Audit comparatif EVA 2023 / Végéstrate 2023

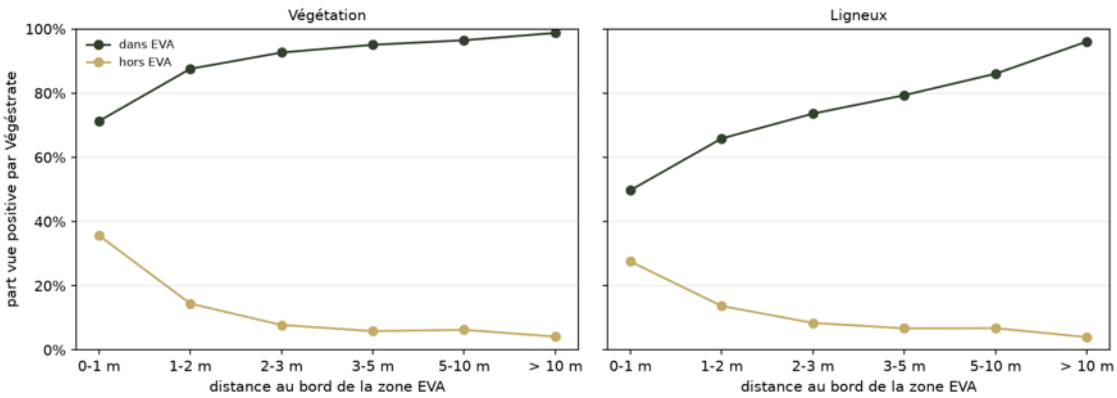
Synthèse

Deux inventaires de la végétation 2023 de la Métropole de Lyon sont comparés sur 527,964 km² : **EVA 2023**, photo-interprétation manuelle en polygones, et **Végéstrate 2023**, raster à 0,2 m issu du LiDAR et des orthophotographies qui est ensuite vectorisé. **L'audit mesure l'accord entre les deux sources** : aucune n'est une vérité terrain. Les pourcentages portent sur la surface comparée ; sauf mention contraire, ils sont calculés au pixel de 0,2 m, EVA étant rastérisée sur la grille Végéstrate.

- **Végétal / non végétal : 92,5 % d'accord.** Les deux sources voient la végétation aux mêmes endroits.
- **Ligneux / herbacé : 84,6 % d'accord ; arbre / arbuste / herbe : 80,2 %.** Plus la comparaison entre dans le détail des strates, plus les définitions diffèrent (arbuste dès 1 m pour EVA, végétation moyenne dès 1,5 m pour la classification LiDAR utilisée par Végéstrate ; ce seuil peut être adapté au besoin, en reclassant le LiDAR brut ou en seuillant directement le modèle de hauteur).
- **Forêts (EVA code 311) : Végéstrate retrouve la couverture ligneuse attendue sur 98,3 % de leur surface**, à l'échelle du polygone EVA.
- **Le désaccord est concentré sur les contours.** À l'intérieur des zones végétales EVA, Végéstrate voit de la végétation sur 71,3 % des pixels à moins de 1 m du bord et sur 98,7 % à plus de 10 m.
- **Végéstrate voit la végétation que l'unité minimale d'EVA (100 m²) écarte** : 37,3 % de la végétation ligneuse Végéstrate située en zone EVA non végétale est en taches de moins de 100 m² (arbres isolés, alignements), soit 2,957 km².



Cinq lieux d'exemple : orthophotographie, EVA 2023, Végétrate 2023 et accord végétation



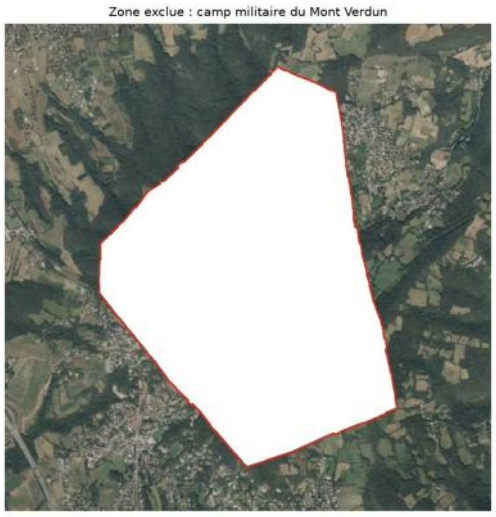
Part des pixels vus végétaux (ou ligneux) par Végétrate selon la distance au bord des zones EVA

