

Apport des réseaux de neurones convolutifs à la cartographie de l'occupation du sol : *Cas d'usage sur le Grand Avignon*

Nicolas Massot

2024-2025

Sous la direction de M. Mounir Redjimi & M. Didier Josselin

Master Géomatique et Conduite de Projets Territoriaux

Résumé

Ce mémoire trouve son origine dans un double contexte. D'abord, il y a une pression à la production de données d'occupation du sol suite aux récentes actions législatives, avec l'objectif Zéro Artificialisation Nette en tête. Deuxièmement, les données en libre accès sur le marché sont perfectibles sur certains points de vue, qui sont développés et justifiés, notamment en ce qui a trait à l'agilité, la précision et la résolution spatiale et temporelle. C'est pourquoi, cette étude a pour ambition de répondre à la problématique suivante : *"Dans quelle mesure les apports de l'apprentissage profond à la cartographie de l'occupation du sol permettent-ils de monitorer les territoires en adéquation avec les besoins métiers des acteurs publics ?"*. La méthode originale conçue dans le cadre de ce mémoire mobilise les réseaux de neurones convolutifs et a été développée pour répondre aux besoins opérationnels des collectivités territoriales, comme le Grand Avignon, partenaire de ce mémoire. Cette chaîne de traitements retourne une segmentation de l'occupation du sol particulièrement précise à un mètre de résolution, à six classes, avec un score F1 moyen de 0.91 et un mIoU de 0.77. Finalement, cette méthode a pour vocation d'être utilisée par les acteurs concernés, et ce faisant, une réflexion a été menée sur les modalités d'utilisation opérationnelle, en plus de l'avoir conçue dans un objectif de reproductibilité des résultats.

Mots clés : Zéro artificialisation nette, Occupation du sol, Méthode territorialisée, Adéquation besoins métiers, Apprentissage profond

Abstract

The background to this study is twofold. Firstly, there is pressure to produce land-cover data following recent legislative actions, with the objective of Zero Net Artificialization in mind. Secondly, open-access data on the market can be improved in certain aspects, which are developed and justified, notably in terms of agility, accuracy and spatial and temporal resolution. The aim of this study is therefore to answer the following question : "To what extent do the contributions of deep learning to land-cover mapping make it possible to monitor territories in line with the professionals requirements of public administrations ? The original method developed for this brief mobilizes convolutional neural networks, and has been modeled to meet the operational needs of local authorities such as Grand Avignon, the partner in this study. This processing chain returns a land cover segmentation at one-meter resolution, with six classes, an average F1 score of 0.91 and a mIoU of 0.77. Finally, this method is intended to be used by stakeholders, and in so doing, thought has been given to the modalities of its operational use, in addition to having designed it with a view to the reproducibility of results.

Key words : Zero net artificialization, Land cover, Territorialized method, Matching professionals requirements, Deep learning

Apport des réseaux de neurones convolutifs à la cartographie de l'occupation du sol :

Cas d'usage sur le Grand Avignon

Étudiant : Nicolas Massot

Directeurs de mémoire : M. Mounir Redjimi & M. Didier Josselin

Année universitaire : 2024-2025

Je soussigné Nicolas Massot, certifie que le contenu de ce mémoire est le résultat de mon travail personnel. Je certifie également que toutes les données, tous les raisonnements et toutes les conclusions empruntés à la littérature sont soit exactement recopiés et placés entre guillemets dans le texte, soit spécialement indiqués et référencés dans une liste bibliographique en fin de volume. Je certifie enfin que ce document, en totalité ou pour partie, n'a pas servi antérieurement à d'autres évaluations, et n'a jamais été publié. Les propos indiqués n'engagent que l'auteur.

Table des matières

Résumé	3
Remerciements	8
Acronymes et sigles	10
Chapitre 1 : Introduction	12
Chapitre 2 : Dispositifs réglementaires pour la sobriété foncière	17
2.1 Le millefeuille administratif pour la protection des ENAF	18
2.2 La stratégie ZAN pour une gestion économe du foncier	20
2.2.1 L'application de l'objectif ZAN est débattue	22
2.3 Consommation d'espaces ou artificialisation, quelles définitions donner ?	25
2.4 Le rapport triennal pour la mesure de la consommation d'espaces	27
2.5 Conclusion	29
Chapitre 3 : État de l'art des données d'occupation du sol	30
3.1 Couches d'occupation du sol continues	31
3.1.1 Analyse qualitative	31
3.1.1.1 Occupation du sol à grande échelle (OCS GE)	31
3.1.1.2 CoSIA	35
3.1.1.3 Corine Land Cover	40
3.1.1.4 ESA World Cover	42
3.1.1.5 Occupation des Sols Opérationnelle (OSO) - Pôle Theia	42
3.1.1.6 ESRI Land Cover	44
3.1.1.7 OpenStreetMap	45
3.1.2 Analyse quantitative de la précision des couches d'occupation du sol continues	46
3.1.2.1 Méthodologie de l'analyse de la précision	46
3.1.2.2 Résultats de l'analyse de la précision	51

3.2 Données contextuelles discontinues	58
3.2.0.1 LiDAR HD IGN	58
3.2.0.2 BD Topo	59
3.2.0.3 Registre Parcellaire Graphique	61
3.2.0.4 BD Forêt	62
3.3 Conclusion	65
3.3.1 Des données de qualité intermédiaire	65
3.3.2 Vers un outil de suivi de l'occupation du sol territorialisé	67
 Chapitre 4 : Méthode pour la création de cartes d'occupation du sol par réseaux de neurones convolutifs	 69
4.1 Réseaux de neurones convolutifs et U-Net : définitions et concepts	70
4.1.1 De la classification aux réseaux de neurones	70
4.1.2 Le modèle U-Net	72
4.2 Présentation de la chaîne de traitements	79
4.3 Présentation et méthode pour la réalisation du masque d'entraînement	82
4.3.1 Pistes de recherche pour la création du masque d'entraînement	83
4.3.1.1 Agrégation de données déjà existantes	83
4.3.1.2 Utilisation de Segment Anything Model 2	84
4.3.1.3 Occupation du sol multi-sources	86
4.3.2 Méthode adoptée	88
4.3.3 Mesure de la précision du masque d'entraînement	92
4.4 Présentation des données pour les images d'entraînement	96
4.5 Études préalables	100
4.5.1 Signature spectrale et séparabilité des classes avec la distance de Jeffries-Matusita	100
4.5.2 Étude et impact de la colinéarité	103
4.6 Pré-traitements et phase inférentielle	106
4.7 Post-traitement	109
 Chapitre 5 : Évaluation des résultats de la classification	 110

5.1 Cartographie du résultat	111
5.2 Conclusion de l'analyse des résultats	118
5.3 Axes d'amélioration et pistes de recherche	119
Chapitre 6 : Conclusion	123
6.1 Résultats et hypothèses	124
6.2 Perspectives, reproductibilité et science ouverte	125
6.3 Une mesure universelle est-elle réalisable ?	126
Bibliographie	128
Table des figures	131
Liste des tableaux	134
Annexes	135
.1 Règles de classification hiérarchique pour l'occupation du sol (méthode rejetée)	135
.2 Hyperparamètres	137
.3 Configuration	137
.4 Comparatif avec CoSIA	138

Remerciements

En premier lieu, je souhaite remercier M. Didier Josselin, co-directeur de ce mémoire, qui m'a apporté ses connaissances et toute sa passion pour les "méthodes". Il a affiné mon regard sur l'adéquation thématique-méthode, et cette vision a permis de me débrider dans la réalisation de ce mémoire, et je l'espère, pour la suite de mon parcours en géographie.

Je souhaite également adresser un grand merci à Samy Khelil, qui pendant nos visios un peu trop tardives, a su me conseiller lors de la conception de certaines parties de la méthode, et qui a également porté un regard complémentaire à celle-ci.

Je remercie aussi le Grand Avignon et M. Gourmelon pour l'intérêt porté au projet, ainsi que pour m'avoir donné l'opportunité de travailler sur un sujet concret, ce qui me tient à cœur. De même, merci à Boris Wattrelos pour m'avoir apporté son regard sur les produits de l'IGN, ainsi que son analyse sur la manière avec laquelle ces derniers répondent aux besoins contemporains.

Ce mémoire s'est également construit en dehors de l'université. C'est pourquoi je tiens à remercier ma famille : à ma mère pour son attention et sa bienveillance, à mon père pour ses conseils avisés et nos confrontations d'idées, et à mon frère pour m'avoir tant inspiré. Je les remercie, eux et Alain Massot, de m'avoir offert l'opportunité d'étudier à l'Université Laval, une expérience qui m'a permis de prendre de la hauteur sur la vision et le rôle que doit occuper la géomatique. Merci également à Gaëlle Bouvard, Daniel Bélanger et Alain Vignal pour leurs remarques éclairées et leurs regards extérieurs.

Merci à mes amis, surtout ceux de la promo, pour avoir rythmé la rédaction de ce mémoire de rires et de blagues en tout genre et de soirées animées.

Je souhaite conclure ces remerciements en exprimant ma plus sincère gratitude à Monsieur Redjimi. Retenons de lui son opiniâtreté et son abnégation au travail, ainsi que son engagement envers les étudiants et envers moi-même, dans l'accomplissement de ce mémoire. Sans son soutien, ce travail n'aurait jamais été réalisé. Au-delà d'avoir été le directeur de ce mémoire, il a été celui qui a initié mon intérêt pour les sciences géographiques ; intérêt qui n'a fait que grandir depuis.

*À mes parents qui m'ont permis de faire des études,
À Monsieur Redjimi qui m'a tant appris,
Merci.*

Acronymes et sigles

ADAM	Adaptive Moment Estimation
ALUR	Accès au Logement et un Urbanisme Rénové
CEREMA	Centre d'études et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement
CGCT	Code général des collectivités territoriales
CIFOR	Center for International Forestry Research
CLC	Corine Land Cover
COS	Coefficient d'occupation des sols
CoSIA	Couverture du Sol par Intelligence Artificielle
DGFiP	Direction générale des Finances publiques
ENAF	Espace naturel, agricole et forestier
EPCI	Établissements publics de coopération intercommunale
ERC	Éviter, réduire, compenser
ESA	European Space Agency
FAO	Organisation pour l'alimentation et l'agriculture
FLAIR	French Land cover from Aerospace ImageRy
GPS	Global Positioning System
IDM	Inverse Difference Moment
IGN	Institut national de l'information géographique et forestière
INRAE	Institut national de recherche pour l'agriculture, l'alimentation et l'environnement
IoU	Intersection over Union
LiDAR	Light Detection And Ranging
MNH	Modèle numérique de hauteur
MNS	Modèle numérique de surface
MNT	Modèle numérique de terrain
NDGI	Normalized Difference Glacier Index
NDVI	Indice de Végétation par Différence Normalisée
NDWI	Normalized Difference Water Index
OBIA	Oriented Based Image Analysis
OCS GE	Occupation du sol à grande échelle
OSM	OpenStreetMap
PAC	Politique agricole commune
PACA	Provence-Alpes-Côte d'Azur
PCI	Plan cadastral informatisé
PLU	Plan local d'urbanisme

POS	Plan d'occupation des sols
PSI	Patch Shape Index
RPG	Registre parcellaire graphique
SAM	Segment Anything Model
SCoT	Schéma de cohérence territoriale
SRADDET	Schéma Régional d'Aménagement, de Développement Durable et d'Égalité des Territoires
TRACE	Trajectoire de Réduction de l'Artificialisation Concertée avec les Élus locaux
UMC	Unité minimale de collecte
ZAN	Zéro artificialisation nette

Chapitre 1 :

Introduction

Ce chapitre introductif a pour vocation de poser le cadre général du mémoire, en contextualisant les enjeux liés à la consommation d'espaces et à l'artificialisation des sols. Il s'appuie sur les évolutions récentes du corpus législatif pour identifier les besoins méthodologiques induits par ces dernières. Il interroge également la nécessité d'une adéquation entre les méthodes mobilisées et les besoins opérationnels des acteurs publics. Enfin, il expose la problématique de recherche, précise les objectifs et formule les hypothèses qui orienteront l'analyse.

La lutte contre l'étalement urbain est devenue l'un des principaux mots d'ordre des politiques d'aménagement. En 2023, 20 000 hectares ont été consommés à l'échelle nationale. Ainsi, c'est l'équivalent de 18 % de la superficie du Vaucluse qui chaque année se voit bétonnée. Derrière ce chiffre important, se cache une dynamique de fond : le rythme de consommation d'espaces a baissé d'un tiers entre 2009 (30 000 ha) et 2023, avec un point d'inflexion visible dès 2010 (Delattre et Lecourt 2024).

Au regard des enjeux environnementaux, sociaux et économiques, la limitation de l'étalement urbain est un sujet qui, depuis 25 ans, fait l'objet de nombreux textes législatifs. En effet, conscients des implications, les pouvoirs publics ont mis en place un corpus de lois qui traite de la consommation d'espaces ou de l'artificialisation. Le point d'orgue de ce phénomène d'environnementalisation des dispositifs réglementaires qui ont trait à la réduction de l'étalement urbain est la loi Climat et résilience (« Loi n° 2021-1104 du 22 août 2021 portant lutte contre le dérèglement climatique et renforcement de la résilience face à ses effets » 2021). Cette loi de rupture introduit l'objectif Zéro Artificialisation Nette (ZAN). Le texte initial¹ vise à diviser par deux le rythme de consommation d'espaces sur la période 2021-2031, par rapport à la période 2011-2021. Ensuite, le solde entre les surfaces artificialisées et renaturées devra être nul sur la période 2031-2050.

La consommation d'espaces et l'artificialisation, sont deux notions distinctes, mais complémentaires, qui sont définies par le décret n° 2023-1096 du 27 novembre 2023 relatif à l'évaluation et au suivi de l'artificialisation des sols. Celui-ci définit la consommation d'espaces comme étant la "création ou extension effective d'espaces urbanisés sur le territoire concerné". Quant à elle, l'artificialisation est "l'altération durable de tout ou partie des fonctions écologiques d'un sol, en particulier de ses fonctions biologiques, hydriques et climatiques, ainsi que de son potentiel agronomique par son occupation ou son usage". Ainsi, la consommation d'espaces est une mesure connue des élus locaux. Elle est toutefois méthodologiquement partielle et imprécise, puisqu'elle est mesurée avec les fichiers fonciers. À l'inverse, la définition de l'artificialisation tend vers une exhaustivité, mais elle laisse une certaine place à l'interprétation.

De ce fait, la mesure de la consommation d'espaces ou de l'artificialisation est un problème principalement méthodologique, qui se doit de répondre aux besoins d'aide à la décision des acteurs. Ces mesures doivent se baser sur des définitions claires et non-ambigües qui font foi. La chaîne de pensée pour établir une méthode de mesure de la consommation d'espaces ou d'artificialisation doit débuter par l'identification d'un besoin métier en lien avec les enjeux identifiés. Ensuite, il faut définir précisément la consommation d'espaces et l'artificialisation, et

1. Qui a évolué, comme nous le verrons plus tard.

enfin, établir la méthode géomatique en lien direct avec la définition et les besoins. Ce processus n'est pas linéaire et nécessite plusieurs itérations à travers les trois étapes mentionnées afin de les améliorer mutuellement.

Étant donné que la mesure de la consommation d'espaces passée détermine les droits à construire à l'avenir, il est crucial de disposer de données pertinentes et harmonisées à l'échelle nationale. Cependant, à ce jour, aucune base de données ne permet d'évaluer finement l'évolution de l'occupation du sol sur l'ensemble du territoire, à l'exception des données cadastrales produites par la DGFiP, lesquelles ne peuvent que mesurer la consommation d'espaces (Delattre et Lecourt 2024). De plus, la loi Climat et Résilience laisse une certaine flexibilité quant à la mesure de la consommation d'espaces, en autorisant les collectivités à se doter d'un outil de mesure de l'occupation du sol territorialisé. Toutefois, depuis quelques années, de nombreux acteurs nationaux ou internationaux (IGN, ESA, ...) produisent des données sur l'occupation et l'usage du sol. La pertinence de ces données peut être décrite avec de nombreux facteurs qualitatifs et quantitatifs. D'un point de vue qualitatif, on peut entre autres citer la résolution spatiale, temporelle et sémantique², ainsi que la fraîcheur de la donnée³. Sur l'aspect quantitatif, de nombreux indicateurs de mesure de la précision peuvent être mobilisés. Ces deux aspects sont centraux pour une mesure fiable. En effet, comment monitorer un territoire avec des données ayant une résolution temporelle et une fraîcheur de plusieurs années ? De plus, l'évolution de la consommation d'espaces se mesure en centièmes de pourcentage chaque année. Sur la période 2020-2023, la région PACA a consommé 0.125 % de sa superficie (Cerema 2024). Ainsi, comment être sûr que cette évolution sur trois ans n'est pas due à un manque de précision de la donnée ?

Au-delà de la connaissance précise de l'occupation du sol, ce mémoire trouve son origine dans un double contexte. D'une part, une demande croissante de données impulsée par les textes législatifs, et d'autre part, des données disponibles sur le marché qui sont perfectibles, comme nous le verrons plus tard. Ainsi, après avoir fait état des dispositifs réglementaires qui ont trait à la régulation de la consommation d'espaces et de l'artificialisation, nous réaliserons un état de l'art des données d'occupation du sol. Ensuite, nous proposerons une chaîne de traitements originale et innovante basée sur les réseaux de neurones convolutifs, qui a pour ambition de combler les lacunes identifiées dans l'état de l'art. À la suite de quoi, un protocole de validation de la précision de la donnée sera mené. Nous finirons par prendre du recul en évoquant les questions liées à la reproductibilité et à l'utilisation opérationnelle,

2. La résolution sémantique désigne le niveau de détail de la composition des classes thématiques utilisées pour décrire l'occupation du sol.

3. Délais entre la date d'acquisition et de publication des données.

avant d'évoquer les axes d'améliorations de la méthode.

Ainsi, ce mémoire aspire à répondre à la problématique suivante :

Dans quelle mesure les apports de l'apprentissage profond à la cartographie de l'occupation du sol permettent-ils de monitorer les territoires en adéquation avec les besoins métiers des acteurs publics ?

Pour répondre à la problématique, nous passerons par les objectifs intermédiaires suivants :

- Objectif 1 : Faire état des besoins métiers pour la mesure de la consommation d'espaces et de l'artificialisation en lien avec les textes législatifs.
- Objectif 2 : Faire état du marché des données d'occupation du sol pour identifier les éventuels axes d'améliorations.
- Objectif 3 : Concevoir une chaîne de traitements précise et reproductible permettant de segmenter l'occupation du sol, l'inférer sur la zone d'étude et mesurer la précision.
- Objectif 4 : Construire une réflexion et porter un regard critique sur l'usage opérationnel de la méthode proposée.

Ainsi, plusieurs hypothèses (H_i) sont formulées. Celles-ci seront vérifiées puis discutées en fonction des résultats obtenus :

- H_1 : Les données d'occupation du sol et les données contextuelles discontinues actuelles sont suffisantes pour mesurer la consommation d'espaces ou l'artificialisation.
- H_2 : La précision est une notion subjective, relative aux besoins du commanditaire.
- H_3 : La colinéarité des images d'entraînement a un impact négatif à la performance du modèle.
- H_4 : La méthode par apprentissage profond incarne une plus-value dans la cartographie de l'occupation du sol à fine échelle spatiale, temporelle, et sur la reproductibilité des résultats sur différentes zones et millésimes.

Enfin, ce mémoire a été mené avec la communauté d'agglomération du Grand Avignon qui, comme toutes les collectivités territoriales, est sujette à l'application de l'objectif ZAN. Ainsi, disposer d'un outil de suivi de l'occupation du sol territorialisé à fine échelle spatiale et temporelle, permet d'établir un diagnostic plus fin des dynamiques territoriales. De ce fait, ce partenariat permet d'avoir un cas d'usage concret. Celui-ci se matérialise à travers les discussions qui permettent

l'identification des besoins métiers, et l'application d'une méthode pour y répondre. Ainsi, la chaîne de traitements sera inférée sur les 16 communes que comporte le Grand Avignon.



Figure 1. Carte de localisation du Grand Avignon et ses communes.
N. Massot, 2025

Chapitre 2 :

Dispositifs réglementaires pour la sobriété foncière

Ce chapitre fait état des documents législatifs qui ont trait à la lutte contre la consommation d'espaces, l'artificialisation et à la protection des espaces naturels, agricoles et forestiers. Nous traiterons plus en détails de la question de l'objectif Zéro Artificialisation Nette, de ses évolutions, ainsi que des éléments de définition inhérents. Ensuite, nous parlerons de deux notions intrinsèques à l'objectif ZAN : la "consommation d'espaces" et "artificialisation". Nous finirons par évoquer le rapport triennal de la mesure de la consommation d'espaces.

2.1 Le millefeuille administratif pour la protection des ENAF

La mise en œuvre des politiques publiques fait de plus en plus appel à des processus de gouvernance multi-niveaux, de recours à un mode de gouverner à distance (Epstein 2006) et de gouvernement par les instruments (Lascombes et Le Galès 2005)). L'artificialisation des sols est un sujet à enjeux majeurs qui fait l'objet de nombreux textes législatifs. Conscients des implications, les pouvoirs publics ont mis en place un corpus de lois qui traitent de la question de la périurbanisation, de la consommation d'espaces ou de l'artificialisation. On assiste depuis quelques décennies à un phénomène d'environnementalisation des documents d'urbanisme dans lequel ces derniers considèrent l'espace non pas comme une ressource à exploiter, mais comme une aménité à protéger, dans la mesure où celle-ci est source d'externalités positives et de services écosystémiques.

En effet, dès 1985, la loi Montagne stipule que l'urbanisation doit se faire en continuité avec les bourgs, villages et hameaux existants, afin d'éviter le mitage et la dispersion des constructions. Par ailleurs, elle vise à protéger le foncier lié aux activités agricoles, pastorales et forestières. Entrée en vigueur l'année suivante, la loi Littoral vise à préserver l'équilibre écologique du littoral en mettant en place une protection graduée en fonction de la proximité avec le rivage. Il est notamment interdit de construire en dehors des espaces déjà urbanisés sur une bande de 100 mètres à compter de la limite haute du rivage. Bien que limitées géographiquement⁴, ces deux lois précurseures ont posé les bases de la politique de lutte contre l'artificialisation des sols et la protection des milieux.

Ensuite, la loi Solidarité et Renouvellement Urbains (2000) est la loi fondatrice en ce qui a trait à la lutte contre l'artificialisation. Elle vise à réduire la fracture sociospatiale en ville, la pénurie de logements sociaux ainsi que le phénomène de périurbanisation en densifiant "de manière raisonnée" les espaces déjà urbanisés (Géoconfluences 2024). D'abord, elle amène la caducité des Plans d'Occupation des Sols (POS) qui, de par leur structure, étaient favorables à l'étalement urbain. Ensuite, elle initialise la hiérarchie des normes en instaurant des outils de planification tels que les Schémas de Cohérence Territoriale (SCoT) et les Plans Locaux d'Urbanisme (PLU), qui eux même intègrent des objectifs de modération de la consommation d'Espaces Naturels, Agricoles et Forestiers (ENAF) à l'échelle du bassin de vie et de la commune. En outre, l'article 55 de cette loi vise à récréer un équilibre social dans les territoires en disposant d'un nombre minimum de logements sociaux, proportionnel à leur parc résidentiel. Les communes de plus de 3 500 habitants⁵ intégrées dans

4. Pour loi Montagne : 4 173 communes sont totalement concernées, et 156 soumises à une obligation partielle.
Pour loi Littoral : 1 212 communes concernées.

5. Ou 1 500 habitants en Île-de-France

des agglomérations ou intercommunalités de plus de 50 000 habitants comprenant au moins une commune de plus de 15 000 habitants, doivent avoir au moins 25 % de logements sociaux parmi leurs résidences principales (Transition écologique 2022). À l'échelle nationale, 60 % des communes ne respectent pas ces quotas (Uhart 2023). Dans le bassin de vie d'Avignon, sur les 19 communes concernées, seule Avignon a atteint cet objectif. Toutefois, la loi SRU a tout de même permis de largement favoriser le développement de logements sociaux, et d'outiller les collectivités territoriales dans la planification en créant les PLU et les SCoT.

Adoptées en 2009-2010, les lois Grenelle I & II ont également introduit des mesures pour lutter contre l'artificialisation à travers la mise en place d'une gouvernance écologique pour gérer les impacts environnementaux. Outre l'instauration du concept de "trames vertes et bleues"⁶, celle-ci renforce l'importance du SCoT en rendant obligatoire le chiffrement de la consommation d'ENAF.

Par ailleurs, la loi pour l'Accès au Logement et un Urbanisme Rénové (ALUR) publiée au Journal Officiel en 2014 a introduit plusieurs mesures pour lutter contre l'étalement urbain. Celle-ci intègre l'analyse de la consommation d'espaces dans les PLU et les SCoT. Toutefois, au moment de sa promulgation, il n'existe pas de définition précise de la notion de consommation d'espace, et aucun dispositif coercitif n'avait été mis en place. En outre, la loi rend caduc le concept de Coefficient d'Occupation des Sols (COS) dans les plans locaux d'urbanisme⁷. Avant la loi ALUR, les communes avaient la possibilité de mettre des COS maximaux, ce qui revenait à limiter la densité constructible des parcelles. Ainsi, même si à l'échelle communale les besoins en logement étaient élevés, les constructions devaient rester limitées en hauteur ou en volume. De plus, la suppression des COS favorise le renouvellement urbain. En effet, elle permet de construire "la ville sur la ville", en créant de nouveaux logements dans les centres urbains, en transformant d'anciens bâtiments ou en requalifiant des friches. Par ailleurs, le texte vise à permettre un "juste dimensionnement" de la réserve foncière et prévoit que les zones classées 2AU⁸ depuis plus de neuf ans et qui ne font l'objet d'aucun projet d'aménagement, seront reclasées en zones naturelles (zone N). Ce principe a pour objectif de réduire la réserve foncière "théorique" en contrecarrant le surdimensionnement des ouvertures à l'urbanisation en périphérie de l'enveloppe urbaine, et ce, sans que cela soit corrélé aux besoins concrets d'urbanisation (Lenormand 2013).

De plus, la loi pour la reconquête de la biodiversité, de la nature et des paysages

6. Réseaux de corridors écologiques destinés à préserver et à restaurer les continuités écologiques et réduire la fragmentation des habitats naturels.

7. Les COS désignent le rapport entre la surface habitable et la surface de la parcelle. Si la valeur est supérieure à 1, on utilise plus que la surface au sol de notre parcelle pour faire des mètres carrés habitables.

8. Nomenclature du PLU. Zones à urbaniser pas encore constructibles et distantes des réseaux de voirie, eau, assainissement et électricité.

(2016) a plusieurs objectifs en ce qui a trait à la réduction de la consommation d'espaces. D'abord, elle apporte de nouvelles dispositions applicables aux centres commerciaux comme la végétalisation des toitures, l'installation de production d'énergies renouvelables ou encore la lutte contre l'artificialisation des sols dans les aires de parking. Elle met également en application le concept de la séquence "éviter, réduire, compenser" (ERC), qui vise à minimiser les impacts environnementaux des projets d'urbanisme. Cette séquence implique d'abord d'éviter les externalités négatives en choisissant par exemple des sites qui ne sont pas des zones écologiques sensibles, puis de réduire les impacts inévitables par l'intégration d'espaces verts et de solutions écologiques, et enfin de compenser les impacts résiduels par la création de nouvelles zones naturelles ou la restauration d'habitats dégradés, ce qui permet de tendre vers un développement urbain durable ([« Loi n° 2016-1087 du 8 août 2016 pour la reconquête de la biodiversité, de la nature et des paysages » s. d.](#)).

Enfin, mis en place par l'article 10 de la loi relative à la nouvelle organisation territoriale de la République (NOTRe), le Schéma Régional d'Aménagement, de Développement Durable et d'Égalité des Territoires (SRADDET) est élaboré à l'échelle régionale. Il fixe les grandes orientations en matière d'aménagement du territoire, de développement durable et de cohésion territoriale.

2.2 La stratégie ZAN pour une gestion économe du foncier

Le SRADDET est intégrateur de l'objectif "Zéro Artificialisation Nette" (ZAN) établi par la loi Climat et Résilience de 2021. Cette loi stipule que l'artificialisation des sols doit être compensée par des actions de renaturation, avec des proportions croissantes, jusqu'à atteindre une compensation totale d'ici à 2050 dans le but de protéger les ENAF. Dans sa version initiale, le ZAN imposait que le "budget artificialisation"⁹ de la période 2021–2031 devait être divisé par deux par rapport à la période 2011–2021. Ensuite, celui-ci devait atteindre le zéro net en 2050. C'est-à-dire que chaque mètre carré construit devra être compensé par la renaturation d'un mètre carré ailleurs (cf. séquence ERC).

Si le SRADDET est intégrateur du ZAN, le "budget artificialisation" doit être en compatibilité avec le celui-ci. Il en va de même pour le SCoT qui doit être en compatibilité avec les règles du SRADDET ainsi que le PLU(i) avec le SCoT. Cette manière de fonctionner donne plus de pouvoir aux collectivités territoriales et EPCI, ce qui va dans le sens de la décentralisation (cf. figure 2). Les objectifs régionaux de réduction du rythme d'artificialisation des sols sont hétérogènes. Les régions ont adopté des stratégies différentes à ce sujet. Tandis que le Centre-Val de Loire, la Bretagne, les Hauts-de-France et l'Occitanie prennent les devants en étant plus ambitieux que

9. Solde entre la superficie artificialisée, et la superficie renaturée pour un territoire donné.

n'est la loi, d'autres régions comme l'Auvergne-Rhône-Alpes ou la région PACA n'ont pas encore inscrit l'objectif ZAN dans leur SRADDET (cf. table 2).

Territorialisation des objectifs ZAN

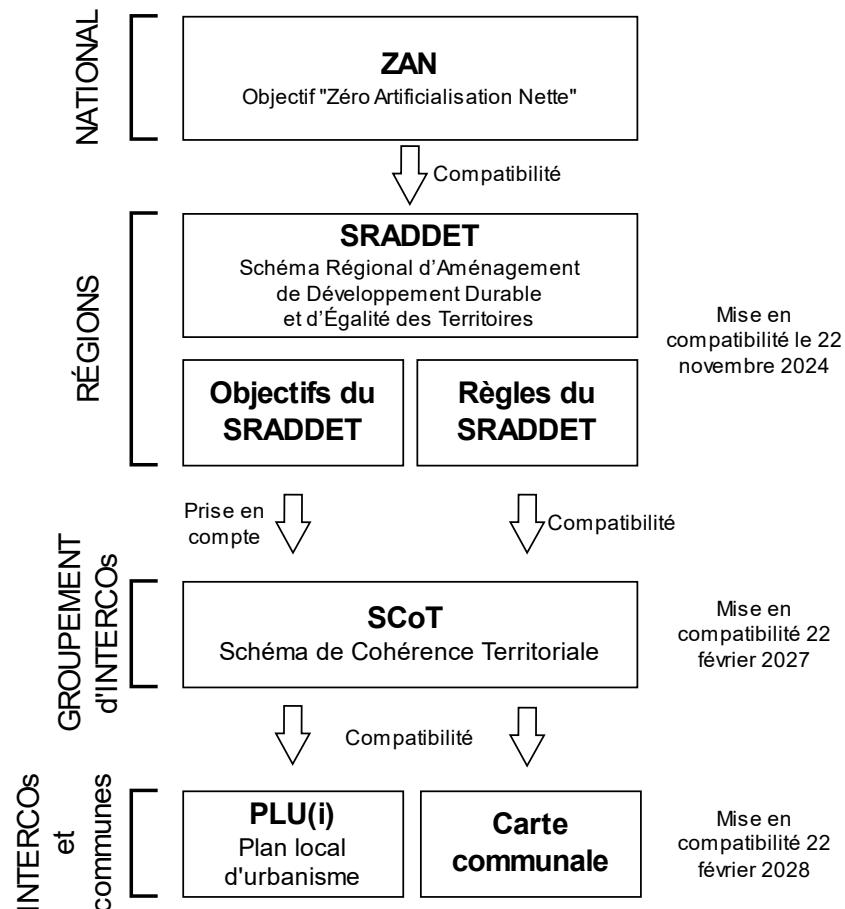


Figure 2. Territorialisation des objectifs ZAN.

Source : La lutte contre l'artificialisation des sols, Notre affaire à tous, 2023, modifié par N. Massot, 2025

	2025	2030	2035	2040	2050
Centre-Val de Loire	50 %		zéro		
Bretagne		50 %	75 %	zéro	
Grand Est		50 %			75 %
Hauts-de-France		divisé par 3		divisé par 4	divisé par 6
Normandie		50 %			zéro
Nouvelle-Aquitaine		50 %			zéro
PACA		50 %			
Occitanie				zéro	
Bourgogne-Franche-Comté			50 %		zéro
Pays de la Loire					zéro
Auvergne-Rhône-Alpes					Pas d'échéancier

Table 2. Objectif de consommation d'espaces des SRADDET
N. Massot, 2025. Source : Doré, G. (2023). *Les SRADDET et l'objectif ZAN en 2023*

2.2.1 L'application de l'objectif ZAN est débattue

La mise en œuvre de la loi Climat et Résilience soulève de nombreux débats entre l'exécutif et les collectivités territoriales. D'abord, certains élus locaux ont pointé des problèmes dans la méthode de mesure de l'artificialisation. Les périodes considérées pour la renaturation, et les délais pour la mise en compatibilité des documents d'urbanisme sont jugés trop courts (architectes 2025). De plus, en réduisant les ouvertures à l'urbanisation, les collectivités locales voient leurs marges de manœuvre en matière de fiscalité locale directe limitées. En effet, si ces dernières veulent augmenter leurs recettes à travers la taxe foncière (dernière survivante des "quatre vieilles"¹⁰), elles ne peuvent plus se fonder sur la base fiscale. Dès lors, leur unique marge de manœuvre devient l'augmentation du taux d'imposition, ce qui revient à augmenter les impôts¹¹. Cela représente un coût politique important pour les élus locaux. En réponse à ces défis, la "loi visant à faciliter la mise en œuvre des objectifs de lutte

10. En fiscalité française, les "quatre vieilles" désignent les quatre anciens impôts locaux principaux (taxe foncière sur les propriétés bâties et non bâties, taxe d'habitation, contribution économique territoriale).

11. La fiscalité locale est basée sur le principe suivant :

$$\text{Recette} = \text{Base} \times \text{Taux d'imposition}$$

Ici, La base correspond à la superficie bâtie tandis que le taux correspond au taux d'imposition.

contre l'artificialisation des sols et à renforcer l'accompagnement des élus locaux", dite ZAN 2, a été adoptée en juillet 2023. Cette loi clarifie le référentiel national pour mesurer l'artificialisation¹² et ses définitions (cf. section 2.3). Elle introduit également le principe de "garantie universelle", permettant à chaque commune, sans condition de densité, d'artificialiser un hectare pour la période 2021-2031. Cette mesure venait pallier ce qui était considéré comme une incohérence de la part des maires¹³, qui, n'ayant pas (beaucoup) artificialisé sur la période 2011-2021, considéraient leur budget artificialisation comme trop faible, voire parfois nul. Plus récemment, la Trajectoire de Réduction de l'Artificialisation Concertée avec les Élus locaux (TRACE) est à l'heure actuelle, une proposition de loi qui vise à simplifier de nouveau la mise en application du ZAN. Cette proposition de loi intègre deux éléments phares. D'une part, elle supprime l'objectif intermédiaire de réduction de 50 % de l'artificialisation des sols d'ici à 2031, en le remplaçant par un objectif à horizon 2034, fixé par les régions. D'autre part, celle-ci ré-allonge à 10 ans le délai de mise en compatibilité des documents d'urbanisme (SRADDET, SCoT, PLU(i)) (cf. Figure 3).

12. Il s'agit de l'Occupation des Sols à Grande Échelle, produit poussé par l'IGN, le CEREMA et l'INRAE (cf. section 3.1.1.1).

13. Principalement issus des communes rurales.

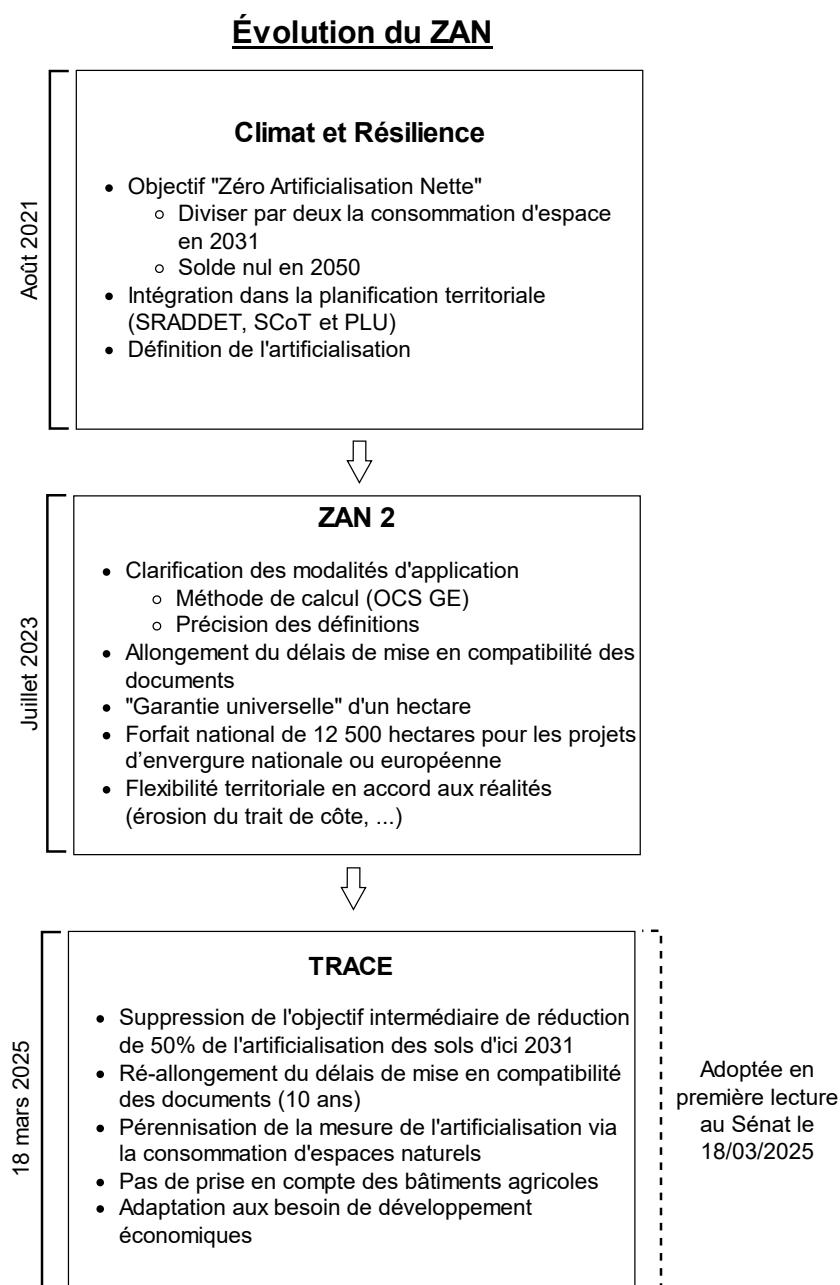


Figure 3. Évolution du ZAN N. Massot, 2025

2.3 Consommation d'espaces ou artificialisation, quelles définitions donner ?

La législation sur la réduction de ce qu'on appelle couramment "l'artificialisation" nécessite d'avoir des métriques fiables, qui elles-mêmes doivent se baser sur des définitions claires et non-ambigües. En ce sens, il est important de différencier deux notions proches, mais qui ne doivent pas être confondues :

- La consommation d'espaces
- L'artificialisation

La loi Climat et Résilience apporte des définitions à ces deux termes que nous pouvons synthétiser de cette manière :

	Consommation d'espaces	Artificialisation
Définition	Création ou extension effective d'espaces urbanisés sur le territoire concerné	Altération durable de tout ou partie des fonctions écologiques d'un sol, en particulier de ses fonctions biologiques, hydriques et climatiques, ainsi que de son potentiel agronomique par son occupation ou son usage
Restrictivité	Faible	Élevé
Période	2031 ou 2034 (si TRACE)	2050
Temporalité	Mesure connue des élus	Mesure d'avenir
Donnée de référence	Fichiers fonciers	OCS GE (occupation et usage)
Précision	-	+

Table 3. Définitions de la consommation d'espaces et de l'artificialisation.

