_n3	Double	0	8	10	136	136	10
lib_n3	String	0	53	52	144	146	52
cod_n4	Double	0	8	10	200	198	10
lib_n4	String	0	53	52	208	208	52
cod_n5	Double	0	8	10	264	260	10
lib_n5	String	0	23	22	272	270	22
lcr_2010	String	0	3	2	295	292	2
niv6_10	Char	0	1	1	298	294	10
cs_ign10	String	0	13	12	299	295	12
surf_m	Double	15	8	24	312	307	10
surf_ha	Double	15	8	24	320	331	10
perimetre	Double	15	8	24	328	355	10
millesime	String	0	10	9	336	379	9
source	String	0	71	70	346	388	70
comment	String	0	161	160	417	458	160
doute	Double	0	8	10	584	618	10

ocsge2_GE_dpt51_2019.dbf

Field Name	Type	Dec	Msize	Dsize	Moffs	Dofts	Dwd
gid	Double	0	8	10	0	1	10
cod_n1	Double	0	8	10	8	11	10
lib_n1	String	0	46	45	16	21	45
cod_n2	Double	0	8	10	64	66	10
lib_n2	String	0	61	60	72	76	60
cod_n3	Double	0	8	10	136	136	10
lib_n3	String	0	53	52	144	146	52
cod_n4	Double	0	8	10	200	198	10
lib_n4	String	0	53	52	208	208	52
cod_n5	Double	0	8	10	264	260	10
lib_n5	String	0	23	22	272	270	22
lcr_2019	String	0	3	2	295	292	2
niv6_19	Char	0	1	1	298	294	10
cs_ign19	String	0	13	12	299	295	12
surf_m	Double	15	8	24	312	307	10
surf_ha	Double	15	8	24	320	331	10
perimetre	Double	15	8	24	328	355	10
millesime	String	0	10	9	336	379	9
source	String	0	71	70	346	388	70
comment	String	0	161	160	417	458	160
doute	Double	0	8	10	584	618	10

ocsge2_GE_dpt51_2021.dbf

Field Name	Type	Dec	Msize	Dsize	Moffs	Dofts	Dwd
gid	Double	0	8	10	0	1	10
cod_n1	Double	0	8	10	8	11	10
lib_n1	String	0	46	45	16	21	45
cod_n2	Double	0	8	10	64	66	10
lib_n2	String	0	61	60	72	76	60
cod_n3	Double	0	8	10	136	136	10
lib_n3	String	0	53	52	144	146	52
cod_n4	Double	0	8	10	200	198	10
lib_n4	String	0	53	52	208	208	52
cod_n5	Double	0	8	10	264	260	10
lib_n5	String	0	23	22	272	270	22
lcr_2021	String	0	3	2	295	292	2
niv6_21	Char	0	1	1	298	294	10
cs_ign21	String	0	13	12	299	295	12
surf_m	Double	15	8	24	312	307	10

Contrôle qualité de la base de données « Occupation des Sols » à grande échelle en Région Grand Est – Millésime 2021

surf_ha	Double	15	8	24	320	331	10
perimetre	Double	15	8	24	328	355	10
millesime	String	0	10	9	336	379	9
source	String	0	71	70	346	388	70
comment	String	0	161	160	417	458	160
doute	Double	0	8	10	584	618	10

ocsge2_GE_dpt51_mutations_2010_2019.dbf

Field Name	Type	Dec	Msize	Dsize	Moffs	Dooffs	Dwd
gid	Double	0	8	10	0	1	10
cod_10_n1	Double	0	8	10	8	11	10
lib_10_n1	String	0	46	45	16	21	45
cod_10_n2	Double	0	8	10	64	66	10
lib_10_n2	String	0	61	60	72	76	60
cod_10_n3	Double	0	8	10	136	136	10
lib_10_n3	String	0	53	52	144	146	52
cod_10_n4	Double	0	8	10	200	198	10
lib_10_n4	String	0	53	52	208	208	52
cod_10_n5	Double	0	8	10	264	260	10
lib_10_n5	String	0	23	22	272	270	22
lcr_2010	Double	0	8	10	296	292	10
niv6_10	Char	0	1	1	304	302	10
cod_19_n1	Double	0	8	10	312	303	10
lib_19_n1	String	0	46	45	320	313	45
cod_19_n2	Double	0	8	10	368	358	10
lib_19_n2	String	0	61	60	376	368	60
cod_19_n3	Double	0	8	10	440	428	10
lib_19_n3	String	0	53	52	448	438	52
cod_19_n4	Double	0	8	10	504	490	10
lib_19_n4	String	0	53	52	512	500	52
cod_19_n5	Double	0	8	10	568	552	10
lib_19_n5	String	0	23	22	576	562	22
lcr_2019	Double	0	8	10	600	584	10
niv6_19	Char	0	1	1	608	594	10
surf_m	Double	15	8	24	616	595	10
surf_ha	Double	15	8	24	624	619	10
perimetre	Double	15	8	24	632	643	10
source2010	String	0	71	70	640	667	70
source2019	String	0	71	70	711	737	70
comment	String	0	161	160	782	807	160
doute	Double	0	8	10	944	967	10

ocsge2_GE_dpt51_mutations_2019_2021.dbf

Field Name	Type	Dec	Msize	Dsize	Moffs	Dooffs	Dwd
gid	Double	0	8	10	0	1	10
cod_19_n1	Double	0	8	10	8	11	10
lib_19_n1	String	0	46	45	16	21	45
cod_19_n2	Double	0	8	10	64	66	10
lib_19_n2	String	0	61	60	72	76	60
cod_19_n3	Double	0	8	10	136	136	10
lib_19_n3	String	0	53	52	144	146	52
cod_19_n4	Double	0	8	10	200	198	10
lib_19_n4	String	0	53	52	208	208	52
cod_19_n5	Double	0	8	10	264	260	10
lib_19_n5	String	0	23	22	272	270	22
lcr_2019	Double	0	8	10	296	292	10
niv6_19	Char	0	1	1	304	302	10
cod_21_n1	Double	0	8	10	312	303	10
lib_21_n1	String	0	46	45	320	313	45
cod_21_n2	Double	0	8	10	368	358	10
lib_21_n2	String	0	61	60	376	368	60
cod_21_n3	Double	0	8	10	440	428	10
lib_21_n3	String	0	53	52	448	438	52
cod_21_n4	Double	0	8	10	504	490	10
lib_21_n4	String	0	53	52	512	500	52

Contrôle qualité de la base de données « Occupation des Sols » à grande échelle en Région Grand Est – Millésime 2021

cod_21_n5	Double	0	8	10	568	552	10
lib_21_n5	String	0	23	22	576	562	22
lcr_2021	Double	0	8	10	600	584	10
niv6_21	Char	0	1	1	608	594	10
surf_m	Double	15	8	24	616	595	10
surf_ha	Double	15	8	24	624	619	10
perimetre	Double	15	8	24	632	643	10
source2019	String	0	71	70	640	667	70
source2021	String	0	71	70	711	737	70
comment	String	0	161	160	782	807	160
doute	Double	0	8	10	944	967	10