Données

Donnée	Description
EVA 2023	Produit par Geofit pour la Métropole par photo-interprétation manuelle. Polygones, champs `COD3` (classe) et `CODH` (classe de hauteur), EPSG:3946. Unité minimale de collecte (UMC) 100 m², largeur minimale de collecte (LMC) 3 m, taux de couverture propre à chaque classe.
Végéstrate 2023, classes	Produit par TelesCoop et le LIRIS. Raster 0,2 m, EPSG: 3946. 0 autre, 1 herbacé (FLAIR-HUB), 2 végétation moyenne (LiDAR 1,5-5 m), 3 végétation haute (LiDAR > 5 m). Les zones « Autre » du LiDAR reconnues comme végétation moyenne ou haute par FLAIR-HUB sont complétées (`src/postprocessing/merge_classifications.py`), puis un filtre modal 3x3 est appliqué pour lisser.
IRIS, communes, LCZ	Base de données data.gouv (IRIS INSEE, zones climatiques locales).

Le domaine est l'union des IRIS de la Métropole : les deux sources y sont découpées.

Zone exclue. L'orthophotographie 2023 est vide sur le camp militaire du Mont Verdun : Végéstrate n'y a pas de donnée. EVA a complété cette zone grâce à des sources secondaires (OSM, IGN, etc.). Le contour est tracé à partir des pixels transparents de l'orthophotographie Grand Lyon 2023 (WMS `ortho_2023`, 0,5 m). Ces 5,865 km² sont retirés pour les deux sources ; leur contenu EVA est dans `zone_exclue.csv`.



Zone exclue

Harmonisation des classes

EVA mêle des classes de couverture et des classes d'usage du sol ; Végéstrate ne décrit que la couverture physique. Chaque classe EVA est rattachée à un groupe de couverture attendue :

groupe	libellé	codes EVA (page du guide)
arbore	Strate arborée (> 3 m)	111 (p. 3)
arbustif	Strate arbustive (1-3 m)	112 (p. 4)
herbace	Herbacé	113 (p. 5), 221 (p. 15), 222 (p. 16), 223 (p. 17), 224 (p. 18), 225 (p. 19), 361 (p. 30), 362 (p. 31), 411 (p. 37)
boise_mixte	Boisé multistrate (forêts, haies, fourrés)	241 (p. 21), 311 (p. 22), 312 (p. 23), 321 (p. 24), 322 (p. 25), 331 (p. 26), 341 (p. 27), 342 (p. 28), 351 (p. 29), 371 (p. 32), 372 (p. 33), 374 (p. 34)
non_vegetal	Non végétalisé	121 (p. 8), 122 (p. 9), 381 (p. 35), 382 (p. 36)
eau	Eau	511 (p. 39), 521 (p. 40), 522 (p. 41), 523 (p. 42), 524 (p. 43), 525 (p. 44), 531 (p. 45), 532 (p. 46), 533 (p. 47)
non_comparable	Non comparable (usage du sol, strates mêlées, cultures)	114 (p. 6), 115 (p. 7), 211 (p. 10), 212 (p. 11), 213 (p. 12), 214 (p. 13), 215 (p. 14), 231 (p. 20), 412 (p. 38)

Les classes « non comparable » (cultures classées à partir du RPG, friches, jardins collectifs, vergers) sont laissées hors des matrices. Comptées comme végétales, l'accord végétal / non végétal passe à 93,4 %.

Recoupage des hauteurs

EVA distingue < 1 m, 1-3 m et > 3 m ; Végéstrate distingue la végétation moyenne (1,5-5 m) et haute (> 5 m). Pour comparer les strates, Végéstrate est recoupé sur les seuils d'EVA (variante principale : lecture littérale sur hauteurs arrondies : < 1 m herbe, 1 à 3 m arbuste, > 3 m arbre, soit arbuste dès 1 m arrondi et arbre dès 4 m arrondi) :

- les pixels herbacés (classe 1, FLAIR-HUB) restent de l'herbe ;
- les zones de végétation moyenne ou haute (classes 2 et 3) sont reclassées selon leur hauteur médiane ;
- une zone de classe 2 ou 3 dont la hauteur médiane est nulle n'a pas de hauteur mesurée : elle garde sa classe Végéstrate native. C'est le cas de la végétation reconnue par FLAIR-HUB là où le

LiDAR ne donne pas de hauteur, notamment autour du camp militaire du Mont Verdun, où la lacune du LiDAR semble plus large que celle de l'orthophotographie.