N. Massot, 2025. Source : Décret n° 2023-1096 du 27 novembre 2023 relatif à l'évaluation et au suivi de l'artificialisation des sols

Cette définition est complétée par le décret n° 2023-1096 du 27 novembre 2023 relatif à l'évaluation et au suivi de l'artificialisation des sols dans lequel on retrouve une catégorisation des sols artificialisés et non-artificialisés (cf. table 4).

Surfaces artificialisées	
1° Surfaces dont les sols sont imperméabilisés en raison du bâti (constructions, aménagements, ouvrages ou installations).	Supérieur ou égal à 50 m ² d'emprise au sol
2° Surfaces dont les sols sont imperméabilisés en raison d'un revêtement (artificiel, asphalté, bétonné, couvert de pavés ou de dalles).	Supérieur ou égal à 2 500 m ² d'emprise au sol ou de terrain
3° Surfaces partiellement ou totalement perméables dont les sols sont stabilisés et compactés ou recouverts de matériaux minéraux, ou dont les sols sont constitués de matériaux composites (couverture hétérogène et artificielle avec un mélange de matériaux non minéraux).	
4° Surfaces à usage résidentiel, de production secondaire ou tertiaire, ou d'infrastructures notamment de transport ou de logistique, dont les sols sont couverts par une végétation herbacée (**).	
5° Surfaces entrant dans les catégories 1° à 4°, qui sont en chantier ou en état d'abandon.	
Surfaces non artificialisées	
6° Surfaces naturelles dont les sols sont soit nus (sable, galets, rochers, pierres ou tout autre matériau minéral, y compris les surfaces d'activités extractives de matériaux en exploitation) soit couverts en permanence d'eau, de neige ou de glace.	Supérieur ou égal à 2 500 m ² d'emprise au sol ou de terrain
7° Surfaces à usage de cultures dont les sols sont soit arables ou végétalisés (agriculture), y compris si ces surfaces sont en friche, soit recouverts d'eau (pêche, aquaculture, saliculture).	
8° Surfaces dont les sols sont végétalisés et à usage sylvicole.	
9° Surfaces dont les sols sont végétalisés et qui constituent un habitat naturel.	
10° Surfaces dont les sols sont végétalisés et qui n'entrent pas dans les catégories précédentes.	

Table 4. Catégories de surfaces et seuils de référence.

N. Massot, 2025

(*) Les infrastructures linéaires sont qualifiées à partir d'une largeur minimale de cinq mètres.

(**) Une surface végétalisée est qualifiée d'herbacée dès lors que moins de vingt-cinq pour cent du couvert végétal est arboré.

Source : Article R. 101-1 du Code de l'Urbanisme

Ces éléments de définition sont centraux, dans la mesure où ils permettent de définir un référentiel commun pour mesurer le solde entre les surfaces artificialisées et renaturées voulu par l'objectif ZAN. De plus, elle est une grille de lecture nationale qui permet l'homogénéisation et la comparaison des mesures d'une collectivité territoriale à une autre. Toutefois, cette définition peut être critiquable selon certains points de vue. En effet, elle est une vision binaire de la question, en opérant une distinction stricte entre surfaces artificialisées et non-artificialisées, sans prendre en compte le gradient d'artificialisation ou la qualité du sol. Par ailleurs, cette liste est déterminante d'un point de vue sectoriel. Par exemple, que statuer sur les jardins individuels, les panneaux photovoltaïques, les serres agricoles ou les carrières ? La question des serres agricoles a été abordée à l'Assemblée Nationale en avril 2023 (Haury 2023). Le fait de ne pas considérer les bâtiments agricoles comme zone artificialisée est présent dans le projet de loi TRACE. Par ailleurs, les catégories de surfaces d'artificialisation (cf. table 4) mentionnent que "Les surfaces à usage résidentiel [...] dont les sols sont couverts par une végétation herbacée" sont considérées comme artificialisées. De ce fait, cela exacerbe les externalités négatives des urbanités en n'encourageant pas les maîtres d'œuvre à aménager des espaces verts, ce qui tend à l'apparition des îlots de chaleur urbains, à la réduction de la faune et de la flore urbaine, de la marchabilité urbaine...

Les définitions de consommation d'espaces et d'artificialisation qui sont la base de tout le processus décisionnel du ZAN sont soumises à des contradictions et des incohérences selon les points de vue et les intérêts de chacun (exécutif, collectivités territoriales, entreprises, ...).

2.4 Le rapport triennal pour la mesure de la consommation d'espaces

"Le maire d'une commune ou le président de l'établissement public de coopération intercommunale doté d'un plan local d'urbanisme, d'un document d'urbanisme en tenant lieu ou d'une carte communale présente au conseil municipal ou à l'assemblée délibérante, au moins une fois tous les trois ans, un rapport relatif à l'artificialisation des sols sur son territoire au cours des années civiles précédentes.", Article L2231-1 du CGCT.

Cet article du Code général des collectivités territoriales issu de la loi Climat et résilience, stipule que les communes ou intercommunalités dotées d'un document d'urbanisme, doivent établir au minimum tous les trois ans un rapport sur le rythme de l'artificialisation des sols et le respect des objectifs de sobriété foncière déclinés au niveau local. Le premier rapport s'effectuera ainsi sur la période 2021-2024. L'article R2231-1 du CGCT précise les indicateurs à mentionner dans ce rapport triennal ¹⁴ :

14. Le rapport peut comporter d'autres indicateurs et données.

- La consommation des espaces naturels, agricoles et forestiers, exprimée en nombre d'hectares, le cas échéant en la différenciant entre ces types d'espaces, et en pourcentage au regard de la superficie du territoire couvert. Sur le même territoire, le rapport peut préciser également la transformation effective d'espaces urbanisés ou construits en espaces naturels, agricoles et forestiers du fait d'une renaturation ;
- Le solde entre les surfaces artificialisées et les surfaces renaturées, telles que définies dans la nomenclature annexée à l'article R. 101-1 du code de l'urbanisme ;
- Les surfaces dont les sols ont été rendus imperméables, au sens des 1° et 2° de la nomenclature annexée à l'article R. 101-1 du code de l'urbanisme ;
- L'évaluation du respect des objectifs de réduction de la consommation d'espaces naturels, agricoles et forestiers et de lutte contre l'artificialisation des sols fixés dans les documents de planification et d'urbanisme. Les documents de planification sont ceux énumérés au III de l'article R. 101-1 du code de l'urbanisme.

De plus, la préfecture du Gard a effectué un travail complémentaire en rédigeant un "Guide à l'attention des maîtres d'ouvrage des documents d'urbanisme et de leurs bureaux d'études prestataires", ce qui inclut le Grand Avignon. Ce guide propose une "méthode visant à homogénéiser les pratiques, pour constituer la base des données nécessaires à l'élaboration d'un document d'urbanisme, répondant aux objectifs de sobriété foncière et de maîtrise de la consommation d'espaces posés par la loi "Climat et Résilience". Alors que l'article L2231-1 du CGCT parle d'artificialisation, le rapport fourni par la Préfecture du Gard veut mesurer la consommation d'espaces.

Selon le document fourni par la Préfecture du Gard, la mesure de consommation d'espaces se ferait en trois temps :

1. Délimitation de l'enveloppe urbaine existante au début de la période considérée

- **Objectif :** Identifier les espaces urbanisés au début de la période considérée.
- Utiliser des photographies aériennes et des indices tels que la densité de l'urbanisation, la continuité du bâti, la structuration par des voies de circulation, et la présence d'équipements collectifs.
- Sur les franges urbaines, ajuster les contours au plus près du bâti existant et en faisant fi du parcellaire.
- Exclure les "vides" de plus de 2 500 m², considérés comme des extensions

générant de la consommation d'espace.

2. Identification des espaces consommés en dehors de cette enveloppe urbaine

- **Objectif :** Repérer les constructions et parcelles consommées en dehors des zones urbanisées.
- Utiliser une couche SIG des bâtiments.
- Ne pas comptabiliser les secteurs d'exploitation de carrières, les bâtiments agricoles isolés, et les champs photovoltaïques réversibles.

3. Calcul de la consommation d'espaces NAF sur la période considérée

- **Objectif :** Quantifier la consommation d'espaces naturels, agricoles et forestiers.
- Calculer la surface des parcelles identifiées comme consommées.

Ainsi, ces éléments permettent de bien saisir les dynamiques spatiales et temporelles de la consommation d'espaces en utilisant des indicateurs pertinents. Toutefois, malgré la présence d'un grand nombre d'acteurs sur le segment de la cartographie de l'occupation du sol, nous avons identifié des axes d'améliorations dans ces dernières (cf. Chapitre 3). Celles-ci doivent permettre la réalisation des analyses nécessaires en se basant sur une méthodologie homogène et reproductible à fine échelle temporelle et spatiale.

Ce rapport triennal déclaratif fait preuve de subjectivité dans l'interprétation, notamment en ce qui a trait à la délimitation de l'enveloppe urbaine, au-delà de la détermination arbitraire du seuil d'exclusion de l'enveloppe urbaine des ENAF dont la superficie est supérieure à 2 500 m² d'emprise au sol.

2.5 Conclusion

Quoi qu'il en soit, que la TRACE soit adoptée ou non, il est probable que cette loi ne soit pas la dernière qui traite de la question du Zéro Artificialisation Nette dont l'objectif est, rappelons-le, fixé à 2050. La situation, les modalités, les métriques ainsi que les définitions pourraient être amenées à évoluer. Toutefois, le millefeuille administratif abordé dans cette partie démontre bien l'intérêt de se doter d'un outil de suivi de l'occupation du sol qui permet d'alimenter les politiques publiques.

Il convient également de souligner que si le cadre légal continue d'évoluer, les efforts méthodologiques déployés pour la cartographie de l'occupation du sol pourraient être fragilisés, voire annulés, par un simple changement de définition ou d'objectif.

Chapitre 3 :

État de l'art des données d'occupation du sol

La partie suivante a pour objectif de faire une étude comparative des différentes couches d'occupation du sol continues en tout point de l'espace sur le territoire du Grand Avignon. Ce benchmark a un double objectif. D'une part, identifier les lacunes dans le marché des données d'occupation du sol, et d'autre part, voir lesquelles sont les plus propices pour réaliser le masque d'entraînement. Nous commençons par analyser ces données d'un point de vue qualitatif, puis nous mesurerons la qualité de ces dernières sur le territoire du Grand Avignon à l'aide de matrices de confusion ainsi que de divers indicateurs de précision. Nous finirons par traiter les données contextuelles discrètes utiles à la réalisation du masque d'entraînement.

3.1 Couches d'occupation du sol continues

3.1.1 Analyse qualitative

3.1.1.1 Occupation du sol à grande échelle (OCS GE)

L'occupation du sol à grande échelle (OCS GE) est la base de données faisant office de référentiel national pour la mesure de l'artificialisation (cf. sous-section 2.2.1). Elle a été commandée par le Ministère de la Transition Écologique à l'IGN, au CEREMA et à l'INRAE. L'objectif de ce produit est de permettre de mesurer et de spatialiser l'artificialisation des sols de manière homogène à l'échelle nationale, ce qui en fait son principal atout. D'abord, en ayant une méthode de calcul appliquée sur tout le territoire national, l'IGN permet une comparaison des données à l'instant T entre les territoires cohérente. De plus, avec l'analyse des séries temporelles, on peut analyser les dynamiques d'artificialisation et de consommation d'espaces inter et intra-territorial d'un millésime à l'autre. Cette approche vient combler un besoin de métrique homogène, à un moment où les collectivités territoriales tendent à intégrer leurs propres définitions de l'artificialisation et leurs propres méthodes de calcul (fichiers fonciers, mode d'occupation du sol, ...).

L'OCS GE est une donnée à double entrée, avec une version "couverture" et une version "usage". La version couverture décrit morphologiquement le terrain en mettant en avant les éléments structurants de ce dernier et en faisant fi de leur fonction. L'usage du sol, pour sa part, renvoie à la dimension économique et fonctionnelle de l'espace. Ainsi, la couverture et l'usage du sol sont deux notions proches, mais distinctes qu'il est nécessaire de disposer pour bien discriminer les surfaces artificialisées de celles qui ne le sont pas. En effet, une surface enherbée telle qu'un pâturage sera considéré comme non-artificialisée, tandis qu'un terrain de football ou un bassin d'infiltration serait cette fois-ci considéré comme artificialisé, bien qu'ils soient également enherbés. En conséquence, sans connaître l'usage, il est difficile de distinguer les zones qui pourraient être artificialisées de celles qui ne le seraient pas, et cette caractéristique rend l'OCS GE particulièrement pertinent, dans la mesure où c'est la seule donnée du *benchmark* qui cartographie l'usage.

Ainsi, en croisant ces deux indicateurs, l'IGN établit une grille de lecture de l'artificialisation à travers une matrice d'occupation des sols (cf. figure 4). Celle-ci permet de définir les éléments artificialisés et ceux qui ne le sont pas à partir de l'occupation et de l'usage des sols. En plus d'avoir une précision sémantique inédite, cette nomenclature précise est modulable et permet de s'adapter aux évolutions législatives en ce qui a trait à la définition de l'artificialisation des sols, qui, nous l'avons vu dans le chapitre 2, ont été nombreuses jusqu'ici. L'OCS GE permet effectivement de mesurer l'artificialisation des sols au sens de l'article R. 101-1 du Code de l'Urbanisme, mais

Apport des réseaux de neurones convolutifs à la cartographie de l'occupation du sol :
Cas d'usage sur le Grand Avignon

également la consommation d'espaces au sens de l'article 194 de la loi Climat et Résilience.

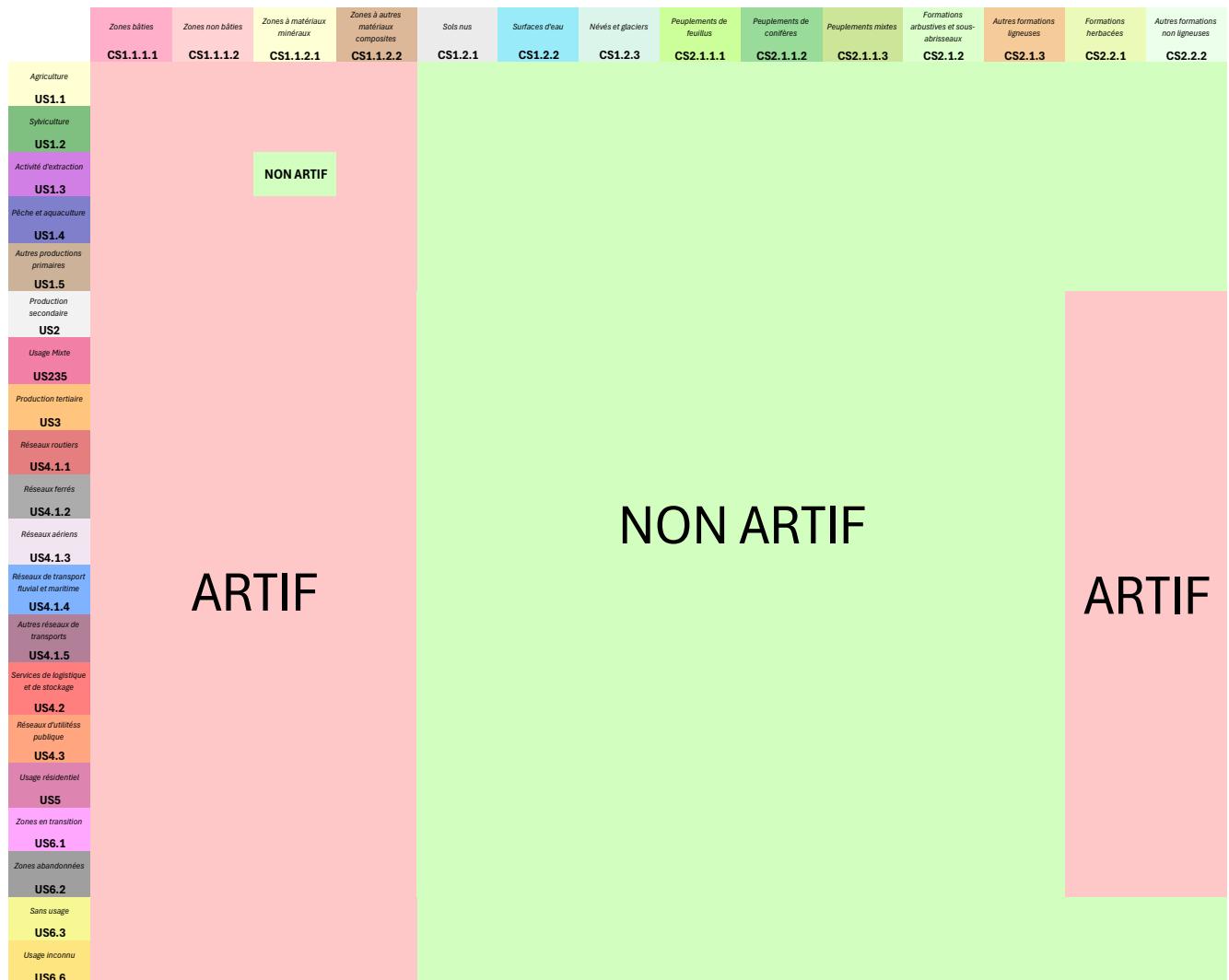


Figure 4. Matrice de l'artificialisation selon l'occupation et l'usage des sols de l'OCS GE.
Source : IGN

OCS GE	
Couverture	Usage
Zones bâties	Agriculture
Zones non bâties	Sylviculture
Zones à matériaux minéraux	Activités d'extraction
Zones à autres matériaux composites	Pêche et aquaculture
Sols nus	Autres productions primaires
Surfaces d'eau	Production secondaire
Névés et glaciers	Usage mixte
Peuplements de feuillus	Production tertiaire
Peuplements de conifères	Réseaux routiers
Peuplements mixtes	Réseaux ferrés
Formations arbustives et sous-arbrisseaux	Réseaux aériens
Autres formations ligneuses	Réseaux de transport fluvial et maritime
Formations herbacées	Autres réseaux de transport
Autres formations non ligneuses	Services de logistique et de stockage
	Réseaux d'utilité publique
	Usage résidentiel
	Zones en transition
	Zones abandonnées
	Sans usage
	Usage inconnu

Figure 5. Classes de la couverture et de l'usage du sol.
N. Massot, 2024. Source : IGN

L'OCS GE se structure en différents millésimes départementaux. Chaque année, l'IGN¹⁵ met à jour une partie de ces données à l'échelle départementale, sans pour autant couvrir toute la France (cf. figure 6). La production de l'OCS GE se base sur la BD Topo¹⁶, le Registre Parcellaire Graphique (RPG)¹⁷, la BD Forêt¹⁸, le RGE ALTI¹⁹ ainsi que sur la photo-interprétation de la BD Ortho par des processus automatisés basés, entre autres, sur l'intelligence artificielle de CoSIA, pour la partie occupation. De ce fait, la BD Ortho²⁰ étant produite tous les trois ans, la production de nouveaux millésimes d'OCS GE ne peut pas aller plus vite que la production de cette dernière.

De plus, l'IGN finance en partie la production de l'OCS GE en partenariat avec différents acteurs européens, nationaux ou territoriaux. Ces partenariats sont à double tranchant, dans la mesure où ils permettent de produire la donnée plus rapidement dans les territoires qui bénéficient de ces financements, au détriment

15. Institut National de l'Information Géographique et Forestière (IGN)

16. La BD Topo est une base de données vectorielles détaillée de l'IGN qui modélise les éléments du territoire français (bâtiments, routes, réseaux, etc.)

17. Le Registre Parcellaire Graphique (RPG) est une base de données géographique qui recense les parcelles agricoles et leur occupation du sol

18. La BD Forêt est une base de données de l'IGN qui cartographie les peuplements forestiers en France, incluant des informations sur les essences, la structure et la gestion des forêts

19. Le Référentiel à Grande Échelle Altimétrique (RGE Altı) est un modèle numérique de terrain de l'IGN qui fournit des données sur le relief et l'altitude du territoire français

20. La BD Ortho est une base de données d'images aériennes ou satellitaires orthorectifiées de l'IGN, offrant une couverture photographique haute résolution du territoire français

des autres²¹.

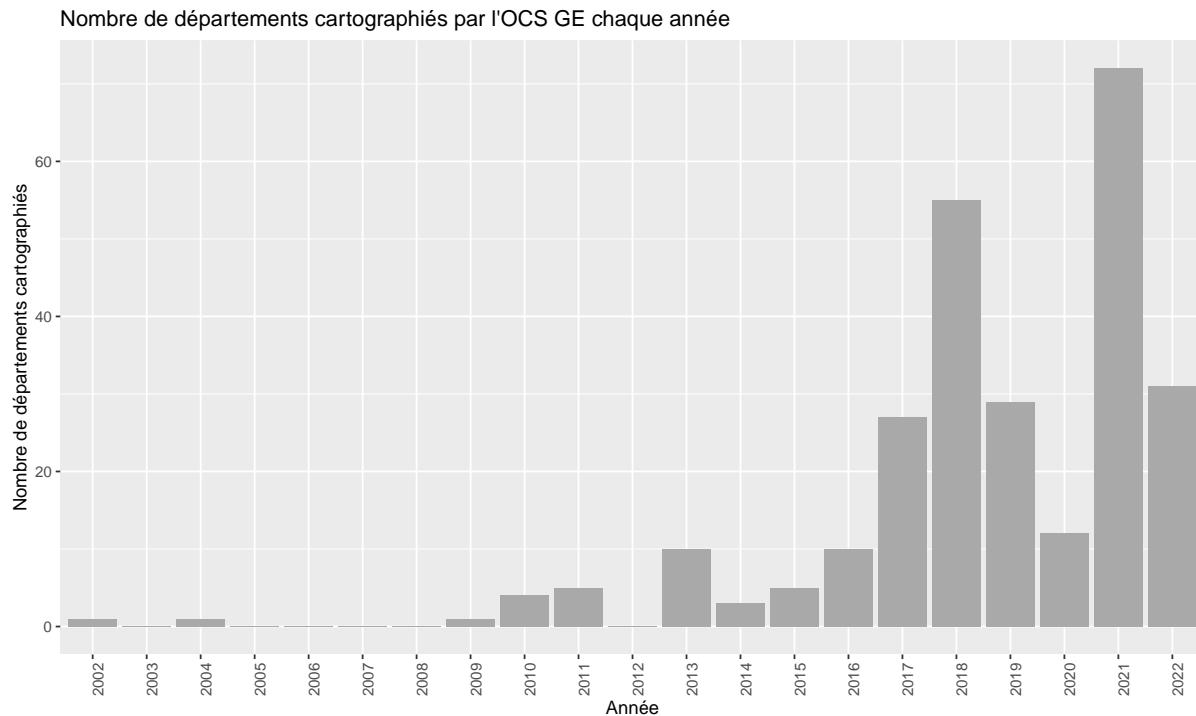


Figure 6. Nombre de départements cartographiés par l'OCS GE chaque année.

N. Massot, 2024. Source : Institut National de l'Information Géographique et Forestière,
OCS GE - Occupation du sol à grande échelle, Consulté le 29 décembre 2024, <https://geoservices.ign.fr/ocsge>

Dans sa documentation, l'IGN précise que l'OCS GE est agrégé à des entités administratives exogènes telles que le PCI Express, les bâtiments issus de la BD Topo, le RPG et les routes. Cette agrégation présente des avantages certains, notamment pour la réduction du bruit (qui est quasi inexistant), pour avoir des résultats plus clairs et lisibles, ou pour considérer l'évolution de l'occupation du sol à l'échelle de la parcelle, ce qui peut être plus facile à interpréter, notamment sur le plan diachronique²². En outre, cette agrégation devrait assurer une compatibilité avec les données cadastrales, favorisant ainsi une gestion foncière et administrative plus efficiente. Cependant, nous avons observé que ce n'était pas toujours le cas. En effet, si l'OCS GE est bien aligné sur les routes issues de la BD Topo, il n'est ni aligné sur les parcelles, ni avec le bâti, comme on peut le voir sur la figure n°7.

21. Voici la liste des départements non-couverts par l'OCS GE : Allier, Ardennes, Bouches-du-Rhône, Calvados, Cantal, Charente, Cher, Corrèze, Drôme, Eure-et-Loir, Indre, Loir-et-Cher, Haute-Loire, Loiret, Manche, Marne, Haute-Marne, Meurthe-et-Moselle, Meuse, Morbihan, Orne, Haute-Savoie, Deux-Sèvres, Vienne, Haute-Vienne, Vosges

22. Modulo les divisions parcellaires.

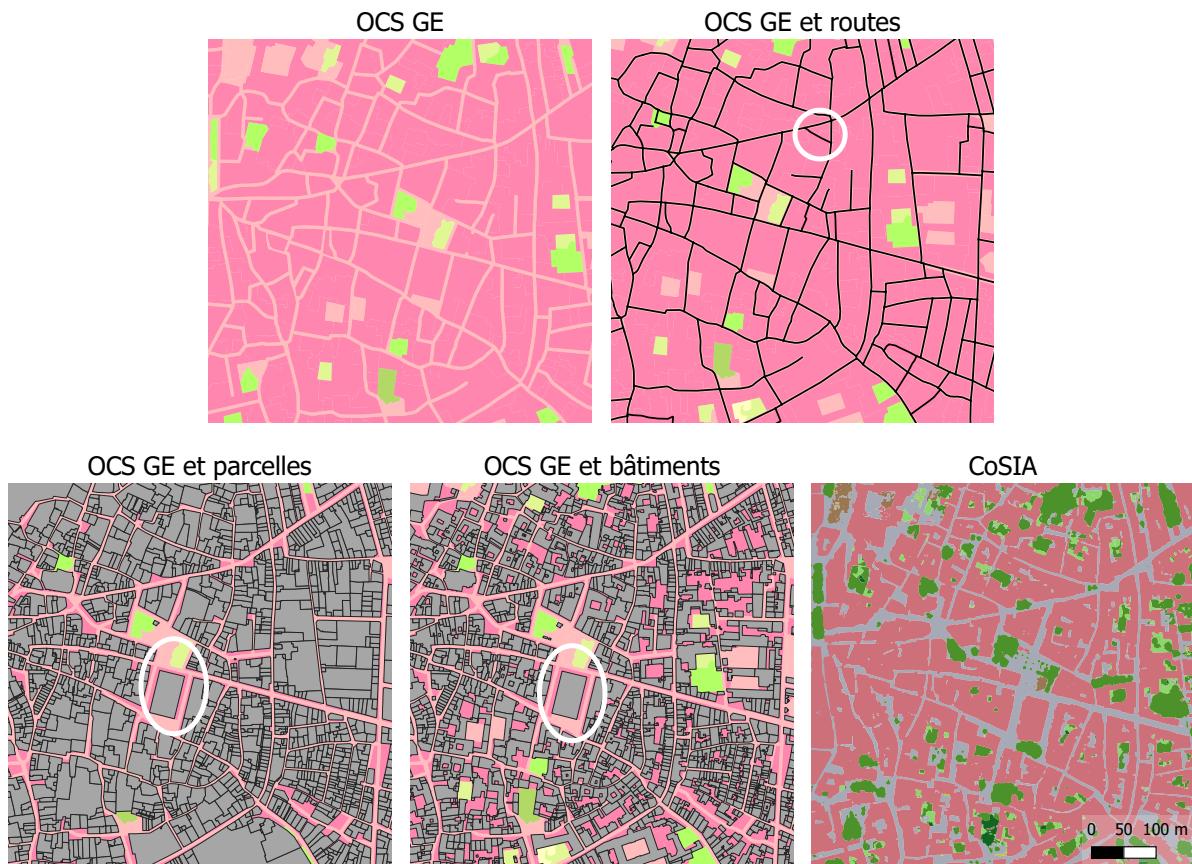


Figure 7. Comparatif de l'OCS GE et CoSIA.

Légende OCS GE : cf. figure 5;

Légende CoSIA : cf. figure 9.

N. Massot, 2024. Source : IGN

3.1.1.2 CoSIA

En 2023, l'IGN a mis à contribution son expertise en IA pour créer un modèle qui permet de cartographier l'occupation du sol à partir d'un corpus de données dont nous préciserons la nature plus tard. Le projet CoSIA survient pour améliorer la production de l'OCS GE à l'aide de l'IA. CoSIA est donc un sous-produit de l'OCS GE. Contrairement à celui-ci, dont la donnée est largement simplifiée, seuillée²³ et agrégée à des entités administratives exogènes, CoSIA est une donnée brute et prédictive dans la mesure où le modèle d'apprentissage profond estime pour chaque pixel (issu de la BD Ortho et du RGE Alti) la probabilité d'appartenance à une classe. De ce fait, le modèle s'affranchit de tous prédictifs extrinsèques. Par exemple, la figure n°7 montre que les routes issues de la BD Topo correspondent aux routes présentes dans l'OCS GE, ce qui n'est pas le cas de CoSIA. De ce fait, elles peuvent ne pas refléter entièrement la réalité du terrain. Ainsi, l'IGN nous met en garde sur son utilisation, en précisant, dans sa documentation technique, que bien que la

23. Seuils de 50 m² pour les zones bâties, et 2 500 m² si non.

finesse sémantique²⁴ soit une des principales forces de CoSIA, la qualité du produit est fortement dépendante de la qualité des données d'entraînement et d'entrées du modèle. Ainsi, les labels ont été réalisés selon la méthode des pyramides de segmentations mobilisée par Guigues et al. 2006. Il s'agit d'une technique utilisée en traitement d'images pour segmenter des images à différentes échelles ou niveaux de détail. Elles reposent sur une segmentation hiérarchique de l'image, où le niveau le plus bas correspond à une segmentation fine, avec des régions homogènes de petite taille, tandis qu'à chaque niveau supérieur, les régions adjacentes et similaires sont fusionnées pour former des régions plus grandes et plus abstraites. Les pyramides de segmentation constituent une méthode fondée sur une approche orientée objet, qui permet ainsi d'analyser une image à différentes échelles, en mettant en évidence les détails locaux aux niveaux fins et les structures globales aux niveaux plus grossiers. La transition entre les niveaux repose sur un critère de similarité entre les régions adjacentes, qui peut être basé sur des caractéristiques visuelles telles que la couleur ou la texture. Cette méthode des pyramides de segmentation a été appliquée sur 50 départements français²⁵, ce qui représente 25×10^{10} pixels de 20 centimètres de côté (IGN 2023).

24. La finesse sémantique fait référence à la précision et au détail des informations sémantiques fournies par CoSIA. Cela signifie que les cartes sont capables de distinguer et de classer avec une grande précision les différentes catégories de couverture du sol.

25. 40 départements pour l'entraînement, et 10 pour la validation.

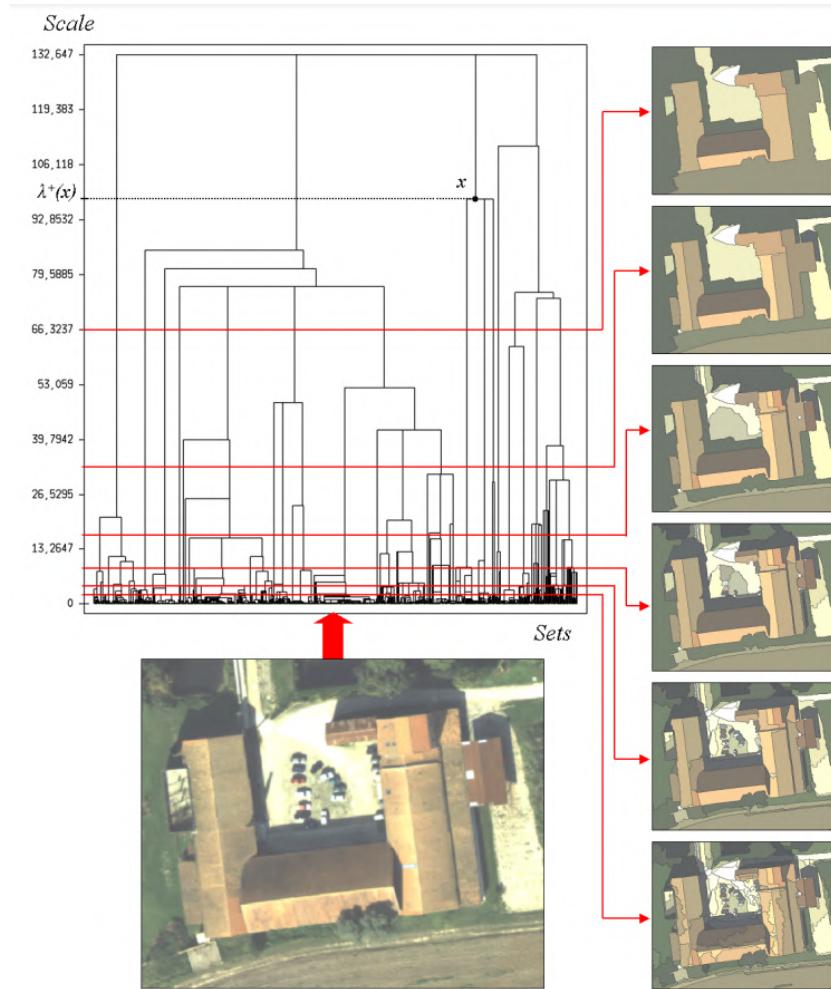


Figure 8. Pyramides de segmentation.

Source : Guigues, L., Cocquerez, J. P., & Le Men, H. (2006). Scale-sets Image Analysis. *International Journal of Computer Vision*, 68(3), 289-317. <https://doi.org/10.1007/s11263-005-6299-0>

Tantôt, nous disions à propos de la finesse sémantique de CoSIA qu'elle était une de ses forces. Avec ses 16 classes (cf. figure 9), elle offre une bonne granularité. Seuls le Corine Land Cover et l'OCS GE Usage proposent plus de classes. Toutefois, comparons ce qui est comparable : les cas d'utilisation de l'OCS GE Usage sont différents de ceux de CoSIA, dans la mesure où l'un décrit les éléments de surface fonctionnels, tandis que l'autre s'affaire à cartographier les éléments morphologiques du sol et du sursol. De plus, la source de donnée de CoSIA étant la BD Ortho, dont la résolution est de 20 centimètres, CoSIA est la donnée d'occupation du sol avec la plus grande résolution spatiale, et dont la résolution sémantique est élevée.

CoSIA	
Bâtiment	#C0392B
Zone imperméable	#8A5A3D
Zone perméable	#8A5A3D
Piscine	#00AEEF
Serre	#4DB6AC
Sol nu	#8A5A3D
Surface eau	#1E8449
Neige	#D9EAD3
Conifère	#1E8449
Feuillu	#4DB6AC
Coupe	#F4A460
Broussaille	#D9EAD3
Pelouse	#4DB6AC
Culture	#F4A460
Terre labourée	#F4A460
Vigne	#C0392B
Autre	#000000

Figure 9. Classes de CoSIA.

N. Massot, 2024. Source : Institut National de l'Information Géographique et Forestière (IGN). (2023). *Cosia - Documentation technique*. Consulté le 29 décembre 2024, sur https://cosia.ign.fr/pdf/Cosia_Documentation_Technique IGN_2023.pdf

L'IGN fait preuve de transparence à propos de la chaîne de traitements de CoSIA, tant sur la nature des données d'entrée que du modèle en lui-même. Cette chaîne de traitement, nommée *French Land cover from Aerospace ImageRy (FLAIR)*, est publiée sur la page GitHub de l'IGN (IGN 2024a). Garioud et al. ont rédigé un *datapaper* expliquant la méthode (Garioud, Bookjans et al. 2022). Ce dernier est accompagné du modèle pré-entraîné disponible en source ouverte sur GitHub (Garioud, Peillet et al. 2022). L'accès ouvert de toute la chaîne de traitements est, dans le contexte de la crise de la reproductibilité²⁶, devenu une prérogative tacite à tous les travaux, notamment ceux financés par le contribuable. De ce fait, nous connaissons en détail la méthode utilisée pour CoSIA, que nous avons retranscrite dans la figure 10. Dans celui-ci, nous voyons les étapes nécessaires à la production de CoSIA. Ce dernier se base sur les prises de vues aériennes de la BD Ortho datant de 2024 ainsi que sur le MNT et MNS du RGE ALTI. Après conception et entraînement du modèle, celui-ci est inféré sur la France entière avant de subir un post-traitement classique.

26. La crise de la reproductibilité est la crise méthodologique dans le domaine des sciences selon laquelle de nombreux résultats scientifiques publiés sont difficiles, voire impossibles à reproduire au cours d'études subséquentes.

Chaîne de traitements - CoSIA

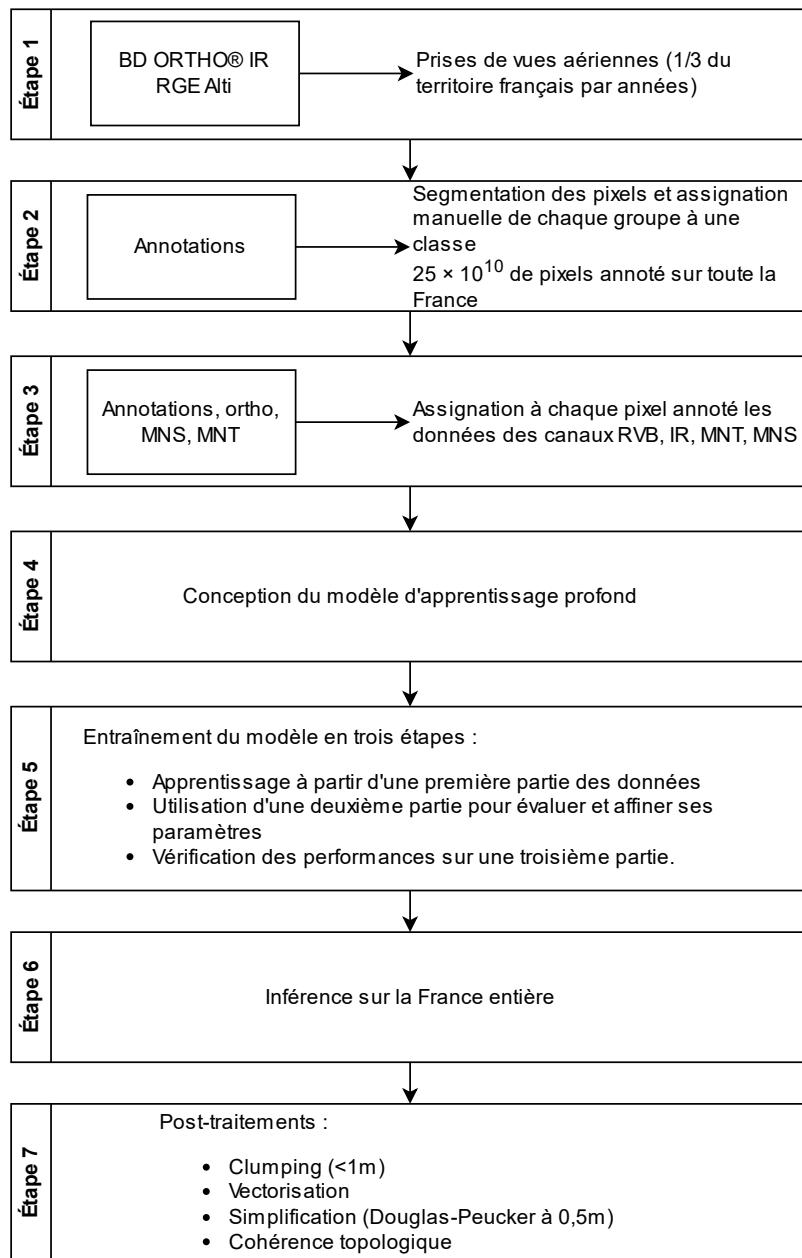


Figure 10. Chaîne de traitements de CoSIA.