ocsge2_GE_dpt51_2010_2019_2021.bdf

Field Name	Type	Dec	Msize	Dsize	Moffs	Dooffs	Dwd
gid	Double	0	8	10	0	1	10
cod_10_n1	Double	0	8	10	8	11	10
lib_10_n1	String	0	46	45	16	21	45
cod_10_n2	Double	0	8	10	64	66	10
lib_10_n2	String	0	61	60	72	76	60
cod_10_n3	Double	0	8	10	136	136	10
lib_10_n3	String	0	53	52	144	146	52
cod_10_n4	Double	0	8	10	200	198	10
lib_10_n4	String	0	53	52	208	208	52
cod_10_n5	Double	0	8	10	264	260	10
lib_10_n5	String	0	23	22	272	270	22
lcr_2010	Double	0	8	10	296	292	10
niv6_10	Char	0	1	1	304	302	10
cs_ign10	String	0	13	12	305	303	12
cod_19_n1	Double	0	8	10	320	315	10
lib_19_n1	String	0	46	45	328	325	45
cod_19_n2	Double	0	8	10	376	370	10
lib_19_n2	String	0	61	60	384	380	60
cod_19_n3	Double	0	8	10	448	440	10
lib_19_n3	String	0	53	52	456	450	52
cod_19_n4	Double	0	8	10	512	502	10
lib_19_n4	String	0	53	52	520	512	52
cod_19_n5	Double	0	8	10	576	564	10
lib_19_n5	String	0	23	22	584	574	22
lcr_2019	Double	0	8	10	608	596	10
niv6_19	Char	0	1	1	616	606	10
cs_ign19	String	0	13	12	617	607	12
cod_21_n1	Double	0	8	10	632	619	10
lib_21_n1	String	0	46	45	640	629	45
cod_21_n2	Double	0	8	10	688	674	10
lib_21_n2	String	0	61	60	696	684	60
cod_21_n3	Double	0	8	10	760	744	10
lib_21_n3	String	0	53	52	768	754	52
cod_21_n4	Double	0	8	10	824	806	10
lib_21_n4	String	0	53	52	832	816	52
cod_21_n5	Double	0	8	10	888	868	10
lib_21_n5	String	0	23	22	896	878	22
lcr_2021	Double	0	8	10	920	900	10
niv6_21	Char	0	1	1	928	910	10
cs_ign21	String	0	13	12	929	911	12
surf_m	Double	15	8	24	944	923	10
surf_ha	Double	15	8	24	952	947	10
perimetre	Double	15	8	24	960	971	10
source2010	String	0	71	70	968	995	70
source2019	String	0	71	70	1039	1065	70
source2021	String	0	71	70	1110	1135	70
comment	String	0	161	160	1181	1205	160
doute	Double	0	8	10	1344	1365	10

Contrôle qualité de la base de données « Occupation des Sols » à grande échelle en Région Grand Est – Millésime 2021

ocsge2_GE_dpt51_mutations_2010_2019_2021.dbf

Field Name	Type	Dec	Msize	Dsize	Moffs	Doffs	Dwd
gid	Double	0	8	10	0	1	10
cod_10_n1	Double	0	8	10	8	11	10
lib_10_n1	String	0	46	45	16	21	45
cod_10_n2	Double	0	8	10	64	66	10
lib_10_n2	String	0	61	60	72	76	60
cod_10_n3	Double	0	8	10	136	136	10
lib_10_n3	String	0	53	52	144	146	52
cod_10_n4	Double	0	8	10	200	198	10
lib_10_n4	String	0	53	52	208	208	52
cod_10_n5	Double	0	8	10	264	260	10
lib_10_n5	String	0	23	22	272	270	22
lcr_2010	Double	0	8	10	296	292	10
niv6_10	Char	0	1	1	304	302	10
cod_19_n1	Double	0	8	10	312	303	10
lib_19_n1	String	0	46	45	320	313	45
cod_19_n2	Double	0	8	10	368	358	10
lib_19_n2	String	0	61	60	376	368	60
cod_19_n3	Double	0	8	10	440	428	10
lib_19_n3	String	0	53	52	448	438	52
cod_19_n4	Double	0	8	10	504	490	10
lib_19_n4	String	0	53	52	512	500	52
cod_19_n5	Double	0	8	10	568	552	10
lib_19_n5	String	0	23	22	576	562	22
lcr_2019	Double	0	8	10	600	584	10
niv6_19	Char	0	1	1	608	594	10
cod_21_n1	Double	0	8	10	616	595	10
lib_21_n1	String	0	46	45	624	605	45
cod_21_n2	Double	0	8	10	672	650	10
lib_21_n2	String	0	61	60	680	660	60
cod_21_n3	Double	0	8	10	744	720	10
lib_21_n3	String	0	53	52	752	730	52
cod_21_n4	Double	0	8	10	808	782	10
lib_21_n4	String	0	53	52	816	792	52
cod_21_n5	Double	0	8	10	872	844	10
lib_21_n5	String	0	23	22	880	854	22
lcr_2021	Double	0	8	10	904	876	10
niv6_21	Char	0	1	1	912	886	10
surf_m	Double	15	8	24	920	887	10
surf_ha	Double	15	8	24	928	911	10
perimetre	Double	15	8	24	936	935	10
source2010	String	0	71	70	944	959	70
source2019	String	0	71	70	1015	1029	70
source2021	String	0	71	70	1086	1099	70
comment	String	0	161	160	1157	1169	160
doute	Double	0	8	10	1320	1329	10

Aucune anomalie de nommage, de type et de dimension n'a été constatée sur les différents fichiers livrés.

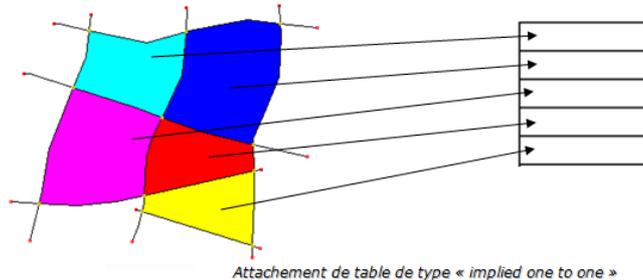
Relation entre polygones et attributs

Il s'agit de vérifier qu'à chaque polygone correspond un enregistrement dans la table attributaire et un seul. Ainsi :

- Il ne peut y avoir un nombre différent d'objets et d'enregistrements dans la table attributaire.
- Il ne peut pas exister des éléments sans attachement à un enregistrement dans la table attributaire.

Contrôle qualité de la base de données « Occupation des Sols » à grande échelle en Région Grand Est – Millésime 2021

- Il ne peut pas exister d'enregistrement dans la table attributaire sans attachement à un polygone.
- Un polygone ne peut être attaché qu'à seul enregistrement dans la table attributaire.
- Un enregistrement dans la table ne peut être attaché qu'à un seul polygone.



Fichiers OCS mono et multi-dates

Aucune anomalie n'a été constatée sur les différents fichiers livrés.

Fichiers EVOLUTION

Après reconstruction topologique, la couche montre la présence de nombreux polygones sans attributs. Ces polygones correspondent à des îles dans d'autres polygones d'évolution. Le shape file ne considère pas les îles (trous) comme des polygones à part entière, c'est donc la reconstruction qui génère ces polygones et il est normal qu'ils ne soient liés à aucun attribut.

Donc, aucune anomalie n'a été constatée sur les différents fichiers livrés.

Contrôles topologiques

TOPOLOGIE DE NIVEAU 1

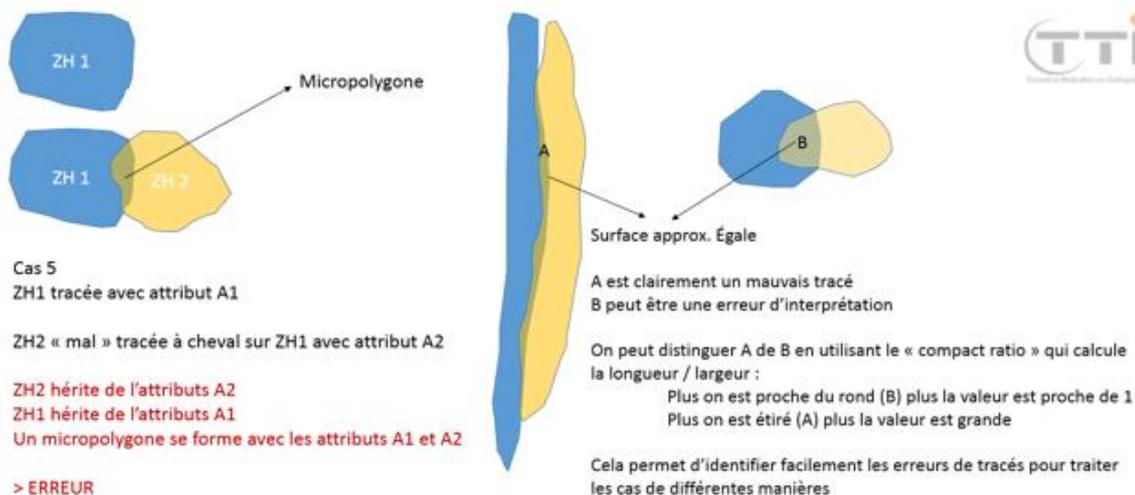
Il s'agit des contrôles élémentaires de la topologie. La structure topologique surfacique est valide lorsque :