Une zone est un ensemble 4-connexe de pixels de même classe, la règle utilisée pour vectoriser Végéstrate. La hauteur d'une zone est la hauteur médiane sur ses pixels, comme l'attribut `median_height` des polygones Végéstrate publiés dans IA.rbre (`compute_zone_elevation` dans `src/postprocessing/apply_vegetation_elevation.py`). Les zones sont calculées par bloc de 2048 pixels avec une marge de 20 m : un massif plus grand prend la médiane de sa partie visible.

Un pixel herbacé de Végéstrate est un pixel où le LiDAR n'a pas vu de végétation et où FLAIR-HUB a reconnu de l'herbe.

Le modèle de hauteur de végétation de Végéstrate est arrondi au mètre (paramétrisable), chaque seuil n'est donc connu qu'à ± 0,5 m. Selon la variante de seuils (`principal`, `seuils_bas`, `seuils_hauts`), l'accord par strate varie de 79,4 % à 80,2 % ; le détail est dans `indicateurs_accord.csv`.

Niveaux, échelles et indicateurs

La comparaison se fait sur trois niveaux : **végétation** (végétal / non végétal), **ligneux (strates hautes et moyennes) / herbacé**, **strates** (arbre / arbuste / herbe, sur les seules classes EVA à strate unique : forêts, haies et fourrés sont exclus de ce niveau).

Les deux données ont des échelles différentes : **le pixel** de 0,2 m pour Végéstrate (qui est ensuite vectorisé) et **le polygone pour EVA**.

Les indicateurs sélectionnés pour les comparaisons sont :

- **Accord global** et **décomposition du désaccord** : désaccord de **quantité** (surfaces totales différentes) et d'**allocation** (surfaces semblables à des endroits différents).
- **Composition des polygones EVA** : part des polygones EVA (et de leur surface) où Végéstrate retrouve la couverture attendue par la définition de leur classe (taux de couverture du guide EVA). C'est la comparaison à l'échelle d'EVA.
- **Accord selon la distance au bord** des zones EVA, dedans et dehors.
- **Petites taches** : végétation Végéstrate en taches 4-connexes de moins de 100 m² (l'UMC d'EVA).

- **Couverture par commune**.

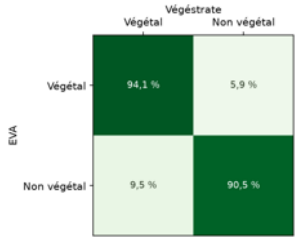
Résultats

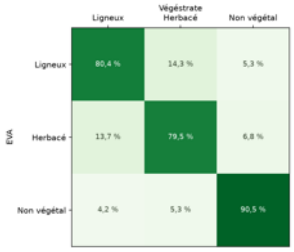
Accord global

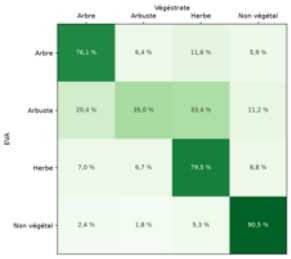
niveau	version_vegestrate	surface_km2	accord	desaccord_quantite	desaccord_allocation	kappa
vegetation	natif	445,823	92,5 %	0,9 %	6,6 %	0,848
ligneux_herbace	natif	445,823	84,6 %	3,0 %	12,4 %	0,762
ligneux_herbace	principal	445,823	84,6 %	3,0 %	12,4 %	0,762
strates	natif	380,897	80,0 %	3,0 %	17,0 %	0,690
strates	principal	380,897	80,2 %	4,3 %	15,4 %	0,694

Matrices de confusion

Répartition de chaque classe EVA (ligne) entre les classes Végéstrate (colonne) ; la diagonale correspond à l'accord.

