

WALOUS

D6.3 : Validation de la carte d'occupation du sol de Wallonie 2018

Subvention financée par le Service Public de Wallonie, DGO3 et le Département de
la Géomatique du Secrétariat Général.

Juin 2020

OBJET DU RAPPORT

Une première partie du rapport fournit une description détaillée de la donnée de la carte d'occupation du sol de la Wallonie 2018 via sa fiche de métadonnées. Ensuite, une évaluation de la qualité de la carte est quantifiée sur base d'un échantillon aléatoire stratifié indépendant des informations utilisées lors de la production de la donnée.

MÉTADONNÉES DE LA CARTE D'OCCUPATION DU SOL

Métadonnées – Carte d'Occupation du Sol de la Wallonie 2018 – format matriciel

Titre : Carte d'Occupation du Sol de la Wallonie (v1.2018) – format matriciel

Type de ressource : Couche de données

Type de représentation spatiale : Raster (GRID)

Accroche :

Cette couche de données matricielle reprend la cartographie de l'occupation du sol de l'ensemble du territoire wallon pour l'année 2018 (COSW2018).

Description :

Cette couche de données matricielle reprend la cartographie de l'occupation du sol de l'ensemble du territoire wallon pour l'année 2018 (COSW2018). Elle est complémentaire de la carte d'utilisation du sol produite conjointement dans le projet WALOUS.

L'occupation du sol se définit comme la « couverture physique et biologique de la surface terrestre, y compris les surfaces artificielles, les zones agricoles, les forêts, les zones (semi-)naturelles, les zones humides et les masses d'eau » (directive européenne INSPIRE 2007/2/CE, 2007).

Les spécifications techniques de la COSW2018 résultent d'une procédure d'analyse des besoins et de co-constructions avec un panel d'acteurs wallons. Cette version de la COSW2018 fournit cette classification sur une grille de 1x1m (grille INSPIRE) et est caractérisée par une unité minimale de cartographie de 15m².

La COSW2018 détaille l'occupation du sol en 11 classes. Certains pixels appartiennent toutefois à deux classes en cas d'étagement. Ces codes à deux chiffres non nuls sont détaillés à la suite des codes principaux:

- 1) Revêtement artificiel sol: cette classe reprend tous les revêtements artificiels du sol de nature peu ou pas perméable (ex béton, le bitume ou les pavés). Ceci comprend le réseau routier, les trottoirs, les terrasses, les parkings (les emplacements de parking semi-perméables non végétalisés sont également dans cette catégorie) et les terrains ou pistes de sport en matériaux synthétiques.
- 2) Constructions artificielles hors sol: cette classe reprend tous les bâtiments et autres constructions s'élevant au-dessus du sol.
- 3) Réseau ferroviaire: cette classe reprend les rails et ballasts des chemins de fer encore en activité (les RaVEL entre dans la classe 1). Cette classe se différencie de la classe 1 par sa grande perméabilité.
- 4) Sols nus: cette classe inclut tout type de roche mère n'étant couverte par des végétaux supérieur à aucun moment de l'année. Ces sols sont soit naturels (roches affleurantes, falaises, berges caillouteuses...), soit générés par l'activité humaine (extraction, sols compactés, coupe à blanc de l'année...)
- 5) Eaux de surface: cette classe comprend toutes les surfaces d'eau libre, naturelles ou artificielles. Ceci inclut donc à la fois les cours d'eau (rivières, fleuves et canaux) et les plans d'eau (mares, étangs, lacs, bassins de décantation, piscines extérieures).

- 6) Alternance sol nu/végétation : cette classe inclut toutes les parcelles temporairement couvertes par de la végétation herbacée ayant fait l'objet d'un labour en début de saison agricole. Elle est donc caractérisée par la présence de sol nu durant une partie de l'année. On y retrouve toutes les cultures annuelles ainsi que les prairies semées dans l'année ou à l'automne 2017.
- 7) Couvert herbacé : cette classe reprend tous les sols recouverts par de la végétation herbacée tout au long de l'année. Cette végétation peut être d'origine naturelle (landes, mégaphorbiaies, tourbières, pelouses naturelles, végétation rudérale recolonisant une friche ou une ancienne coupe à blanc...), agricole (prés et prairies non labourés dans la saison) ou artificielle (jardin, terrains de sport, parcs...)
- 8) Arbres résineux : cette classe comprend tous les arbres de plus de 3 m (isolés, en haie ou en peuplement) du groupe des résineux (gymnosperme).
- 9) Arbres feuillus : cette classe comprend tous les arbres de plus de 3 m (isolés, en haie ou en peuplement) du groupe des feuillus (angiosperme).
- 80) Arbustes résineux : cette classe comprend tous les arbres et arbustes de moins de 3 m (isolés, en haie ou en peuplement) du groupe des résineux (gymnosperme).
- 90) Arbustes feuillus : cette classe comprend tous les arbres et arbuste de moins de 3 m (isolés, en haie ou en peuplement) du groupe des feuillus (angiosperme).

