

OSO : Une occupation du sol annuelle sur toute la France métropolitaine

Grâce aux nombreuses images acquises par Sentinel-2, une carte d'occupation du sol de la France métropolitaine est produite chaque année depuis 2016. Fruit de processus entièrement automatiques mis au point au [Cesbio](#), elle offre une vision globale des surfaces naturelles, agricoles et artificialisées. En plus, elle est disponible en open data. Visite guidée.

Mots clés : occupation du sol, urbain, cultures, Sentinel-2, open data



Une vue de l'OSO 2017

Extension urbaine, construction d'infrastructures, mutations agricoles, reconquête d'espaces naturels... le paysage évolue en permanence. L'observation et la mesure de ces transformations sont au cœur de nombreuses politiques publiques. Depuis quelques années, les bases de données d'occupation du sol sont devenues les instruments privilégiés de cette mesure. Car elles donnent, en tout point du territoire, une information sur la couverture physique et/ou biologique du sol (prairies, routes...), bon reflet des activités humaines.

Mais derrière le terme d'occupation du sol, se cachent des niveaux de détails, des modes et des coûts de production très différents. La taille des éléments les plus fins observés (aussi appelée unité minimale de collecte - UMC) peut aller de quelques dizaines de m² à plusieurs dizaines d'hectares. La nomenclature (nombre de postes de la légende), peut varier de quelques items (zones artificialisées, zones naturelles, zones cultivées) à plus de 80 postes comme c'est le cas pour le [MOS d'Île-de-France](#). Elle peut intégrer une dimension « usage » plus ou moins importante, distinguant par exemple différents types d'activités économiques ou agricoles.

OSO, kézako ?

Dans ce paysage complexe, l'OSO occupe une place à part. Depuis 2016, elle couvre chaque année l'ensemble de la France métropolitaine (des travaux sont en cours pour adapter les chaînes de production aux territoires ultramarins) avec une unité minimale de collecte allant de 100 à 400 m² selon les catégories et une nomenclature en 17 postes (2016 et 2017), étendue à 23 postes en 2018. La seule autre base de données d'occupation du sol couvrant tout le territoire est Corine Land Cover, bien moins détaillée spatialement (UMC à 25 hectares) et mise à jour uniquement tous les six ans (voir fiche Corine Land Cover).

Son mode de constitution est également unique : C'est la première base de données produite de façon entièrement automatisée à partir d'images satellitaires Sentinel-2 traitées par le [pôle Theia](#) (corrections atmosphériques et masque de nuages – niveau 2A).

La fréquence de revisite exceptionnelle du couple de satellites (tous les 5 jours) et le fait qu'ils produisent des données en continu, permet d'obtenir chaque année entre une dizaine et une cinquantaine d'images exploitables (sans nuages) en chaque point du territoire. Cette profondeur temporelle est très utile pour comprendre l'occupation du sol, notamment dans les milieux naturels et cultivés. Si feuillus et conifères se confondent sur une image acquise au printemps, ils sont bien différenciables sur une image d'hiver. Idem pour les cultures qui ont chacune leur dynamique temporelle. Afin de disposer d'un maximum d'images sur une même zone, le CESBIO a mis au point une méthode pour interpoler les dalles nuageuses à partir des dalles précédentes et suivantes.

Avec des mesures dans 13 longueurs d'onde, allant du moyen infrarouge au bleu, il est possible de réaliser de nombreuses combinaisons de canaux, permettant de distinguer par exemple des rizières, des cultures en jachère, etc.

Dans le cadre du Centre d'expertise scientifique (CES) du pôle Theia dédié à l'occupation du sol, le [CESBIO](#) a rassemblé les méthodes développées dans de nombreux travaux de recherche au sein d'une même chaîne de traitements, baptisée Iota2, capable d'exploiter ces séries temporelles. Des données de référence (zones fiables pour lesquelles l'occupation du sol est connue, prises dans la BD Topo ou Corine Land Cover par exemple) guident les algorithmes et garantissent la qualité du résultat. Ainsi, pour chaque pixel, vu à travers une série d'images, une prédiction de classement dans la catégorie principale de l'année est effectuée.

Un produit directement exploitable

OSO se présente sous deux formes.

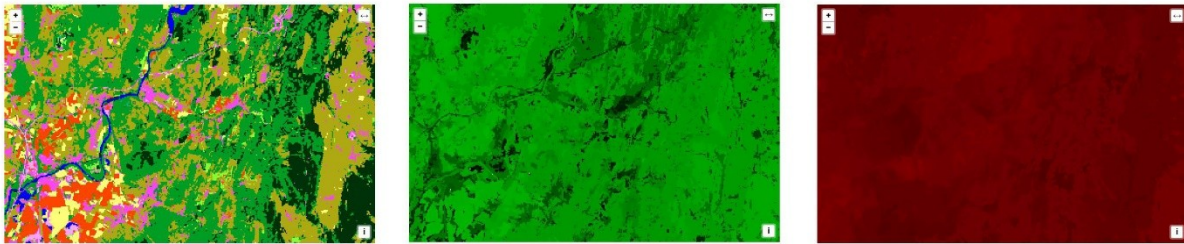
Trois fichiers raster (pixel de 10 m de base) couvrent toute la France, en format GeoTIFF. Chacun comprend plus de 10 milliards de pixels.