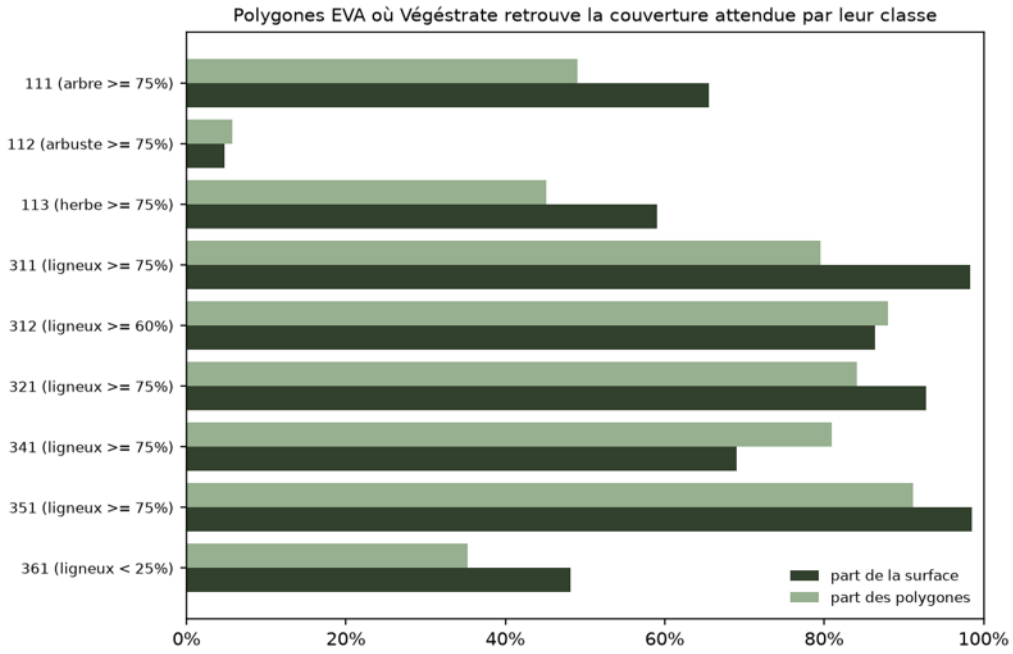






À l'échelle des polygones EVA

Pour chaque classe EVA, part des polygones (et de leur surface) où Végéstrate retrouve la couverture attendue par la définition de la classe (« ligneux » = arbre + arbuste), variante principale. Végéstrate retrouve cette couverture sur 98,3 % de la surface des forêts et 65,6 % de la strate arborée, mais sur 4,7 % seulement de la strate arbustive, ce qui est cohérent avec la différence de définition de l'arbuste (dès 1 m pour EVA, 1,5 m pour la classification LiDAR). Détail par classe dans `regles_couverture.csv`.



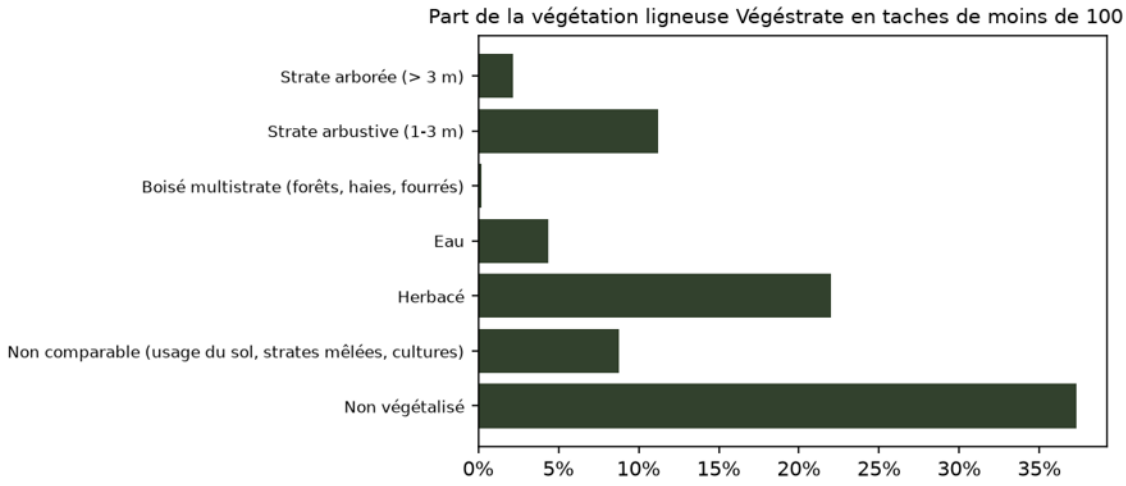
Couverture attendue retrouvée par Végéstrate dans les polygones EVA

Accord selon la distance au bord des zones EVA

Voir la figure de la synthèse. Dans les zones EVA végétales, Végéstrate voit de la végétation sur 71,3 % des pixels à moins de 1 m du bord, 98,7 % à plus de 10 m. Hors des zones EVA végétales, il en voit sur 35,6 % des pixels à moins de 1 m du bord, 4,1 % à plus de 10 m. Le désaccord se concentre dans les premiers mètres autour des contours.