N. Massot, 2024. Source : Source : IGN. (2023). Cosia - Documentation technique. Consulté le 29 décembre 2024, sur https://cosia.ign.fr/pdf/Cosia_Documentation_Technique_IGN_2023.pdf

3.1.1.3 Corine Land Cover

Le Corine Land Cover (CLC) est un inventaire biophysique de l'occupation des sols et de son évolution, produit dans le cadre du programme européen d'observation de la terre Copernicus. Il s'agit d'une base de données commune à 39 États européens et est gérée en France par le ministère de la Transition Écologique, de l'Énergie, du Climat et de la Prévention des Risques, avec l'appui de l'IGN depuis 2018. Le fait d'avoir une donnée comparable homogène à l'échelle de l'Europe est un atout majeur. De plus, le CLC est la donnée d'occupation du sol la plus ancienne, et est régulièrement mis à jour. Les millésimes commencent en 1990 et à partir de 2000, les données sont mises à jour tous les six ans. Toutefois, l'objectif annoncé du CLC est de suivre l'état et l'évolution de l'occupation des sols sur des zones homogènes d'au moins 25 ha (cf. figure 11). Ce seuil de résolution géométrique incarne une première limite, dans la mesure où celui-ci impacte fatalement sa précision globale et rend impossible le suivi des phénomènes d'occupation des sols à fine échelle, comme le mitage ou l'évolution des bocages. Conscient de ses limites, Copernicus travaille sur le CLC+ qui prévoit une unité minimale de collecte (UMC)²⁷ de 0,5 à 5 ha et une résolution métrique de 10 m, avec une fréquence de mise à jour améliorée.

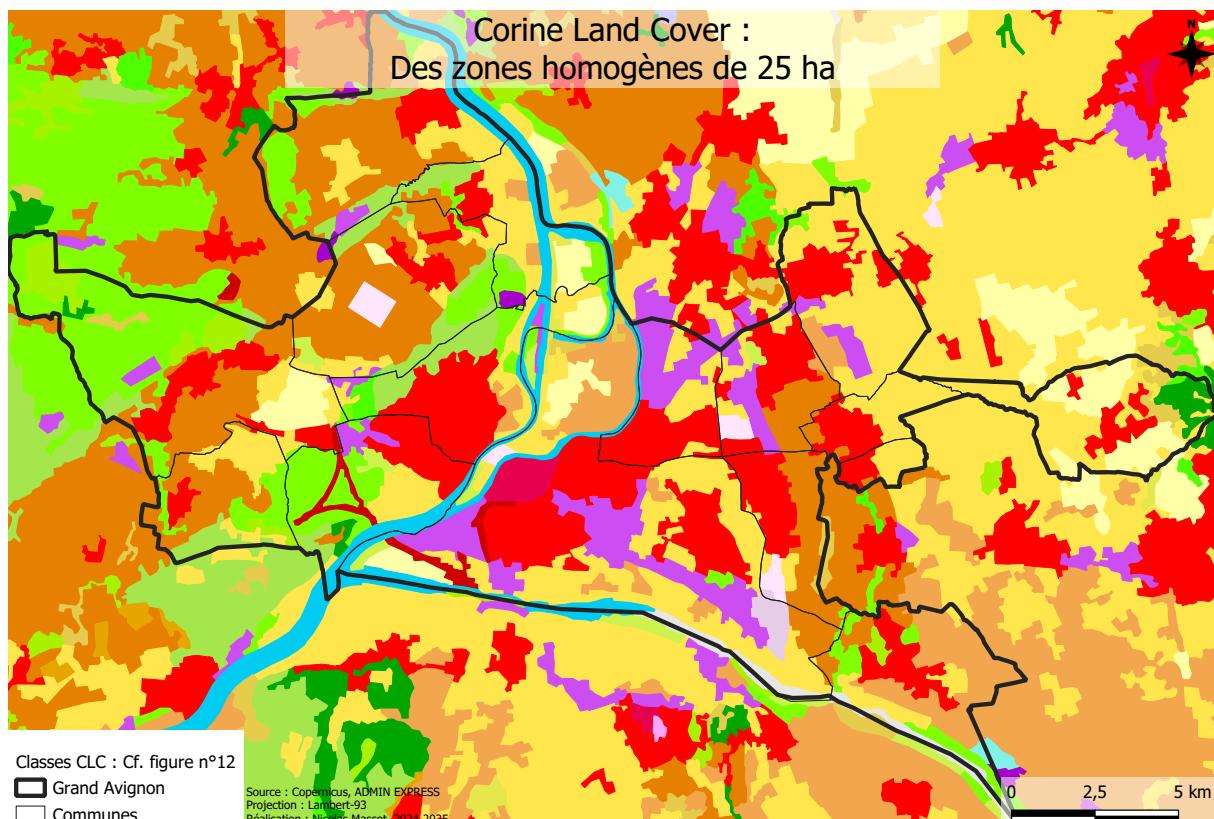


Figure 11. Le Corine Land Cover sur Avignon : des zones homogènes de 25 ha.
N. Massot, 2024. Source : Copernicus

27. L'unité minimale de collecte (UMC) : plus petite superficie homogène qui peut être identifiée et analysée avec précision à partir de données de satellitaire

Malgré sa résolution spatiale réduite, le CLC dispose d'une grande variété de classes. La classification est répartie en trois niveaux de complexité emboités. Le niveau un comprend cinq classes, le niveau deux en compte quinze, tandis que le niveau trois en a quarante-quatre (cf. figure 12). Notons également la désignation Département et Régions d'Outre-Mer qui en comprend cinquante. Ce nombre de classes plus élevé se justifie par la présence de cultures qui ne sont pas présentes en Europe continentale telles que les bananeraies ou les caféiers.

Corine Land Cover	
Tissu urbain continu	
Tissu urbain discontinu	
Zones industrielles ou commerciales et installations publiques	
Réseaux routier et ferroviaire et espaces associés	
Zones portuaires	
Aéroports	
Extraction de matériaux	
Décharges	
Chantiers	
Espaces verts urbains	
Équipements sportifs et de loisirs	
Terres arables hors périmètres d'irrigation	
Périmètres irrigués en permanence	
Rizières	
Vignobles	
Vergers et petits fruits	
Oliveraies	
Prairies et autres surfaces toujours en herbe à usage agricole	
Cultures annuelles associées à des cultures permanentes	
Systèmes culturels et parcellaires complexes	
Surfaces essentiellement agricoles, interrompues par des espaces naturels importants	
Territoires agroforestiers	
Forêts de feuillus	
Forêts de conifères	
Forêts mélangées	
Pelouses et pâturages naturels	
Landes et broussailles	
Végétation sclérophylle	
Forêt et végétation arbustive en mutation	
Plages, dunes et sable	
Roches nues	
Végétation clairsemée	
Zones incendiées	
Glaciers et neiges éternelles	
Marais intérieurs	
Tourbières	
Marais maritimes	
Marais salants	
Zones intertidales	
Cours et voies d'eau	
Plans d'eau	
Lagunes littorales	
Estuaires	
Mers et océans	

Figure 12. Classes du CLC niv. 3.
N. Massot, 2024. Source : Copernicus

3.1.1.4 ESA World Cover

Les données d'occupation du sol de l'Agence Spatiale Européenne sont caractérisées par une couverture spatiale mondiale et par sa grande résolution de 10 mètres. Ce sont les satellites Sentinel-1 et 2 qui sont à l'origine de ces données. Ces dernières sont disponibles en 2020 et 2021 et sont composées de 11 classes (cf. figure 13). Les cas d'usage sont environnementaux et en lien avec la gestion des ressources, notamment agricoles. Cette dominante agricole s'explique par la présence de certains financeurs clés du projet, tels que la FAO²⁸, la CIFOR²⁹, le *World Ressources Institute* ou encore les Nations Unies à propos de la lutte contre la désertification.

ESA World Cover	
Couvert arboré	Green
Arbustes	Orange
Prairies	Yellow
Terres cultivées	Purple
Bâti	Red
Végétation nue / clairsemée	Grey
Neige et glace	
Plans d'eau permanents	Blue
Zones humides herbacées	Cyan
Mangroves	Green
Mousses et lichens	Yellow

Figure 13. Classes de l'ESA World Cover.
N. Massot, 2024. Source : ESA

3.1.1.5 Occupation des Sols Opérationnelle (OSO) - Pôle Theia

Tout comme le Corine Land Cover, le produit Occupation des Sols Opérationnelle (OSO) fait preuve d'une grande finesse sémantique, surtout en ce qui a trait aux composantes agricoles (cf. figure 14). Cette donnée produite par le Centre d'Expertise Scientifique sur l'occupation des sols (CES OSO), composante du Pôle Theia³⁰ est réalisée annuellement, entre 2016 et 2020 à partir de données Sentinel-2 A et Sentinel-2 B. Ces sources des données leur permettent d'avoir une résolution spatiale de 10 mètres en mode raster, et 20 en mode vecteur. Comme dit précédemment, avec ces 23 classes, l'OSO possède une nomenclature précise, surtout dans les espaces liés à l'agriculture. En effet, les données satellitaires sont acquises tout au long de l'année, ce qui permet de considérer l'évolution des formes végétales qui diffèrent au gré des saisons. De ce fait, ces données font preuve de "fraîcheur" dans la mesure où leur diffusion de ces données se fait seulement trois mois après l'acquisition du millésime de référence. À titre de comparaison, le CLC est mis à jour tous les six ans avec un délai de trois ans.

Au sujet du traitement, le CES OSO utilise la chaîne de traitements iota2 (Iota2-

28. Food and Agriculture Organization

29. Center for International Forestry Research

30. Le Pôle Theia est composé de plusieurs groupes de recherche : Cesbio, l'UMR Ispa, Dynafor, CNRM, UMR Tetis, IGN-Matis, Costel et Sertit ICUBE

project 2024)³¹ basée sur Orfeo ToolBox. Elle est centrée autour de l'algorithme de classification supervisée *Random Forest* pour lequel les annotations sont issues de la BD Topo, du Corine Land Cover, de l'Urban Atlas et du Registre Parcellaire Graphique. Cette chaîne de traitements se base sur les séries temporelles annualisées de Sentinel-2 A et Sentinel-2 B et est entièrement automatisée, ce qui explique la capacité à produire une classification annuellement sur l'ensemble du territoire métropolitain³².

Les classes du produit OSO sont particulièrement précises en ce qui a trait aux composantes agricoles. Elles permettent de bien caractériser les essences des cultures. La précision sémantique de la végétation est également intéressante. De ce fait, OSO est très pertinent pour le suivi de l'évolution de l'agriculture et de la végétation, tant d'un point de vue qualitatif que quantitatif. Toutefois, bien qu'un effort ait été produit pour faire la distinction entre les zones urbaines denses et diffuses, les zones industrielles et commerciales et les routes, le pôle Theia met en garde sur le taux d'erreur de commission élevé entre ces classes. Ainsi, ce dernier recommande de fusionner ces quatre classes pour réduire ces confusions, ce qui permettrait d'avoir, selon eux, une précision globale de 90 % sur ces quatre classes.

31. Infrastructure pour l'Occupation des sols par Traitement Automatique

32. Notons que l'outil est libre puisqu'il est publié sous la licence AGPL.

Occupation des Sols Opérationnelle	
Urbain dense	■ Magenta
Urbain diffus	■ Light Magenta
Zones industrielles et commerciales	■ Cyan
Routes	■ Cyan
Colza	■ Yellow
Céréales à paille	■ Light Green
Protéagineux	■ Green
Soja	■ Orange
Tournesol	■ Light Green
Maïs	■ Red
Riz	■ Light Blue
Tubercules / Racines	■ Brown
Prairies	■ Olive
Vergers	■ Purple
Vignes	■ Dark Red
Forêts de feuillus	■ Dark Green
Forêts de conifères	■ Black
Pelouse	■ Yellow
Landes	■ Teal
Surfaces minérales	■ Red
Plages et Dunes	■ Orange
Glaciers et neiges éternelles	■ Grey
Eau	■ Blue
Serres	■ Dark Blue

Figure 14. Classes du produit OSO.
N. Massot, 2024. Source : Pôle Theia

3.1.1.6 ESRI Land Cover

L'ESRI Land Cover est une carte de la couverture du sol élaborée à partir des données Sentinel-2 et disponible dans le monde entier à une résolution de 10 mètres. De plus, elle possède une résolution temporelle annualisée allant de 2017 à 2023. Ces millésimes sont produits à partir d'algorithmes d'apprentissage profond, ce qui permet de pouvoir comparer les données d'une année sur l'autre, contrairement aux algorithmes d'apprentissage supervisé. Des "milliards" de pixels annotés manuellement ont permis d'entraîner le modèle avec une résolution spectrale de six bandes.

ESRI Land Cover	
Eau	Blue
Arbres	Green
Végétation inondée	Cyan
Cultures	Yellow
Zone bâtie	Red
Sol nu	Light Gray
Neige/Glace	Light Blue
Nuages	Gray
Terrain de parcours	Orange

Figure 15. Classes de ESRI Land Cover.
N. Massot, 2024. Source : ESRI

3.1.1.7 OpenStreetMap

Les contributeurs OpenStreetMap (OSM) s'affairent également à cartographier l'utilisation du sol. La clé '*landuse*' est utilisée pour désigner l'occupation du sol, qu'il s'agisse de zones résidentielles, commerciales, industrielles, agricoles ou naturelles. Les contributeurs ajoutent ces tags en s'appuyant sur des observations locales, des données gouvernementales ou des images satellites, ce qui enrichit la base de données avec des informations précises et à jour.

Cependant, bien que ces données fournissent des informations essentielles, elles ne couvrent pas l'ensemble du territoire. Par exemple, dans le Grand Avignon, 27 % de la superficie n'est pas référencée en matière d'occupation du sol. Néanmoins, cette région compte 40 tags d'occupation du sol différents, offrant une certaine précision localisée sur la typologie du sol. Toutefois, parmi ces 40 tags, seuls 12 représentent plus d'un pour cent de la superficie du Grand Avignon, avec une forte concentration sur les vignobles, les terres agricoles, les forêts et les zones résidentielles (cf. figure 16).

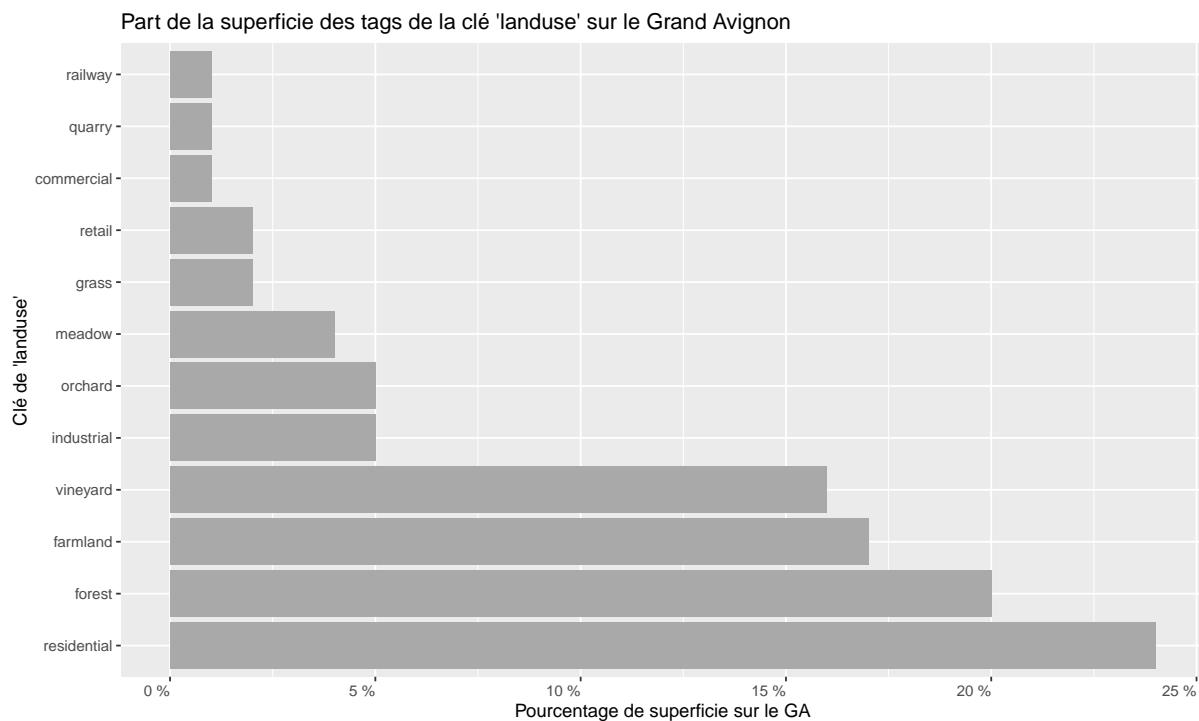


Figure 16. Part de la superficie des tags de la clé 'landuse' sur le Grand Avignon.
N. Massot, 2024. Source : OpenStreetMap

3.1.2 Analyse quantitative de la précision des couches d'occupation du sol continues

3.1.2.1 Méthodologie de l'analyse de la précision

Après avoir réalisé une analyse qualitative des données d'occupation du sol, nous proposons une analyse quantitative de ces dernières. Celle-ci mobilise pour chaque couche des matrices de confusion, et différents indicateurs quantifiant la précision des données. L'objectif de cette étude comparative est de mettre en place un référentiel commun, basé sur une méthode standardisée, permettant de comparer la précision de chaque couche d'occupation du sol. Comme mentionné dans la sous-partie précédente, et comme on peut le constater sur la figure n°17, chaque couche d'occupation du sol possède un nombre de classes différent. Face à ce constat, nous avons choisi de retirer les classes qui occupent moins de 1 % de la superficie du Grand Avignon, et de reclasser les classes restantes en quatre macro-classes (cf. 18) :

- Artificialisé
- Eau
- Herbe / sol nu / champs
- Végétation

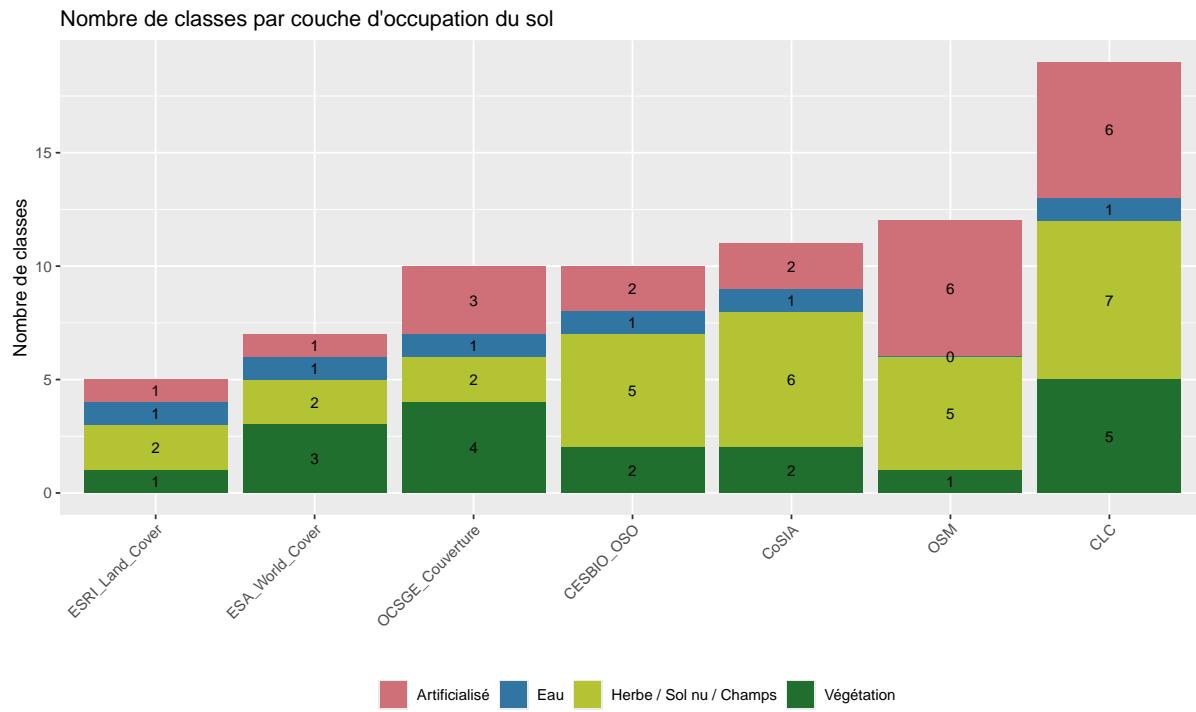


Figure 17. Nombre de classes par couche d'occupation du sol.
N. Massot, 2024. Source : IGN, Copernicus, ESRI, OSM, ESA, Pôle Théia

Apport des réseaux de neurones convolutifs à la cartographie de l'occupation du sol :
Cas d'usage sur le Grand Avignon

Macro-classes	OCSGE Couverture	CoSIA	CLC n° 3	ESA World Cover	CESBIO OSO	ESRI Land Cover OSM Landuse Landcover
Artificialisé	Zones bâties	Bâtiment	Zones industrielles ou commerciales et installations publiques	Zones bâties	Serris	Zones bâties
	Zones non bâties	Zone perméable	Tissu urbain discontinu	Urban diffus	Carrière	Commerce commercial
	Zones à matériaux minéraux		Réseaux routier et ferroviaire et espaces associés		Chemin de fer	Commerce de détail
			Équipements sportifs et de loisirs		Industriel	Industriel
			Tissu urbain continu		Résidentiel	Résidentiel
			Aéroports			
			Broussaille	Surfaces essentiellement agricoles, interrompues par des espaces naturels importants	Tourne-sol	Champs
			Culture	Prairies	Prairies	Pâturages
Herbe / Sol nu / Champs	Autres formations ligneuses	Terres arables hors périmètres d'irrigation	Terres cultivées	Landes	Landes	Prairie
		Vergers et petits fruits		Vignes	Vignes	Vergier
		Vignobles		Vergers	Vergers	Vignoble
		Systèmes cultureux et parcelaires complexes				Terre agricole
		Zone imperméable	Pelouses et pâtures naturels			
			Oliveraie			
Végétation	Peuplements de feuillus	Conifère	Forêt mélangée	Couverture arboree	Foêts de conifères / Arbres	Forêt
		Feuillu	Forêt et végétation arbustive en mutation	Arbustes	Foêts de feuillus	
	Peuplements mixtes		Forêts de conifères	végétation nue / clairsemée		
	Formations arbustives et sous arbisseaux		Végétation sclérophylle			
			Forêts de feuillus			
Eau	Surface d'eau		Cours et voies d'eau	Plans d'eau permanents	Eau	Eau

Figure 18. Reclassification des couches d'occupation du sol.
N. Massot, 2024. Source : IGN, Copernicus, ESRI, OSM, ESA, Pôle Théia

Afin de valider la précision des données étudiées, nous utilisons des matrices de confusion. Ces dernières sont un tableau croisé entre valeurs observées et valeurs modélisées et permettent de comparer la classe assignée à chaque pixel à partir des couches d'occupation du sol après reclassement dans les macro-classes avec la vraie classe extraite de la vérité-terrain. Pour ce faire, nous effectuons un échantillonnage systématique qui consiste à distribuer un nombre défini d'échantillons selon un schéma répétitif régulier. Ici, nous utilisons une grille de points d'une résolution de 750 mètres sur le territoire du Grand Avignon. Cela représente un total de 673 échantillons qui ont été manuellement classés selon les quatre macro-classes décrites précédemment, en se basant sur la photo-interprétation de la BD Ortho. Les données d'occupation du sol étudiées ont des dates de référence différentes les unes des autres. De ce fait, afin d'éviter un anachronisme, nous créons les vérités terrain en classant les échantillons en se basant sur le millésime de la BD Ortho de l'année correspondante. Seuls l'Occupation des Sols Opérationnelle (OSO) et l'ESRI Land Cover sont millésimés en 2023, alors qu'aucun millésime de la BD Ortho n'est disponible pour cette année. Pour ces deux couches, nous nous sommes alors basés sur la vérité terrain de 2024.

- BD Ortho 2018 : Corine Land Cover
- BD Ortho 2021 : Occupation du Sol à Grande Échelle, ESA World Cover
- BD Ortho 2024 : CoSIA, OpenStreetMap (clé '*landuse*'), Occupation des Sols Opérationnelle (OSO), ESRI Land Cover

Cette méthode a pour principal avantage de s'assurer que les échantillons manuellement labellisés ne comportent pas d'erreur de classification, en plus de réduire l'anachronisme entre la vérité terrain et la couche d'occupation du sol évaluée.

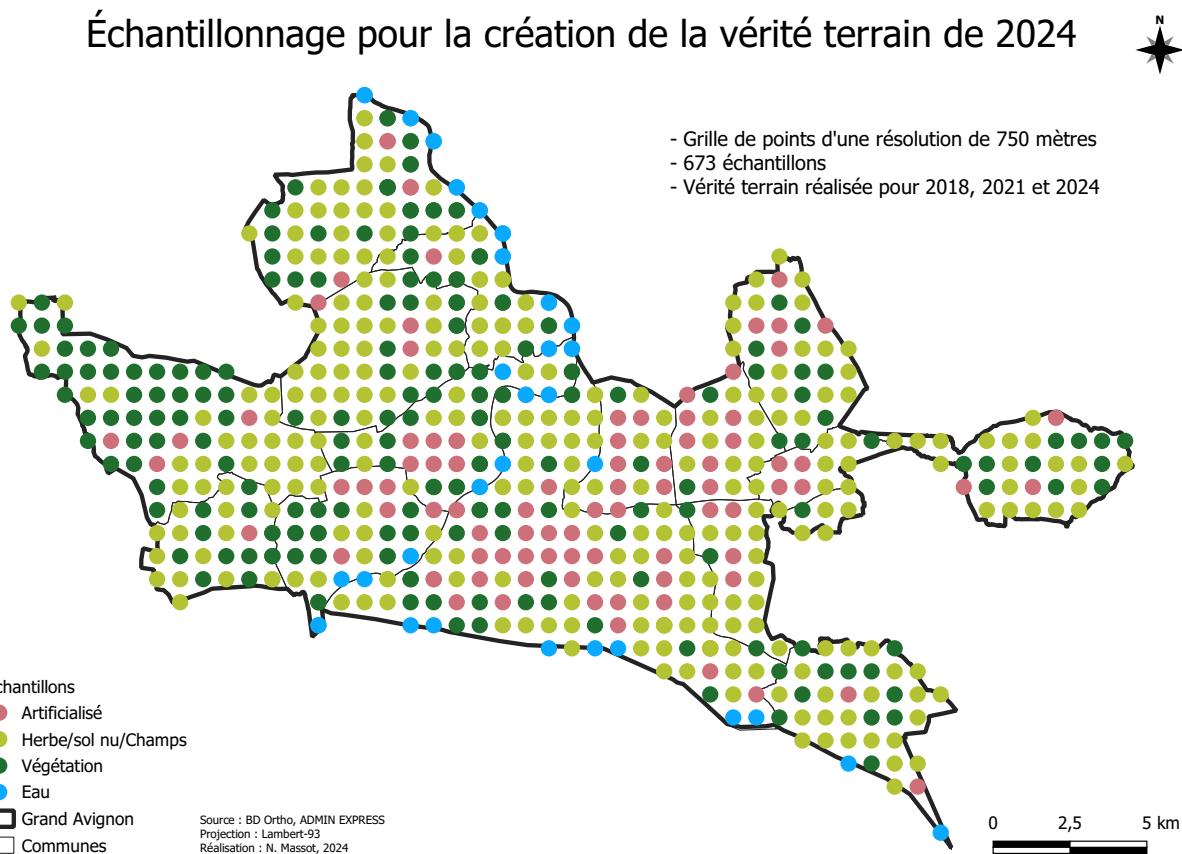


Figure 19. Échantillon pour la création de la vérité terrain de 2024.
N. Massot, 2024

En comparant la vérité terrain ainsi produite et les prédictions des différentes données, nous calculons des matrices de confusion. Ces matrices de confusion sont accompagnées des indicateurs de précision globale, du coefficient de Kappa et des erreurs de commission et d'omission³³.

33. La précision globale est la part moyenne des pixels bien classés ($VP+VN/N$). Le coefficient de Kappa indique le degré d'accord (concordance) entre deux juges concernant la classification d'éléments en tenant compte de l'accord dû au hasard. Les erreurs de commission représentent la part des pixels mal classés et se lisent horizontalement. Les erreurs d'omission représentent la part des pixels non classés et se lisent verticalement. Ces deux derniers sont mesurés à l'échelle globale.

3.1.2.2 Résultats de l'analyse de la précision

L'OCS GE présente une précision globale de 78,95 % et un coefficient Kappa de 0,68, ce qui indique une performance générale correcte (cf. figure 20). Toutefois, la matrice de confusion révèle des disparités dans les performances des classes "Artificialisé", "Végétation" et "Herbe/Sol nu/Champs", bien que ces dernières restent encore convenables.

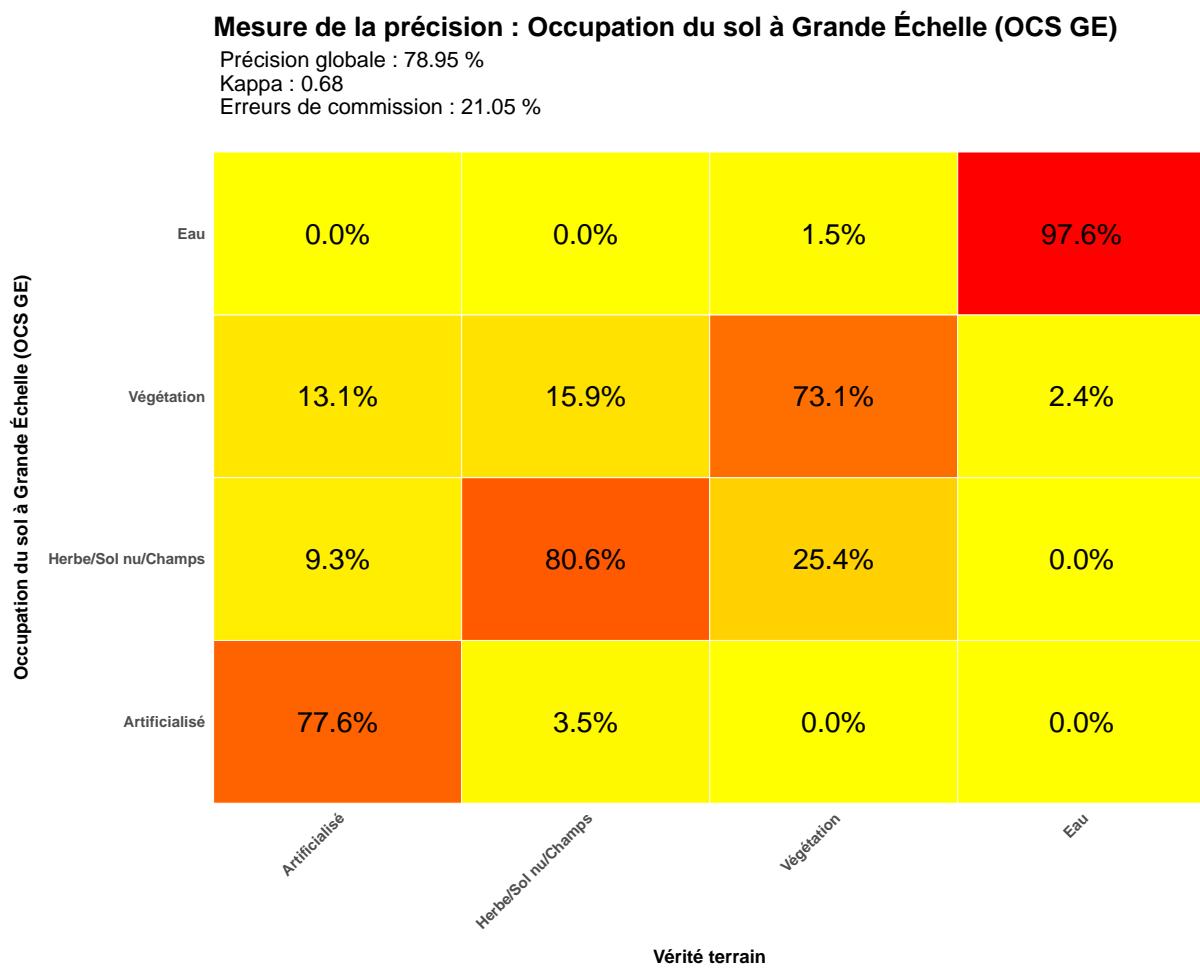


Figure 20. Mesure de la précision : Occupation du sol à Grande Échelle (OCS GE).
N. Massot, 2024

Quant à lui, CoSIA présente des résultats en deçà de ceux de l'OCS GE. En effet, les indicateurs de précision globale et de Kappa sont plus faibles, et on remarque des erreurs de commission prononcées entre l'artificialisé et l'herbe/sol nu/champs. Cependant, la végétation est modélisée de manière très précise, ce qui n'est pas le cas de l'eau qui fait preuve d'erreurs de commission avec la végétation et l'herbe/sol nu/champs. La principale marge de progression pour CoSIA se situe au niveau des surfaces artificialisées et de l'eau.

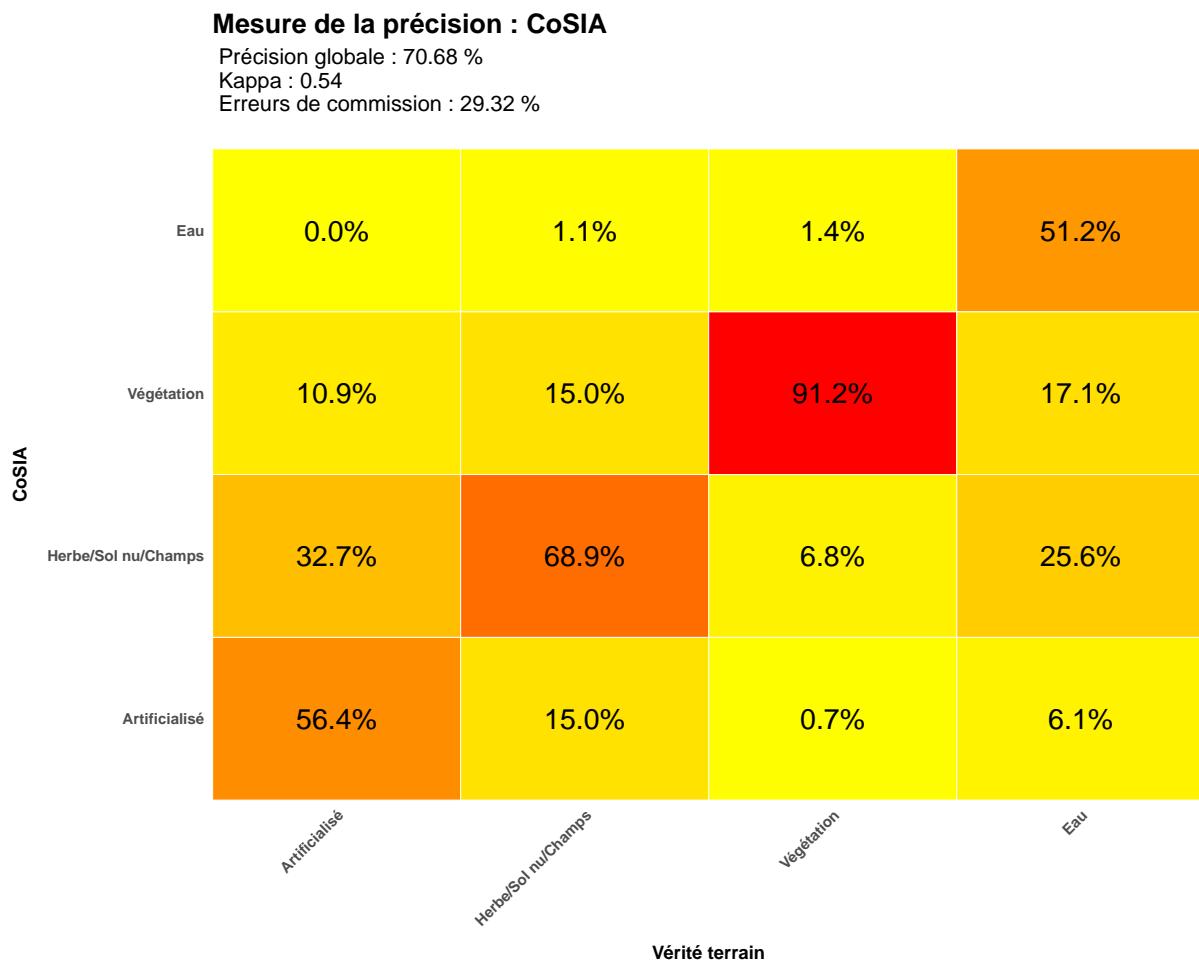


Figure 21. Mesure de la précision : CoSIA.
N. Massot, 2024

Le Corine Land Cover aurait présenté un bon score de précision globale, s'il n'était pas dégradé par des erreurs de commission, de 54 %³⁴ entre la classe "artificialisé" et les classes "herbe/sol nu/champs" et "végétation". Ces erreurs seraient dues à l'agrégation des données dans des zones homogènes d'au moins 25 hectares.