- Aucun arc n'est dupliqué.
- Toutes les surfaces sont fermées.
- Aucun arc n'est « pendant » : un arc « pendant » est un arc dont l'une au moins de ses extrémités n'est connectée à aucun autre arc.
- Aucun polygone ne se chevauche avec un autre.
- Aucun vide ne peut être observé entre deux polygones voisins
- Deux arcs ne peuvent se croiser sans être interconnectés.
- Aucun point sur une même ligne n'est dupliqué

Présence de micro-intersections de précision

En fonction des logiciels utilisés pour la validation topologique, il arrive que des micro-polygones soient générés lors de la reconstruction en raison de la précision (dimension des nombres flottants après la virgule) de chaque logiciel. Ce problème est fréquent

Contrôle qualité de la base de données « Occupation des Sols » à grande échelle en Région Grand Est – Millésime 2021

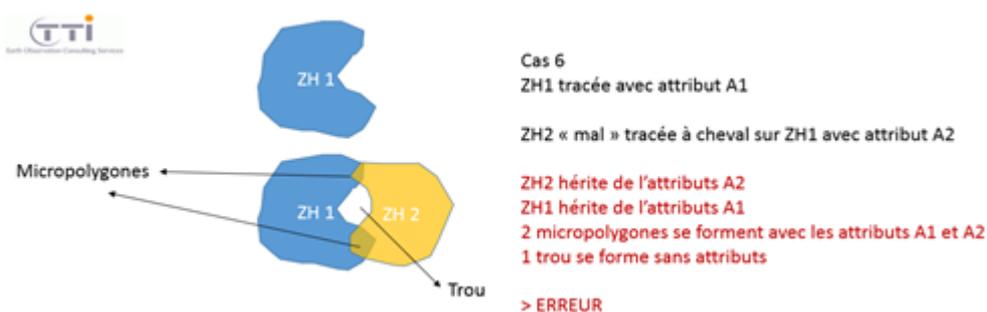


ocsge2_GE_dpt51_2010.shp	ok
ocsge2_GE_dpt51_2019.shp	ok
ocsge2_GE_dpt51_2021.shp	ok
ocsge2_GE_dpt51_2010_2019_2021_v2.shp	ok
ocsge2_GE_dpt51_mutations_2010_2019.shp	ok
ocsge2_GE_dpt51_mutations_2019_2021.shp	ok
ocsge2_GE_dpt51_mutations_2010_2019_2021.shp	ok

Aucun micro-polygone n'a été repéré sur les couches livrées

Absence de vides entre les polygones

Comme pour les superpositions, des espaces entre les polygones normalement jointifs peuvent générer des micro-polygones à la reconstruction topologique, Ces vides forment de nouveaux polygones qui ne sont reliés à aucun élément dans la table attributaire et il est alors très facile de les identifier par requête.

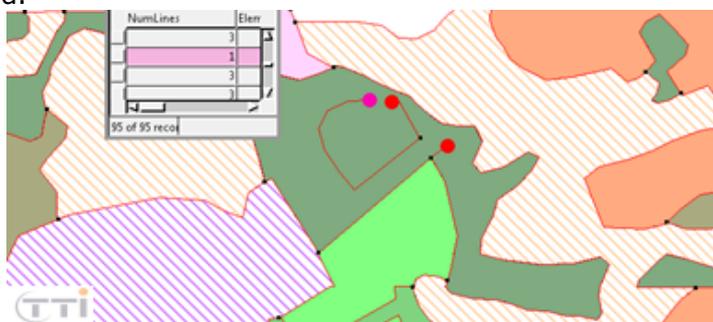


ocsge2_GE_dpt51_2010.shp	ok
ocsge2_GE_dpt51_2019.shp	ok
ocsge2_GE_dpt51_2021.shp	ok
ocsge2_GE_dpt51_2010_2019_2021_v2.shp	ok
ocsge2_GE_dpt51_mutations_2010_2019.shp	ok
ocsge2_GE_dpt51_mutations_2019_2021.shp	ok
ocsge2_GE_dpt51_mutations_2010_2019_2021.shp	ok

Aucun espace vide entre polygones n'a été repéré sur les fichiers livrés.

Absence d'arc pendant ou de polygones ouverts

Les polygones non fermés produisent ce que l'on appelle des arcs pendants. Il s'agit de lignes dont l'un des nœuds (initial ou final) n'est relié qu'à elles-mêmes alors que dans une topologie propre, ils devraient être reliés à au moins trois lignes. Ainsi, en SIG le repérage des arcs pendants est très simple par une requête sur le nombre de lignes partant d'un nœud.



Exemple d'arc pendant et de polygone ouvert.

Aucun arc pendant ou polygone ouvert n'a été repéré sur les couches livrées.

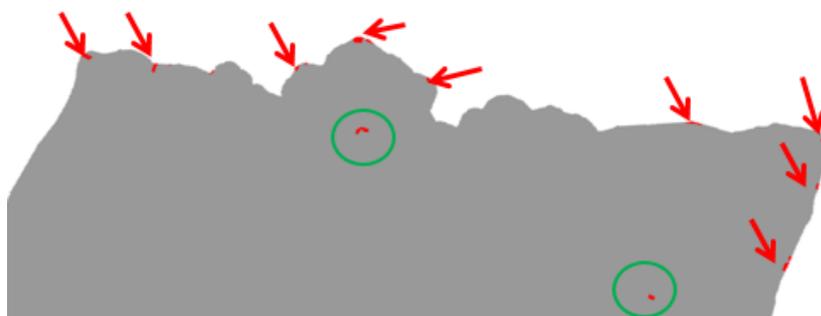
Contrôles géométriques

CONTRAINTES DE PRODUCTION

Respect des surfaces minimales

L'**unité minimale de collecte** (UMC) correspond à la plus petite unité spatiale cartographiée. Normalement, aucune parcelle ne doit être de surface inférieure à celle-ci et dans le même temps, un objet supérieur à cette surface doit être délimité. Dans la pratique, ce n'est pas toujours vrai. Le photo-interprète est parfois obligé de « tricher » afin de prendre en compte une entité importante (notamment dans les postes urbains) plus petite que l'UMC ou la continuité des réseaux.

Le respect des UMC prend en compte deux paramètres : la surface et la nomenclature. Ce repérage des polygones inférieurs à la surface de l'UMC en fonction de son code de nomenclature est fait sur SIG via une requête automatisée.



De plus, le découpage lié aux bordures de la zone de travail ne suit pas la réalité de l'occupation du sol. C'est une limite administrative. De fait, la délimitation de l'occupation du sol doit s'arrêter sur cette limite et implicitement, cela peut engendrer de petits polygones de bordure inférieurs à l'UMC. Ces polygones ne sont évidemment pas à prendre en compte.