Végétation en petites taches (< 100 m²)

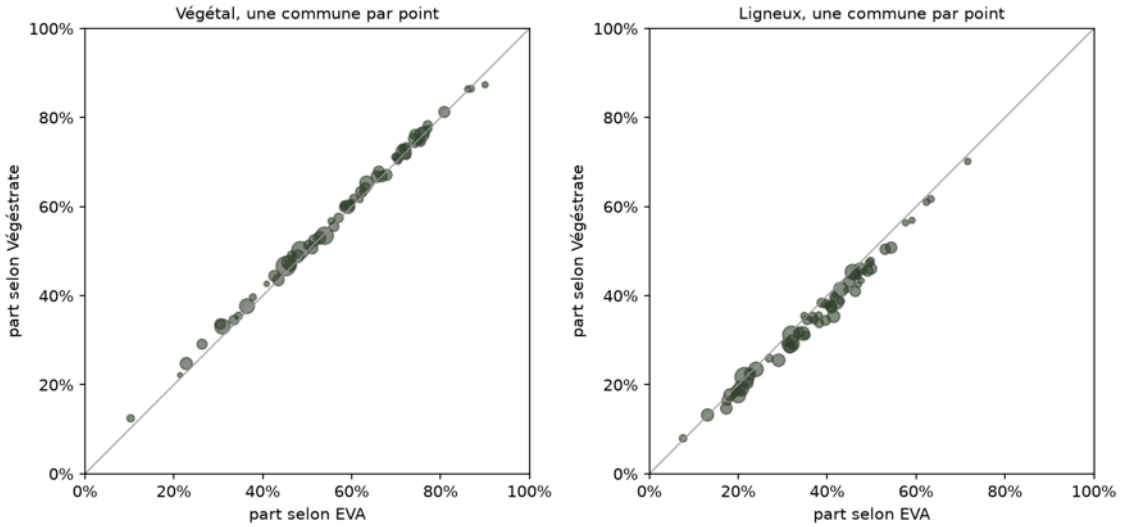
Sur toute la végétation Végéstrate, 1,1 % est en taches de moins de 100 m². Dans les zones EVA non végétales, cette part monte à 37,3 % pour la végétation ligneuse : arbres isolés et alignements qu’EVA n’a pas à cartographier.



Petites taches

Couverture par commune

Parts calculées sur la surface comparable de chaque commune, une commune par point, taille proportionnelle à la surface. Détail dans `couverture_communes.csv` et `couverture_iris.csv`.



Couverture par commune

La correspondance est globalement la même pour toutes les communes : aucune ne se démarque nettement.

Limites

- Accord n’est pas exactitude : sans vérité terrain, un désaccord peut venir de l’une, de l’autre ou des deux sources.
- EVA cartographie la couverture dominante de polygones d’au moins 100 m² avec un taux de couverture de 60 à 75 % : jusqu’à un quart d’un polygone « strate arborée » peut légitimement être

autre chose, et ses contours englobent la végétation. Au pixel, ces écarts admis par la définition d'EVA comptent comme des désaccords ; la composition des polygones EVA donne l'accord à l'échelle d'EVA.
- Le modèle de hauteur de canopée (nDSM Végéstrate) est arrondi au mètre et ne porte que sur les pixels que Végéstrate classe en végétation : là où EVA voit un arbre et Végéstrate rien, aucune hauteur ne permet de distinguer un écart de définition d'un écart de détection.

Reproduire

Ce document est généré à partir de `audit/methodology_template.md` et de `audit/config.yaml` ; les chiffres proviennent des tableaux CSV produits dans le dossier de sortie de l'audit.

```
python -m audit.run --config audit/config.yaml
```

Étapes : `audit.prepare` (zone exclue, export IRIS et LCZ depuis la base IA.rbre), `audit.compute` (comptages pixel par blocs), `audit.examples` (rasters des lieux d'exemple), `audit.report` (tableaux et ce document), `audit.figures` (figures), `audit.publish` (PDF et copies vers la documentation et le site). `python -m audit.selfcheck` vérifie les fonctions de calcul sur des cas simples.

Tableaux complets dans le dossier de sortie :

- `indicateurs_accord.csv` : accord et désaccord pour tous les niveaux et toutes les variantes de seuils ;
- `matrices.csv` : matrices en km², parts par ligne et par colonne ;
- `regles_couverture.csv`, `polygones_composition.csv.gz` : couverture attendue retrouvée par classe et composition de chaque polygone ;
- `distance_bord.csv`, `petites_taches.csv` ;
- `couverture_communes.csv`, `couverture_iris.csv`, `accord_lcz.csv` ;
- `zone_exclue.csv`, `exemples_stats.csv` et les rasters `exemples/<lieu>_*.tif` avec leurs styles QGIS `.qml`.