Mesure de la précision : Corine Land Cover

Précision globale : 66.72 %

Kappa : 0.5

Erreurs de commission : 33.28 %

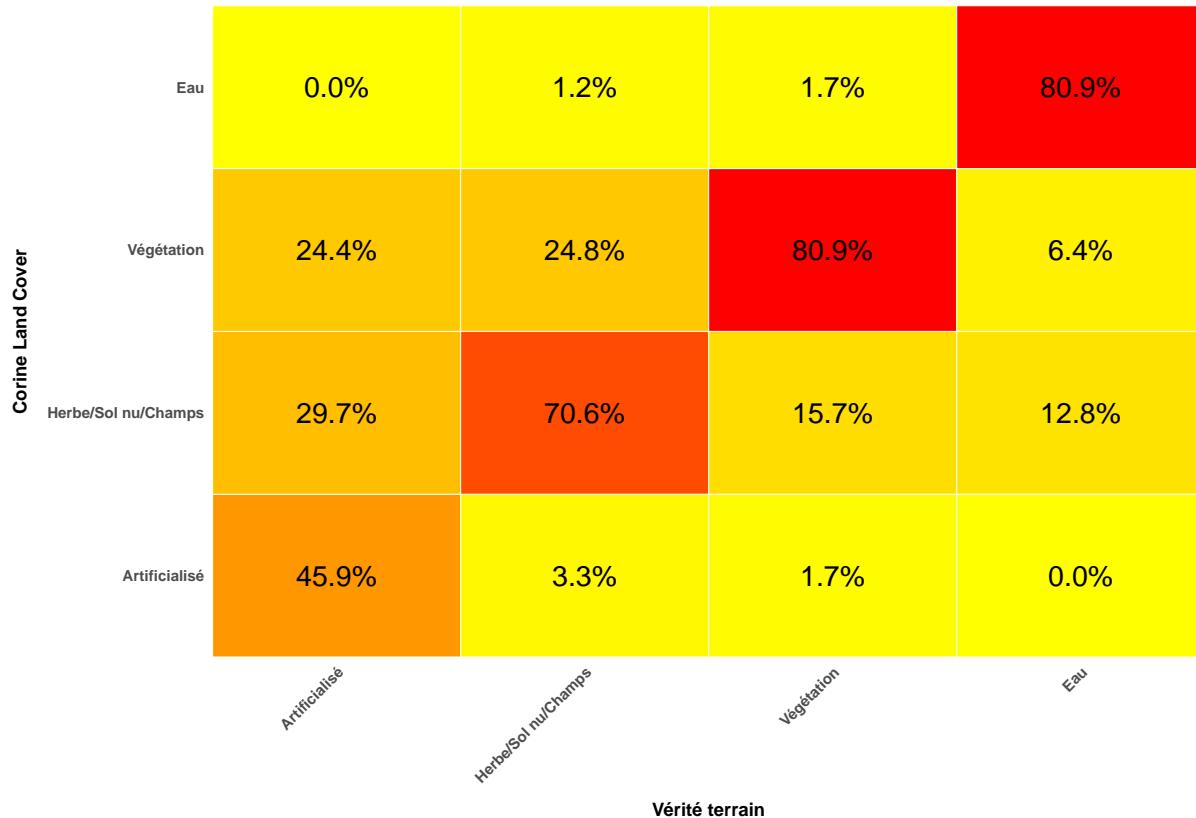


Figure 22. Mesure de la précision : Corine Land Cover.
N. Massot, 2024

34. 24,4+29,7

L'ESA World Cover possède des bons scores de précision pour l'eau et pour l'herbe/sol nu/champs. Toutefois, ceux-ci sont à contraster au regard de la performance de la végétation et de l'artificialisé. Nous voyons que l'erreur de commission est à 48,3 %³⁵, ce qui est anormalement élevé.

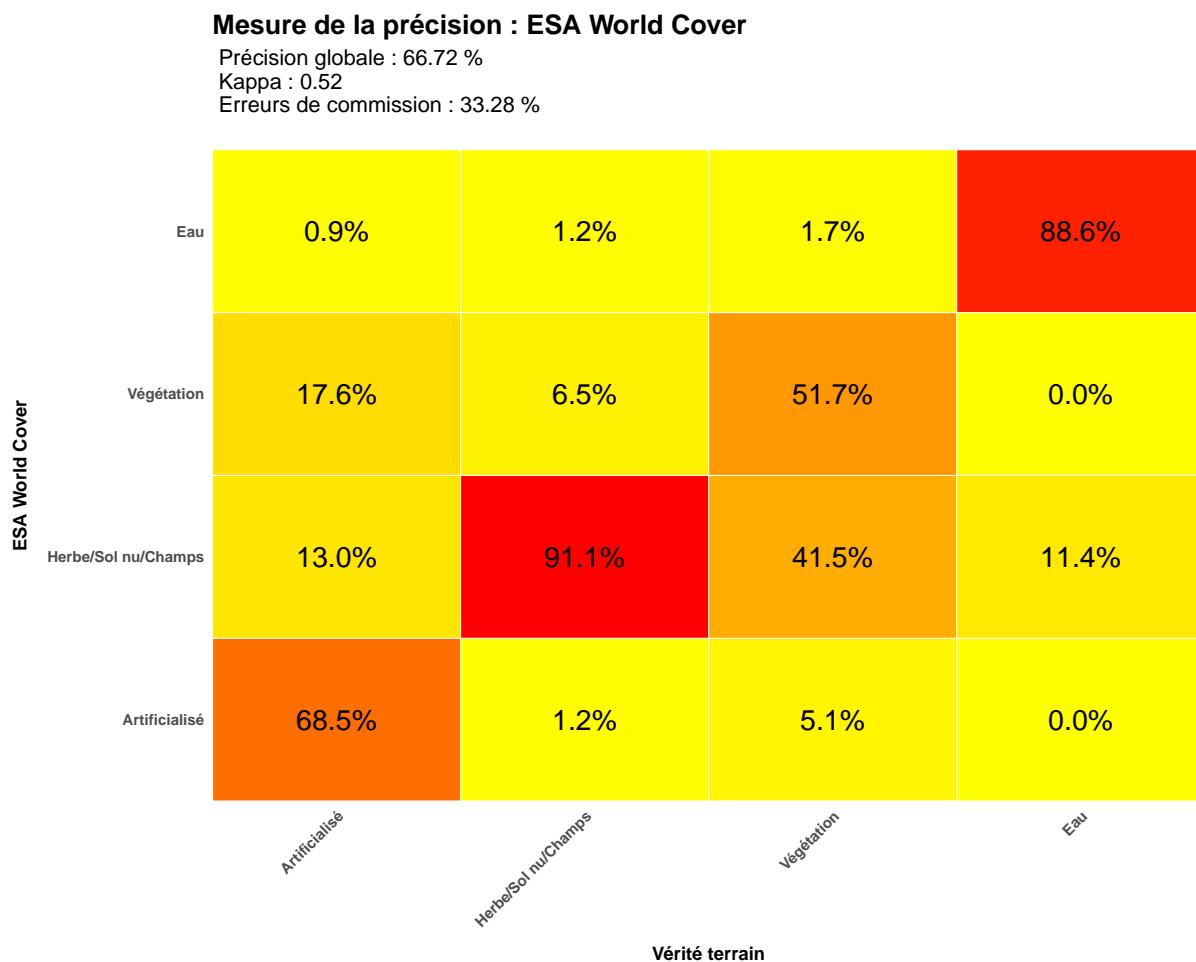


Figure 23. Mesure de la précision : ESA World Cover.
N. Massot, 2024

35. 41,5+1,7+5,1

L'Occupation des Sols Opérationnelle du Pôle Théia est performante en ce qui a trait à l'eau et à l'herbe/sol nu/champs, mais les résultats sont en demi-teinte pour les classes artificialisées et la végétation. Dans l'analyse qualitative, nous mettions en avant le fait que cette donnée est particulièrement fine sémantiquement parlant sur les composantes agricoles. Au-delà de cet aspect qualitatif, nous voyons que la précision de cette dernière est proche de 80 %, ce qui représente un bon score du point de vue quantitatif.

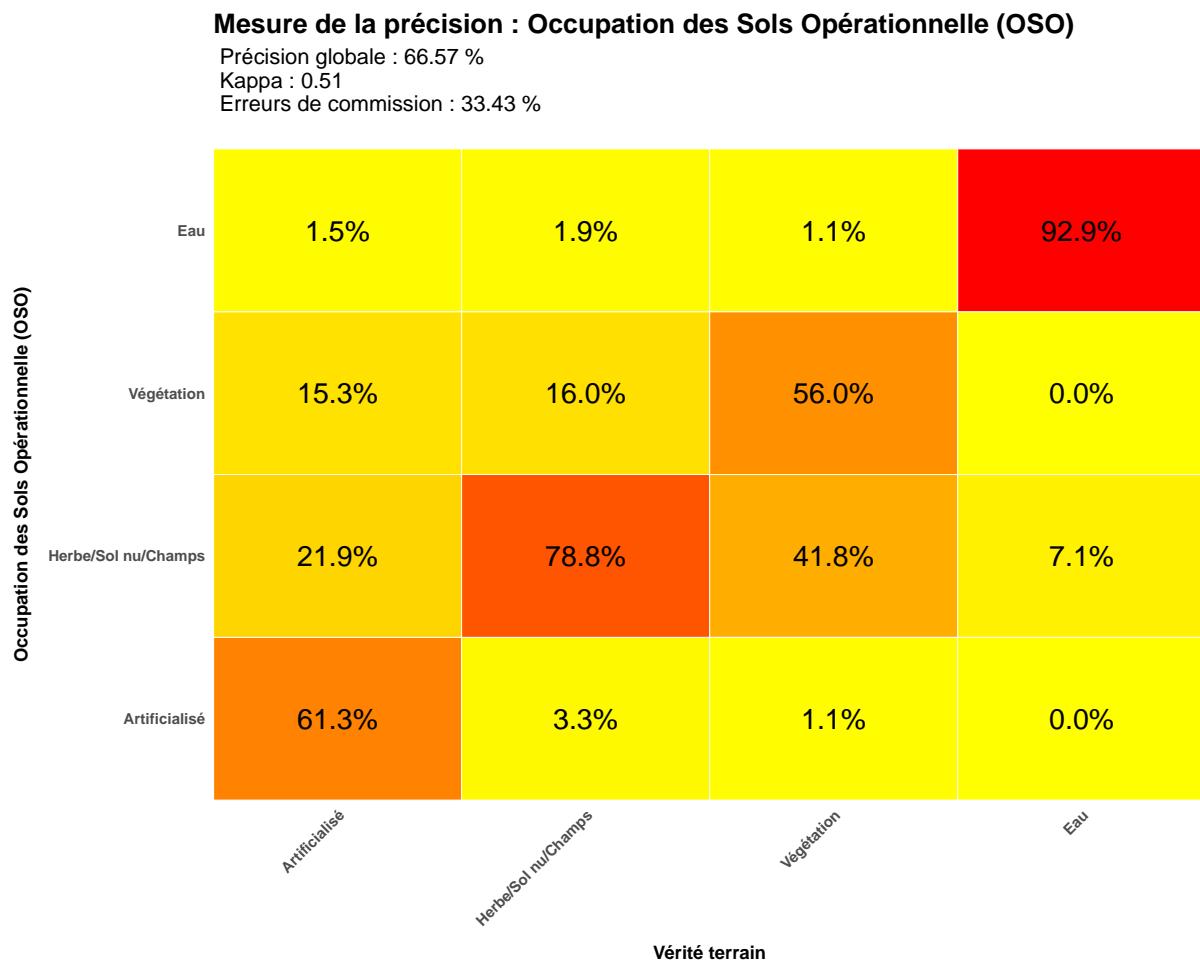


Figure 24. Mesure de la précision : Occupation des Sols Opérationnelle (OSO).
N. Massot, 2024

L'ESRI Land Cover est la couche d'occupation du sol qui a la précision globale la plus faible. Elle ne dépasse pas les 60 %. Elle fait tout de même preuve d'une bonne précision en ce qui a trait à l'eau et à l'herbe/sol nu/champs, mais au même titre que l'OSO et l'ESA World Cover, on observe un taux important d'erreurs de commission pour l'artificialisé et la végétation.

Mesure de la précision : ESRI Land Cover

Précision globale : 59.88 %
Kappa : 0.52
Erreurs de commission : 40.12 %

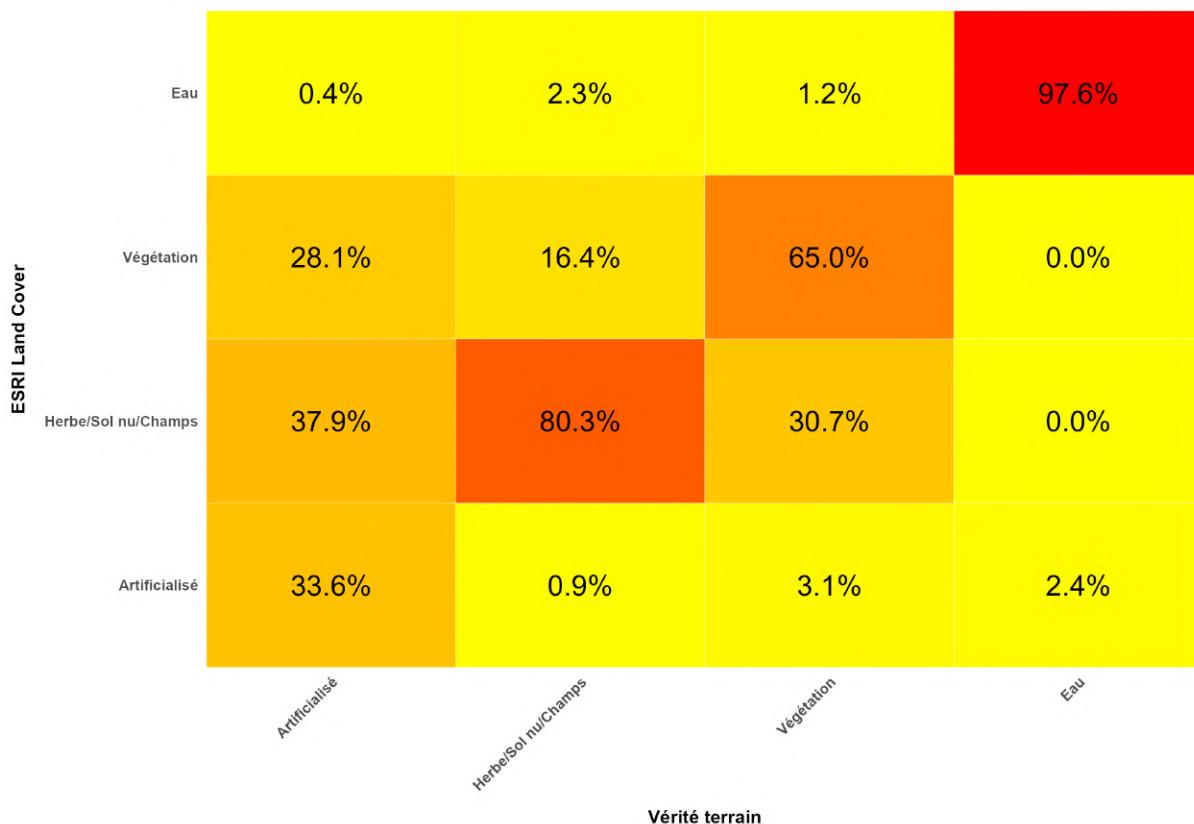


Figure 25. Mesure de la précision : ESRI Land Cover.
N. Massot, 2024

Les dernières données d'occupation du sol de l'étude sont issues d'OpenStreetMap pour la clé '*landuse*'. Contrairement à toutes les données précédentes, les données d'OSM ne sont pas présentes en tout point de l'espace. 27 % de la superficie du Grand Avignon n'est pas cartographié. Malgré cela, les données OSM se distinguent par leur précision pour toutes les classes, modulo l'*artificialisé*, qui présente des erreurs de commission de 46.5 %. La précision des trois autres classes est autour des 90 %, ce qui est un très bon score. De ce fait, la précision globale est de 80 % et le coefficient de kappa est à 0.71, ce qui représente un degré d'accord d'un bon niveau. Ces deux indicateurs sont les plus élevés de notre étude comparative. Cela met en évidence la puissance de la cartographie libre et participative qui surpassé de ce point de vue les productions des plus grandes agences (ESRI, ESA, IGN, ...)

Mesure de la précision : OpenStreetMap

Précision globale : 80.13 %
Kappa : 0.71
Erreurs de commission : 19.87 %

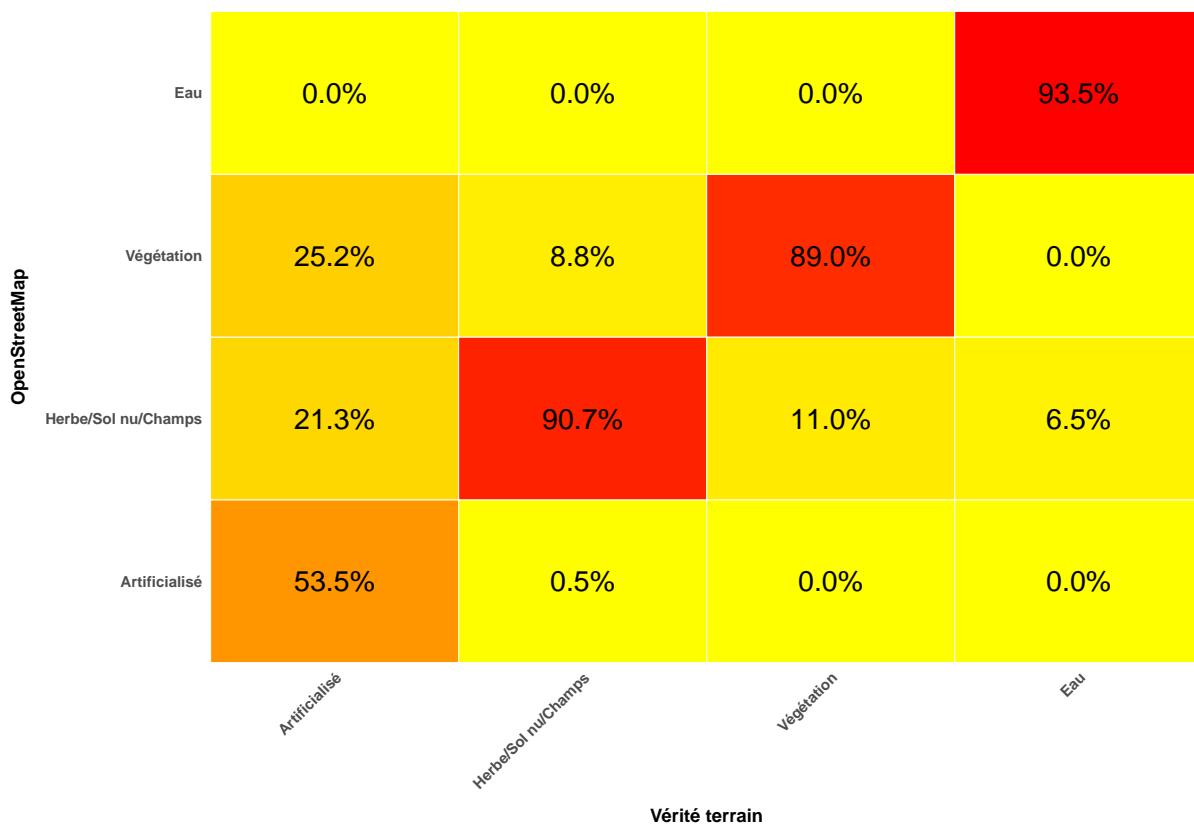


Figure 26. Mesure de la précision : OpenStreetMap.
N. Massot, 2024

3.2 Données contextuelles discontinues

Cette section vise à établir un état de l'art des données contextuelles discontinues dans l'espace, c'est-à-dire des informations qui ne sont pas uniformément réparties sur l'ensemble du territoire. Ces données, bien que non continues, pourraient s'avérer précieuses pour enrichir et guider la réalisation du masque d'entraînement, en apportant des informations contextuelles ou spécifiques sur certaines typologies de zones.

3.2.0.1 LiDAR HD IGN

Contrairement à toutes les données vues jusqu'ici, le LiDAR HD de l'IGN ne se base pas sur des données issues de l'imagerie satellitaire. En 2023, l'IGN s'est lancé dans la production du LiDAR Haute Densité, qui permet une cartographie 3D du sol et du sursol de l'ensemble du territoire français, en utilisant un système de balayage par laser aéroporté qui mesure les distances entre un point situé sur un avion et des points au sol. Associé à un GPS et à une centrale inertuelle, il permet de générer un nuage de points géoréférencé avec une densité de dix impulsions lumineuses (dans le proche infrarouge) par mètre carré. À date, le LiDAR HD est disponible pour environ $\frac{2}{3}$ des départements français, et ceux-ci sont principalement localisés dans le sud et l'est de la France.

Outre les coordonnées x, y et z qui définissent la position spatiale du point dans un système de référence donné, ce produit a pour particularité d'être classifié. Cela veut dire que pour chaque point, on connaît sa typologie selon les classes mentionnées dans la figure 27.

LiDAR HD	
Non classé	
Sol	
Végétation faible	
Végétation moyenne	
Végétation forte	
Bâtiement	
Eau	
Tablier du pont	
Sursol pérenne	
Points virtuels	
Divers - bâti	

Figure 27. Classes du LiDAR HD.
N. Massot, 2024. Source : IGN

Ainsi, dans l'optique de réaliser le masque d'entraînement, le LiDAR HD pourrait être utile pour plusieurs raisons. Premièrement, sa classification associée à sa haute

densité de points (10 pulsations par m²) permettrait de détourer les éléments thématiques avec une grande précision. Ces données altimétriques pourraient être utilisées en entrée du modèle sous forme de MNT ou MNS, comme cela a été fait entre autres par Kwan et al. 2020, et Garioud, Peillet et al. 2022.

3.2.0.2 BD Topo

La BD Topo est une base de données vectorielle développée par l'IGN depuis 2009, et mise à jour tous les trimestres depuis 2019. Elle est structurée en 60 couches, dont 57 couches géographiques qui couvrent de manière exhaustive et cohérente l'ensemble du territoire national. Ces données permettent de représenter et d'analyser une grande variété d'éléments géographiques, répartis dans les catégories suivantes :

- **Administratif** : Limites et unités administratives (communes, départements, régions, etc.);
- **Bâti** : Constructions, incluant les bâtiments résidentiels, industriels et publics;
- **Hydrographie** : Éléments liés à l'eau (rivières, lacs, canaux, etc.);
- **Lieux nommés** : Lieux ou lieux-dits possédant un toponyme, décrivant un espace naturel ou un lieu habité ;
- **Occupation du sol** : Végétation, estran, haies et autres types de couverture du sol ;
- **Services et activités** : Services publics, stockage et transport des sources d'énergie, lieux et sites industriels ;
- **Transport** : Infrastructures du réseau routier, ferré et aérien, ainsi que les itinéraires ;
- **Zones réglementées** : Zonages soumis à des réglementations spécifiques (parcs naturels, zones protégées, etc.).

Dans notre cas, la BD Topo pourrait être utile pour créer le masque d'entraînement, notamment sur les bâtiments, l'hydrographie ou encore les tronçons routiers. Toutefois, comme on peut le constater dans la figure 28, on observe de nombreuses inexacititudes quant à la précision géométrique de ces données. Ces dernières font preuve de décalages plus ou moins prononcés, ou tout simplement d'oublis. Ainsi, dans la réalisation du masque d'entraînement, il faut être attentif à l'intégrité des données utilisées. Notons que l'IGN rédige des rapports de contrôle qualité pour chaque millésime départemental de la BD Topo, et les taux d'accords sont souvent proches ou égaux des 100 % (IGN 2024c), ce qui questionne la pertinence de la métrique utilisée pour quantifier la précision des données.

Incohérences de la BD Topo

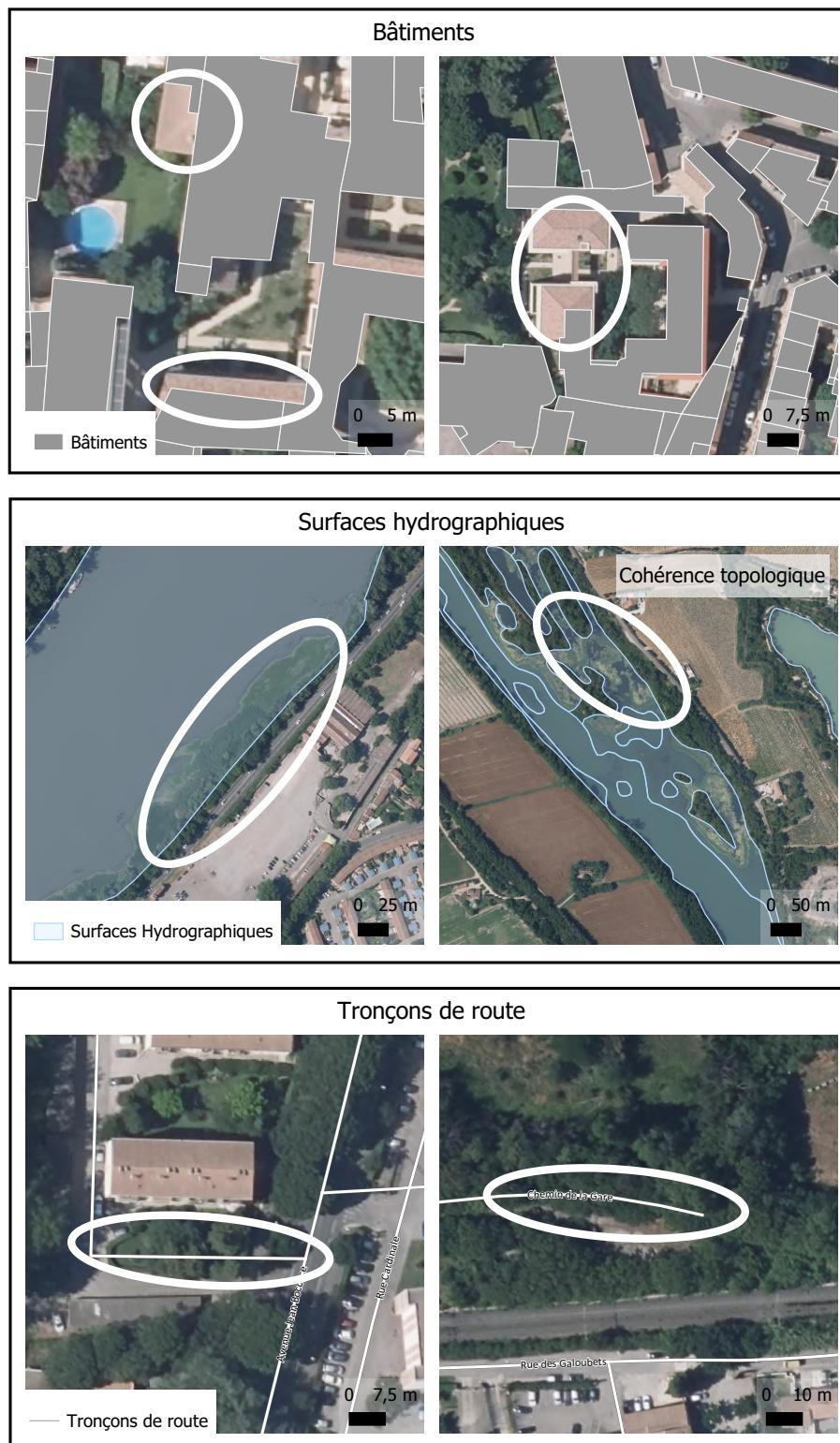


Figure 28. Incohérences de la BD Topo en comparaison avec la BD Ortho 2024 sur Avignon.
N. Massot, 2024. Source : IGN

3.2.0.3 Registre Parcellaire Graphique

Le Registre Parcellaire Graphique (RPG) constitue le référentiel géographique principal pour l'identification des parcelles agricoles sur le territoire français. Depuis 2007, il fait l'objet d'une mise à jour annuelle, s'appuyant sur les déclarations des exploitants agricoles transmises à l'Agence de Services et de Paiement (ASP)³⁶. Cette mise à jour régulière permet d'intégrer les évolutions dans l'utilisation des terres, en distinguant 664 types de cultures différents. Ces données jouent un rôle central dans la mise en œuvre de la Politique Agricole Commune (PAC)³⁷, en permettant le suivi et le contrôle des surfaces agricoles déclarées par les exploitants, dans le but de percevoir des aides financières européennes. De ce fait, les exploitants ne percevant pas d'aides ne sont pas présents dans la base de données, ce qui constitue un biais majeur.

En effet, les données déclaratives du RPG ne sont pas exhaustives et intègres. Comme le montre la figure 29, on observe d'une part un certain nombre de parcelles manquantes, non référencées dans la base de données, et d'autre part, des parcelles regroupées sous un seul polygone, ce qui peut entraîner des imprécisions dans la représentation des surfaces agricoles et dans le décompte des parcelles. Ces limites soulignent la nécessité de compléter les données déclaratives par des vérifications terrain et des outils de télédétection pour améliorer la fiabilité du référentiel.

36. L'ASP est un établissement public administratif chargé de la gestion des aides publiques et de l'accompagnement à l'emploi et à la formation professionnelle.

37. La Politique Agricole Commune (PAC) est une politique européenne visant à soutenir les agriculteurs, garantir la sécurité alimentaire, promouvoir des pratiques agricoles durables et assurer le développement rural.

Incohérences du registre parcellaire graphique

Données manquantes (champs et serres)



Un polygone mais plusieurs parcelles



Figure 29. Incohérences dans le Registre Parcellaire Graphique.

N. Massot, 2024. Source : IGN

L'IGN travaille actuellement sur un nouveau produit nommé "Occupation du sol agricole". Ce projet financé par le Ministère de l'Agriculture mobilisera des méthodes à base d'IA, la BD Ortho et des images Sentinel-2. Il permettra de référencer l'ensemble de terres agricoles à l'échelle nationale.

3.2.0.4 BD Forêt

Finalement, la BD Forêt recense toutes les formations forestières homogènes de même essence dont la superficie excède 0,5 hectare. Ces formations sont qualifiées par une nomenclature à 32 classes qui qualifient le type de forêt (IGN 2024c). L'IGN a adopté la définition internationale de la forêt de la FAO, en respectant les critères structurels suivants :

- Taux de couvert arboré : supérieur à 10 %
- Largeur : supérieure à 20 mètres
- Surface : supérieure à 0,5 hectare
- Hauteur potentielle des arbres à maturité : supérieure à 5 mètres

Au-delà de cette définition morphologique, la FAO s'appuie sur l'aspect fonctionnel des forêts. En effet, selon cette définition, une forêt est bien plus qu'un simple ensemble d'arbres ou une surface mesurable. C'est un "écosystème complexe, vivant,

un espace de vie, culturel et spirituel, et un bien commun qui doit être protégé" (Movement 2024). De ce fait, si les zones agroforestières entrent dans les clous des critères structurels, elles ne répondent pas à la définition fonctionnelle de la forêt de la FAO, et sont de ce fait exclues de la BD Forêt. Toutefois, dans sa documentation, l'IGN appuie sur le fait qu'il existe tout un gradient entre les zones exclusivement forestières et les zones agroforestières. Ainsi, l'institut met en garde sur le fait que "des zones de forêt où l'usage agricole n'est pas prédominant ont donc pu être exclues du masque livré" (IGN 2024b), tout en indiquant que les prochaines BD Forêt auront un attribut "Forêt à usage agricole", et même "Forêt urbaine" qui ne sont pour l'instant pas pris en compte. En outre, les surfaces anthropisées de moins de 5 000 m² (maisons, parkings, champs ...) qui sont entièrement entourées par des zones forestières ont été incluses dans les masques.

L'IGN vise une actualisation de la BD Forêt tous les trois ans. Si pour le Vaucluse, la dernière version a été produite en 2022 à partir de la BD Ortho de 2021, pour le Gard, elle date de 2018. À l'échelle du Grand Avignon qui s'étend sur ces deux départements, cela constitue donc un anachronisme à prendre en compte dans la réalisation du masque d'entraînement, dans la mesure où les couverts forestiers évoluent relativement rapidement.

En outre, comme on peut le constater sur la figure 30, quelques incohérences subsistent dans la BD Forêt.

- polygones plus petits que 0.5 hectare et de ce fait, ne respectent pas la définition internationale de la forêt de la FAO (en haut à gauche)
- erreurs topologiques (en haut à droite)
- incohérences dans la mesure où ce qui semble être des landes sont référencées, tandis que les forêts ouvertes situées à côté ne le sont pas (en bas à gauche)
- espaces anthropisés entourés de forêt inclus dans le masque. Malgré le fait que la documentation de l'IGN justifie ce choix, cela constitue une source d'erreurs de commission

En résumé, il semble que la BD Forêt soit de bonne qualité à petite échelle, mais à grande échelle, nous remarquons un certain nombre d'incohérences.

BD Forêt & BD Ortho : de nombreuses incohérences



Figure 30. Incohérences de la BD Forêt en comparaison avec la BD Ortho 2021 pour le Vaucluse et 2018 pour le Gard sur le Grand Avignon.
N. Massot, 2024. Source : IGN

N.B. Les exemples d'incohérences présentés dans cette section ont tous été identifiés sur le territoire du Grand Avignon.

3.3 Conclusion

3.3.1 Des données de qualité intermédiaire

En conclusion, notre étude comparative des couches d'occupation du sol révèle une qualité en demi-teinte, autant pour les données contextuelles discontinues que pour les couches d'occupation du sol continues. La précision globale de ces dernières se situe fréquemment autour de 70 %, ce qui témoigne d'une performance intermédiaire, ce qui peut biaiser les analyses d'occupation du sol en séries temporelles (cf. table 5). Cependant, cette précision globale masque d'importantes disparités locales. En effet, tandis que certaines classes, comme la classe "eau", ou dans une moindre mesure celle de l'"herbe/sol nu/champs" affichent une bonne précision, les classes "artificialisé" ou "végétation" sont caractérisées par des métriques de précision situées dans la moyenne haute. En effet, nous observons des erreurs de commission récurrentes pour ces deux dernières, ce qui souligne des difficultés dans la séparabilité de ces classes.

Rang	Donnée	Précision globale
1	OpenStreetMap	80,13 %
2	OCS GE	78,95 %
3	CoSIA	70,68 %
4	Corine Land Cover	66,72%
5	ESA World Cover	66,72%
6	Occupation des Sols Opérationnelle	66,57 %
7	ESRI Land Cover	59,88 %

Table 5. Précision globale des différentes données de couverture du sol.
N. Massot, 2025.

Par ailleurs, sur le plan qualitatif, les résolutions spatiales et temporelles ainsi que la fraîcheur des données apparaissent comme des facteurs limitants majeurs. Ces limitations peuvent entraver le suivi précis et régulier de l'évolution de l'occupation du sol, notamment dans des contextes dans lesquels les changements sont rapides ou à fine échelle. Pour améliorer la pertinence de ces données, il est essentiel de disposer de données à plus haute résolution, tant spatiale que temporelle, avec des unités minimales de collecte plus petites, et de développer des méthodes de classification reproductibles. Au niveau des données contextuelles discontinues, nous avons vu qu'elles proviennent de différentes sources : acquisition télémétrique et satellitaire, fichiers fonciers et déclaratifs. En dépit du fait qu'elles soient considérées comme des référentiels communs, ces données révèlent un certain nombre de limites dans la qualité, la régularité et l'intégrité des données, malgré le fait qu'elles soient produites par un même organisme institutionnel.

Les couches d'occupation du sol actuelles offrent une base utile pour un large panel d'analyses, mais des améliorations techniques et méthodologiques sont nécessaires pour répondre aux besoins d'une gestion et d'une planification territoriale éclairée. Parmi elles, deux axes de progression sont les plus importants dans l'optique de répondre aux besoins du Grand Avignon :

- La reproductibilité de la méthode pour permettre un suivi régulier de l'évolution de l'occupation du sol
- La résolution spatiale pour mener des analyses à grande échelle

Ainsi, au regard des résultats qualitatifs et quantitatifs de cet état de l'art, l'hypothèse H₁ qui suggère que les données d'occupation du sol et les données contextuelles discontinues actuelles sont suffisantes pour mesurer la consommation d'espaces ou l'artificialisation est rejetée.

Malgré tout, dans l'optique de réaliser le masque d'entraînement, ces données constituent une base intéressante, mais rappellent que la vérification de "l'intelligence humaine" est indispensable pour garantir la qualité des données d'entrées du modèle, dans l'optique de pouvoir automatiser la production de couches d'occupation du sol par réseaux de neurones convolutifs.

Somme toute, ces résultats montrent que la problématique de la cartographie de l'occupation du sol est largement étudiée par de nombreux organismes (IGN, CEREMA, INRAE, Agence Spatiale Européenne, Pôle THEIA, ESRI, ...). Toutefois, les résultats qualitatifs et quantitatifs de l'étude comparative mènent à se demander s'il existe une mesure "parfaite" qui ferait foi. En faisant part de mes résultats à Boris Wattrelos, responsable de projet en innovation numérique à l'IGN, celui-ci a ajouté, à propos de l'OCS GE, que l'important, c'est d'avoir un outil homogène à l'échelle française qui soit validé et adopté par tous les acteurs. Cela permettra d'avoir une grille de lecture commune et de pouvoir différencier les territoires qui artificialisent le plus, sans se buter à des prédictats méthodologiques. Les données géographiques font toujours preuve de biais. Néanmoins, la présence supposée uniforme de ces biais sur l'ensemble du territoire national permet de comparer les territoires entre eux, et d'étudier l'évolution de l'occupation du sol en séries temporelles pour analyser les dynamiques spatiales. De plus, l'évolution des coûts de production est exponentielle à la précision de la classification. Passer de 90 % à 95 % de précision coûtera bien plus cher que de passer de 80 % à 85 %. Par ailleurs, certains élus locaux s'arrêtent à la mesure de la précision de la donnée qui est très dépendante des définitions fixées par la loi. Ainsi, toujours selon lui, le fait de s'embourber dans des débats stériles sur la définition, la métrique ou la méthode peut servir de refuge, car cela peut permettre

de gagner du temps pour ne pas agir sur la réduction du rythme d'artificialisation des sols.

3.3.2 Vers un outil de suivi de l'occupation du sol territorialisé

La cartographie de l'occupation du sol n'est pas une science exacte. Comme tout classifieur, elle possède des biais (cf. section 4.1.1). De plus, choisir à quelle échelle administrative l'outil est destiné à être déployé est central. Les méthodologies de cartographie de l'occupation du sol peuvent être déployées à l'échelle nationale ou de la collectivité territoriale. Les deux échelles de couvertures possèdent chacune des avantages et des inconvénients qui sont résumés dans le tableau 6. Les outils à l'échelle nationale, notamment ceux de l'IGN, ont pour principale qualité la finesse sémantique. En effet, tandis que CoSIA possède 16 classes³⁸, l'OCS GE en possède 14 pour la version "couverture", et 20 pour la version "usage". En outre, la plus grande force des outils nationaux est le fait d'être basés sur une même méthodologie qui permet :

- la comparaison des territoires entre eux,
- la comparaison d'un même territoire en séries temporelles.

Ces deux principaux avantages sont notamment permis par la capacité à disposer d'une puissance de calcul en adéquation avec l'objectif de couverture nationale. À l'inverse, les outils territorialisés sont plus agiles et permettent de millésimer à fine résolution temporelle. Enfin, ils sont plus adaptatifs en permettant de personnaliser la composition des classes d'occupation du sol en fonction des besoins territoriaux. Idéalement, la meilleure solution est de combiner les données nationales de l'IGN (OCS GE et CoSIA) avec un outil de mesure de l'occupation du sol territorialisé, afin de tirer parti des avantages des deux approches, qui sont souvent antagonistes. De ce fait, la combinaison d'un outil national avec un outil territorialisé permet de diversifier les usages et de répondre plus précisément aux besoins des collectivités territoriales. Face à ces constats, des critères de qualité sont définis en fonction des besoins spécifiques du commanditaire.