XX PARAGRAPHE A METTRE A JOUR

Observation des tests automatisés

UMC OCS_GE2_D51_2010_NIV4,

Pas significatif, car continuité réseau ferré/routier/hydro :

cod_n1	cod_n4	surf
1	1411	66.85
5	5110	230.11
5	5110	474.76
5	5110	305.26
5	5110	316.30
5	5110	452.76
5	5110	173.79
5	5110	387.79
5	5110	247.54
5	5110	255.23

Erreurs :

cod_n1	cod_n4	surf
1	1510	190.43

1 erreur d'UMC repérée

UMC OCS_GE2_D51_2019_NIV4

Pas significatif, car continuité réseau ferré/routier/hydro :

cod_n1	cod_n4	surf
1	1411	66.85
5	5110	173.79
5	5110	474.76
5	5110	305.26
5	5110	316.30
5	5110	452.76
5	5110	230.11
5	5110	387.79
5	5110	247.54
5	5110	255.23

Erreurs :

cod_n1	cod_n4	surf
1	1510	190.43

1 erreur d'UMC repérée

UMC OCS_GE2_D51_2021_NIV4,

Pas significatif, car continuité réseau ferré/routier/hydro :

cod_n1	cod_n4	surf
1	1411	66.85
5	5110	316.30
5	5110	305.26
5	5110	474.76
5	5110	452.76
5	5110	255.23
5	5110	247.54
5	5110	387.79
5	5110	173.79
5	5110	230.11

Erreurs :

cod_n1	cod_n4	surf
1	1510	190.43

1 erreur d'UMC repérée

LCR	1	2	3	4	5	6	7	8	9	10
umc	50	250	250	250	250	250	250	250	250	250

Exception, combiné au niveau4 cod1 = 1, l'umc du 10 passe à 2500 m²

UMC OCS_GE2_D51_2010_LCR

Pas significatif, car continuité réseau ferré/routier/hydro :

cod_n1	cod_n4	lcr_2010	surf
1	1411	3	66.85
5	5110	6	173.79
5	5110	6	247.54
5	5110	6	230.11

Erreurs :

cod_n1	cod_n4	lcr_2010	surf
1	1340	2	62.06
1	1340	2	2.64
1	1311	4	39.32
1	1123	4	166.74
1	1213	4	190.60

Erreurs sur l'exception :

5 erreurs d'UMC repérées

UMC OCS_GE2_D51_2019_LCR

Pas significatif, car continuité réseau ferré/routier/hydro :

cod_n1	cod_n4	lcr_2019	surf
1	1411	3	66.85
5	5110	6	230.11
5	5110	6	247.54
5	5110	6	173.79

Erreurs :

cod_n1	cod_n4	lcr_2019	surf
1	1340	2	2.64
1	1123	4	168.62
1	1311	4	39.32
1	1123	4	166.74
1	1213	4	190.60
1	1220	4	140.59

Erreurs sur l'exception :

6 erreurs d'UMC repérées

UMC OCS_GE2_D51_2021_LCR

Pas significatif, car continuité réseau ferré/routier/hydro :

cod_n1	cod_n4	lcr_2019	surf
1	1411	3	66.85
5	5110	6	247.54
5	5110	6	230.11
5	5110	6	173.79

Erreurs :

cod_n1	cod_n4	lcr_2021	surf
1	1123	4	166.74

Contrôle qualité de la base de données « Occupation des Sols » à grande échelle en Région Grand Est – Millésime 2021

1	1123	4	168.62
1	1213	4	190.60
1	1213	4	37.03
1	1213	4	51.94
1	1220	4	140.59
1	1311	4	39.32
1	1330	4	109.19
1	1340	2	2.64

Erreurs sur l'exception :

9 erreurs d'UMC repérées

UMC OCS_GE2_D51_2021_CS

UMC OCS_GE2_D51_2019_CS

UMC OCS_GE2_D51_2010_CS

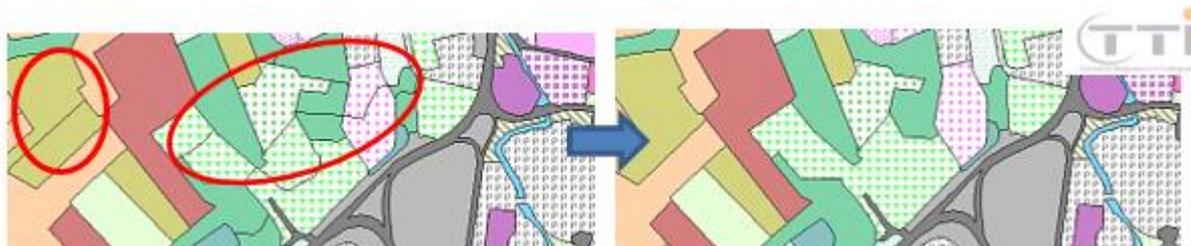
Attention, le champ CS a été ajouté pendant le marché, mais n'est pas contractuel par rapport au CCTP

cs_ign	umc
CS1.1.1.1	50
CS1.1.1.2	250
CS1.1.2.1	250
CS1.1.2.2	250
CS1.2.1	300
CS1.2.2	250
CS2.1.1	250
CS2.1.1.1	500
CS2.1.1.2	1000
CS2.1.1.3	1000
CS2.1.2	1000
CS2.1.3	1000
CS2.2.1	250

Quelques polygones < UMC (CS) ont été repérés dans les couches 2010/2019/2021. Cela reste anecdotique sur le nombre total de polygones.

Contrôle des polygones voisins identiques

Ce contrôle automatique est effectué sur SIG. La routine permet de prendre en compte plusieurs tables et plusieurs champs pour détecter les lignes dont les attributs du polygone sélectionné sont identiques de chaque côté de celle-ci.



La couche de niveau 5 présente plusieurs polygones voisins de code identique. Cela s'explique par la présence des champs « comment » et « indice » (doute). Si deux polygones voisins ont la même OS, mais que sur l'un d'eux un doute a été émis, la limite est préservée. L'OS est alors identique à droite et à gauche de la limite commune.

La fusion des polygones de niveau 5 pour donner la couche de niveau 4 a utilisé les champs cod_n4, mais aussi les champs « comment » et « indice » (doute). De ce fait, il persiste également au niveau 4 quelques polygones voisins ayant le même code.

Aucune erreur de codes voisins identiques hors présence de doutes ou commentaires n'a été relevée.

Contrôles sémantiques

REPLISSAGE DE LA TABLE ATTRIBUTAIRE

Contrôle des ID

Le numéro maximal doit correspondre au nombre de polygones de la couche. Par ailleurs, il ne doit pas exister deux ID avec le même nombre. Cette validation est automatisée.

Les ID sont conformes aux exigences.

Contrôle des intitulés dans les tables attributaires

Il s'agit de contrôler que le remplissage de la table attributaire et plus particulièrement celui des intitulés de classe est conforme à la nomenclature. Néanmoins, il s'agit également de contrôler qu'il n'y a pas de codes (niveaux 1 à 5) présents qui n'existent pas dans la nomenclature.

Pour ce contrôle, un processus automatisé compare les valeurs ou les intitulés des tables attributaires avec ceux de la nomenclature et signale toute anomalie. Les anomalies signalées sont :

- Incohérence des codes de niveaux 1 à 5 au sein d'un même polygone
- Présence d'un code qui n'existe pas dans l'un des champs
- Absence de code dans l'un des champs
- Absence d'intitulé dans l'un des champs
- Présence d'un intitulé qui n'existe pas dans l'un des champs
- Présence d'un intitulé non conforme (exemple au pluriel alors qu'il est au singulier dans la nomenclature).
- Incohérence entre les codes et les intitulés correspondants

Pas d'incohérences observées entre les codes et les intitulés des niveaux 1 à 5.

Contrôle des surfaces de polygones

Les champs contenant les surfaces en m² et ha ont été recalculés et comparés à ceux de la couche. Généralement, si une erreur est détectée, cela traduit un oubli de mise à jour du champ. Il sera alors demandé au prestataire de corriger cela. Le calcul étant lié au logiciel, il est normal de trouver de petites variations.

Nous avons donc ignoré les différences de +/- 5m² et les écarts jusqu'à 25 m² pour les très grands polygones (> 8 km²) et/ ou avec de nombreuses îles (> 40) ou très complexes du réseau routier.

Pas d'erreurs significatives observées dans les calculs de surfaces des autres couches.