Les étiquettes doubles sont des codes à deux chiffres dont le premier chiffre indique l'élément inférieur et le second chiffre indique l'élément supérieur.

La classe 11 indique la présence de d'une route sous un pont. Les classes 18 et 19 sont utilisées lorsqu'un arbre, respectivement résineux ou feuillus, masque une route.

Les classes 28 et 29 sont utilisées lorsqu'un arbre, respectivement résineux ou feuillus, masque un bâtiment.

Les classes 31, 38 et 39 indiquent la présence de rails sous un pont, des résineux ou des feuillus.

Les classes 51,58 et 59 indiquent la présence d'eau sous un pont, des résineux ou des feuillus.

La classe 62 reprend les cultures sous un bâtiment, c'est-à-dire les serres agricoles ou horticoles.

La COSW2018 a été construite par la fusion de classifications orientée-objet GEOBIA et orienté-pixel exploitant les données sources suivantes :

- Orthophotos 2018 de la Région wallonne ;
- Modèle Numérique de Hauteur généré par la différence du Modèle Numérique de Surface photogrammétrique associé aux orthophotos 2018 et du Modèle Numérique de Terrain LiDAR de 2013-2014 ;
- Données satellitaires Sentinel-2 (programme européen Copernicus) ;
- Squelette vectoriel construit par intégration des données du réseau routier et délinéation du bâti du PICC 2018 et des cours d'eau navigables et des chemins de fer de l'IGN TOP10vGIS (2018).

La COSW2018 a été consolidée par une approche de règles automatisées et par consolidation manuelle par des experts.

La COSW2018 a été validée à l'aide d'un échantillon aléatoire de 1430 points photointerprétés indépendamment du processus de classification et analysés selon les règles de l'art.

La mise en œuvre de la COSW2018 s'inscrit dans le cadre du projet WALOUS (2017-2020) subventionné par la Région wallonne et réalisé par un consortium de deux universités (UCLouvain, ULB) et d'un centre de recherche public de type UAP (ISSeP).

La couche est présentée avec deux symbolologies : une symbolologie pour les éléments supérieurs et une symbolologie pour les éléments inférieurs. La première symbolologie se rapproche de ce qui est visible du ciel sur les orthophotos 2018, alors que la seconde permet de voir ce qui est au sol.

Généalogie :

Les étapes principales de production de la donnée COSW2018 sont les suivantes :

1. Acquisition des données sources – missions aériennes et satellitaires (orthophotos, modèles numériques, Sentinel-2) ;
2. Création d'un squelette vectoriel au départ des données de référence en ce qui concerne la délimitation de l'emprise au sol des bâtiments, du réseau routier, des chemins de fers et des cours d'eau navigables ;
3. Classifications orientée-objet et pixel des données d'observation de la Terre ;
4. Fusion des approches par méthode algorithmiques (Random Forest) ;
5. Consolidation par règles de reclassification automatiques et manuelles ;
6. Validation ;
7. Publication.

Références et crédits :

https://www.researchgate.net/publication/335368069_CREATING_WALLONIA'S_NEW_VERY_HIGH_RESOLUTION_LAND_COVER_MAPS_COMBINING_GRASS_GIS_OBIA_AND_OTB_PIXEL-BASED_RESULTS

Avec le soutien de la Région Wallonne

Date de création : dd/mm/yyyy

Date de dernière révision : dd/mm/yyyy

Fréquence de mise à jour : Non planifiée/~~Ponctuelle~~/Régulière

Contact :

Type de contact	Organisme	Email
Point de contact première ligne	Helpdesk carto du SPW (SPW)	helpdesk.carto@spw.wallonie.be
Gestionnaire de la donnée	Consortium ?	contact@walous.be
Propriétaire	Service public de Wallonie (SPW)	
Contact pour les métadonnées		
Distributeur	Service public de Wallonie (SPW)	helpdesk.carto@spw.wallonie.be
A l'origine de		

Ressources associées :

- Orthophotos 2018
- Modèle Numérique de Surface photogrammétrique associé aux orthophotos 2018
- Relief de la Wallonie - Modèle Numérique de Terrain (MNT) 2013-2014
- Données satellitaires Sentinel-2 (programme européen Copernicus)
- PICC 2018
- IGN TOP10vGIS 2018
- Lifewatch 2015

Conditions d'accès et d'utilisation :

ACCÈS :

Les conditions générales d'accès s'appliquent
(<http://geoportail.wallonie.be/files/documents/ConditionsSPW/DataSPW-CGA.pdf>)

UTILISATION :

Les conditions générales d'utilisation s'appliquent
(<http://geoportail.wallonie.be/files/documents/ConditionsSPW/DataSPW-CGU.pdf>)

Source à mentionner :

Géométries métiers utilisées : données cadastrales, unités statistiques, ...