Cela permet de valider l'hypothèse H₂, selon laquelle la précision est une notion subjective relative aux besoins du commanditaire.

38. plus la classe "Autre."

Avantages	Outil national	Outil territorialisé
Résolution temporelle	-	+
Résolution spatiale	= (CoSIA)	=
Résolution sémantique	+	-
Agilité	-	+
Comparaison entre les territoires	+	-
Comparaison en séries temporelles	=	+

Table 6. Comparaison des concepts d'outils nationaux et territorialisés.
N. Massot, 2025

Chapitre 4 :

Méthode pour la création de cartes d'occupation du sol par réseaux de neurones convolutifs

Ce chapitre permettra au lecteur d'appréhender la méthode conçue et mobilisée dans le but de segmenter l'occupation du sol du Grand Avignon en utilisant les réseaux de neurones convolutifs. Cette chaîne de traitements multi-étapes mobilise de nombreux outils, méthodes et concepts (logiciels, scripts Python et R, données, ...). Son objectif est de combler certaines lacunes identifiées dans le chapitre 3 au regard des besoins métiers des collectivités comme le Grand Avignon. De plus, les pistes de recherches qui n'ont pas été fructueuses sont également évoquées. Enfin, à travers la conception de cette chaîne de traitements, une réflexion éclairée sur les moyens d'optimisation de la méthodologie a été menée. Ces considérations ont nécessité de nombreux essais et erreurs, impliquant de multiples itérations dans la chaîne de traitements pour la rendre plus efficace, efficiente et cohérente.

4.1 Réseaux de neurones convolutifs et U-Net : définitions et concepts

4.1.1 De la classification aux réseaux de neurones

Une classification est un système organisé et hiérarchisé de catégorisation d'objets. Elle est également définissable comme étant le regroupement d'objets ou de nombres selon leurs caractéristiques ou propriétés communes, pour faciliter l'organisation, la compréhension et la résolution de problèmes. Les classements sont des outils essentiels pour organiser les données au sein d'un ensemble (Wikipedia 2025a). Ils sont utilisés dans de nombreux domaines, allant de la catégorisation d'espèces vivantes, en passant par les maladies, les éléments chimiques, etc. La classification d'individus est aussi et surtout un domaine des mathématiques et des statistiques. Dans leur papier nommé "*Do we Need Hundreds of Classifiers to Solve Real World Classification Problems ?*", Fernández-Delgado et al. 2014 ont réalisé un *benchmark* de 179 classifieurs³⁹. En mettant en avant le fait que l'algorithme de classification *Random forest* est le meilleur de leur étude avec une *accuracy* de 94 %, ils rappellent toutefois que le classifieur idéal n'existe pas, et que la méthode doit s'adapter aux besoins métiers. En effet, de manière schématique, chaque classifieur a sa propre finalité. Selon (Pelletier 2017), le choix entre différents classifieurs se base sur :

- le paramétrage
- le temps de calcul
- la stabilité
- la robustesse
- la précision

De nombreuses méthodes plus ou moins robustes sont utilisées pour la cartographie de l'occupation du sol (cf. figure 31). En télédétection, on distingue traditionnellement les analyses d'images orientées objet (OBIA) des approches pixels à pixels. Ces dernières considèrent les pixels comme primitive⁴⁰, et étudient leurs valeurs radiométriques à travers les signatures spectrales en faisant fi des valeurs des pixels avoisinants. Cette approche vise à regrouper les pixels selon leur valeur de réflectance à l'aide d'un ensemble limité de facteurs explicatifs : les bandes spectrales (Körting et al. 2013). À l'inverse, les méthodes d'analyses d'images orientées objet consistent à segmenter une image en objets distincts, en utilisant des techniques de classification basées sur l'aspect spectral, textural, contextuel et morphologique. Les techniques de segmentation employées pour extraire des objets à partir d'images doivent être conçues non pas pour identifier des objets statiques, mais pour détecter

39. Hors apprentissage profond.

40. Une primitive est l'unité de base utilisée pour l'analyse, comme un pixel dans une image satellite.

des zones susceptibles de changer par rapport au reste de l'image (Smith 1995). L'algorithme de Baatz and Schäpe est un bon exemple de cette méthode. Il minimise l'hétérogénéité spectrale (écart-type des reflectances) et spatiale (compacité et lissage) en agrandissant les régions de manière itérative (Nigri Happ et al. 2010).

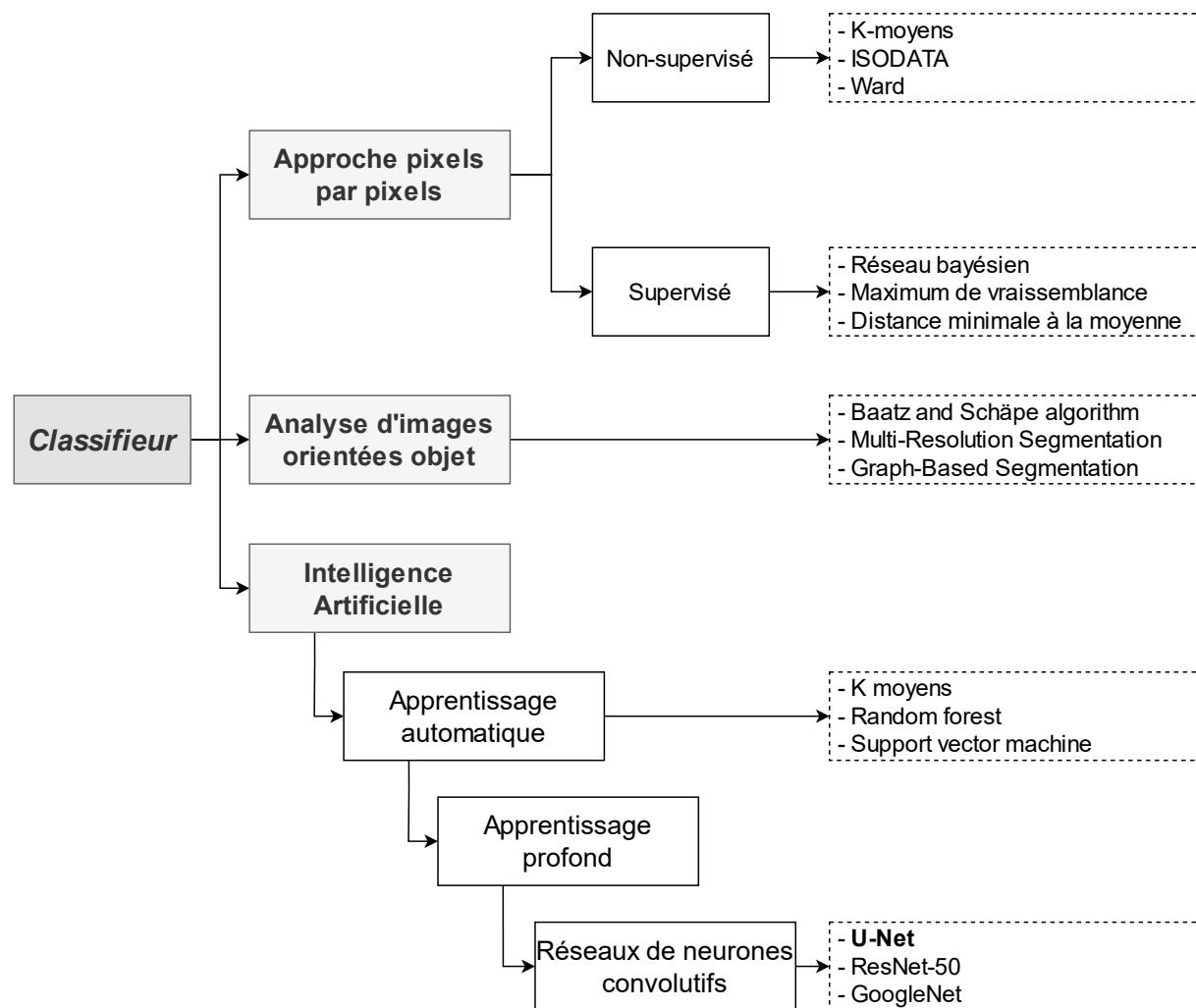


Figure 31. Typologie non-exhaustive des classificateurs pour la cartographie de l'occupation du sol.

N. Massot, 2025

Les classificateurs basés sur l'analyse des pixels, et ceux sur l'analyse d'images orientées objet ont tous deux des avantages propres à chacun. **La valeur ajoutée de l'apprentissage profond à la cartographie de l'occupation du sol réside dans la capacité de cette famille de méthodes à combiner les avantages des approches pixels à pixels et des approches orientées objets.** À ces avantages, s'ajoute le fait que les structures spatiales parfois complexes de l'occupation du sol sont apprises, enregistrées et généralisées dans les poids du réseau de neurones lors de la phase d'apprentissage, à la manière des neurones biologiques. Cette singularité apporte une troisième plus-value à ce type de méthodes. Ainsi, l'apprentissage profond contribue à la cartographie de l'occupation du sol de trois manières :

1. en exploitant les valeurs radiométriques des pixels,
2. en adoptant une approche orientée objet,
3. en apprenant automatiquement des structures spatiales complexes difficiles à modéliser.

4.1.2 Le modèle U-Net

Les réseaux de neurones convolutifs, inventés par Kunihiko Fukushima en 1980 (Fukushima 1980), et largement popularisés par la suite par Yann LeCun dans les années 1990 avec les modèles LeNet (Wikipedia 2025b), ont d'abord été utilisés pour des applications théoriques basiques telles que la reconnaissance de caractères. Par la suite, Ronneberger et al. 2015 ont introduit le modèle U-Net qui avait pour finalité de segmenter les cellules biologiques à partir d'une image de microscope optique. À partir de ce moment-là, les cas d'usages des réseaux de neurones convolutifs se sont multipliés, et sont aujourd'hui utilisés dans de nombreux domaines. Parmi eux, on peut citer la segmentation des routes, des trottoirs et des obstacles pour la voiture autonome, la détection de défauts de fabrication sur des pièces mécaniques, l'identification des maladies des plantes et la segmentation de l'occupation du sol.

Comme mentionné dans la figure 31, il existe un certain nombre d'architecture de réseaux de neurones convolutifs. Dans ce mémoire, nous utiliserons U-Net, car c'est le plus adapté à nos besoins. En effet, il est particulièrement efficace dans les contextes où les données annotées et où la puissance de calcul sont limitées. De plus, son architecture symétrique, qui sera décrite par la suite permet de combiner les informations globales et les détails locaux. Elle facilite la récupération des contours fins et des structures précises, souvent perdues lors des opérations de convolution et de *pooling*. Grâce à ces caractéristiques, U-Net offre une grande précision

dans la délimitation des objets, ce qui le rend particulièrement adapté à nos besoins.

Le modèle U-Net permet de faire une segmentation multi-classes de l'occupation du sol. Pour ce faire, lors de la phase d'entraînement, il prend en entrée deux données :

- les images d'entraînement
- le masque d'entraînement

Ensuite, lors de la phase d'inférence, il ne prend qu'un raster d'inférence qui a les mêmes caractéristiques⁴¹ que l'image d'entraînement, ainsi que le modèle entraîné.

Les images d'entraînement sont des facteurs explicatifs, qui peuvent être des bandes spectrales provenant d'images satellites ou d'orthophotos. Traditionnellement, elles incluent les canaux rouge, vert et bleu, mais il est possible d'ajouter davantage de bandes. Cet aspect-là sera détaillé plus tard. En revanche, le masque d'entraînement est une image mono-bande, où chaque pixel porte une valeur catégorielle entière allant de 1 à n, avec n représentant le nombre de classes que le modèle doit apprendre à prédire. Chaque classe est ainsi identifiée par un numéro. L'association de ces deux jeux de données (les images et le masque d'entraînement) permet au modèle de trouver les bons poids pour associer les valeurs des images d'entraînement (facteurs explicatifs) à la vérité terrain (masque d'entraînement). Cela lui permet, au fil des itérations dans le réseau, d'apprendre à effectuer les bonnes prédictions en mettant à jour les poids.

Le modèle U-Net repose sur une architecture en forme de "U" qui allie un chemin contractant (encodeur) et un chemin expansif (décodeur) (cf. figure 32). Tandis que la phase d'encodage permet de capturer le contexte de l'image, la phase de décodage permet de localiser précisément les individus tout en gardant le contexte de l'image fourni par la phase d'encodage. Ainsi, ces deux phases de descente et de remontée permettent d'extraire à la fois les caractéristiques fines de l'occupation du sol, ainsi que les informations contextuelles plus larges⁴².

41. Composition, ordre des bandes, et résolution spatiale.

42. À l'image d'un peintre qui dessine les grandes lignes d'un paysage (l'encodage) avant de repasser au pinceau plus fin pour ajouter les détails (décodage).

L'architecture U-Net

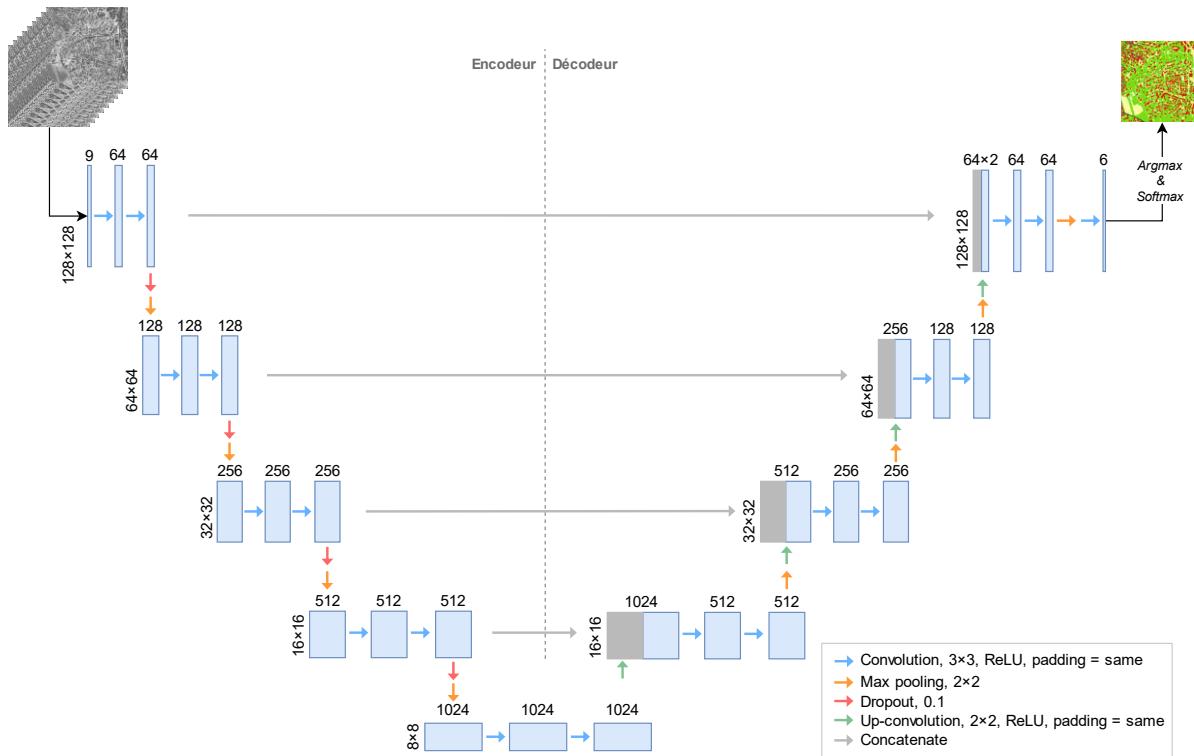


Figure 32. Architecture U-net. Chaque boîte bleue correspond à une image multi-canaux. Le nombre de canaux est marqué au-dessus de chaque boîte. Les boîtes grises représentent les images concaténées. La dimension des images est marquée en bas à gauche de chaque boîte.

N. Massot, 2025

Dans un premier temps, l'image satellite multi-canaux est soumise à une série de convolutions successives (cf. figure 33). L'opération de convolution consiste, pour chaque position du noyau de convolution (*kernel*), à calculer la somme des produits entre les valeurs d'un voisinage local de l'image (ici une fenêtre de 3×3 pixels) et les poids correspondants du noyau, à laquelle on ajoute un biais. C'est l'opération de convolution qui permet d'étudier le voisinage local de l'occupation du sol ; le résultat est ensuite passé à la fonction d'activation *Rectified Linear Unit* (ReLU)⁴³. L'équation d'une convolution s'écrit ainsi :

$$y = \sigma \left(\sum \text{input} \otimes \text{kernel} + b \right)$$

Où :

y est la sortie de la convolution,

σ est la fonction d'activation ReLU,

input est le tenseur d'entrée (ici, la valeur du pixel de l'image à segmenter),

kernel est le noyau de convolution appliqué à l'*input*,

b est le biais ajouté à la somme,

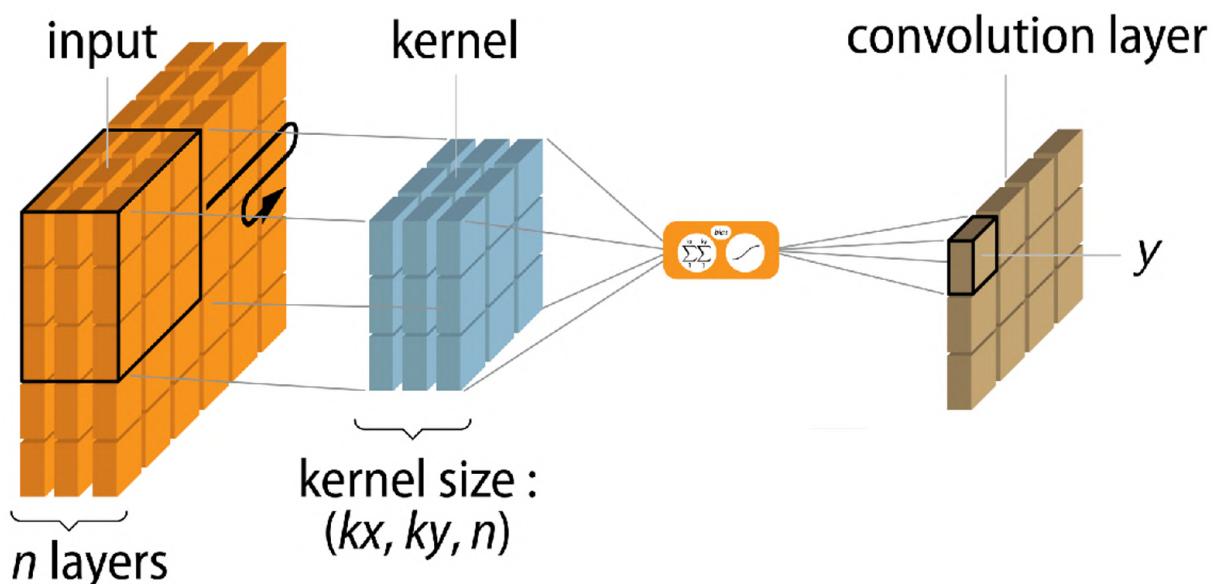


Figure 33. Convolution multi-canaux.

Source : FIDLE, 2024. Programme FIDLE 2024/25. CNRS. <https://fidle.cnrs.fr/w3/programme.html>

43. La fonction ReLU (*Rectified Linear Unit*) est une fonction d'activation définie par $\text{ReLU}(x) = \max(0, x) = \begin{cases} x & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$. Elle introduit de la non-linéarité dans le réseau. En activant uniquement les valeurs positives, elle permet une propagation efficace du gradient et évite certains problèmes liés à d'autres fonctions comme la sigmoïde, notamment l'atténuation du gradient.

Ensuite, la couche de convolution obtenue passe dans le *max-pooling* (cf. figure 34). Cette étape permet de simplifier l'information géographique en divisant par deux la taille de l'image tout en gardant l'information la plus importante. Pour ce faire, une fenêtre de 2×2 balaye la couche de convolution, et garde la valeur la plus haute parmi les quatre pixels considérés. Ici nous utilisons le maximum qui permet de faire apparaître la saillance, mais il est possible d'utiliser la moyenne ou la médiane. Par ailleurs, le modèle utilise le *dropout*, qui est une fonction de régularisation qui consiste à désactiver aléatoirement un certain pourcentage des neurones d'une couche pendant l'entraînement. Cette technique permet de réduire le sur-apprentissage (*overfitting*), en forçant le réseau à ne pas dépendre d'un petit nombre de connexions et à ne pas apprendre "par cœur". Le *dropout* améliore donc sa capacité à généraliser sur de nouvelles données (cf. figure 35). Cette phase d'encodage construit des représentations de plus en plus abstraites, capables de distinguer les motifs caractéristiques des différentes catégories d'occupation du sol (bâtiments, arbres, routes, etc.).

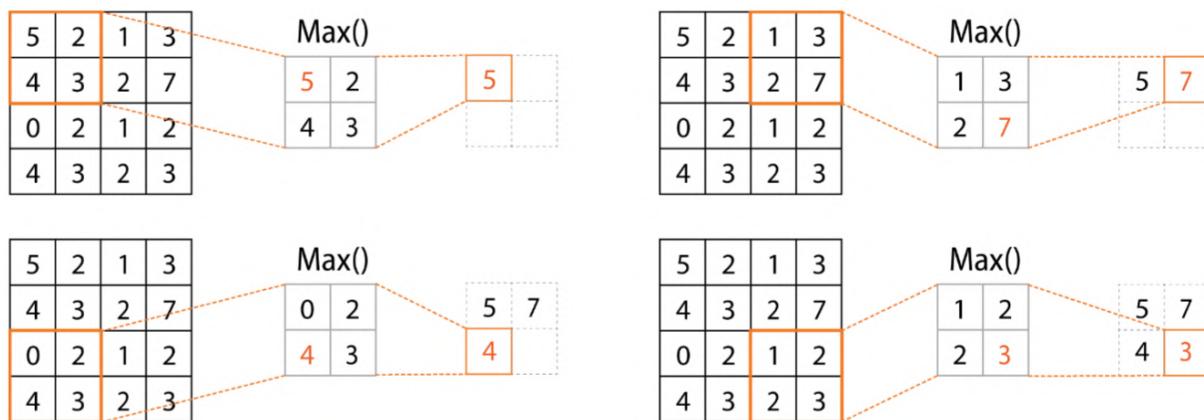


Figure 34. Max-pooling.

Source : FIDLE, 2024. Programme FIDLE 2024/25. CNRS. <https://fidle.cnrs.fr/w3/programme.html>

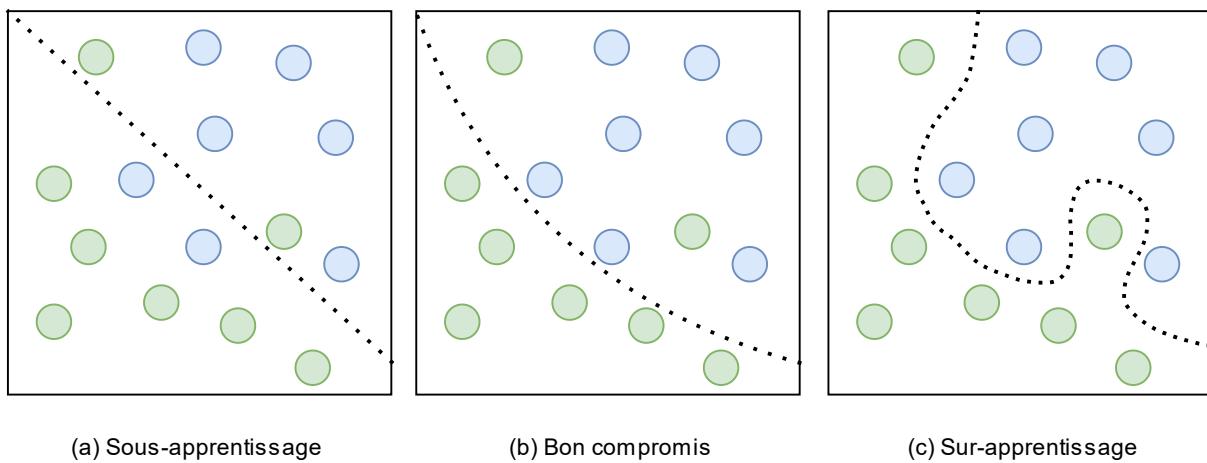


Figure 35. Sous-apprentissage, apprentissage approprié et sur-apprentissage en apprentissage profond.
N. Massot, 2025

La zone centrale de l'architecture du modèle est l'espace latent. Dans celui-ci, l'image est représentée par un très grand nombre de canaux à faible résolution spatiale, qui concentrent l'information la plus riche en raison de la succession des opérations de convolution et de *max-pooling*. Vient ensuite la phase de décodage, où chaque étape commence par une opération de convolution transposée (*up-convolution*) qui double la taille spatiale tout en réduisant le nombre de canaux. À chaque palier d'encodage est associé un palier de décodage. Ainsi, la phase de décodage est alimentée par la convolution transposée précédente, ainsi que par la concaténation du palier d'encodage associé. La concaténation (représentée en gris sur la figure 32) entre les cartes de convolution transposées, ainsi qu'avec celles issues de la phase d'encodage à résolution correspondante, réintroduit les détails spatiaux perdus lors de la réduction. Cette stratégie de "skip connections" garantit que les contours des objets segmentés soient préservés.

En sortie, le modèle retourne une dernière couche de convolution dont la taille est de $128 \times 128 \times 6$. À ce stade, chacune des six bandes et une carte de probabilité d'appartenance à une classe d'occupation du sol. Elles décrivent, pour chaque pixel, la probabilité d'appartenance à la classe en question. Ces probabilités sont obtenues en appliquant une fonction softmax, qui transforme les scores bruts de probabilités en une distribution normalisée dont la somme est égale à 1. Enfin, la fonction *argmax* génère l'image segmentée finale en assignant à chaque pixel le numéro de la classe ayant la probabilité maximale.

Lors de la phase d'inférence, le modèle est inféré sur une image qui comprend toute la zone d'étude (ici, le Grand Avignon), comme nous le verrons plus en détails par la suite. Cette image est découpée en patchs de 128×128 pixels, avec une

superposition entre chacun de ces patchs de 64 pixels. Après inférence de chacun des patchs, l'algorithme implémenté ici fusionne ces patchs pour reconstituer l'image segmentée de même dimension que celle que l'on a fournie au modèle. Pour ce faire, ces patchs sont fusionnés en utilisant la fenêtre de Hann. Cette méthode permet de pondérer la partie périphérique des patchs superposée lors de la fusion des cartes de probabilité (cf. figure 36). Elle attribue moins de poids aux pixels situés en périphérie qu'à ceux situés au centre. Cela permet d'avoir des transitions douces entre les patchs et d'annuler les artefacts de bordure (effets de bord). Enfin, pour produire la segmentation finale, on applique également une fonction *argmax* qui crée l'image finale en attribuant à chaque pixel la classe ayant la probabilité la plus élevée.

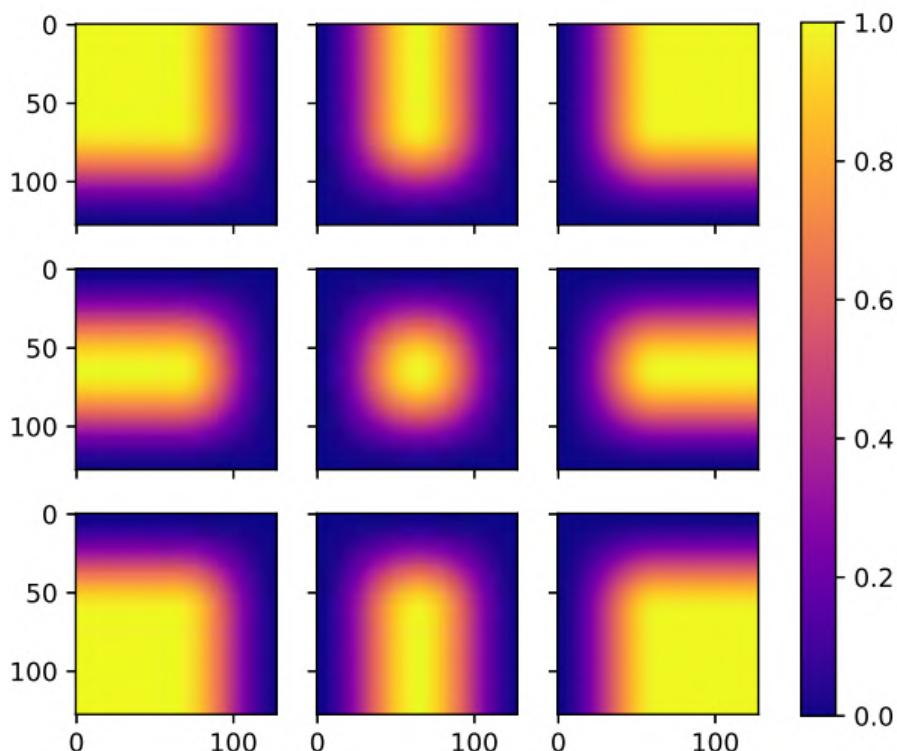


Figure 36. Pondération des pixels selon la fenêtre de Hann.

Source : N. Pielawski, and C. Wählby (2019), Introducing Hann windows for reducing edge-effects in patchbased image segmentation, <https://arxiv.org/pdf/1910.07831>

Enfin, l'entraînement du modèle repose sur l'optimiseur Adam (*Adaptive Moment Estimation*), un algorithme de descente de gradient⁴⁴ qui adapte dynamiquement le taux d'apprentissage⁴⁵ pour chaque paramètre, en tenant compte à la fois de la moyenne et de la variance des gradients passés. Cet optimiseur permet une

⁴⁴. La descente de gradient est une méthode d'optimisation qui consiste à ajuster progressivement les paramètres d'un modèle dans le sens opposé au gradient de la fonction de perte, afin de minimiser celle-ci.

⁴⁵. Le *learning rate*, ou taux d'apprentissage, est un hyperparamètre qui détermine la vitesse à laquelle un modèle ajuste ses paramètres lors de l'entraînement. Une valeur trop élevée peut provoquer une divergence de l'apprentissage en sautant le minimum de la fonction de perte, tandis qu'une valeur trop faible ralentit la convergence inutilement.

convergence rapide et stable vers le minimum de la fonction de perte⁴⁶. La fonction de perte utilisée est l'entropie croisée catégorielle (*categorical cross-entropy*), qui évalue la différence entre la distribution de probabilité prédictive et la distribution réelle. Elle pénalise fortement les erreurs de classification, ce qui renforce les prédictions proches de la vérité issue du masque d'entraînement.

4.2 Présentation de la chaîne de traitements

La cartographie de l'occupation du sol à l'aide du modèle U-Net ne se résume pas uniquement au modèle en lui-même tel que décrit précédemment. Plusieurs traitements, plus ou moins complexes, sont nécessaires en amont et en aval de celui-ci. Cette section vise à présenter la chaîne de traitements dans sa globalité, tandis que les suivantes ont pour objectif de détailler certains éléments de cette dernière. La chaîne de traitements comprend une vingtaine d'étapes qui sont conçues pour tendre autant que possible vers une automatisation à 100 %. L'ontologie de la figure 37 est le diagramme d'activité qui résume l'ensemble de la chaîne de traitements imaginée dans le cadre de ce mémoire. Cette section et les suivantes ont pour but d'expliquer ces étapes pas à pas.

Comme mentionné préalablement, le modèle utilise le masque d'entraînement comme vérité terrain, ainsi que les images d'entraînement en tant que facteurs explicatifs de l'occupation du sol. La création du masque d'entraînement est automatisée à l'aide d'un script FME. Les images, quant à elles, sont composées d'éléments spectraux et morphologiques calculés semi-automatiquement avec FME et QGIS. Les méthodes choisies pour créer ces derniers sont décrites dans les deux sections suivantes. Ensuite, si la précision du masque le permet, et si les images permettent de bien discriminer les classes, le masque et les images d'entraînements sont découpés en patchs de 128×128 . Les images sont ensuite normalisées avec des valeurs comprises entre 0 et 1 (`01_PREPROCESSER.py`). Par la suite, le modèle est entraîné avec ces données à l'aide d'un *notebook Python* (`02_ENTRAINER.ipynb`). Après l'entraînement, celui-ci calcule la fonction de perte (*loss*) ainsi qu'une mesure de la précision basée sur la comparaison avec le masque d'entraînement (*overall accuracy* et matrice de confusion). C'est à ce stade que des itérations ont été effectuées pour perfectionner les divers traitements faits en amont.

Ensuite, vient la partie opérationnelle. C'est celle-ci qui sera mobilisée pour inférer le modèle afin d'obtenir de nouveaux millésimes. Cette partie fait appel au modèle

46. La fonction de perte (*loss*) permet de mesurer à quel point les prédictions du modèle s'éloignent de la vérité terrain.

entraîné à l'étape précédente. Tout d'abord, pour des questions de ressources mémoire, il est nécessaire de découper le raster d'inférence (qui couvre les communes du Grand Avignon) en sept sous-ensembles égaux, dont les dimensions de largeur et la hauteur en nombre de pixels sont divisibles par 128, et qui se superposent de 32 pixels (`03_DECOUPER_SOUS_ENSEMBLE.py`). Ensuite, chacun de ces sous-ensembles est inféré à l'aide du script `04_INFÉRER.py`⁴⁷. Enfin, nous refusionnons les sept sous-ensembles avec la fenêtre de Hann (`05_FUSIONNER_SOUS_ENSEMBLE.py`). La chaîne de traitements se termine par l'application d'un filtre de majorité de 3×3 et la mesure de la précision.

47. L'algorithme d'inférence est composé des sous-étapes suivantes : découper le sous-ensemble d'inférence en patchs de $128 \times 128 \times 9$ avec une superposition de 64px, normaliser, inférer, fusionner avec la fenêtre de Hann.

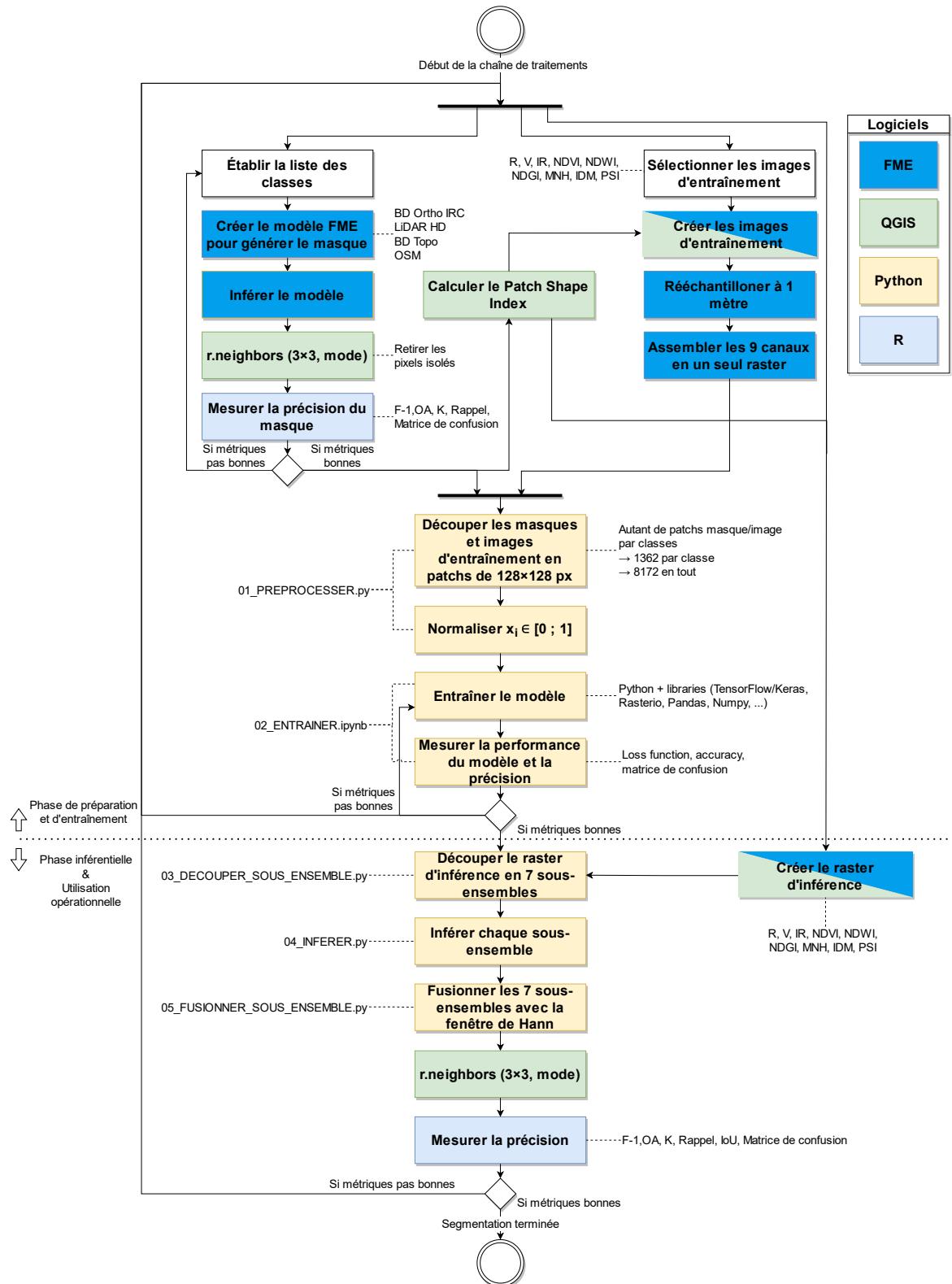


Figure 37. Chaîne de traitements mobilisée.
N. Massot, 2025

4.3 Présentation et méthode pour la réalisation du masque d'entraînement

L'objectif de la méthodologie est d'avoir une cartographie de l'occupation du sol à six classes et à un mètre de résolution. Pour que le modèle U-Net puisse prédire l'appartenance d'un pixel d'un mètre carré à l'une de ces six classes, le masque d'entraînement doit avoir ces mêmes caractéristiques (1 m^2 et six classes). Fournir au modèle un masque d'entraînement de très bonne qualité est primordial, dans la mesure où celui-ci va en grande partie conditionner la capacité du modèle à prédire l'occupation du sol. Par ailleurs, pour une puissance de calcul donnée, plus le nombre de classes augmente, plus il devient exponentiellement difficile de les distinguer les unes des autres. De ce fait, dans notre contexte, une composition à six classes semble être le meilleur arbitrage entre finesse sémantique et séparabilité des classes. C'est pourquoi le modèle prédira l'appartenance des pixels aux classes suivantes :

Classe d'occupation du sol	Couleur
Bâtiment	Red
Arbre	Green
Bitume	Grey
Eau	Blue
Herbe	Light Green
Sol nu	Yellow

Table 7. Composition des classes du modèle.
N. Massot, 2025

Par ailleurs, si la donnée de référence n'est pas assez précise, la mise à jour des poids du modèle sera faite à partir de données potentiellement incohérentes, ce qui affecte grandement la précision du modèle. Ainsi, ce travail fut l'occasion de tester plusieurs méthodologies pour faire le masque d'entraînement qui se sont caractérisées par des niveaux de pertinence variés. Avant de décrire la méthodologie finale adoptée pour la réalisation du masque d'entraînement, les trois pistes de recherches seront brièvement décrites.