L'IMPERMÉABILISATION AU NIVEAU 5

Le producteur procède à des requalifications de certains éléments du socle et à une resegmentation du niveau 4 sur les postes urbains par photo-interprétation. Pour le contrôle, il s'agit donc de contrôler par PIAO un échantillon de points. Le contrôle, bien que plus simple, s'apparente alors au contrôle thématique.

Il est à noter que les zones imperméabilisées et bâties sont l'héritage de l'intégration du bâti dans le socle. Il ne s'agit donc pas de contrôler de façon exhaustive cette classe « imperméable bâti », mais plutôt d'essayer de relever des erreurs de la BD-Topo : bâti indiqué, mais n'existant plus ou pas encore construit. Ce contrôle est réalisé pour le millésime 2021, car les autres millésimes ont déjà été contrôlés dans le précédent marché.

Échantillonnage :

Un échantillon aléatoire a été fait sur le département. Il n'est pas nécessaire de contrôler un nombre de points important pour estimer la qualité, car cette information a déjà été largement contrôlée sur les anciens millésimes du département en cours de production et que les résultats ont toujours été en dessous du taux d'erreur maximal attendu. 500 points ont été contrôlés sur le millésime 2021.

Taux d'erreurs sur le Niveau 5 de la nomenclature

Il est important de noter que ce contrôle ne s'attache qu'à la (im) perméabilité de l'occupation du sol. La nature du niveau 4 n'est pas vérifiée durant ce contrôle.

Valeurs	Échantillon	Refusé	Fiabilité
1 – bâti imperméable	50	1	98.0 %
2 – imperméable non bâti	200	11	94.5 %
3 - Perméable	200	14.5	92.75 %
0 – (hors urbain)	50	5	90.0 %
TOTAL 500 polygones		31.5	93.7 %

L'analyse des erreurs montre surtout qu'il s'agit essentiellement d'erreurs sur la perméabilité :

- Dans classes difficiles : zones en mutation, exploitations agricoles, etc. du fait de la nature du sol parfois difficile à estimer
- Des subdivisions possibles non réalisées
- Erreurs d'omission dans la mise à jour 2019

NOTE : aucun contrôle sur StreetView n'a été fait.

Conclusion :

Le taux d'erreur sur le volet imperméabilisation est de **6.3 %** ce qui est au-delà de la fiabilité attendue.

Le résultat est donc excellent.

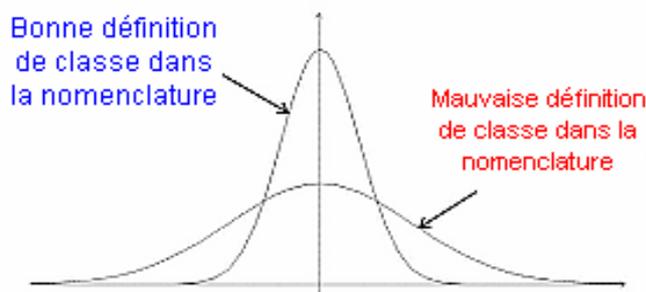
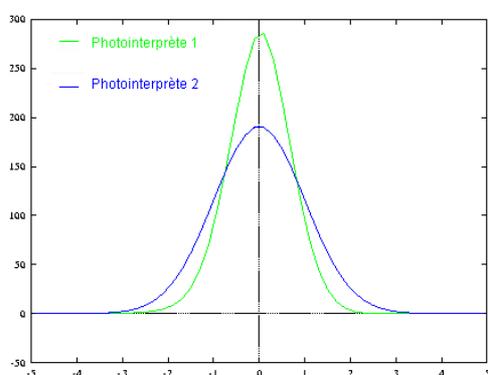
Contrôles thématiques

ÉCHANTILLONNAGE

Le contrôle de la qualité thématique vise à identifier le taux d'erreurs, mais aussi à tenter d'en trouver l'origine.

Origine des erreurs :

- Les clartés des définitions de classes dans la nomenclature
- L'expérience du photo-interprète ou sa compréhension de la nomenclature
- L'erreur de saisie du code
- L'erreur de segmentation
- L'erreur propagée d'une mauvaise donnée exogène
- L'erreur induite de la qualité des ortho-photos
- L'erreur d'omission par manque de donnée exogène
- Le contrôleur peut se tromper en validant ou refusant un polygone à tort



Exemple de la répercussion du taux de réussite en fonction à gauche, des photo-interprètes, à droite de la définition de la nomenclature

À cela s'ajoutent deux paramètres :

- Certaines classes sont moins représentées que d'autres
- Le tirage doit être représentatif de toute la zone et l'échantillonnage doit donc être correctement réparti sur toute la zone.

L'échantillonnage des polygones est donc fait suivant une méthode probabiliste, il s'agit d'un tirage aléatoire en grappes et stratifié. C'est-à-dire un tirage aléatoire orienté qui prend en compte à la fois l'occurrence des classes et la répartition.

Méthodologie

Il s'agit donc de tirer aléatoirement des polygones en respectant ces **4 règles** :

- Inventorier avec au minimum 1000 polygones.
- La répartition par classe doit correspondre à l'occurrence réelle de la classe dans la couche/nombre total de polygones.
- Pour les postes présentant moins de 12 polygones, l'intégralité des polygones de la classe est vérifiée.
- Répartition équitable dans le secteur analysé.

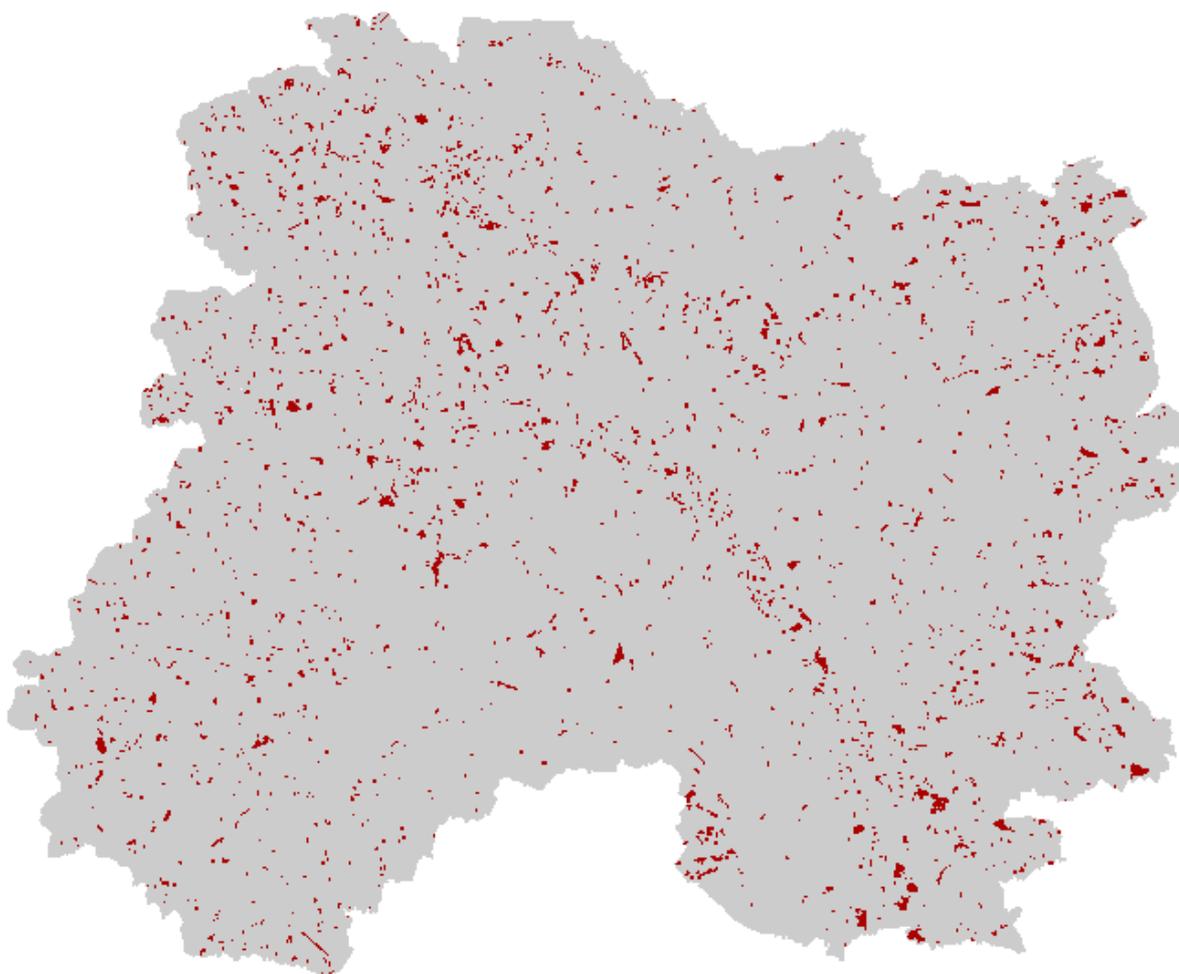
Contrôle qualité de la base de données « Occupation des Sols » à grande échelle en Région Grand Est – Millésime 2021