EVALUATION DE LA QUALITÉ DE LA CARTE D'OCCUPATION DES SOLS

La qualité est estimée via la matrice de confusion construite par rapport à un set de 1460 points de validation photointerprétés. Le plan d'échantillonnage stratifié et aléatoire utilisé pour évaluer la qualité de la fusion et des données d'entrée de la fusion (livrable WP5) permet d'évaluer la qualité de la consolidation également. Cependant, le jeu de points a été légèrement corrigé à cause d'erreurs de photo-interprétation sur 46 points. Le livrable WP5 a donc également été mis à jour suite à ces corrections. Le nombre de points utilisés dans le calcul des matrices pour chaque donnée est repris dans le tableau 1, ainsi que le poids attribué à chaque point (explication des poids, cfr. livrable WP5).

Tableau 1 : Nombre de points par strate, choisis sur base des exactitudes des classes en sortie de fusion sur les 4 zones d'étude afin de minimiser la variance de prédiction. Les poids servent à corriger le biais de stratification en attribuant à chaque classe sa probabilité d'observation lors de la construction de la matrice de confusion.

Classe	Nombre de points	Poids de chaque point de la classe
Constructions artificielles au sol	186	1,5
Constructions artificielles hors sol	98	2,5
Eau	62	0,7
Herbage	533	5,5
Alternance sol nu/végétation basse sur une année	220	14,7
Sols nus	92	1,6
Feuillus	104	22,7
Résineux	205	5,6

1.1.1 MATRICE DE CONFUSION

Tout comme pour l'évaluation des données d'entrée de la fusion et de la fusion elle-même, la matrice de confusion pour calculer l'exactitude globale de la carte est créée via l'outil d'analyse spatiale « extract values to points » de ArcGIS. Un équivalent de cet outil existe également dans QGIS : « add raster values to points ». Le calcul de précision globale a ensuite été effectué via un tableau dynamique croisé dans Excel en additionnant les poids de chaque point pour chaque combinaison référence/classification.

Le calcul de l'exactitude globale de la carte d'occupation des sols a été effectué sur base du rapport entre la somme des valeurs de la diagonale de la matrice de confusion et la somme des valeurs de toutes les cellules de la matrice de confusion.

1.1.2 QUALITÉ DES DONNÉES DE LA FUSION

Suite à la correction de 46 points, les exactitudes globales de données de la fusion sur base de 1430 points de validation sont reprises dans le tableau 2. Les points tombant dans les prairies temporaires (32) ont été retirés du jeu de validation car leur photo-interprétation sur base des orthophotos 2018 était peu cohérente. En effet, il n'est pas possible, sur base d'une seule image, de déterminer si la valeur du point correspond à une classe « herbage » ou une classe « alternance sol nu/végétation basse sur une année ».

Tableau 2 : Exactitude globale des données de la fusion et de la fusion

Classification	Exactitude globale [%]
Orthophotos 2018 par objets – sans distinction des terres arables	87,2
Orthophotos 2018 par pixels – sans distinction des terres arables	90
Orthophotos 2018 par pixels – avec distinction des terres arables	73,9
Sentinel - occupation du sol 2018 (10 m de résolution)	65,3
Sent2Agri – masque et typologie des terres arables 2018	94,2
Masque forestier 2016 en deux classes	93,3
Données du SIGEC 2018 en deux classes	86,7
FUSION BRUTE DES DONNEES (sans consolidation)	87,4

1.1.3 QUALITÉ DE LA CARTE D'OCCUPATION DES SOLS (VERSION 1)

Les résultats de la première version de la carte d'occupation des sols en date du 11/06/2020 ont été estimés sur base des mêmes points que les résultats d'évaluation de la qualité des données d'entrée de la fusion et de la fusion elle-même. L'amélioration de l'exactitude globale grâce à la consolidation est très significative. On obtient une nouvelle exactitude globale de 89.7%. Ceci est insuffisant pour conclure que l'objectif de qualité de 90% a été atteint avec une confiance de plus de 95%, mais reste très encourageant lorsque la consolidation sera terminée.

Tableau 3 : Matrice de confusion des données brutes de la fusion. Les classes 80 et 90 sont fusionnées aux classes 8 et 9, respectivement. La classe 3 est regroupée à la classe 1. Seules les étiquettes supérieures sont prises en compte.