4.3.1 Pistes de recherche pour la création du masque d'entraînement

4.3.1.1 Agrégation de données déjà existantes

En premier lieu, il paraît pertinent de mettre à profit les couches d'occupation du sol que les différents acteurs ont produit. De ce fait, l'idée première est de se baser sur l'état de l'art des données d'occupation du sol (cf. Chapitre 3). Dans ce chapitre, une partie du *benchmark* est consacré à l'étude de la précision de ces dernières. Son objectif est d'identifier les données les plus précises afin de déterminer la ou les meilleures d'entre elles de ce point de vue. Celui-ci nous indique que la donnée d'occupation du sol la plus précise provient de OpenStreetMap avec le tag "landuse", avec une précision globale de 80 %. Cependant, celle-ci n'est pas continue en tout point de l'espace⁴⁸, et la composition des classes d'occupation du sol n'est pas des plus adéquates (cf. figure 16). Ensuite, c'est l'OCS GE qui est deuxième de la liste avec une précision globale de 78.5 %. Malgré la finesse sémantique (cf. figure 5), l'OCS GE n'est pas une donnée brute. La version "occupation" se base en partie sur les données de CoSIA. Seules les géométriques de plus de 50 m² pour les bâtiments, et de plus de 2 500 m² pour le reste sont conservées, et celles-ci passent ensuite par une étape de régularisation, et sont agrégées à des données contextuelles discontinues qui proviennent par exemple de la BD Topo⁴⁹ ou du RPG. De ce fait, l'OCS GE n'est pas intéressant pour la réalisation du masque d'entraînement dans la mesure où celui-ci doit décrire l'occupation du sol à l'échelle du pixel. La troisième donnée d'occupation du sol en matière de précision est CoSIA, qui, malgré le fait qu'elle soit une donnée brute, a une précision globale de 70.6 %, ce qui est trop faible et aurait induit trop d'erreurs.

Par ailleurs, une autre possibilité réside dans l'identification de la précision des couches d'occupation du sol à l'échelle des classes. Il serait ainsi possible d'agrégier les classes les plus adéquates des différentes données pour créer une nouvelle couche d'occupation du sol composite. À titre d'exemple, admettons que nous partions sur une classification à quatre classes, comme celle utilisée pour mesurer la précision des couches d'occupation du sol (cf. Chapitre 3), il est possible de procéder comme ceci :

48. 27 % du Grand Avignon non-couvert.

49. Surfaces hydrographiques, tronçons de route, registre parcellaire graphique.

Classe	Source	Précision de la classe
Artificialisé	OCS GE	77.6 %
Herbe/sol nu/champs	ESA World Cover	91.1 %
Végétation	CoSIA	91.2 %
Eau	ESRI Land Cover/OCS GE	97.6 %

Table 8. Agrégation des classes d'occupation du sol les plus précises.

Les chiffres sont issus des matrices de confusion de la partie 3.1.2.2.

N. Massot, 2025

Toutefois, au-delà du fait que certaines données soient en mode vecteur et d'autres en mode raster, plusieurs autres facteurs limitants persistent. D'abord, le fait de fusionner ces quatre classes induirait fatallement la présence de zones sans données, ce qui n'est pas acceptable dans la mesure où le masque d'entraînement doit être continu en tout point de l'espace. Une interpolation des valeurs catégorielles est techniquement possible⁵⁰, mais l'occupation du sol est un entremêlement de structures complexes qui ne sont pas spatialement corrélées. Cela aurait induit des erreurs préjudiciables.

À ce premier frein, s'ajoute le fait que les résolutions spatiales de ces quatre données sont différentes. En effet, l'ESA World Cover et l'ESRI Land Cover sont tous deux à 10 mètres de résolution en mode raster. À l'inverse, l'OCS GE, étant en mode vecteur, est composée de polygones régularisés et agrégés à des données contextuelles qui ne permettent pas une rasterisation à un mètre qui soit assez précise.

4.3.1.2 Utilisation de Segment Anything Model 2

En deuxième lieu, l'utilisation de Segment Anything Model 2 (Kirillov et al. 2023) pour la réalisation du masque d'entraînement a été testé. SAM est un modèle de segmentation d'images dont les cas d'usage en traitements d'images sont nombreux. Le module python `samgeo` (Osco et al. 2023 & Wu et al. 2023) est un modèle optimisé spécifiquement pour les applications en segmentation d'images satellitaires. Celui-ci est utilisable de trois manières différentes :

50. Notamment avec les polygones de Voronoï.

Nom	Description
<i>Point prompts</i>	Segmentation d'une image en pointant les objets avec une couche de points
<i>Box prompts</i>	Segmentation d'une image à partir de boîtes englobantes
<i>Text prompts</i>	Segmentation d'une image avec un prompt textuel et deux paramètres d'exhaustivité

Table 9. Les trois manières d'utiliser samgeo.
N. Massot, 2025

Ces trois méthodes permettent toutes de segmenter un type d'élément du sol ou du sursol, mais leur mise en œuvre diffère. Tandis que la première a besoin d'une couche de points qui pointe les objets à segmenter, la deuxième prend en entrée une boîte englobante. Ces deux méthodes sont limitées dans la mesure où il est compliqué de pointer ou d'encadrer individuellement chaque type d'occupation du sol (cf. tableau 7) du Grand Avignon de manière exhaustive.

La troisième manière d'utiliser `samgeo` est d'utiliser la méthode "*text prompts*". Celle-ci se distingue par la possibilité d'entrer des prompts textuels, ce qui permet de décrire précisément les éléments à segmenter. Pour affiner davantage le type d'occupation du sol souhaité, il est également essentiel d'expérimenter avec divers prompts. Aussi, la méthode prend deux paramètres d'exhaustivité :

- *box_threshold* : $\in [0 ; 1]$. Une valeur plus élevée rend le modèle plus sélectif, identifiant uniquement les instances d'objets les plus sûrs, ce qui réduit le nombre total de détections.
- *text_threshold* : $\in [0 ; 1]$. Une valeur plus élevée nécessite une association plus forte entre l'objet et le texte, ce qui conduit à des associations plus précises, mais potentiellement moins nombreuses.

Outre le temps de traitement trop élevé pour inférer ce modèle sur tout le Grand Avignon, les résultats laissent à désirer. La figure 38 montre une tentative de segmentation des bâtiments et des arbres à partir de la BD Ortho RVB. Divers prompts ont été testés, et les segmentations ont été générées pour l'ensemble des combinaisons des deux coefficients avec un incrément de 0,1. La figure 38 montre la meilleure segmentation pour les arbres et pour les bâtiments. Dans celle-ci, nous constatons que certaines régions d'intérêt (bâtiments ou arbres) sont omises. D'autre part, le masque de segmentation manque de précision à l'échelle du pixel, et enfin, certaines zones ne faisant pas partie des régions d'intérêt sont incluses dans le masque. Une hypothèse pouvant expliquer le manque de précision de l'inférence de `samgeo` sur la zone d'étude résiderait dans le fait qu'il a été entraîné sur des paysages urbains et une végétation typique des États-Unis. Les formes urbaines et les types de végéta-

tion sont différents entre les deux pays. De ce fait il est possible que cela réduise la précision du modèle. De plus, cette imprécision pourrait également être liée à l'image utilisée. En effet, si l'étalonnage des canaux RVB et la résolution spatiale de l'image d'entraînement diffèrent de ceux de l'image d'inférence, cela pourrait affecter les résultats. Enfin, l'agrégation de plusieurs masques de segmentation en un seul laisse toujours deux problèmes : que faire lorsque deux masques se superposent, et que faire lorsqu'aucun masque ne couvre une zone ? Pour ces raisons, Segment Anything Model 2 utilisé à travers samgeo n'est pas adapté à la réalisation du masque d'entraînement.

Prompt : "buildings"
box_threshold : 0.2
text_threshold : 0.4



Prompt : "green vegetation that ressembles trees"
box_threshold : 0.2
text_threshold : 0.4



Figure 38. Tentative de segmentation des bâtiments et des arbres avec samgeo.
N. Massot, 2025. Source : BD Ortho, Google satellite

4.3.1.3 Occupation du sol multi-sources

La troisième piste de recherche pour réaliser le masque d'entraînement consiste à agréger quatre types de données : l'OCS GE, les bâtiments de la BD TOPO, un MNH⁵¹ et un indice NDVI calculé depuis la BD Ortho IRC. Pour ce faire, une grille vectorielle d'un mètre a été faite sur le Grand Avignon. À cette grille, une jointure spatiale concaténait les valeurs du NDVI et du MNH dans le champ "**Classe_NDVI_Hauteur**" comme on peut le voir sur les tableaux 10 et 11.

Ensuite, une intersection est faite entre L'OCS GE, les bâtiments de la BD TOPO et la grille (avec le champ Classe_NDVI_Hauteur). Enfin, une requête de conditions en cascade a été rédigée pour assigner une classe à chaque polygone selon les valeurs des champs associés. Cette requête se base sur le champ Classe_NDVI_Hauteur et

51. MNS-MNT

Table 10. Structure de la variable
Classe_NDVI_Hauteur.
N. Massot, 2025

Position	Valeurs	Signification
1er chiffre	1, 2, 3	Classe de hauteur : 1 = basse, 2 = moyenne, 3 = haute
2e chiffre	0	Chiffre fixe (séparateur)
3e chiffre	0, 1	Végétation selon NDVI : 1 = végétation présente, 0 = absente

Table 11. Interprétation de
Classe_NDVI_Hauteur
N. Massot, 2025

Valeur	Interprétation
100	Pas de végétation
101	Végétation basse
201	Végétation moyenne
301	Végétation haute

sur certaines classes de l'OCS GE. Elle abouti à une classification à 10 classes⁵² (cf. Annexe .1).

Toutefois, malgré le fait que ce masque soit continu en tout point de l'espace, cette méthode présente des limites (cf. figure 39). D'abord, les transitions entre les trois niveaux de végétation sont trop fréquentes (a). Ensuite, le détourage des bâtiments de la BD Topo ne suit pas la vérité terrain (b). Par ailleurs, le détourage des panneaux solaires est trop fin (c). Enfin, les réseaux ferrés pouvaient être soit des lignes de TGV, soit des faisceaux de triage ou encore du tramway, qui lui-même pouvait être composé de surfaces minérales ou d'herbe. Le modèle U-Net apprend des valeurs radiométriques, mais également des formes. Le fait d'avoir des détourages trop fins, des décalages entre la vérité terrain et le masque, ou plusieurs structurations différentes pour une même classe est un frein à l'apprentissage. Cette méthode de cartographie de l'occupation du sol multi-source est, des trois présentées jusqu'ici, la plus prometteuse. Toutefois, nous avons entraîné le modèle U-Net sur ce masque, et les résultats n'étaient pas exploitables, avec une sur-représentation du sol nu et un mauvais détourage de presque toutes les classes.

52. Bâtiment, Immeuble, Photovoltaïque au Sol, Réseaux ferrés, Réseaux routiers, Sol nu, Surfaces d'eau, Végétation basse, Végétation haute, Végétation moyenne

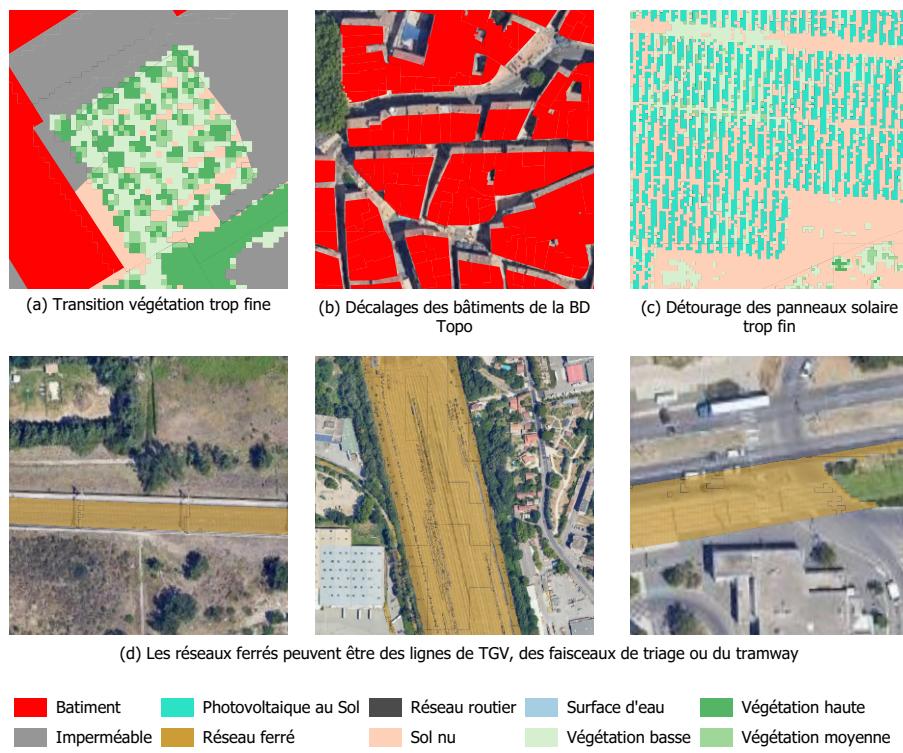


Figure 39. Limites de l'occupation du sol multi-sources.
N. Massot, 2025

4.3.2 Méthode adoptée

La méthode présentée dans cette section est celle qui a été utilisée pour réaliser le masque d'entraînement. Il s'agit d'une chaîne de traitements faite sous FME de manière automatique, standardisée et surtout, reproductible. Seuls quelques pré-traitements ont dû être faits sur QGIS. Cette méthode se base sur l'agrégation de plusieurs sources de données à une grille d'un mètre, dans le but d'assigner à chaque pixel l'appartenance à une classe (cf. tableau 7). Les sources de données sont multiples, et sont synthétisées dans le tableau 12. L'ensemble de la méthode est décrit dans le diagramme d'activité de la figure 41. La figure 12 synthétise la ou les source(s) de données utilisées pour le masque d'entraînement.

En premier lieu, les recherches effectuées montrent que le LiDAR HD de l'IGN se distingue par sa capacité à bien classifier les éléments du sursol, à une résolution spatiale, particulièrement fine de dix points par mètre carré. C'est cette source de données qui a été choisie pour les bâtiments (classification = 6) et pour les arbres (classification = 4 ou 5)⁵³. Les dalles LiDAR sont distribuées par carreaux d'un kilomètre de côté, qui contiennent chacun environ 21 millions de points. De ce fait, avant de commencer à traiter cette donnée, la première étape est d'affiner le nuage de points (boîte `PointCloudThinner` sur la figure 41). Cet algorithme ne garde qu'un

53. La classe n°4 correspond à la végétation moyenne, et la classe n°5 correspond à la végétation forte.

point par rayon de 30 cm, et permet de diviser par quatre le poids du nuage, tout en gardant suffisamment d'informations. Ensuite, on filtre le nuage pour ne garder que les bâtiments ou les arbres (dans deux couches séparées), avant de faire la jointure spatiale avec la grille d'un mètre en comptant les points par carreaux. À ce niveau-là, si un seul point classé bâtiment est contenu dans un carreau, tout le carreau est considéré comme bâti. Cela a pour effet de créer une zone tampon indésirable autour de tous les bâtiments et les arbres en englobant des types d'occupation du sol différents. C'est pourquoi, nous ne considérons comme bâtiment que les carreaux qui ont plus de deux points LiDAR classés comme bâtiment. Pour les arbres, le seuil est fixé à quatre points LiDAR. Ces seuils ont été fixés par vérifications visuelles. La figure 40 illustre qu'en appliquant ces seuils, on élimine des carreaux qui ne sont pas majoritairement composés de bâtiments (ou d'arbres), mais qui auraient été classés à tort comme tels en l'absence de ce filtrage.

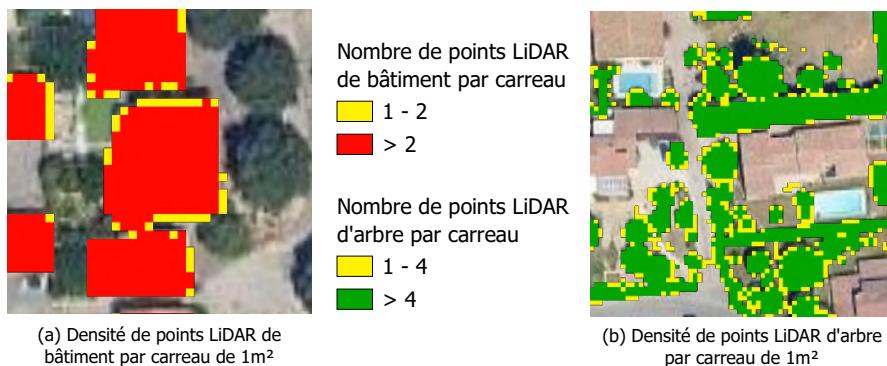


Figure 40. Effet du seuil de densité de point minimal sur le détourage des éléments du sursol.

N. Massot, 2025

Il est possible qu'un carreau puisse être classifié comme bâtiment et comme arbre. C'est pourquoi la série de trois `clipper` permet de fusionner les bâtiments et les arbres en donnant la priorité aux arbres (cf. figure 41). Lorsqu'un pixel est simultanément classé comme "arbre" et "bâtiment", la priorité est donnée à la classe "arbre", car les arbres apparaissent au-dessus des bâtiments sur les images satellitaires.

Pour classifier le bitume, nous agrégeons trois sources de données :

- Les tronçons de route de la BD Topo (`buffer` sur le champ *largeur_de_chaussee*),
- Les parkings d'OpenStreetMap (*amenity=parking*),
- Une classification supervisée par plus proches voisins avec un seuil de 10.

Les tronçons de la BD Topo sont triés pour ne garder que les routes. Ensuite, comme elles sont diffusées sous forme de lignes, un tampon a été effectué selon le

champ "*largeur_de_chaussee*". Parallèlement, une classification supervisée par plus proche voisin a été faite avec un seuil de 10, afin de classifier le bitume. Ensuite, nous fusionnons à ces deux premiers les parkings d'OpenStreetMap. Enfin, la priorité est donnée à la couche des bâtiments et des arbres avant de supprimer les polygones de moins de 8m².

Quatrièmement, les surfaces hydriques sont calculées à partir d'un seuil NDWI calculé depuis la BD Ortho (surfaces hydriques $\in [0.7; 1]$). Le résultat du seuil est ensuite fusionné avec les tronçons hydrographiques de la BD Topo. Cette couche est finalement fusionnée avec le bitume en lui donnant la priorité, avant de supprimer les polygones de moins de 3 m².

Pour l'herbe, un seuil NDVI est calculé : herbe $\in [0.005; 1]$. La borne maximale de ce seuil est volontairement permissive, dans la mesure où, cette donnée va être fusionnée avec les bâtiments, arbres, bitume et eau en leur donnant la priorité. Par exemple, un pixel classé comme "bâtiment" et "herbe", sera considéré comme bâti.

Enfin, après avoir tout fusionné, les pixels qui ne sont pas encore assignés à l'une des cinq premières classes sont considérées comme étant du sol nu. Cette méthode par déduction permet de combler tous les trous, et est particulièrement précise et apte à refléter la réalité terrain, comme le montre la matrice de confusion à la figure 42.

Classe	Source	Ordre de priorité
Bâtiment	LiDAR (classification = 6)	2
Arbre	LiDAR (classification = 4 ou 5)	1
Bitume	BD Topo (tronçon de routes), OSM (parking), Classification supervisée	3
Eau	BD Topo (surface hydrographique), BD Ortho (seuil NDWI)	4
Herbe	BD Ortho (seuil NDVI)	5
Sol nu	Le reste	6

Table 12. Sources des classes d'occupation du sol et ordre de priorité. Ce tableau synthétise le diagramme d'activité de la figure 41.

N. Massot, 2025

Apport des réseaux de neurones convolutifs à la cartographie de l'occupation du sol :
Cas d'usage sur le Grand Avignon

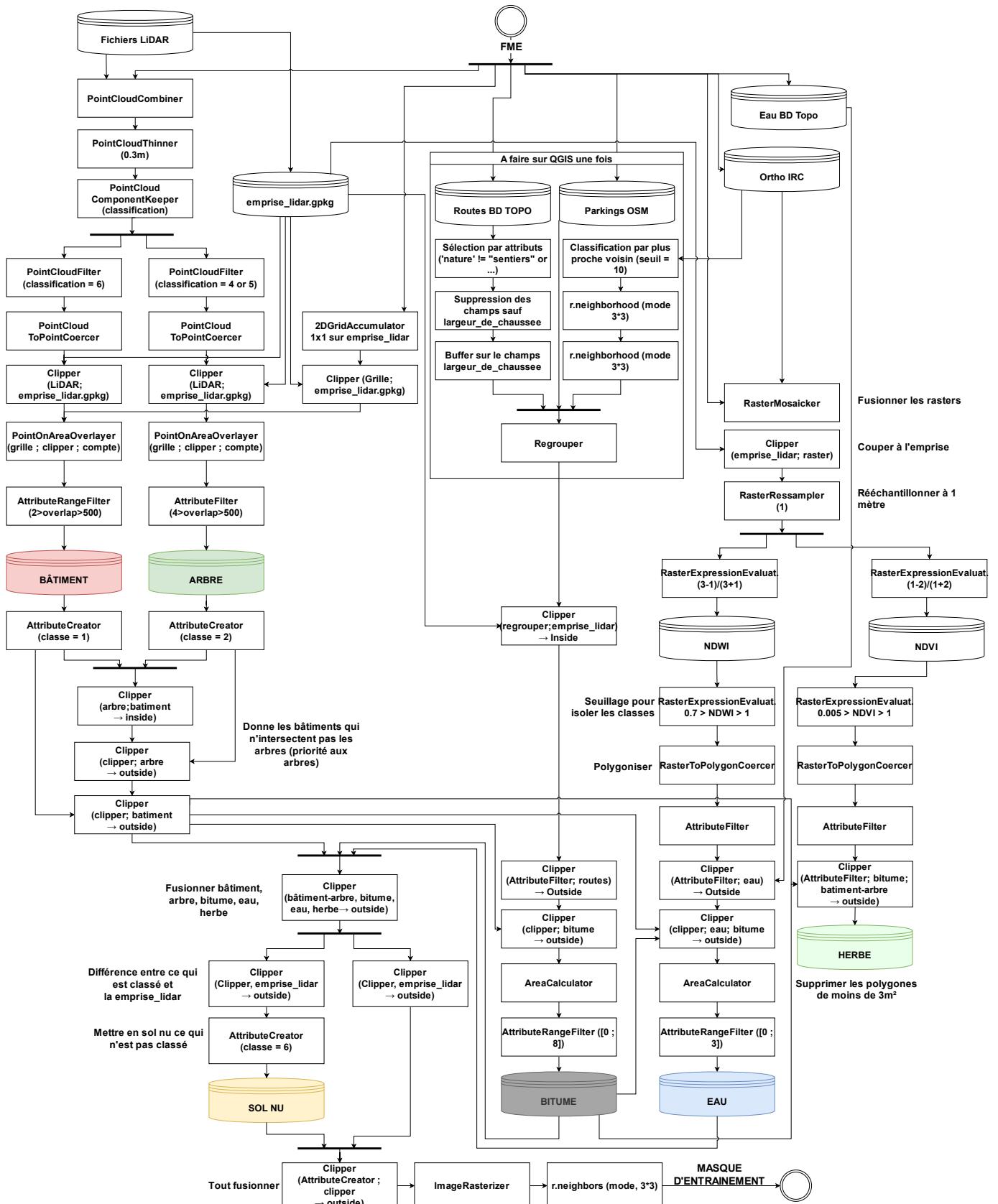


Figure 41. Chaîne de traitements pour la réalisation du masque d'entraînement avec FME.

N. Massot, 2025

4.3.3 Mesure de la précision du masque d'entraînement

Les métriques d'évaluation ainsi que la matrice de confusion (cf. figure 42) révèlent que la classification réalisée sur le masque d'entraînement présente une performance élevée. L'accuracy⁵⁴ atteint 93 %, ce qui traduit une forte capacité du modèle à reproduire les classes présentes dans les données. Cependant, cette métrique peut être trompeuse dans le cas où la fréquence d'apparition des classes serait déséquilibrée. C'est le cas ici, où les classes "arbre" et "sol nu" sont sur-représentées. Cela peut gonfler artificiellement l'accuracy sans refléter la performance réelle du modèle sur les classes minoritaires.

Dans ce contexte, le score F1 constitue un indicateur plus pertinent. Il correspond à la moyenne harmonique entre la précision⁵⁵ et le rappel⁵⁶, en pénalisant davantage les déséquilibres entre ces deux dimensions. Ainsi, le score F1 permet d'évaluer plus équitablement la performance du modèle sur l'ensemble des classes, en particulier celles moins représentées dans le jeu de données. Dans notre cas, le score F1 atteint 0.92, ce qui confirme la robustesse du modèle.

Enfin, le coefficient de kappa, qui mesure le degré d'accord entre les prédictions et les observations corrigées du hasard, s'élève à 0.91. Cette valeur élevée témoigne d'un accord très fort entre les classes prédites et les classes réelles, au-delà de ce qui pourrait être attendu par simple coïncidence.

$$\text{Overall Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{Précision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Rappel} = \frac{TP}{TP + FN} \quad (3)$$

$$\text{Score F1} = 2 \cdot \frac{\text{Précision} \cdot \text{Rappel}}{\text{Précision} + \text{Rappel}} \quad (4)$$

Où :

- TP : vrais positifs (*true positives*)
- TN : vrais négatifs (*true negatives*)
- FP : faux positifs (*false positives*)
- FN : faux négatifs (*false negatives*)

54. Proportion de prédictions correctes (positives et négatives) parmi l'ensemble des prédictions effectuées.

55. Proportion de vrais positifs parmi toutes les prédictions positives : elle mesure la fiabilité des prédictions positives.

56. Proportion de vrais positifs correctement prédits parmi toutes les occurrences positives : elle mesure la capacité à ne pas omettre de cas positifs.

$$\text{Kappa} = \frac{p_o - p_e}{1 - p_e} \quad (5)$$

Où :

- p_o : exactitude observée
- p_e : exactitude attendue par hasard

Mesure de la précision : Masque d'entraînement (FME)

Score-F1 : 0.92
Overall Accuracy : 93.3 %
Precision (moyenne) : 91.75 %
Rappel (moyenne) : 92.99 %
Kappa : 0.91

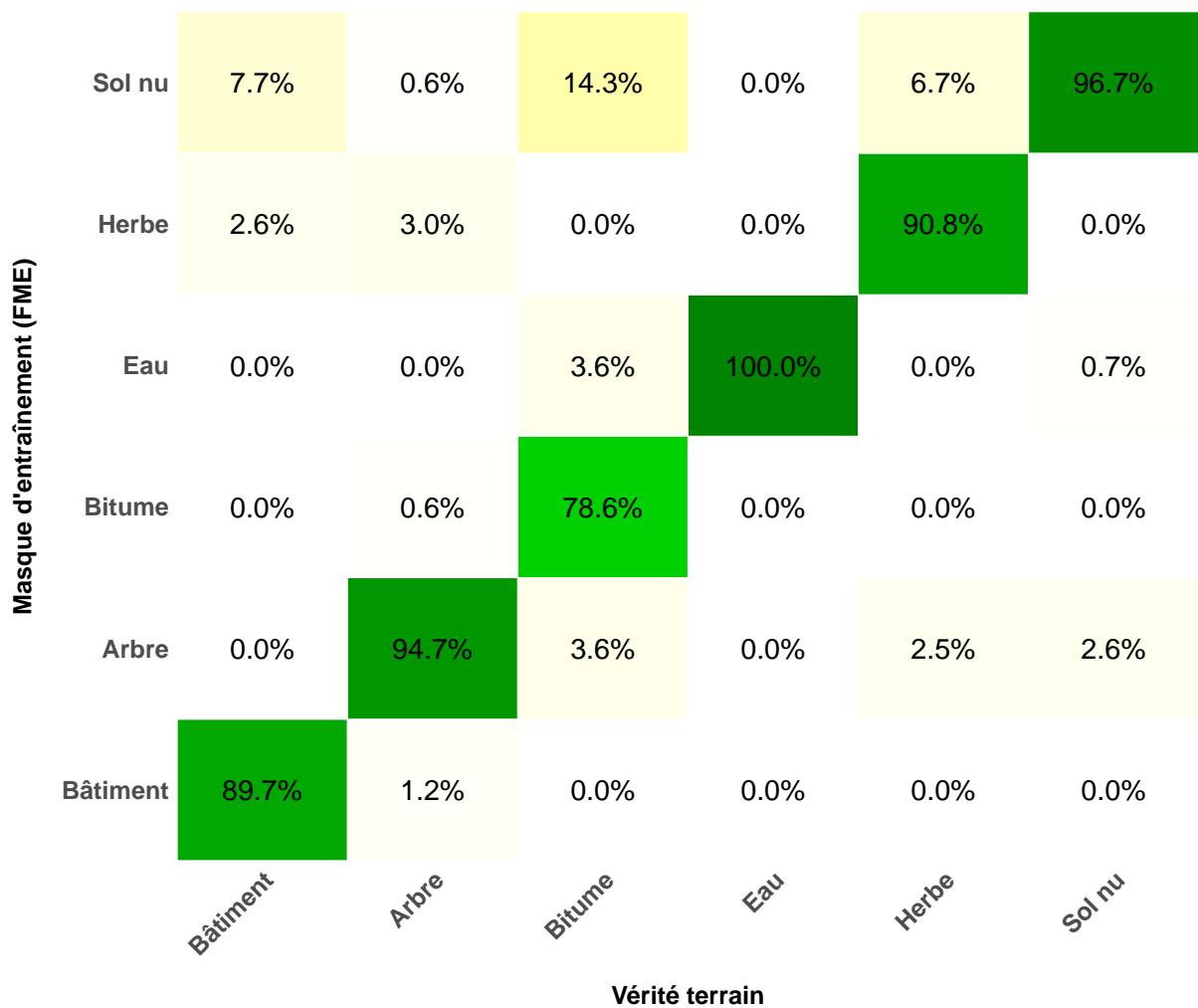


Figure 42. Matrice de confusion du masque d'entraînement.
N. Massot, 2025

la suivante (cf. figure 43) permet de visualiser le masque d'entraînement, et de le comparer à la vérité terrain et à l'OCS GE. On voit que le masque d'entraînement détouche bien les types d'occupation du sol et qu'il est visuellement fidèle à l'image satellite. Par ailleurs, il est bien plus fin spatialement que l'OCS GE.



Figure 43. Extraits du masque d'entraînement, et comparaison avec la vérité terrain et l'OCS GE.

N. Massot, 2025

La répartition des classes dans le masque d'entraînement présente un déséquilibre notable. L'écart-type de la fréquence d'apparition des classes s'élève à 12.9 % (cf. figure 44). Or, pour être efficace, le modèle U-Net nécessite une distribution des classes aussi équilibrée que possible. Dans notre cas, la classe "arbre" est la plus représentée avec 31.6 % des pixels, tandis que la classe "eau" est la moins représentée avec seulement 3.2 %.

De plus, le modèle ne traite pas l'ensemble du masque d'entraînement d'un seul tenant, mais le divise en patchs de 128×128 pixels. Pour limiter les effets du déséquilibre entre classes, le script `01_PREPROCESSER.py` (cf. diagramme d'activité figure 37) découpe le masque en patchs, identifie la classe majoritaire dans chacun d'entre eux, et ajuste le découpage de façon à obtenir autant de patchs pour chacune des six classes d'occupation du sol. Cette méthode permet d'atténuer la sous-représentation de certaines classes dans l'apprentissage. De ce fait, tandis que l'écart-type de la fréquence d'apparition des classes d'occupation du sol dans le masque d'entraînement est de 12.9 %, celui des patchs sélectionnés n'est plus que de 7.2 %. Ainsi, ce découpage permet de réduire les sur/sous-représentations des classes d'occupation du sol et d'avoir une variété plus équilibrée. Cette méthode n'est pas une distorsion de la réalité, mais une sélection des zones permettant une répartition des classes optimale. Ici, **8 172 patchs** sont découpés et utilisés pour l'entraînement du modèle, ce qui représente **134 millions de pixels labellisés** et 1 362 patchs par classe. Notons que pour chaque patch de masque est associé un patch d'image d'entraînement. La section suivante vise à présenter la composition de ces images d'entraînement.

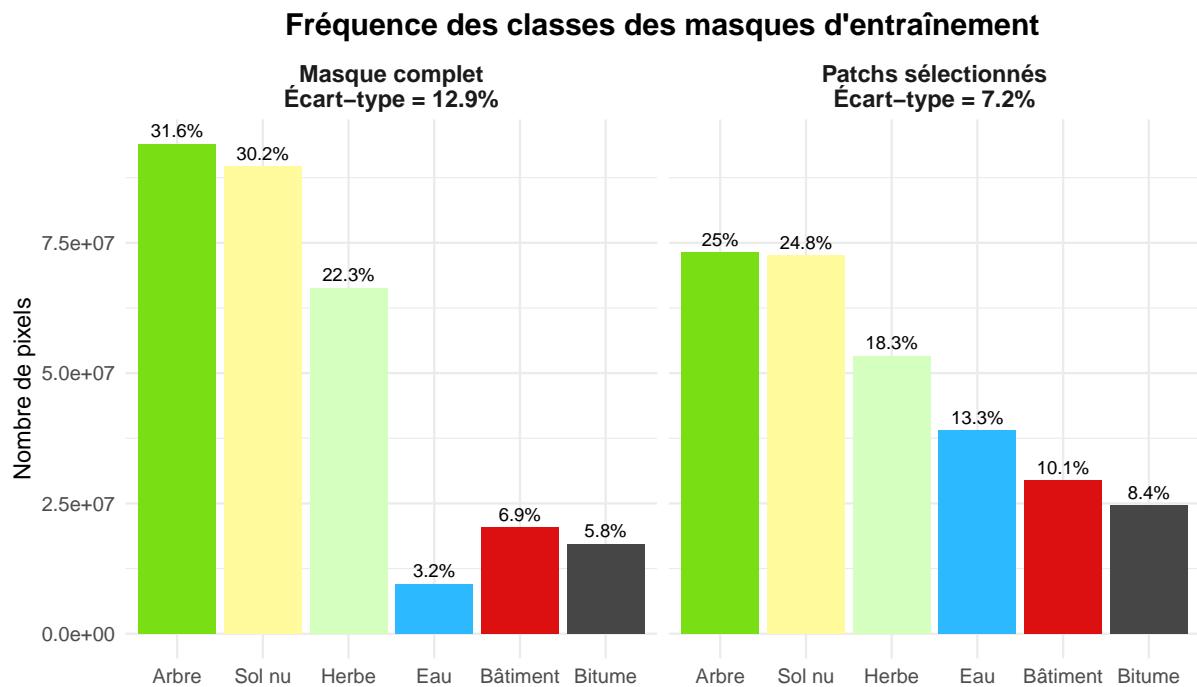


Figure 44. Fréquence d'apparition des classes dans le masque d'entraînement avant et après découpe en patchs.

N. Massot, 2025

4.4 Présentation des données pour les images d'entraînement

Les images d'entraînement servent de facteurs explicatifs de l'occupation du sol et sont utilisées, en association avec le masque, lors de la phase d'apprentissage du modèle. Un travail de recherche a été effectué pour trouver la meilleure combinaison d'images d'entraînement qui permettent d'expliquer au mieux l'occupation du sol. La figure 45 synthétise les éléments spectraux et morphologiques des images d'entraînement. Les éléments spectraux sont les bandes ou les indices qui ont trait à la radiométrie, tandis que les éléments morphologiques sont des données qui décrivent les structures de l'occupation du sol.

Pour les éléments spectraux, les canaux rouge, vert et infra-rouge de la BD Ortho IRC sont inclus dans les images d'entraînement. La bande du bleu n'est présente que dans la BD Ortho RVB, dont la période de capture diffère de la BD Ortho IRC, ce qui crée un diachronisme entre les deux jeux de données. De plus, la longueur d'onde du bleu fait partie du spectre visible, lesquelles ne sont pas les plus représentatives. Ces éléments ont conduit au choix de ne pas utiliser cette bande-là.

Ensuite, en se basant sur les trois bandes spectrales citées précédemment, trois indices spectraux ont été calculés (cf. table 13). Le NDVI est utilisé pour caractériser la

végétation. De plus, le NDWI est utilisé pour délimiter l'eau, mais il a aussi un pouvoir explicatif pour délimiter le bitume⁵⁷. Enfin, le NDGI est habituellement utilisé pour caractériser les glaciers et donc les surfaces blanches. En opposition, il est ici utilisé pour expliquer les surfaces foncées telles que le bitume.

Indice	Acronyme	Définition
Indice de végétation par différence normalisée	NDVI	Mesure la santé et la vigueur de la végétation
Indice de l'eau par différence normalisée	NDWI	Délimiter les zones d'eau. Peut également expliquer les surfaces bitumineuses
Indice de la glace par différence normalisée	NDGI	Délimiter les glaciers (blanc). Utile pour expliquer les surfaces bitumineuses (noire)

Table 13. Liste des indices spectraux et leur définition.
N. Massot, 2025

1. Indice de Végétation par Différence Normalisée (NDVI)

$$NDVI = \frac{IR - Rouge}{IR + Rouge} = \frac{B3 - B1}{B3 + B1}$$

2. Indice de l'Eau par Différence Normalisée (NDWI)

$$NDWI = \frac{Vert - IR}{Vert + IR} = \frac{B2 - B3}{B2 + B3}$$

3. Indice de la Glace par Différence Normalisée (NDGI)

$$NDGI = \frac{Vert - Rouge}{Vert + Rouge} = \frac{B2 - B1}{B2 + B1}$$

Avec, pour la BD Ortho IRC :

- $B1$: Bande rouge
- $B2$: Bande verte
- $B3$: Bande infra-rouge

Outre les éléments spectraux, les éléments morphologiques sont également utiles pour décrire l'occupation du sol. Ici, ils sont au nombre de trois. D'abord, en

57. Dans la première version du modèle FME destiné à la création du masque d'entraînement, un seuil NDWI avait été employé pour délimiter les zones bitumées.

soustrayant le modèle numérique de terrain (MNT) au modèle numérique de surface (MNS) de l'IGN, nous obtenons le modèle numérique de hauteur (MNH). Celui-ci est un facteur explicatif plus important que le MNS ou encore le MNT. En effet, le point commun entre une maison située sur une colline et une autre située sur une plaine est sa hauteur, pas son altitude. Ce raisonnement marche également pour les arbres.

Ensuite, un indice de texture est utilisé. L'*Inverse Difference Moment* (IDM) est une métrique basée sur une matrice de co-occurrence des niveaux de gris (Gray Level Co-occurrence Matrix) (Rault et al. 2016). Ici, elle est calculée à partir du MNH. Elle mesure la similarité des pixels voisins. Une valeur élevée désigne une texture homogène avec peu de variations locales (eau, sol nu, ...). À l'inverse, une valeur faible désigne des textures hétérogènes avec de fortes variabilités locales (bâti, arbres, ...).

$$IDM = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{P(i,j)}{1 + (i - j)^2}$$

où :

- $P(i,j)$: élément de la matrice de co-occurrence représentant la probabilité conjointe des niveaux de gris i et j ,
- N : nombre de niveaux de gris dans l'image,
- i et j : indices correspondant aux niveaux de gris.