Le socle est contrôlé à part, et donc l'échantillonnage exclut les réseaux routiers et ferrés.

Total 2021	4019 polygones
------------	----------------

Une fois l'échantillonnage effectué, un script extrait automatiquement les polygones 2021 concernés. Les polygones 2010 et 2019 ont déjà été contrôlés lors du précédent projet.

Ci-dessous, sont présentés les polygones contrôlés, sans limites de taille.



Millésime 2021- niveau 4

Méthodologie de contrôle

Une fois l'échantillonnage effectué, un script ajoute les informations suivantes dans une table attributaire : l'ID du polygone, la position du polygone par rapport aux bordures (pour les UMC), le millésime, et les codes d'occupation du sol niveau 4 et 5 du polygone ou les code LCR et CS pour cette couche spécifique.

Sur l'échantillon de contrôle, une table attributaire spécifique au contrôle est ajoutée. Cette dernière permet d'optimiser le contrôle, de faciliter le travail à la fois du contrôleur et des producteurs lorsque **le CQE va leur fournir son jeu d'échantillon noté et commenté.**

Par ailleurs, il permet de bien encadrer le classement des erreurs et aussi de pondérer en fonction de ce qui est observé. Ainsi, il prend en compte le contexte (petit polygone ou très grand (un écart de tracé est alors acceptable ou non), son voisinage, l'importance de l'erreur, sa redondance, la divergence d'opinion, la différence de l'efficacité du contrôle.

Une erreur redondante et une erreur de non compréhension d'une classe, une fois signalée en erreurs plusieurs fois, il n'est pas nécessaire de mettre en erreur les suivantes, juste de les signaler, sinon les statistiques générales sont faussées. Bien évidemment, à la charge du producteur de reconstruire et corriger toute cette classe)

Un CQE pousse toujours le contrôle plus loin que ce que le producteur à l'obligation de faire (regarder sur google earth, street view, des données que le producteur n'utilise). Cette méthode permet donc de signaler des erreurs sans les noter mais on peut avoir une idée de la fiabilité réelle (par rapport à la vérité terrain).

En parallèle du contrôle, le CQE produit un fichier de point pour cibler les erreurs. En effet, dans notre méthodologie, nous ne contrôlons pas un maillage aléatoire de points, mais des polygones dans leur intégralité, qu'il fasse 500m² ou 300000m². Le contrôle n'est donc pas localisé sous un point mais de façon plus cohérente sur le polygone pour en juger la pertinence, le tracé, les omissions à l'intérieur ou à l'inverse non utiles.

La table attributaire s'organise comme suit :

ok	GEOM1	GEOM2	THEM1	THEM2	tot
<input checked="" type="checkbox"/>	Limite(s) à corriger (1)	▼	▼	▼	1.
<input checked="" type="checkbox"/>	Subdivision à faire (1)	▼	▼	▼	1.
<input checked="" type="checkbox"/>	Subdivision à faire (1)	▼	▼	▼	1.
<input checked="" type="checkbox"/>	Etendre voisins (1)	▼	▼	▼	1.
<input checked="" type="checkbox"/>	Subdivision à faire (1)	▼	▼	▼	1.
<input checked="" type="checkbox"/>	▼	Etendre voisins (1/2)	Erreur d'identification (1)	▼	1.
<input checked="" type="checkbox"/>	▼	▼	Erreur d'identification (1)	▼	1.
<input checked="" type="checkbox"/>	Limite(s) à corriger (1)	▼	▼	▼	1.
<input checked="" type="checkbox"/>	Etendre voisins (1)	▼	▼	▼	1.

Le contrôleur appelle chaque polygone un par un, sans se soucier de scanner la zone ou savoir où il se trouve afin d'avoir un avis neutre s'il connaît en partie la région. Une fois le polygone contrôlé, il coche le champ « OK » et passe au suivant.

Les champs déroulants sont préconfigurés avec une liste d'erreurs possibles et une note associée. Les erreurs peuvent être géométriques ou thématiques et elles peuvent être multiples. Si la somme des erreurs d'un même polygone dépasse 1, elle est ramenée à 1

Contrôle qualité de la base de données « Occupation des Sols » à grande échelle en Région Grand Est – Millésime 2021

dans un champ calculé afin de ne compter que pour 1 (la note maximale d'erreur pour un polygone).

ok	GEOM1	GEOM2	THEM1	THEM2
<input checked="" type="checkbox"/>				
<input checked="" type="checkbox"/>				
<input checked="" type="checkbox"/>				
<input checked="" type="checkbox"/>	# revenir dessus / verif exo			
<input type="checkbox"/>	Divergence d'opinions (0)			
<input type="checkbox"/>	Ecart(s) - gros poly (0)			
<input type="checkbox"/>	Ecart(s) - gros poly (1/4)			
<input type="checkbox"/>	Epine(s) / rétrécissement(s) (0)			
<input checked="" type="checkbox"/>	Epine(s) / rétrécissement(s) (1/4)			
<input type="checkbox"/>	Etendre voisins (1)			
<input type="checkbox"/>	Etendre voisins (1/2)			
<input type="checkbox"/>	Limite incohérente / exogène (1)			
<input type="checkbox"/>	Limite incohérente / exogène (1/2)			
<input type="checkbox"/>	Limite incohérente / exogène (1/4)			
<input type="checkbox"/>	Limite(s) à corriger (1)			
<input type="checkbox"/>	Limite(s) à corriger (1/2)			
<input type="checkbox"/>	Limite(s) à corriger (1/4)			
<input type="checkbox"/>	Problèmes multiples (1)			
<input type="checkbox"/>	Remarque (0)			
<input type="checkbox"/>	Subdivision possible (0)			
<input type="checkbox"/>	Subdivision à faire (1)			
<input type="checkbox"/>	Subdivision à faire (1/2)			
<input type="checkbox"/>	Subdivision à faire (1/4)			
records	Tracé(s) à reviser (0)			

Listes déroulantes d'erreurs géométriques et thématiques

Ces champs sont doublés afin de pouvoir signaler au moins deux erreurs de chaque type. Au-delà, c'est toute la zone qui est à reprendre.

Concernant les erreurs thématiques, trois champs préremplis permettent au contrôleur de suggérer des codes de remplacement, normalement un seul, éventuellement deux en laissant le choix au producteur de garder celui qu'il préfère

THEM1	THEM2	CODE1	CODE2	CODE3
Erreur d'identification (1)		1130		
Erreur d'identification (1)		5120		
Erreur d'identification (1)		1121		
Segment. injustifiée: grouper (1/2)		1124		

Ici, liste des codes pour l'OCSGE2. La liste est réadaptée à chaque projet.