		Classes estimées par la fusion brute								
		1	2	4	5	6	7	8	9	PA [%]
Classes réelles	1 (rev. art. au sol)	352	7	59	9	2	17	6	6	76,9
	2 (cons. hors sol)	32	162	0	0	0	0	3	0	82,2
	4 (sol nu)	66	2	63	3	5	45	0	6	33,2
	5 (eau)	1	0	0	58	0	0	0	0	98,3
	6 (alt. sol nu/herb)	0	0	0	0	2107	502	0	0	80,8
	7 (herbage)	27	0	12	0	40	3139	0	79	95,2
	8 (résineux)	0	0	0	0	0	49	762	123	81,6
	9 (feuillus)	0	0	0	0	0	92	80	2210	92,8
	UA [%]	73,6	94,7	47,0	82,9	97,8	81,7	89,5	91,2	

Tableau 4 : Matrice de confusion de la version consolidée de la carte d'occupation des sols en date du 11 juin 2020. Les classes 80 et 90 sont fusionnées aux classes 8 et 9, respectivement. La classe 3 est regroupée à la classe 1. Seules les étiquettes supérieures sont prises en compte.

		Classes estimées par la carte consolidée d'occupation des sols								
		1	2	4	5	6	7	8	9	PA [%]
Classes réelles	1 (rev. art. au sol)	400	7	12	0	2	39	12	0	84,7
	2 (cons. hors sol)	10	183	0	0	0	6	3	0	90,6
	4 (sol nu)	76	1	59	7	0	34	0	11	31,4
	5 (eau)	0	0	0	59	0	0	0	0	100,0
	6 (alt. sol nu/herb)	5	0	0	0	2363	166	0	6	93,0
	7 (herbage)	14	0	6	0	54	3144	40	79	94,2
	8 (résineux)	0	0	0	0	0	32	845	62	90,0
	9 (feuillus)	5	0	0	0	0	92	265	2083	85,2
	UA [%]		78,4	95,8	76,6	89,4	97,7	89,5	72,5	92,9

La comparaison des matrices de confusion (tableaux 3 et 4) met en évidence les bénéfices du travail de consolidation. Les exactitudes par classe permettent une analyse plus détaillée des erreurs relatives. L'exactitude pour l'utilisateur (UA) est la probabilité qu'un pixel d'une catégorie donnée de la carte soit correctement classé ; l'exactitude du producteur (PA) est la probabilité qu'un élément du terrain se retrouve au bon endroit sur la carte.

L'UA moyenne après consolidation est de 87%, soit une amélioration de 3% par rapport à la fusion brute. On remarque une très nette (presque 30%) amélioration de la classe « sol nu » mais une surestimation des résineux au détriment des feuillus à cause de la méthode automatique de densification des résineux. L'UA des autres classes est soit meilleure, soit équivalente à celle de la fusion brute.

La PA moyenne après consolidation est de 84%, soit 4% de mieux que la fusion brute. On peut voir que la détection des bâtiments a été fortement améliorée (près de 14%). Par contre, la PA des feuillus se voit réduite à cause de la surestimation de résineux (dont la PA augmente en contrepartie). La classe des sols nus reste sous-estimée malgré les premières actions de consolidation. La plupart des erreurs proviennent cependant d'une confusion avec les revêtements artificiels, or même la photointerprétation de ces points reste peu fiable, d'autant que la limite entre un sol compacté et certains revêtements est assez floue.

En valeur absolue, il ressort que l'erreur la plus fréquente est la surestimation des résineux au détriment des feuillus. L'amélioration du visuel de la carte basée sur une généralisation automatique a en effet été contre-productive en terme d'exactitude (diminution de 3% par rapport à la fusion brute entre ces deux classes). La consolidation de ces erreurs sera une de nos priorités. Cet exercice de consolidation sera grandement facilité quand les données de la carte des types de forêt 2018 seront disponibles. L'analyse statistique des conflits entre bases de données par zone nous permettra par ailleurs une consolidation mieux ciblée. La deuxième erreur la plus fréquente est la confusion entre les terres labourées et les herbages. Un effort particulier sera également fourni dans ce sens, mais il convient de noter que, d'une part, l'information sur le labour des prairies ou le non-labour des cultures n'est disponible dans aucune autre base de donnée sur le territoire wallon et, d'autre part, les résultats actuels sont déjà très bons en terme relatif (96%). Enfin, la troisième erreur la plus fréquente concerne les confusions entre feuillus et herbages. Il s'agit ici d'une erreur principalement liée à la précision du modèle numérique de hauteur, donc améliorable quand le MNH LIDAR sera disponible. Un effort supplémentaire permettra la consolidation finale des classes les plus faibles en terme de qualité relative, pour autant que cette consolidation puisse se faire par photo-interprétation.

Il faut par toutefois noter que les matrices de confusion ci-dessus sont basées sur une référence construite à partir des orthophotos 2018. Ceci engendre quelques fausses détection d'erreurs pour certains points de bordure (bords de bâtiment ou de forêt vus à la place du sol). L'exactitude globale de la carte est donc légèrement sous-estimée, principalement pour les couples « 1 vs 2 » et « 7 vs 9 ».