Enfin, le dernier élément morphologique est le *Patch Shape Index* (PSI) (FRAGSTATS s. d.), qui est un indice de l'écologie du paysage. Celui-ci décrit la compacité des patchs issus du masque d'entraînement. Un patch est un ensemble continu de pixels spatialement connectés, qui appartiennent à la même classe d'occupation du sol. Si le patch est carré, valeur la plus compacte dans un mode raster, alors l'indice de forme vaudra 1. À l'inverse, plus le patch est irrégulier, plus la valeur tend vers $+\infty$. Cet indice a été choisi car il est indépendant de la taille du patch comme les indices périmètres/surface. Cela permet de comparer des formes de tailles différentes de manière équitable.

$$\text{Patch Shape Index} = \frac{0.25 \cdot P_{ij}}{\sqrt{a_{ij}}}$$

où :

- P_{ij} : le périmètre de l'entité ij ,
- a_{ij} : l'aire de l'entité ij .

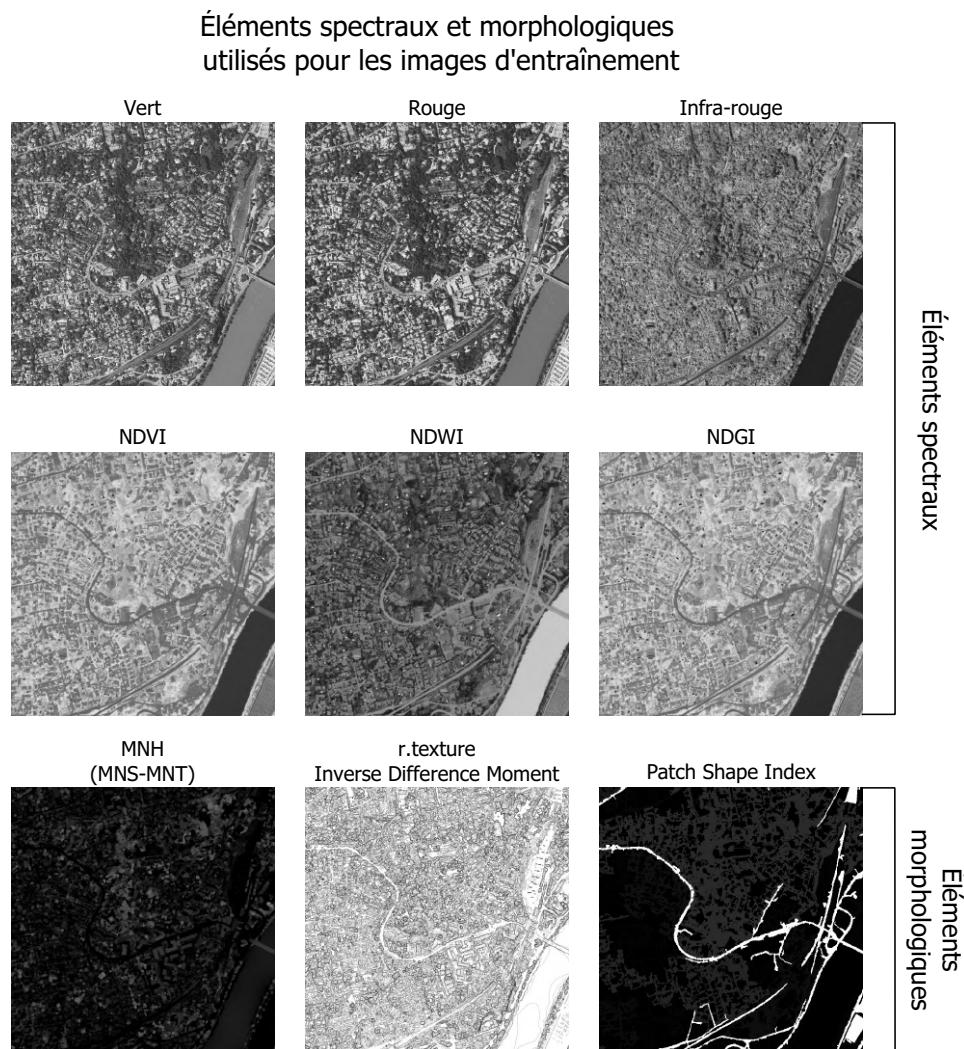


Figure 45. Éléments spectraux et morphologiques utilisés pour les images d'entraînement.
N. Massot, 2025

Par ailleurs, une étape de prétraitement explorée est celle de l'augmentation des données (cf. figure 46). Cela consiste à créer de nouvelles images d'entraînement à partir d'images existantes en leur appliquant des transformations géométriques (rotation, symétrie, zoom, etc.) ou radiométriques (variation de luminosité, contraste, bruit, etc.). Cela permet de multiplier artificiellement le volume d'entraînement, ce qui est particulièrement utile pour améliorer la segmentation d'images, surtout quand le jeu de données est limité. Ici, l'augmentation des données a été testée mais pas adoptée. En effet, le facteur limitant est la puissance de calcul, ce qui ne nous permet pas de mettre plus d'images⁵⁸. De ce fait, le modèle est 4 % plus performant du point de vue de l'accuracy sur l'ensemble des données, plutôt que sur un sous-ensemble de données augmentées.

58. La configuration de la machine utilisée pendant ce mémoire figure à l'annexe 20.

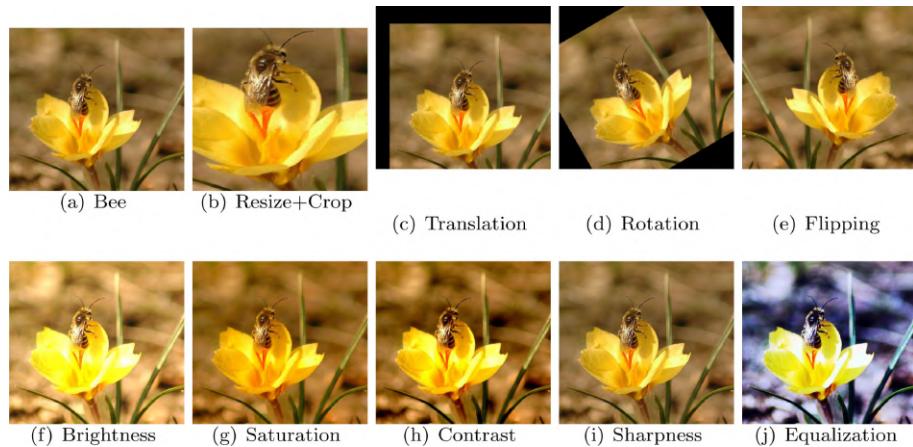


Figure 46. Techniques d'augmentation.

Source : Tian, Y., & Zhang, Y. (2022). A comprehensive survey on regularization strategies in machine learning. *Information Fusion*, 80, 146–166. <https://doi.org/10.1016/j.inffus.2021.11.005>.

4.5 Études préalables

4.5.1 Signature spectrale et séparabilité des classes avec la distance de Jeffries-Matusita

Après avoir acquis les images et le masque d'entraînement, il est nécessaire d'étudier les signatures spectrales pour repérer d'éventuelles ressemblances entre elles. La figure 47 montre que chacune des bandes a des pouvoirs explicatifs variés sur toutes les classes. Ainsi, la combinaison des neuf bandes spectrales et morphologiques permet de bien distinguer toutes les classes, ce qui nous conforte dans l'idée que la composition des bandes est adéquate et adaptée aux besoins.

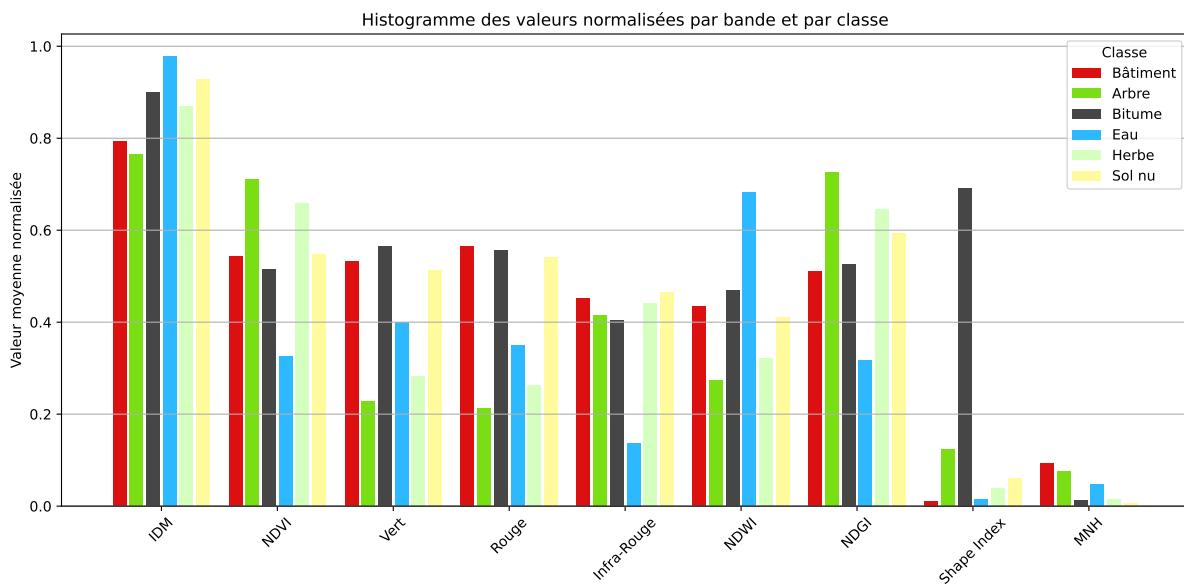


Figure 47. Signatures spectrales des classes selon la bande spectrale (valeurs normalisées).

N. Massot, 2025

Une fois que les signatures ont été déterminées pour chaque classe, il est important de vérifier leur pouvoir séparateur, c'est-à-dire leur capacité à effectuer une bonne discrimination entre les classes sur l'image avant de procéder à la classification. Pour ce faire, nous utilisons la distance de Jeffries-Matusita (Dabboor et al. 2014), qui est une mesure statistique utilisée pour évaluer la séparabilité des classes deux à deux dans un espace multi-dimensionnel. Elle permet d'analyser à quel point deux distributions sont distinctes l'une de l'autre. Les valeurs de séparabilité, comprises entre 0 et 2, sont d'autant plus proches de 2 qu'elles sont élevées. D'un point de vue conceptuel, plus il y a de classes, plus la valeur de séparabilité peut tendre à être faible, et inversement. En complément, plus il y a de bandes spectrales, plus la séparabilité peut tendre à être élevée, et inversement. Cette explication conceptuelle est illustrée dans la figure 48. Calculer la séparabilité des classes est doublement utile. D'abord, elle permet de nous indiquer si le nombre de classes est trop élevé ou si elles ne sont pas assez séparables. Ensuite, elle permet de décrire le pouvoir explicatif des éléments spectraux ou morphologiques : plus il y a de bandes, plus la séparabilité des classes est élevée. La figure 49 montre que, mis à part la combinaison "bâtiment" - "sol nu" qui reste moyenne, le calcul de la distance de Jeffries-Matusita démontre une séparabilité des classes très satisfaisante.

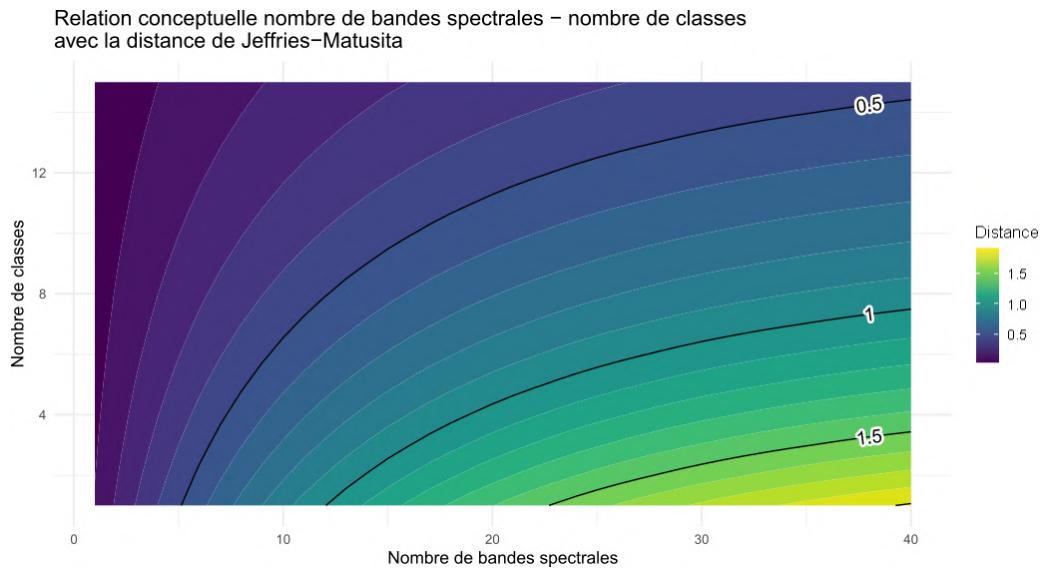


Figure 48. Relation conceptuelle nombre de bandes spectrales – nombre de classes avec la distance de Jeffries–Matusita.

N. Massot, 2025

Matrice de séparabilité des classes

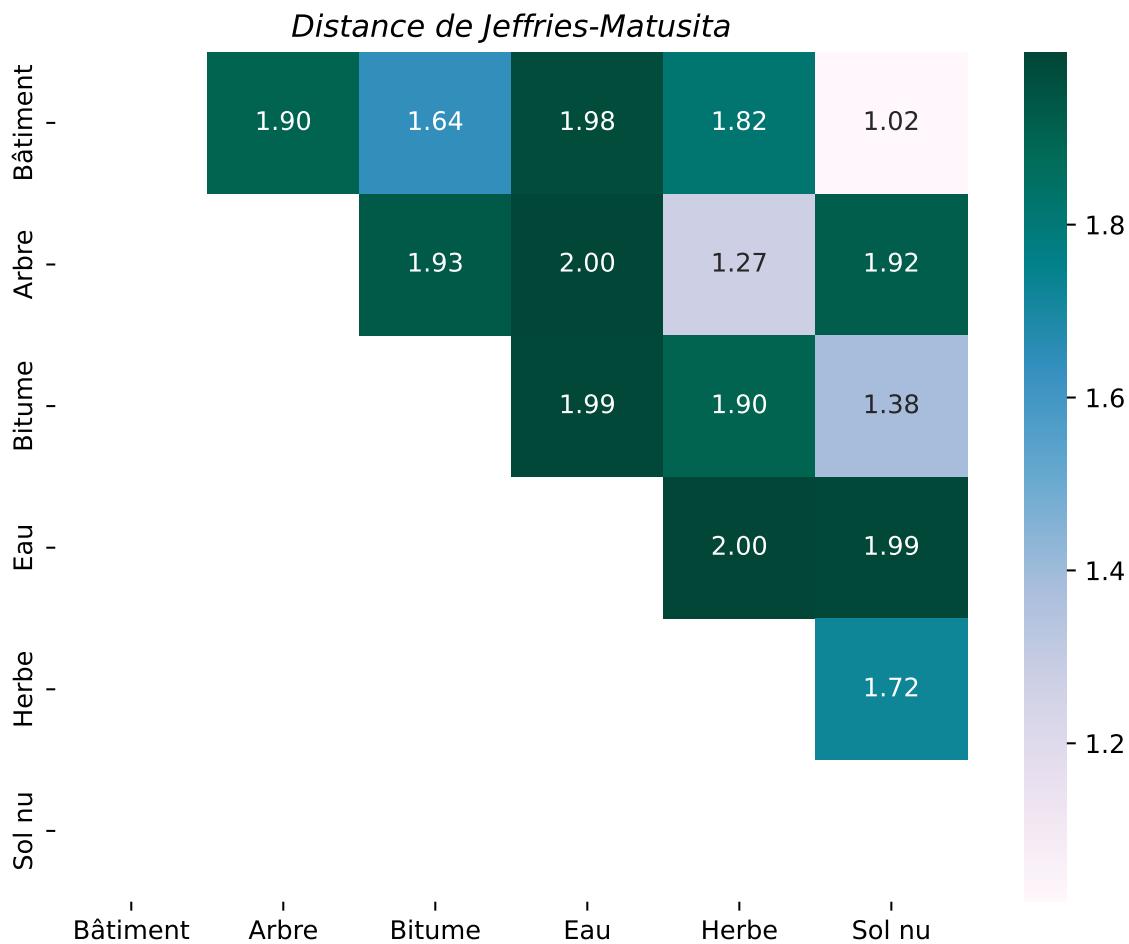


Figure 49. Séparabilité des classes selon la distance de Jeffries-Matusita pour les canaux spectraux et morphologiques utilisés pour les images d'entraînement.

N. Massot, 2025

4.5.2 Étude et impact de la colinéarité

Dans les méthodes de classification de pixels à partir de bandes spectrales, il est d'usage d'étudier la colinéarité des bandes, pour ne garder que celles qui sont le moins corrélées entre elles. L'objectif étant d'éviter la redondance d'information et d'optimiser la puissance de calcul. Ainsi, nous posons l'hypothèse H₃ que, pour les réseaux de neurones convolutifs, la colinéarité des images d'entraînement a un impact négatif à la performance du modèle. De ce fait, comme le montre la figure 50, sur les 36 combinaisons possibles, quatre sont corrélées deux à deux sont :

- Rouge & Vert
- NDWI & NDVI
- NDGI & NDVI
- NDGI & NDWI

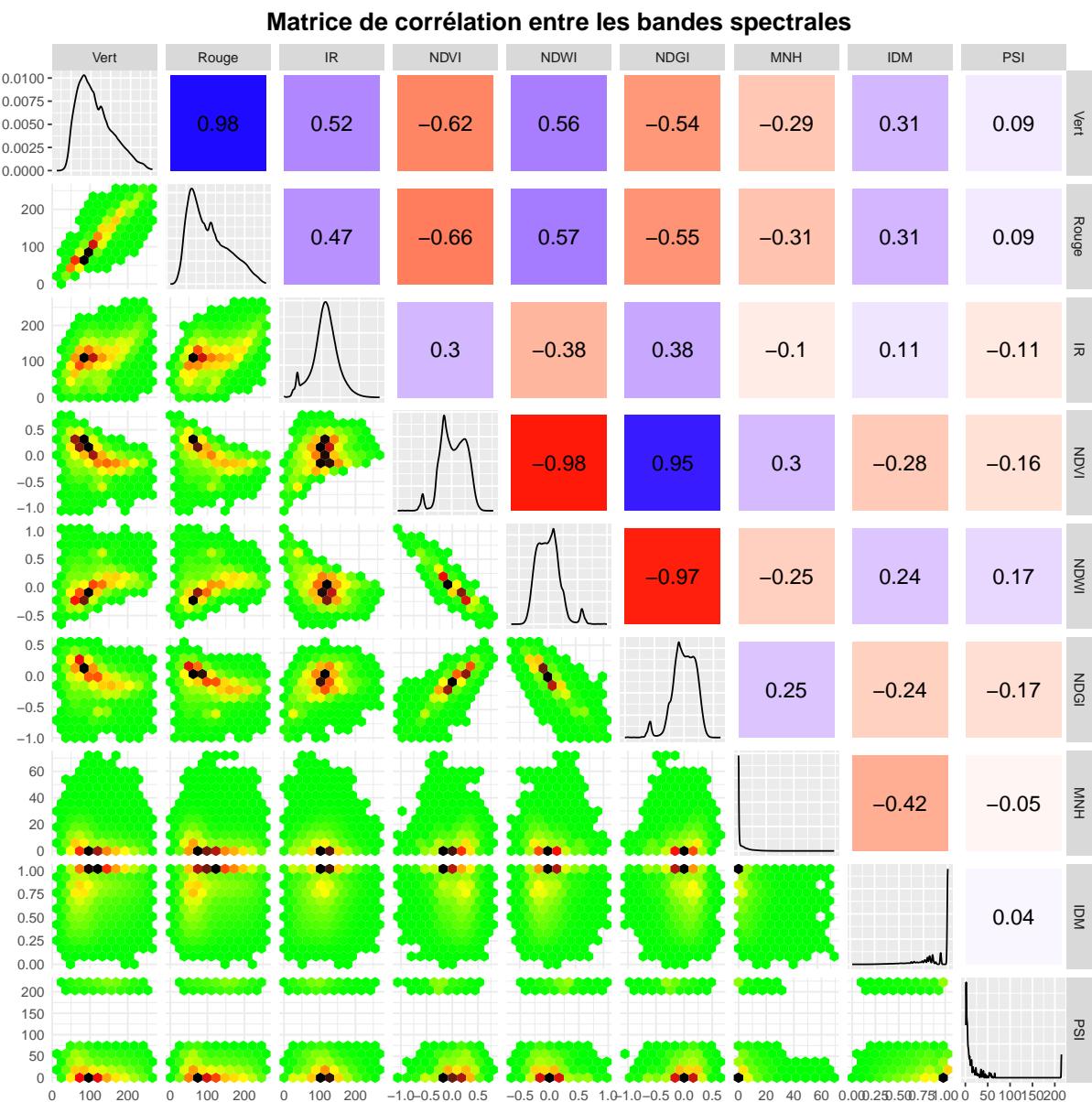


Figure 50. Matrice de corrélation des canaux spectraux et morphologiques utilisés pour les images d'entraînement.

N. Massot, 2025

Ainsi, en retirant deux des trois indices spectraux (NDVI, NDWI, NDGI) et le rouge ou le vert, la colinéarité serait fortement réduite. Alors, pour vérifier la viabilité de l'hypothèse qui suggère que la colinéarité a un impact négatif aux performances du modèle, des tests de performance unitaires ont été faits, en entraînant le modèle après avoir retiré les bandes une à une. Le tableau 14 montre, pour chaque bande retirée, trois métriques clés automatiquement générées par le notebook 02_ENTRAINER.ipynb (cf. figure 37.) D'abord, la *train accuracy* représente le taux de bonne classification en comparaison du masque d'entraînement qui fait office de vérité terrain. Ensuite, le *loss* est la fonction de perte qui permet de mesurer à quel point les prédictions du modèle s'éloignent de la vérité terrain (ici, le masque d'en-

traînement). Quand cette valeur est calculée sur des données d'entraînement : c'est le *train loss*, tandis que quand elle est calculée sur des données de validation, c'est la *validation loss*. Le *train loss* est obtenu à partir des données qui ont servi pour l'entraînement. À l'inverse, les données de validation sont le masque et les images sur lesquelles le modèle ne s'est jamais entraîné. Ainsi, comme il ne les connaît pas, il n'est pas possible que le modèle ait "appris par cœur" à reproduire l'occupation du sol de ces zones spécifiquement. De ce fait, le *validation loss* est donc de ce point de vue une métrique plus pertinente, puisqu'elle quantifie la capacité du modèle à généraliser. Ici, 90 % des données sont destinées à l'entraînement, tandis que les 10 % restants sont destinés à la validation⁵⁹.

Le tableau 14 montre que le fait de retirer une à une les bandes, quelle que soit la bande retirée, n'a aucun impact positif sur les métriques de performance du modèle, toutes choses égales par ailleurs. Ainsi, les valeurs d'*accuracy* varient entre 0.89 et 0.91 lorsque l'on retire une bande. À l'inverse, lorsqu'un entraînement est lancé avec les neuf canaux spectraux et morphologiques, la métrique d'*accuracy* monte à 0.97 au regard du masque d'entraînement. Le même constat s'applique au *train loss* et au *validation loss* : les valeurs les plus faibles sont obtenues lorsque l'entraînement utilise l'ensemble des neuf bandes. Ces tests unitaires démontrent que la colinéarité n'a pas d'impacts négatifs sur la performance prédictive du modèle. Mieux, la combinaison de facteurs explicatifs permet de rendre le modèle plus performant, malgré la colinéarité. Cela s'explique, car le réseau de neurones convolutifs fonctionne par apprentissage par renforcement en ajustant les poids pour maximiser la capacité de prédiction. Ce processus conduit parfois à une survalorisation de certaines caractéristiques des images d'entraînement, tandis que d'autres sont négligées.

59. Notez que ce ne sont pas ces 10 % qui serviront à la validation finale du modèle, et qu'un protocole de mesure de la précision sera présenté au Chapitre 5.

Bandes retirées	Train accuracy	Train loss	Validation loss
Vert	0,89	0,16	0,32
Rouge	0,90	0,15	0,28
NDVI	0,91	0,14	0,28
NDWI	0,91	0,14	0,26
NDGI	0,90	0,13	0,32
IDM	0,86	0,18	0,27
Patch Shape Index	0,92	0,10	0,26
∅	0,97	0,07	0,20

Table 14. Tests de performance unitaires en enlevant les bandes unes par unes.
N. Massot, 2025

Ainsi, l'hypothèse H_3 qui suggère que la colinéarité des images d'entraînement a un impact négatif à la performance du modèle est rejetée.

4.6 Pré-traitements et phase inférentielle

Après avoir acquis et étudié le masque et les images d'entraînement, il est temps de les pré-traiter pour les conformer aux besoins du modèle U-Net. Lors de la phase inférentielle, le modèle prend en entrée les images et le masque d'entraînement découpés en patch dont la taille, en largeur et en hauteur est pair. Le plus souvent, la taille des patchs varie entre 64 et 1024 pixels de côté. Ces dimensions jouent un rôle crucial, car elles influencent directement le comportement du modèle. D'un point de vue conceptuel, les patchs plus petits permettent au modèle de se concentrer sur des détails locaux, ce qui est bénéfique pour segmenter les caractéristiques fines de l'occupation du sol. À l'inverse, des patchs plus grands permettent de capturer davantage de contexte spatial. C'est donc un arbitrage à faire entre segmentation des caractéristiques fines et compréhension du contexte global. Par ailleurs, les petits patchs demandent moins de ressources computationnelles. Ainsi, le choix a été fait de travailler avec des patchs de 128×128 pixels.

Ensuite, comme mentionné dans la partie 4.3.3 et illustré dans la figure 44, les patchs de masque sont découpés de façon à obtenir un nombre équilibré de

patches pour chaque classe d'occupation du sol. Chaque patch de masque est alors associé à un patch d'image d'entraînement. Par la suite, les valeurs des patchs d'images sont normalisées entre 0 et 1. Ces deux étapes sont réalisées par le script `01_PREPROCESSER.py`.

Le script `02_ENTRAINER.ipynb` permet de lancer l'entraînement du modèle U-Net. Celui-ci est configuré par divers hyperparamètres (cf. annexe [.2](#)). Après entraînement, le modèle donne des métriques d'*accuracy* et les fonctions de *loss* d'entraînement et de validation, qui sont très bonnes (cf. figure [51](#)). Les fonctions de perte (*loss*) présentées permettent de suivre l'évolution de la performance et de l'apprentissage du modèle au cours de l'entraînement. On observe une diminution progressive et régulière des erreurs pour l'ensemble d'entraînement (*train_loss*) ainsi que pour l'ensemble de validation (*val_loss*), ce qui indique un apprentissage cohérent et stable. L'absence de remontée ou de divergence significative de la courbe de perte de validation suggère que le modèle généralise correctement aux données qu'il n'a pas vues, sans signes de surapprentissage.

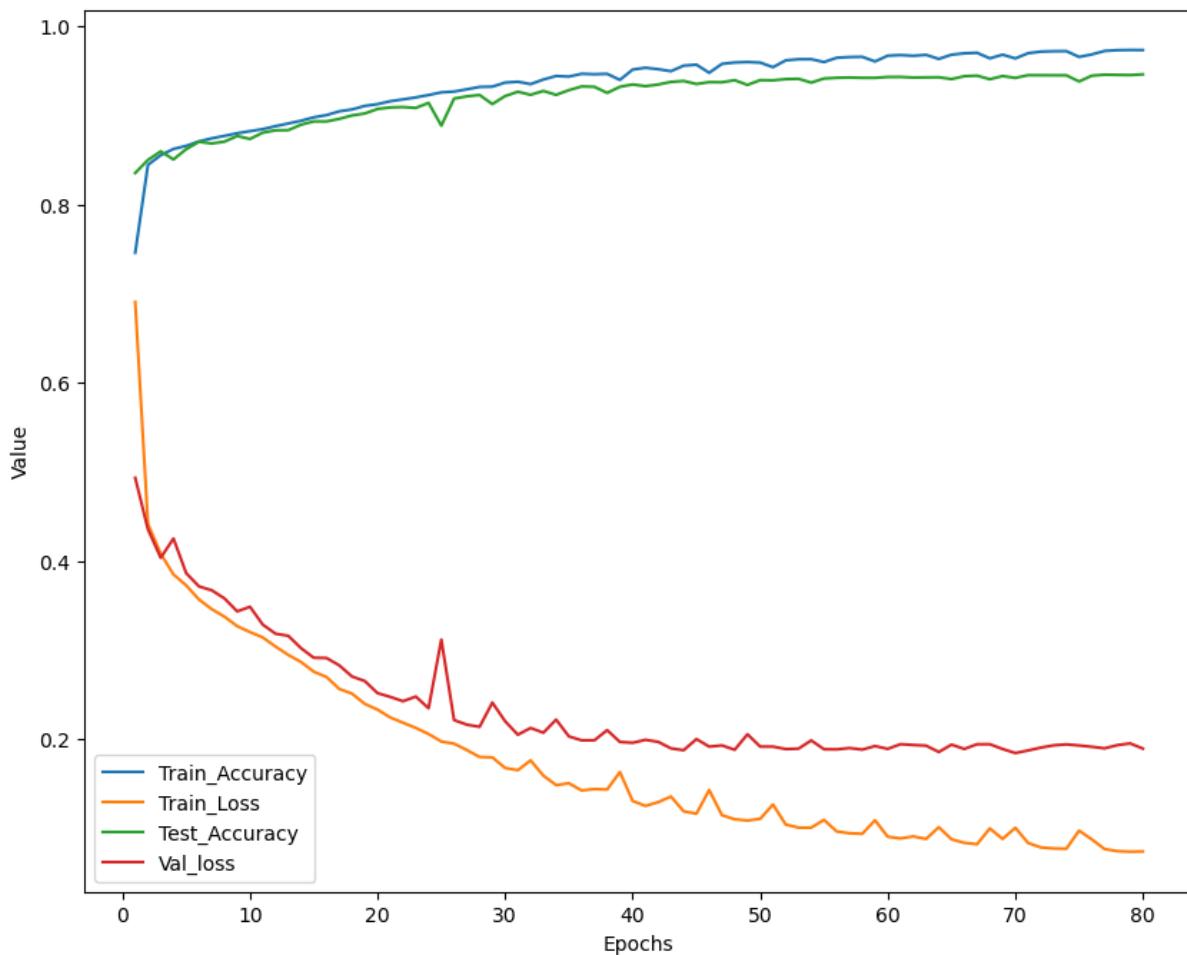


Figure 51. Indicateurs de performance du modèle après entraînement.
N. Massot, 2025

À partir du moment où le modèle est entraîné, la phase inférentielle débute. En effet, comme illustré dans le diagramme d'activité de la figure 37, cette partie de la chaîne de traitements est délimitée avec le trait horizontal en pointillé. La phase inférentielle commence avec la création du raster d'inférence. Ce fichier doit être composé des mêmes neuf bandes spectrales et morphologiques que celles mentionnées dans la figure 45. Ensuite, trois scripts python sont mobilisés. Le premier (03_DECOUPER_SOUS_ENSEMBLE.py) permet de découper le raster d'inférence en sept sous-ensembles⁶⁰ pour des raisons de puissance de calcul limitée. En se basant sur ces derniers, le script 04_INFERER.py va inférer le modèle entraîné précédemment sur chacun des sous-ensembles après les avoir découpés en patchs de 128 × 128. Enfin, le script 05_FUSIONNER_SOUS_ENSEMBLE.py permet de fusionner les sept segmentations prédites du modèle en une seule avec la fenêtre de Hann (cf. figure 36).

60. Les sous-ensembles sont découpés verticalement en parties de taille égale, avec une largeur multiple de 128 et un chevauchement de 64 pixels.

4.7 Post-traitement

À présent, se pose la question du post-traitement. Pendant la phase de développement du modèle, la question de l'agrégation de la segmentation à des données contextuelles (fichiers fonciers, RPG, ...) s'est posée. À l'image de l'OCS GE, cela permet de sortir une carte plus intelligible. Cependant, après avoir discuté avec le Grand Avignon de leurs besoins métiers, ceux-ci estiment que la donnée brute est celle qui est la plus appropriée. En effet, pourquoi agréger la donnée à un découpage parcellaire, étant donné que l'occupation du sol s'affranchit de ces prédictifs ? De plus, le parcellaire est, par définition, amené à être divisé ou fusionné au cours du temps. De ce fait, les analyses en séries temporelles pourraient être biaisées. C'est pour ces raisons que le seul post-traitement utilisé dans la chaîne de traitements est l'application d'un filtre de majorité de 3×3 , qui permet de lisser légèrement les résultats de la segmentation, et de retirer les pixels isolés.

Chapitre 5 :

Évaluation des résultats de la classification

Ce chapitre fait état des différents résultats obtenus suite à la chaîne de traitements. Une visualisation interactive de la segmentation est proposée. Ensuite, la qualité de cette dernière est évaluée à travers différentes métriques et méthodes, avant de faire un comparatif avec CoSIA. Nous finirons par conclure l'analyse des résultats et d'évoquer les potentiels axes d'amélioration de la méthode.

5.1 Cartographie du résultat

La dernière phase de la chaîne de traitements consiste à évaluer l'exactitude de la segmentation (cf. figure 37). Cette étape est essentielle, car elle permet de valider la qualité des données et de déterminer le niveau de confiance que l'on peut y accorder. Avant d'évaluer quantitativement la segmentation, cartographier les résultats permet une vérification visuelle de la segmentation prédictive (b), en comparaison à une image satellitaire (a) (cf. figure 52). Aussi, une cartographie interactive a été faite pour pouvoir visualiser en détails la segmentation résultante de la chaîne de traitements :

https://mercatorien.github.io/MEMOIRE_MASSOT/

Dans ces cartes, nous observons visuellement que la segmentation décrit particulièrement bien les structures complexes des éléments du sol et du sursol, et que la structure spatiale est cohérente. Par ailleurs, la résolution spatiale d'un mètre permet de très bien délimiter les éléments, aussi petits soient-ils.

De plus, en étudiant la carte des probabilités, nous pouvons quantifier l'incertitude de la précision. Ici, la probabilité moyenne est de 97.6 % alors que la médiane est de 100 (cf. figure 54). Cela traduit une asymétrie à gauche. Cela nous conforte dans l'idée que les prédictions sont robustes. En spatialisant ces données, nous constatons que les zones faisant l'objet d'une probabilité réduite se situent dans les espaces de transition entre deux types d'occupation du sol (cf. figure 53). Les cas les plus courants sont situés à la limite entre du bâti et une autre classe de sol (herbe, sol nu ou bitume), et surtout dans les champs, dans lesquelles il est difficile de distinguer l'herbe du sol nu.