Contrôle qualité de la base de données « Occupation des Sols » à grande échelle en Région Grand Est – Millésime 2021

Des scripts post-contrôle permettent de vérifier quelques points élémentaires, par exemple que le CQE ne préconise pas une subdivision d'un polygone alors que celui-ci ne le permet pas, par exemple un polygone d'urbain de 750 m², ne peut pas être subdivisé en deux polygones de 500m² (UMC). Même s'il fait attention pendant son contrôle, l'opérateur n'est pas à l'abri de ce genre d'erreur. En fonction de la surface initiale et des codes proposés (avec chacun leur UMC), le script s'assure donc que la proposition est faisable.

	RMQ	area
	s'il est englobé dans le 450, alors c'est aussi du 450, sinon il faut le raccrocher à l'entreprise	356
0	ripisylve	8667
	probleme de codage, le 322 est uniquement sous les points le reste est en parc	404
	plutot parking vu le marquage	870
	redessiner la vraie enveloppe	486
2		5595
		2670
	ce n'est pas de l'extraction mais du dépôt	3948
		33514
	Pour une route moyenne, soit l'espace est enclavé et perdu, soit il ne l'est pas. Au S-E, pas enclavé	22529
		4195
0	petite cabane de 30 m2 insignifiante, mettre le terrain comme le voisinage en sans usage	1575
	pas différent de 2000	18859

En complément, le contrôleur ajoute un commentaire, si nécessaire, pour préciser l'action à faire ou justifier sa décision

ECH_N4 \ Polygon \ EXT_CQE_V2_DPT88_2021					
V	ech	err	pct	code_N4	lib_N4
<input checked="" type="checkbox"/>	25	1.00	0.96	1220	Equipements eau, énergies, T.I.C. et déchets
<input checked="" type="checkbox"/>	48	11.50	0.76	1313	Emprises d'activité à dominante mixte ou tertiaire
<input checked="" type="checkbox"/>	23	1.00	0.96	1340	Zones d'extraction
<input checked="" type="checkbox"/>	27	5.00	0.81	1510	Espaces verts urbains
<input checked="" type="checkbox"/>	180	15.00	0.92	2110	Cultures annuelles et pluri-annuelles
<input checked="" type="checkbox"/>	22	1.00	0.95	2120	Cultures spécifiques
<input checked="" type="checkbox"/>	22	4.50	0.80	2210	Vignes
<input checked="" type="checkbox"/>	236	15.00	0.94	3140	Coupes à blanc et jeunes plantations
<input checked="" type="checkbox"/>	108	24.50	0.77	3230	Surfaces enherbées semi-naturelles

Le logiciel produit les statistiques automatiquement sur la base du nombre de polygones contrôlés et la somme des erreurs. Sur l'exemple ci-dessus, pour le 2110, 15 erreurs pour 180 polygones, soit une fiabilité de 92%. De même, sur la base des codes d'origine, et des trois colonnes de codes de remplacement/omis, il calcule une matrice de confusion.

RÉSULTATS

Pour mémoire, le CCTP demande un taux de fiabilité de 80 % par classe et 90 % au total.

Taux d'erreurs sur le Niveau 4 de la nomenclature

DEPARTEMENT 51		V2		CQE	
		2021			
code_niv4	typo_niv4	Occurrence	ech.	refusés	pct
1111	Bâti continu dense	1528	20	3.5	82.50
1112	Bâti continu aéré	3543	65	5	92.31
1121	Bâti collectif	1254	26	1	96.15
1122	Bâti mixte	2	2	0	100.00
1123	Bâti individuel dense	11931	252	29	88.49
1124	Bâti individuel lâche	14079	269	24	91.08
1130	Bâti isolé en zone agricole ou naturelle	1313	189	26.25	86.11
1140	Espaces libres en milieu urbain	2257	36	2	94.44
1211	Emprises scolaires et universitaires	555	16	1.75	89.06
1212	Emprises hospitalières	175	12	0	100.00
1213	Equipements sportifs et de loisirs, campings	1461	41	7.5	81.71
1214	Cimetières	697	14	2	85.71
1215	Autres équipements collectifs	1904	46	4	91.30
1220	Équipements eau, énergies, T.I.C. et déchets	2112	23	0	100.00
1311	Emprises d'activités à dominante industrielle	491	23	3	86.96
1312	Emprises d'activités à dominante commerciale	657	21	1	95.24
1313	Emprises d'activité à dominante mixte ou tertiaire	1540	55	6.75	87.73
1314	Anciennes emprises d'activité	80	12	2.25	81.25
1320	Emprises militaires	547	12	1	91.67
1330	Exploitations agricoles	6301	104	16	84.62
1340	Zones d'extraction	93	12	0	100.00
1413	Espaces associés aux réseaux routiers et ferrés	5026	127	12	90.55
1420	Emprises aéroportuaires	8	8	0	100.00
1430	Emprises portuaires	11	11	0	100.00
1510	Espaces verts urbains	1726	31	4	87.10
1610	Espaces en transition	6539	79	5	93.67
1710	Places	34	14	1	92.86
2110	Cultures annuelles et pluri-annuelles	7153	206	21	89.81
2120	Cultures spécifiques	502	20	1.5	92.50
2210	Vignes	1400	12	0	100.00
2221	Vergers traditionnels	1708	131	14.5	88.93
2222	Vergers intensifs	76	18	1	94.44
2223	Pépinières	17	17	0	100.00
2310	Prairies, friches et délaissés agricoles	16139	269	33	87.73
2320	Bosquets et haies	4179	269	32	88.10
3110	Forêts de feuillus	9197	255	24.5	90.39
3120	Forêts de conifères	2232	251	18.75	92.53
3130	Forêts mixtes	895	164	21.5	86.89
3140	Coupes à blanc et jeunes plantations	4608	269	26.5	90.15
3150	Peupleraies et sapinières	3675	12	0.5	95.83
3210	Pelouses et pâturages de montagne	0	0	0	!
3220	Formations pré-forestières	8674	190	26.75	85.92
3230	Surfaces enherbées semi-naturelles	2538	124	23.5	81.05
3310	Plages et sables	37	6	0.5	91.67

Contrôle qualité de la base de données « Occupation des Sols » à grande échelle en Région Grand Est – Millésime 2021

3320	Roches nues	13	13	0	100.00
3340	Zones de sinistre (incendie, tempête)	348	20	4	80.00
4110	Ripisylves et rivulaires	3119	102	5	95.10
4120	Autres milieux humides	1099	13	2	84.62
5110	Cours d'eau et canaux	540	16	0	100.00
5120	Plans d'eau	3158	110	0	100.00
5130	Bassins artificiels	1014	12	0	100.00
	GLOBAL	138185	4019	414.5	89.69

1411	Emprise réseau ferré		Non contrôlé : socle		
1412	Emprise réseau routier		Non contrôlé : socle		

Conclusion

La fiabilité globale de 89.69 % est inférieure à celle attendue (**10.31 % d'erreurs en trop**) ce qui sur 4019 polygones correspond à 12.5 erreurs en trop.

Par classe, la fiabilité globale est en dessus des 80% attendus pour toutes les classes.

Retours du CQC

Des échantillons sur le millésime 2021 ont été proposés pour ce contrôle. Le tirage aléatoire s'est fait sur le niveau 4 du millésime 2021 hors polygones sélectionnés pour le CQE. Le tirage tient compte de l'occurrence de chaque classe de la nomenclature.

Chaque jeu est représentatif de l'ensemble de la donnée sans distinction de classes. L'ensemble des jeux de données du millésime 2021 représente environ 100 polygones chacun.

Les tableaux ci-dessous indiquent les jeux disponibles et ceux qui ont été contrôlés.