- Occupation du sol (en 17 ou 23 classes)
- Carte de validité (nombre d'images utilisées pour produire ce pixel)
- Carte de confiance (indice de confiance du classificateur sur ce pixel)

Un service WMS

Une couche vectorielle est également disponible, par département, en format Shape. Elle est dérivée du raster simplifié (pixel de 20 m) avec lissage pour éviter les effets de crénelage. Chaque polygone a plusieurs attributs : Sa classe majoritaire, le nombre moyen d'images sans nuages utilisées pour la

classification ainsi que l'écart-type, l'indice de confiance de la classe du polygone et le pourcentage de la surface couverte par chaque classe de la nomenclature.



OSO fournit 3 informations : Classification, confiance et validité

Même si ces couches de données sont directement exploitables dans un SIG, elles ne sont pas orthorectifiées sur le RGE de l'IGN et peuvent présenter de légers décalages de l'ordre du pixel (10 à 20 mètres).

Produites par le Cesbio en 2016 et 2017, les couches de l'OSO sont désormais réalisées par l'équipe Theia du CNES. Elles sont à disposition en open data sur le site <https://theia.cnes.fr> ainsi que sur le serveur du Cesbio. Le téléchargement nécessite d'être enregistré sur Theia.

Le millésime 2018 de l'OSO est disponible mars 2019, tandis que sa transformation et son découpage en format vectoriels sont prévus à l'été 2019.

La méthodologie utilisée pour produire l'OSO a été élaborée sur la base d'images Landsat (avec une résolution inférieure). Le Cesbio dispose donc [de cartes d'occupation du sol](#) pour les années 2009, 2010, 2011 et 2014.

Usages et précautions

OSO fournit une vue d'ensemble de l'occupation du sol, assez détaillée sur le milieu agricole, qui peut servir à produire des indicateurs variés et à rendre compte de grandes tendances. Les données sont utilisées dans le domaine de la recherche ainsi que par des entreprises (Bureau d'études Biotopie sur les zones humides par exemple) et des acteurs publics. Ainsi, le [CITEPA](#) a évalué le potentiel d'OSO pour son rapportage annuel pour le GIEC par exemple. Son échelle d'analyse est cohérente avec le RGE de l'IGN, même s'il peut y avoir des décalages de localisation.

Certaines classes sont mieux reconnues que d'autres. Ainsi, à cause principalement de la faible résolution spatiale des capteurs (de 10 m à 60 mètres selon les longueurs d'onde), la décomposition selon les différents postes urbains est plus ou moins fiable : un pixel dans l'urbain diffus a des chances d'être confondu avec de l'urbain dense. En revanche, les surfaces artificialisées prises dans leur ensemble, sans les décomposer en sous-catégories, sont bien discriminées.

Le classement de la végétation naturelle, aussi détaillé soit-il, a également ses limites. Dans la réalité, le passage entre une pelouse et une végétation basse arbustive se fait progressivement et ces classes ne sont pas toujours bien définies dans les bases qui ont servi de référence pour constituer les échantillons d'apprentissage.

Enfin, OSO fournit une vue annuelle, ce qui permet de s'appuyer sur les cycles végétatifs pour différencier par exemple les cultures. Pour les formes d'occupation du sol stables, c'est la valeur

majoritaire au cours de l'année qui est représentée. Si un pixel a changé en milieu d'année, il peut être mal classé.

Ainsi, avant de faire des comparaisons de pixel à pixel, il vaut mieux faire des regroupements de classes de façon à éviter de prendre trop de pixels en limite de classe. OSO n'est par exemple pas adapté pour étudier la transition d'une année sur l'autre entre deux classes proches comme urbain dense et diffus, mais la base fournira des données robustes sur la consommation des terres agricoles en regroupant tous les postes urbains et tous les postes de culture.

Pour aller plus loin :

Le [Cesbio](#) ou Centre d'Etudes Spatiales de la BIOSphère est une unité mixte de recherche basée à Toulouse associant des chercheurs du CNRS, de l'Université Paul Sabatier, du CNES, de l'IRD et de l'INRA.

Pour en savoir plus sur OSO : <http://osr-cesbio.ups-tlse.fr/~oso/>

La nomenclature OSO :

	Nomenclature 2016 et 2017	Nomenclature 2018
Cultures annuelles	Culture d'été	Soja
		Tournesol
		Maïs
		Riz
		Tubercules ou racines
		Céréales à pailles
	Culture d'hiver	Colza
		Protéagineux
		Prairies
		Vignes
Cultures pérennes	Prairies	Vergers et arboriculture
		Vergers et arboriculture
		Vergers et arboriculture
Forêts	Forêt feuillus	Forêt feuillus
		Forêt résineux
Formation naturelle basse	Pelouses et pâturages naturels	Pelouses et pâturages naturels
		Lande ligneuse
Urbain	Bâti dense	Bâti dense
		Bâti diffus
		Zones industrielles et commerciales
		Routes
		Routes
Surfaces minérales	Surfaces minérales naturelles	Surfaces minérales naturelles
		Plages et dunes
Autres	Eau	Eau
		Glaciers et neiges éternelles
		Glaciers et neiges éternelles
Total	17 classes	23 classes

Contact :

Jordi Inglada

CESBIO

jordi.inglada@cesbio.eu