Jeu de données 2021	Plygones	Personne(s) faisant le contrôle	Organisme
DPT51_2021_CQC_GLOBAL_1.zip	100	Jean-Marc HANCZYK	DDT 51
DPT51_2021_CQC_GLOBAL_2.zip	100	Stéphane MARTIN-HUFFSCHMITT	ADEUS
DPT51_2021_CQC_GLOBAL_3.zip	100	Elodie Monnet	CU Grand Reims
DPT51_2021_CQC_GLOBAL_4.zip	100	Elodie Monnet	CU Grand Reims
DPT51_2021_CQC_GLOBAL_5.zip	100	Jean-Marc HANCZYK	DDT 51
DPT51_2021_CQC_GLOBAL_6.zip	100	Jean-Marc HANCZYK	DDT 51
DPT51_2021_CQC_GLOBAL_7.zip	100	Jean-Marc HANCZYK	DDT 51
DPT51_2021_CQC_GLOBAL_8.zip	100	Jean-Marc HANCZYK	DDT 51
DPT51_2021_CQC_GLOBAL_9.zip	100	Laurent WEIBEL	SGARE
DPT51_2021_CQC_GLOBAL_10.zip	100	Jean-Marc HANCZYK	DDT 51
DPT51_2021_CQC_GLOBAL_11.zip	100	Jean-Marc HANCZYK	DDT 51
DPT51_2021_CQC_GLOBAL_12.zip	100		
DPT51_2021_CQC_GLOBAL_13.zip	100		
DPT51_2021_CQC_GLOBAL_14.zip	100		
DPT51_2021_CQC_GLOBAL_15.zip	100		
DPT51_2021_CQC_GLOBAL_16.zip	100		

Avant de regarder les résultats du contrôle CQC, il est important de relever certains points :

- Des classes n'ont pas été échantillonnées, car elles étaient déjà entièrement contrôlées par le CQE du fait de la faible occurrence des objets de ces classes.
- De la même manière, beaucoup de classes présentent un échantillonnage faible. Il est alors important de relativiser le pourcentage d'erreur qui n'est pas forcément significatif.
- Les problèmes topologiques n'ont pas été contrôlés, car déjà entièrement contrôlés par le CQE. On a donc un contrôle essentiellement thématique et dans une moindre mesure de précision géométrique du tracé.
- Enfin, le CQE a passé en revue les remarques faites par les membres du CQC pour les valider ou non. En revanche, il n'a pas validé les polygones où aucune erreur n'était reportée. Cela peut introduire un biais, car l'erreur peut venir d'une non-identification d'une erreur effective.

L'ensemble de ces facteurs explique en partie les différences de chiffres entre le contrôle du CQE et celui du CQC.

Ne sont reportées ici que les erreurs non repérées par le CQE puisque l'échantillonnage prenait en compte ce que le CQE avait déjà contrôlé. En rouge sont indiquées les classes particulièrement sous représentées pour lesquelles le calcul de l'erreur n'est pas significatif.

DEPARTEMENT 51

V2

CQP

		2021			
code_niv4	typo_niv4	Occurrence	ech.	refusés	pct
1111	Bâti continu dense	1528	8	0	100.00
1112	Bâti continu aéré	3543	23	4	82.61
1121	Bâti collectif	1254	10	0	100.00
1122	Bâti mixte	2	0	0	
1123	Bâti individuel dense	11931	85	3	96.47
1124	Bâti individuel lâche	14079	68	6	91.18
1130	Bâti isolé en zone agricole ou naturelle	1313	28	2	92.86
1140	Espaces libres en milieu urbain	2257	13	0	100.00
1211	Emprises scolaires et universitaires	555	10	0	100.00
1212	Emprises hospitalières	175	4	0	100.00
1213	Equipements sportifs et de loisirs, campings	1461	14	2	85.71
1214	Cimetières	697	7	0	100.00
1215	Autres équipements collectifs	1904	9	1	88.89
1220	Équipements eau, énergies, T.I.C. et déchets	2112	6	0	100.00
1311	Emprises d'activités à dominante industrielle	491	7	0	100.00
1312	Emprises d'activités à dominante commerciale	657	6	0	100.00
1313	Emprises d'activité à dominante mixte ou tertiaire	1540	20	2	90.00
1314	Anciennes emprises d'activité	80	2	0	100.00
1320	Emprises militaires	547	7	0	100.00
1330	Exploitations agricoles	6301	27	0	100.00
1340	Zones d'extraction	93	2	1	50.00
1413	Espaces associés aux réseaux routiers et ferrés	5026	40	3	92.50
1420	Emprises aéroportuaires	8	0	0	
1430	Emprises portuaires	11	0	0	
1510	Espaces verts urbains	1726	12	2	83.33
1610	Espaces en transition	6539	23	2	91.30
1710	Places	34	0	0	
2110	Cultures annuelles et pluri-annuelles	7153	52	3	94.23
2120	Cultures spécifiques	502	3	0	100.00
2210	Vignes	1400	10	4	60.00
2221	Vergers traditionnels	1708	22	1	95.45
2222	Vergers intensifs	76	2	0	100.00
2223	Pépinières	17	1	0	100.00
2310	Prairies, friches et délaissés agricoles	16139	60	6	90.00
2320	Bosquets et haies	4179	52	3	94.23
3110	Forêts de feuillus	9197	58	3	94.83
3120	Forêts de conifères	2232	69	5	92.75
3130	Forêts mixtes	895	10	1	90.00
3140	Coupes à blanc et jeunes plantations	4608	76	7	90.79
3150	Peupleraies et sapinières	3675	4	2	50.00
3210	Pelouses et pâturages de montagne	0	0	0	
3220	Formations pré-forestières	8674	49	6	87.76
3230	Surfaces enherbées semi-naturelles	2538	34	5	85.29
3310	Plages et sables	37	0	0	
3320	Roches nues	13	0	0	
3340	Zones de sinistre (incendie, tempête)	348	0	0	

Contrôle qualité de la base de données « Occupation des Sols » à grande échelle en Région Grand Est – Millésime 2021

4110	Ripisylves et rivulaires	3119	24	0	100.00
4120	Autres milieux humides	1099	7	1	85.71
5110	Cours d'eau et canaux	540	5	0	100.00
5120	Plans d'eau	3158	27	0	100.00
5130	Bassins artificiels	1014	4	1	75.00
	GLOBAL	138185	1000	76	92.40

Le contrôle du CQC, par ses observations et remarques, permet en de confirmer les observations thématiques principales faites par le CQE.

Par ailleurs, le contrôle du CQC a été intégré à la matrice de confusion qui recense les différents types d'erreurs rencontrés.

Matrice de confusion

C'est un tableau qui recense le nombre et le type d'erreurs entre la réalité terrain et l'occupation du sol photo-interprétée. Elle est calculée sur la base du contrôle qualité thématique. La matrice de confusion indique le type de confusion sur un tableau à double entrée en indiquant le nombre.

La matrice est fournie au format Excel. Elle ne distingue pas 2010 et 2019. Elle se lit de la gauche vers la droite, c'est-à-dire qu'en colonne A on a l'occupation du sol identifiée dans la couche vectorielle et en colonnes C à AX, l'occupation du sol réelle.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
		avec													
		Bâti continu	Bâti collectif	Bâti mixte	Bâti individuel dense	Bâti individuel lâche	Bâti isolé en zone agricole ou naturelle	Espaces libres en milieu urbain	Emprises scolaires et universitaires	Emprises hospitalières	Équipements sportifs et de loisirs ; campings	Cimetières	Autres équipements collectifs	Équipements eau, Assainissement et déchets	
1															
2	Confusion de	1110	1121	1122	1123	1124	1130	1140	1211	1212	1213	1214	1215	1220	
3	Bâti continu	1110													
4	Bâti collectif	1121													
5	Bâti mixte	1122													
6	Bâti individuel dense	1123													
7	Bâti individuel lâche	1124													
8	Bâti isolé en zone agricole ou naturelle	1130													
9	Espaces libres en milieu urbain	1140													
10	Emprises scolaires et universitaires	1211													
11	Emprises hospitalières	1212													
12	Équipements sportifs et de loisirs ; campings	1213													
13	Cimetières	1214													
14	Autres équipements collectifs	1215													