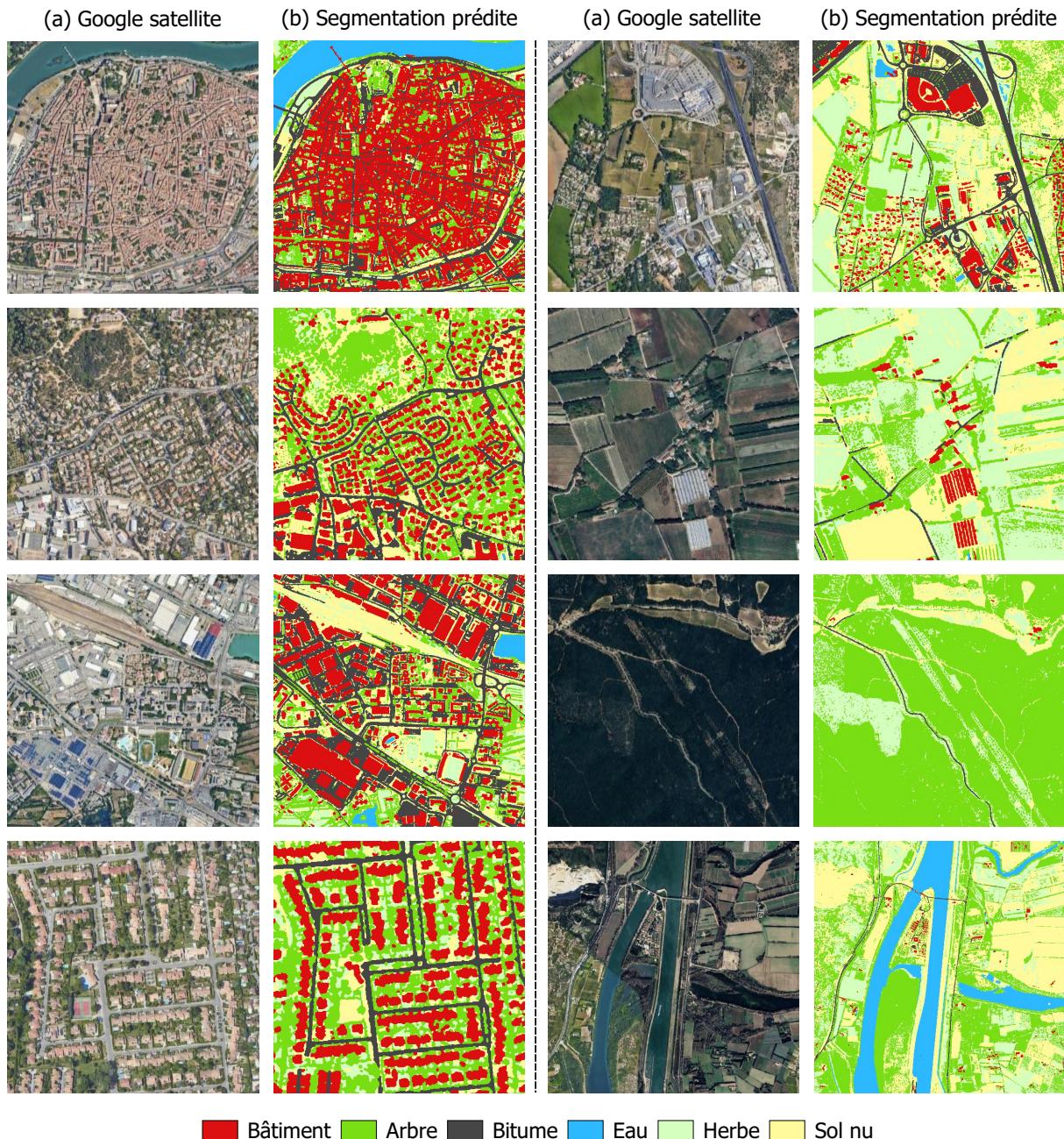


Figure 52. Cartographie de la segmentation (b) en comparaison de l'image satellitaire (a)

N. Massot, 2025

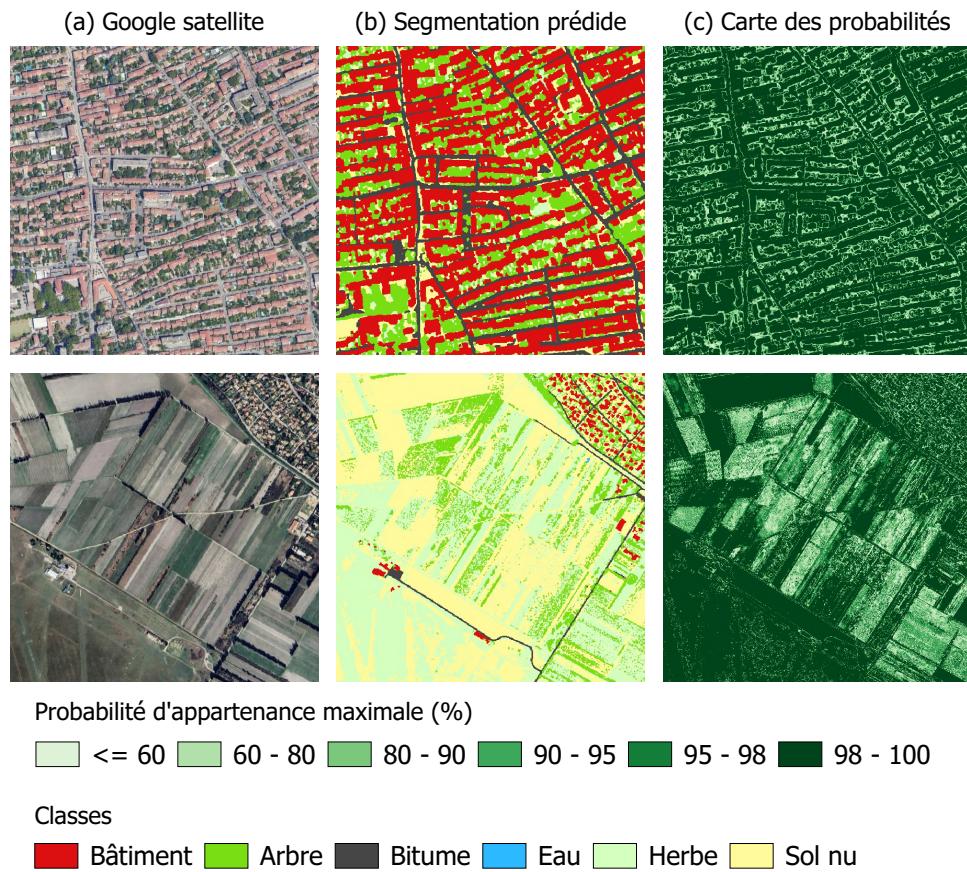


Figure 53. Carte des probabilités (c).
N. Massot, 2025

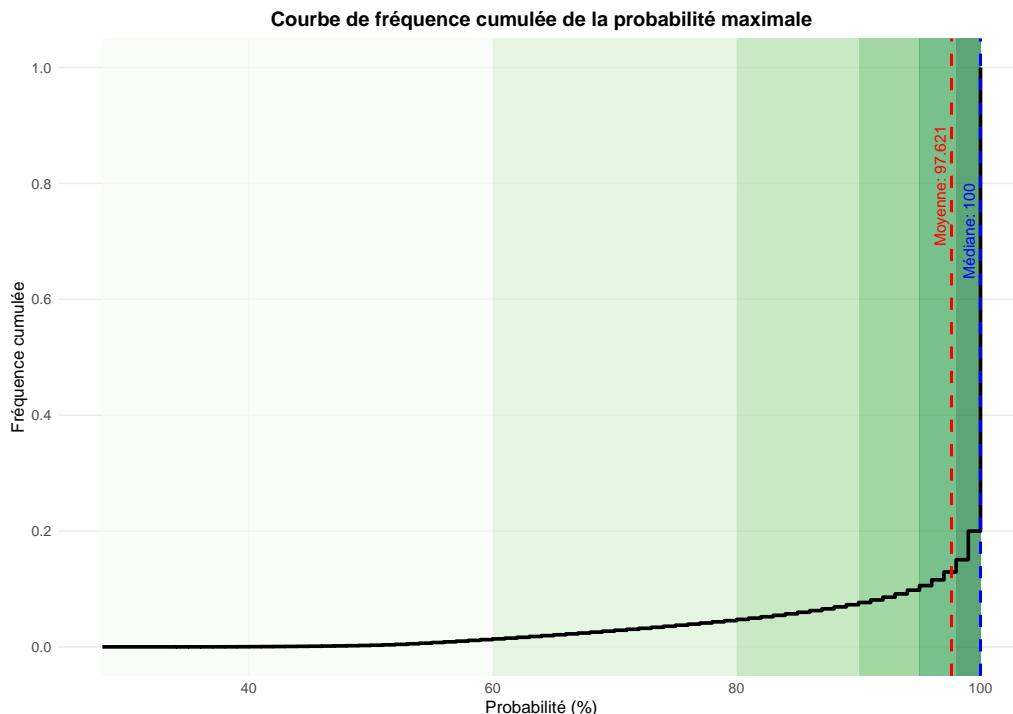


Figure 54. Fréquence cumulée des probabilités maximales.
N. Massot, 2025

Pour évaluer qualitativement les résultats de la segmentation, plusieurs métriques sont mobilisées. En effet, il est important de varier les mesures pour des questions de complémentarité de l'information, et de robustesse de l'évaluation. D'abord, la matrice de confusion est utilisée pour décrire le taux de bonne classification entre chaque classe. Les erreurs d'omission et de commission sont également étudiées. Ensuite, nous calculons le rappel sous forme de matrice. Celui-ci détermine la capacité du modèle à identifier la proportion de vrai positif correctement classifié. Les coefficients de Kappa (accord entre la prédiction et la vérité terrain ajusté du hasard), et le Score F1 (moyenne harmonique de la précision et du rappel) sont également étudiées. La manière de calculer ces indicateurs est précisée dans la sous-section 4.3.3. Enfin, l'*intersection over union* (IoU) est utilisé pour mesurer le degré de chevauchement entre la donnée prédite (sortie du modèle U-Net), et la vérité terrain (masque d'entraînement).

Hormis l'IoU, toutes ces métriques sont générées à partir du même échantillon manuellement labellisé par photo-interprétation de la BD Ortho 2024, que celui utilisé pour l'état de l'art qualitatif des données d'occupation du sol (cf. Chapitre 3). De ce fait, nous disposons d'un référentiel commun qui fait que les résultats sont comparables entre les données évaluées pendant le benchmark et celles de la segmentation U-Net. L'échantillonnage répétitif régulier est une grille de point de 750 mètres de résolution, qui est composé de 678 échantillons (cf. tableau 15). Malgré le déséquilibre des classes, la représentativité spatiale est atteinte, car la loi des grands nombres permet aux mesures de converger rapidement vers l'espérance mathématique. De plus, le score F1, étant une moyenne harmonique de la précision et du rappel, il donne une mesure plus équilibrée de la performance du modèle, indépendamment de sa fréquence dans l'ensemble de données.

Classe	Effectif d'échantillon
Bâtiment	38
Arbre	220
Bitume	34
Eau	30
Herbe	137
Sol nu	219

Table 15. Effectifs des échantillons pour chaque classe.
N. Massot, 2025

La segmentation de l'occupation du sol du Grand Avignon révèle une très bonne performance, tant sur le plan des indicateurs de précision et de rappel que de l'*Intersection over Union*. Le score F1 global s'élève à 0,91, ce qui traduit un bon compromis entre exactitude et exhaustivité, et l'accuracy dépasse les 92 %. Cela indique que plus de neuf pixels sur dix sont correctement classés. Le coefficient Kappa de 0,90 confirme que cette performance excède largement ce qu'un simple tirage aléatoire pourrait produire.

En examinant plus précisément la matrice de précision (cf. figure 55), on observe que l'eau est parfaitement reconnue (accuracy de 100 %), le sol nu à 95,8 %, les arbres à 94,7 % et le bitume à 86,2 %. La classe bâtiment est celle dont les pixels sont les moins bien classifiés. Elle atteint 76,2 % de bonne classification, ce qui signifie qu'un peu moins d'un quart des pixels prédicts bâtiments sont en réalité d'une autre nature. Cela correspond à un léger effet arrondi observable sur certains bâtiments. Toutefois, le rappel des bâtiments indique que 94,1 % des bâtiments prédicts sont correctement classifiés (cf. figure 56). Cette classe est donc, après celle de l'eau (100 %), celle dont les erreurs d'omissions sont les plus basses. Le rappel s'élève à 90,9 % pour les arbres et le sol nu est à 93,4 %. L'herbe et le bitume obtiennent respectivement 93,5 % et 80,6 %, ces deux catégories subissant quelques confusions entre elles et avec les surfaces environnantes.

Les erreurs de commission (cf. figure 55) illustrent enfin que pas ou très peu de pixels d'eau, de sol nu et d'arbres sont omis (0 %, 4,2 % et 5,3 %), tandis que l'herbe, le bitume ou les bâtiments souffrent d'un oubli plus marqué (11,5 %, 13,8 % et 23,8 %). Les bâtiments, les arbres et les sols nus présentent de faibles taux d'omission, de l'ordre de 6 à 7 %, ce qui traduit une bonne sensibilité.

Matrice de confusion : Précision

Score-F1 : 0.91
 Overall Accuracy : 92.2 %
 Rappel (moyen) : 92.1 %
 Kappa : 0.9

	Erreurs d'omission	9.1	19.4	0.0	6.5	6.6		
Précision du producteur	94.1	90.9	80.6	100.0	93.5	93.4		
Moyenne en colonne	12.9	19.3	15.3	16.7	15.4	20.4		
Prédiction U-Net	Sol nu	0.5	1.6	0.5	0.0	1.6	95.8	16.7
	Herbe	0.8	7.7	1.5	0.0	88.5	1.5	16.7
	Eau	0.0	0.0	0.0	100.0	0.0	0.0	16.7
	Bitume	0.0	0.0	86.2	0.0	0.0	13.8	16.7
	Arbre	0.0	94.7	1.1	0.0	2.6	1.6	16.7
	Bâtiment	76.2	11.9	2.4	0.0	0.0	9.5	16.7
	Bâtiment		Arbre	Bitume	Eau	Herbe	Sol nu	Moyenne en ligne
	Vérité terrain						Précision de l'utilisateur	Erreur de commission

Figure 55. Mesure de Précision du modèle U-net.
N. Massot, 2025

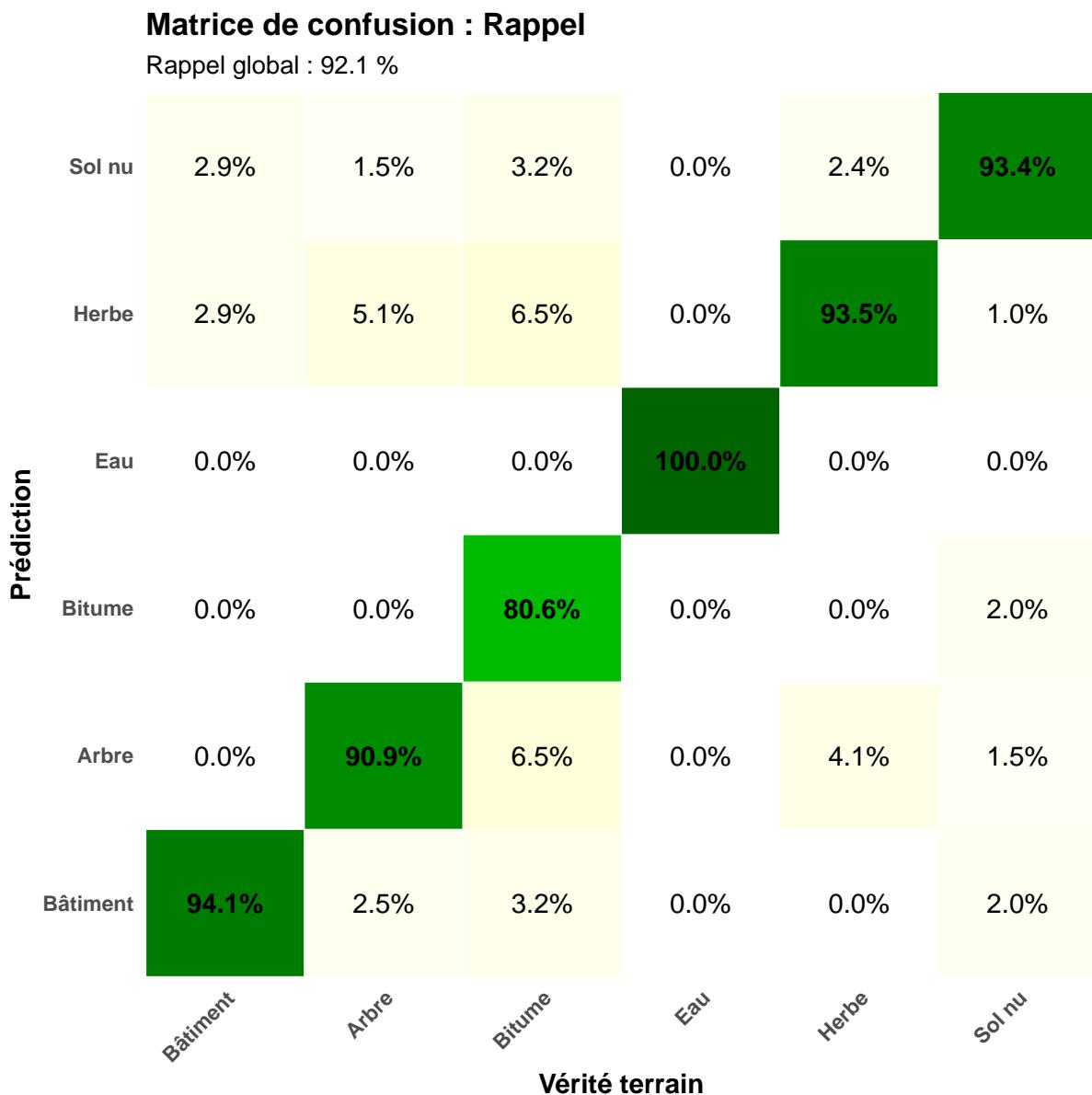


Figure 56. Mesure du rappel du modèle U-net.
N. Massot, 2025

L'analyse de l'IoU, qui mesure la zone commune entre la prédiction et la vérité terrain par rapport à leur union, confirme ces constats. L'*Intersection over Union* (IoU) est définie comme :

$$IoU = \frac{|U \cap V|}{|U \cup V|} = \frac{TP}{TP + FP + FN}$$

où :

- $U \cap V$: intersection entre la prédiction et le masque de vérité terrain,
- $U \cup V$: union entre la prédiction et le masque de vérité terrain,
- TP : vrais positifs (*True Positives*),
- FP : faux positifs (*False Positives*),

- *FN* : faux négatifs (*False Negatives*).

L'IoU moyen de 0,77 atteste d'une très bonne qualité spatiale de la segmentation (cf. figure 57)⁶¹. L'eau, avec un score de 0,89, se distingue par son contour net et sa détection fiable, comme la matrice de confusion le laissait entendre. Les classes bâtiment (0,81), arbre (0,79) et sol nu (0,79) conservent des valeurs élevées, ce qui traduit une délimitation précise. L'herbe, avec 0,77, et le bitume, avec 0,61, restent légèrement en retrait.

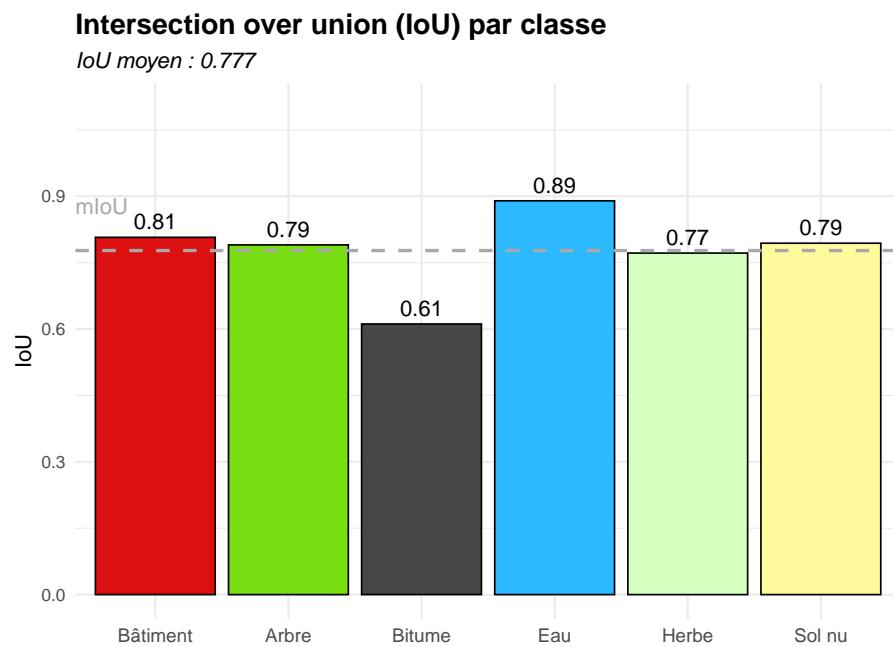


Figure 57. IoU global, par classe et moyen obtenu par l'architecture U-Net basé sur le masque d'entraînement.
N. Massot, 2025

5.2 Conclusion de l'analyse des résultats

En somme, la chaîne de traitements originale proposée permet de produire une cartographie dont la précision est inédite au regard du *benchmark*, et avec une résolution spatiale remarquable d'un mètre. Les résultats obtenus sont satisfaisants pour les classes "bâtiments" et "bitume", et excellents pour les autres catégories.

Le protocole de validation appliqué (échantillonnage et métriques d'évaluation) est identique à celui utilisé pour le *benchmark* du chapitre 3 ainsi que pour la validation des résultats dans ce chapitre, ce qui garantit la comparabilité des analyses. Pour rappel, le *benchmark* a révélé que les données d'OpenStreetMap

61. Un IoU est considéré comme correct quand il est autour de 0.5.

présentaient le meilleur taux de bonne classification (80,13 %), suivies par l'OCS GE (78,95 %) et CoSIA (70,68 %). Dans ce contexte, la méthode de segmentation développée dans ce mémoire affiche un taux de bonne classification supérieur de 13 % à celui de l'OCS GE et de 22 % à celui de CoSIA. Par ailleurs, 134 millions de pixels ont été labellisés pour la phase d'entraînement. À titre de comparaison, l'IGN a labellisé 20.3 milliards de pixels pour CoSIA, ce qui représente un volume d'entraînement 150 fois plus élevé, le tout rendu possible par une puissance de calcul importante⁶². Malgré tout, il faut souligner que là où le modèle réalisé dans ce mémoire affiche un IoU moyen de 0.77, CoSIA affiche dans leur article de données (Garioud, De Wit et al. 2023) un IoU moyen de 0.56. Idem pour l'*overall accuracy* qui est de 92.2 % ici, contre 72.8 % pour FLAIR #2⁶³. Un comparatif avec CoSIA est plus détaillé dans l'annexe .4.

Ces performances témoignent de l'adéquation entre les données issues de la méthode de ce mémoire aux **besoins métiers du Grand Avignon**, ce qui permet une exploitation directe des résultats.

En définitive, l'hypothèse H₄ qui suggère que la méthode par apprentissage profond incarne une plus-value dans la cartographie de l'occupation du sol à fine échelle spatiale, temporelle, et sur la reproductibilité des résultats sur différentes zones et millésimes, est validée.

5.3 Axes d'amélioration et pistes de recherche

En dépit d'une méthode précise, homogène et reproductible, il subsiste quelques axes d'amélioration, qui n'ont pas pu être comblés pour des raisons temporelles ou matérielles. Ces axes d'améliorations présentés ci-dessous permettent, par ailleurs, de porter un regard critique à l'ensemble du projet.

Premièrement, les 8 172 patchs d'images et de masque d'entraînement sont tous issus du Grand Avignon, ce qui est un échantillonnage relativement restreint d'un point de vue quantitatif, mais aussi dans la diversité de la typologie des paysages. Outre le fait que certaines classes demeurent sous-représentées (cf. figure 44), certaines configurations spatiales le sont également. Par exemple, les

62. CoSIA a été entraîné et inféré sur le supercalculateur Jean Zay.

63. Toutefois, cette comparaison est délicate pour deux raisons principales. D'abord, CoSIA dispose de 13 classes, là où le modèle proposé dans ce mémoire n'en segmente que six. Deuxièmement, de toute évidence, segmenter l'occupation du sol sur la totalité du territoire français, et ce, à 20 centimètres de résolution, n'implique pas la même difficulté que de faire la même chose à un mètre pour le Grand Avignon, et ce, de tout point de vue.

bâtiments industriels ou commerciaux, les centres urbains très denses (comme Avignon *intramuros*) sont sous-représentés par rapport aux maisons individuelles. Le constat est le même pour les petits cours d'eau (Sorgue) vis-à-vis des cours d'eau plus gros (Rhône ou Durance). Ainsi, en fournissant plus de données lors de la phase d'entraînement, le modèle pourrait consolider son apprentissage des configurations spatiales, ce qui aurait un effet bénéfique sur la performance globale du modèle.

Par ailleurs, comme mentionné plus haut, on constate un léger effet arrondi, principalement autour des bâtiments, ce qui peut légèrement biaiser le calcul des surfaces bâties. Ce phénomène est intrinsèque au modèle U-Net à cause des couches de déconvolution. Toutefois, une manière de minimiser cet effet serait de travailler à une résolution spatiale plus petite, de 20 ou 50 centimètres, comme le permet la BD Ortho, mais comme ne le permet pas la machine à notre disposition.

Ensuite, une piste de recherche consisterait à pondérer les cartes de probabilités avant l'application de la fonction `argmax()`⁶⁴. En appliquant un coefficient de pondération aux probabilités d'appartenance à une classe, on pourrait tendre vers une apparition plus fréquente de certaines classes. Par exemple, en multipliant par 1,1 les probabilités d'appartenance à la classe "bitume", certaines zones partiellement recouvertes de graviers (petites routes, chemins ou parkings) pourraient être reclassées comme telles, plutôt que d'être assimilées à du sol nu. Le choix du coefficient multiplicateur doit être justifié par une méthode statistique, qui pourrait être la mesure de l'IoU pour tous les coefficients compris entre 0.5 et 1.5 par pas de 0.1.

Finalement, le dernier axe d'amélioration réside dans la capacité à itérer, pour produire des millésimes à fine résolution temporelle. Une des conclusions du *benchmark* du Chapitre 3 est que les données étudiées sont produites à minima tous les trois ans, et avec une fraîcheur⁶⁵ de plusieurs années. Face à ce constat, la méthode proposée dans ce mémoire a pour ambition de répondre à ce besoin de monitorer les territoires à un pas de temps réduit. Toutes les données utilisées pour les images d'entraînement et d'inférence proviennent de l'IGN (BD Ortho IRC, RGE Alti, MNS). De ce fait, nous sommes trop dépendants de la mise à jour de ces données de la part de l'institut pour itérer dans la chaîne de traitements et produire des millésimes (cf. tableau 16).

Il est néanmoins possible d'acquérir des données équivalentes autrement qu'en faisant appel à l'IGN. Les données de la BD Ortho IRC peuvent-être remplacées

64. Pour rappel, la carte des probabilités est ce que le modèle U-Net retourne. Il s'agit d'une image ayant autant de bandes que de classes à prédire. Chacune des bandes décrit la probabilité d'appartenance à une classe.

65. Laps de temps entre l'acquisition et la diffusion de la donnée.

par des images aériennes ou satellitaires (notamment Pléiades Neo). Le modèle numérique de hauteur peut être acquis par relevé LiDAR ou par Pléiades-1A, et permettrait également de calculer l'indice de texture *inverse different moment*.

Finalement, le canal morphologique le plus problématique est celui du *patch shape index*. En effet, celui-ci est, contrairement aux huit autres canaux, produit à partir du masque d'entraînement, qui lui-même est dépendants de nombreuses sources de données (LiDAR HD, BD Ortho IRC, BD Topo, OSM). C'est celui qui est le moins reproductible. Dès lors, deux options se présentent :

- Assumer un anachronisme entre les huit premiers canaux (R, V, IR, NDVI, NDWI, NDGI, MNH, IDM) et le *patch shape index* (PSI). Il serait intéressant d'étudier les résultats de l'inférence du modèle dont un des neuf canaux d'entrée n'est pas de la même année.
- Entraîner un modèle en se passant du PSI. Comme montré dans le tableau 14, cela impliquera une perte de performance du modèle, qui se compensera par une capacité à produire des millésimes plus fréquemment.

Il y a donc un arbitrage à faire entre agilité et précision. Par ailleurs, l'acquisition de données pour se passer de la dépendance à l'IGN à l'échelle d'une collectivité territoriale est envisageable. En effet, comme mentionné, les enjeux de l'artificialisation à l'échelle locale sont omniprésents. Le recours à des levés par drone, à l'imagerie satellite ou aérienne haute résolution ou encore à des campagnes de relevés LIDAR permettrait d'obtenir des données adaptées aux besoins spécifiques du modèle, tout en réduisant la dépendance aux données de l'IGN dont la fréquence de mise à jour est limitée. Cependant, la mise en place de ces acquisitions représente un investissement financier. Cette question relève désormais d'un choix politique qui dépasse le cadre de ce mémoire, mais elle mérite d'être intégré aux réflexions stratégiques des élus locaux dans le contexte actuel de transition écologique et de lutte contre l'artificialisation des sols.

Donnée	Masque d'entraînement	Image d'inférence	Fréquence de mise à jour
LiDAR HD	×		Non communiqué
RGE Alti		×	Non communiqué
BD Ortho	×	×	3 ou 4 ans
MNS Correl		×	3 ou 4 ans (issu de BD Ortho)
BD Topo	×		Trimestrielle
OSM	×		En continu

Table 16. Résumé des données, des usages et fréquences de mise à jour.
N. Massot, 2025

Chapitre 6 :

Conclusion

Ce chapitre conclusif récapitule les éléments contextuels du mémoire ainsi que ses objectifs initiaux. Ensuite, il ré-examinera l'analyse des hypothèses formulées dans l'introduction. Nous prendrons également du recul pour évaluer la méthode employée, en discutant de ses limites, des perspectives d'amélioration, et de la reproductibilité des résultats obtenus.

À la lumière des résultats obtenus, nous pouvons dire que nous avons atteint les objectifs de la présente étude. En effet, ce mémoire fut l'occasion de développer une chaîne de traitements qui ambitionne de répondre à la problématique suivante : "Dans quelle mesure les apports de l'apprentissage profond à la cartographie de l'occupation du sol permettent-ils de monitorer les territoires en adéquation avec les besoins métiers des acteurs publics ?". Cette problématique intervient dans un contexte de besoins grandissants en données fiables, homogènes et reproductibles, suite aux récentes actions législatives, et au fait que les données du marché présentent des axes d'améliorations notoires, notamment en ce qui a trait à la précision et à la résolution spatiale et temporelle.

6.1 Résultats et hypothèses

D'abord, nous avons commencé par détailler le cadre législatif, l'objectif ZAN, les définitions de la consommation d'espaces et de l'artificialisation des sols et les implications que ces éléments induiraient d'un point de vue méthodologique. Ensuite, un état de l'art des données d'occupation du sol traitait de la pertinence des données continues en tout point de l'espace, ainsi que des données contextuelles discontinues. En conséquent, cet état de l'art a permis de **rejeter l'hypothèse H₁**⁶⁶, puisqu'un certain nombre d'axes d'améliorations ont été relevés, notamment en ce qui a trait à la résolution spatiale et temporelle, ainsi qu'à la fraîcheur de la donnée. Par ailleurs, en comparant les avantages et inconvénients des méthodes nationalisées ou territorialisées, nous rappelons que ce sont les critères des besoins métiers qui définissent la notion de qualité. De ce fait, **nous validons l'hypothèse H₂**⁶⁷. La partie principale de ce mémoire consistait à concevoir une chaîne de traitements qui permet de segmenter aussi souvent que nécessaire l'occupation du sol à un mètre de résolution et à six classes. Cette méthode qui tend vers une automatisation à 100 % intègre le modèle de réseaux de neurones convolutifs U-Net. Ce modèle se caractérise par une capacité à prendre en compte les valeurs radiométriques, les formes, et à "apprendre" les structures spatiales pour mieux prédire leur typologie. Nous avons démontré que la colinéarité des images d'entraînement n'avait pas d'effets néfastes sur les métriques de performances du modèle. Nous pouvons donc **refuser l'hypothèse H₃**⁶⁸. Par ailleurs, cette chaîne de traitements se démarque par sa capacité à segmenter l'occupation du sol avec un haut niveau de précision. En effet, elle affiche un **score F1 de 0.91** et un **overall accuracy de 92.2 %**, là où le

66. H₁ : Les données d'occupation du sol et les données contextuelles discontinues actuelles sont suffisantes pour mesurer la consommation d'espaces ou l'artificialisation.

67. H₂ : La précision est une notion subjective relative aux besoins du commanditaire.

68. H₃ : La colinéarité des images d'entraînement a un impact négatif à la performance du modèle.

taux de bonne classification des données du *benchmark* ne dépasse pas les 80 %. Rappelons que sur la période 2020–2023, la région PACA a consommé 0,125 % de son territoire (Cerema 2024). Ainsi, **un écart d'accuracy de 13,25 % avec l'OCS GE est un changement de paradigme et de vision radical des dynamiques de consommation d'espaces**⁶⁹. À la lumière de ces éléments et au regard de la reproductibilité de la méthode, nous pouvons **adopter l'hypothèse H₄**⁷⁰ qui suggère que **la méthode par apprentissage profond incarne une plus-value dans la cartographie de l'occupation du sol à fine échelle spatiale, temporelle, et sur la reproductibilité des résultats sur différentes zones et millésimes.**

6.2 Perspectives, reproductibilité et science ouverte

L'ambition de cette méthode est d'être utilisée. De ce fait, une première perspective pour ce projet serait de se pencher sur l'implémentation de celle-ci au sein du Grand Avignon. Cette implémentation passe par le fait de transmettre le savoir, la méthode et la manière de s'en servir au commanditaire. Dans un contexte de crise de la reproductibilité, deux dépôts (Zenodo et GitHub) permettent de garantir la transparence et la réplicabilité des résultats. Cela permet à tout un chacun d'utiliser la chaîne de traitements et d'y apporter des améliorations. Finalement, une carte interactive permet de visualiser les résultats.

Tous les livrables sont mis à disposition en libre accès via un dépôt Zenodo qui contient :

- Le modèle FME pour la création du masque d'entraînement,
- Les images et le masque d'entraînement découpés en patchs 128 × 128 pixels (deux fois 18 000 patchs),
- Les six algorithmes python développés,
- Le modèle U-Net entraîné,
- La segmentation de l'occupation du sol du Grand Avignon,
- La carte des probabilités.

Parallèlement à cela, les six scripts python ainsi que le modèle FME pour la création du masque d'entraînement sont disponibles sur GitHub. Cela permet de pouvoir les maintenir et les mettre à jour si besoin.

Les ressources :

- Dépôt Zenodo : <https://doi.org/10.5281/zenodo.15591285> DOI 10.5281/zenodo.15591285

69. Accuracy de U-Net – Accuracy de l'OCS GE : 92,2 % – 78,95 % = 13,25 %

70. H₄ : La méthode par apprentissage profond incarne une plus-value dans la cartographie de l'occupation du sol à fine échelle spatiale, temporelle, et sur la reproductibilité des résultats sur différentes zones et millésimes.

- Dépôt GitHub : https://github.com/Mercatorien/MEMOIRE_MASSOT
- Site internet pour la visualisation du résultat :
https://mercatorien.github.io/MEMOIRE_MASSOT/

Fichier	Dépôt	Licence
01_PREPROCESSEUR.py	Zenodo + GitHub	GPL v3 + MIT
02_ENTRAINER.py	Zenodo + GitHub	GPL v3 + MIT
03_DECOUPER_SOUS_ENSEMBLE.py	Zenodo + GitHub	GPL v3
04_INFERRER.py	Zenodo + GitHub	GPL v3
05_FUSIONNER_SOUS_ENSEMBLE.py	Zenodo + GitHub	GPL v3
06_CARTE_PROBA.py	Zenodo + GitHub	GPL v3
SCRIPT_MASQUE_ENTRAINEMENT.fmw	Zenodo + GitHub	GPL v3
lc.json	Zenodo + GitHub	GPL v3
lcs.rar (<i>masque d'entraînement</i>)	Zenodo	LICENCE OUVERTE / OPEN LICENCE 2.0
images.rar (<i>image d'entraînement</i>)	Zenodo	LICENCE OUVERTE / OPEN LICENCE 2.0
UNET_MASSOT.tif	Zenodo	CC BY SA
STYLE_SEGMENTATION.qml	Zenodo	CC BY SA
MODEL_UNET_MASSOT.keras	Zenodo	CC BY SA
CARTE_PROBA.tif	Zenodo	CC BY SA

Table 17. Synthèse des dépôts et des licences.
N. Massot, 2025

6.3 Une mesure universelle est-elle réalisable ?

L'objectif Zéro Artificialisation Nette, instauré en 2021, ambitionne de mesurer la consommation d'espaces et l'artificialisation au mètre carré près en remontant jusqu'à 2011 pour définir le budget artificialisation des communes, dans l'optique de le réduire. La question de l'artificialisation des sols est, a minima, autant un problème qualitatif que quantitatif. Toutefois, le postulat assumé de la loi Climat et Résilience est exclusivement centré sur la réduction arithmétique de la consommation d'espaces d'un point de vue quantitatif, en faisant fi de la qualité des sols. Néanmoins, la cartographie de l'occupation du sol est un domaine systémique qui est lié à un certain nombre d'éléments (lois, régulateurs, élus locaux, lobbies, enjeux économiques, sociaux et environnementaux, méthodologie, ...). Ainsi, lorsque

nous définissons la chaîne de pensée pour établir une mesure de l'artificialisation qui fait foi, nous évoquons le fait qu'elle débute par l'identification des besoins, puis la définition claire et non ambiguë, en se terminant par l'élaboration d'une méthode géomatique en lien direct avec la définition et les besoins. Ici, l'étape des éléments de définition est fatallement bloquante, puisqu'en plus de considérer la consommation d'espaces ou l'artificialisation comme des éléments purement quantitatifs, celle-ci adopte une vision binaire de la naturelité des sols. De ce fait, les définitions tendent à alimenter les débats autour de la mesure de la consommation d'espaces. De plus, la consommation d'espaces ou l'artificialisation se jouent à des échelles particulièrement fines. La mesure de ces dernières n'échappe pas au *modifiable area unit problem* (MAUP) (Andresen 2021), qui démontre que les résultats d'une analyse spatiale dépendent de la taille des unités (c'est l'effet d'échelle), et de la forme de ces dernières (c'est l'effet de zonage). Ainsi, une cartographie de l'occupation du sol basée sur des zones homogènes de 2 500 m², comme celle de l'OCS GE, produira des résultats très différents de ceux obtenus à une résolution d'un mètre. Cette dernière différera elle-même d'une cartographie encore plus fine, à l'échelle de la branche, du brin d'herbe ou de l'atome.

Alors, une fois l'ensemble de ces éléments pris en compte, la possibilité d'une mesure universelle demeure incertaine. Pour autant, **la chaîne de traitements mobilisant l'intelligence artificielle, telle que décrite ici, se révèle comme une alternative sérieuse pour résoudre le problème la cartographie de l'occupation du sol.**

Références

- Andresen, M. (2021). « Modifiable areal unit problem ». In : *CrimRxiv*. doi : [10.21428/cb6ab371.5c28c076](https://doi.org/10.21428/cb6ab371.5c28c076).
- architectes, Ordre des (2025). « Ce que change la nouvelle loi ZAN ». In : url : <https://www.architectes.org/actualites/ce-que-change-la-nouvelle-loi-zan-91738>.
- Cerema (2024). Analyse de la consommation d'espaces : période du 1er janvier 2009 au 1er janvier 2023. Rapp. tech. Ministère de la Transition écologique. url : <https://doc.cerema.fr/Default/doc/SYRACUSE/601069/analyse-de-la-consommation-d-espaces-periode-du-1er-janvier-2009-au-1er-janvier-2023>.
- Dabboor, M., S. Howell, M. Shokr et J. Yackel (2014). « The Jeffries–Matusita distance for the case of complex Wishart distribution as a separability criterion for fully polarimetric SAR data ». In : *International Journal of Remote Sensing* 35.19, p. 6859–6873. url : <https://www.researchgate.net/publication/267396998>.
- Delattre, R. et T. Lecourt (juin 2024). « Faut-il pouvoir mesurer la consommation d'espace pour la réduire ? » In : Communication présentée à Géopoint – Mesurer l'espace et après ?, Avignon, France. url : <https://hal.science/hal-04801687v1>.
- Epstein, Renaud (2006). « Gouverner à distance : Quand l'État se retire des territoires ». In : *Esprit* (Paris, France : 1932), p. 96–111.
- Fernández-Delgado, M., E. Cernadas et S. Barro (2014). « Do we Need Hundreds of Classifiers to Solve Real World Classification Problems ? » In : *Journal of Machine Learning Research* 15, p. 3133–3181. url : <https://dl.acm.org/doi/10.5555/2627435.2697065>.
- FRAGSTATS (s. d.). P2 – Patch Shape Index. url : <https://fragstats.org/index.php/fragstats-metrics/patch-based-metrics/shape-metrics/p2-shape-index>.
- Fukushima, K. (1980). « Neocognitron : A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position ». In : *Biological Cybernetics* 36, p. 193–202. doi : [10.1007/BF00344251](https://doi.org/10.1007/BF00344251).
- Garioud, A., S. Bookjans, E. Peillet, S. Giordano et B. Wattrelos (2022). « FLAIR : French Land cover from Aerospace ImageRy ». In : url : <https://arxiv.org/abs/2211.12979>.
- Garioud, A., A. De Wit, M. Poupée, M. Valette, S. Giordano et B. Wattrelos (2023). « FLAIR 2 : textural and temporal information for semantic segmentation from multi-source optical imagery ». In : arXiv. url : <https://arxiv.org/abs/2305.14467>.
- Garioud, A., S. Peillet, E. Bookjans, S. Giordano et B. Wattrelos (2022). *FLAIR 1 : Semantic segmentation and domain adaptation dataset*. url : https://ignf.github.io/FLAIR/index_fr.html.
- Géoconfluences (2024). Loi SRU (loi relative à la solidarité et au renouvellement urbains). url : <https://geoconfluences.ens-lyon.fr/glossaire/loi-sru-france>.
- Guigues, L., J.P. Cocquerez et H. Le Men (2006). « Scale-sets Image Analysis ». In : *International Journal of Computer Vision* 68.3, p. 289–317. doi : [10.1007/s11263-005-6299-0](https://doi.org/10.1007/s11263-005-6299-0).
- Haury, Y. (2023). Question orale n° 317 : Zéro artificialisation nette et serres agricoles. url : <https://questions.assemblee-nationale.fr/q16/16-317QOSD.htm>.
- IGN (2023). Cosia - Documentation technique. Rapp. tech. url : https://cosia.ign.fr/pdf/Cosia_Documentation_Technique IGN_2023.pdf.

- IGN (2024a). *FLAIR : Artificial Intelligence challenges organised around geo-data and deep learning*. url : https://ignf.github.io/FLAIR/index_fr.html.
- (2024b). *Masque Forêt – Version Beta v2*. url : https://geoservices.ign.fr/sites/default/files/2024-07/Masque-Forêt-version-Beta_v2.pdf.
 - (2024c). *Rapport de contrôle qualité de la BD Topo*. Rapp. tech. url : https://geoservices.ign.fr/sites/default/files/2024-01/RCQ_BDTOPO_D84.pdf.
- Iota2-project (2024). *iota2 - GitLab*. url : https://framagit.org/iota2-project/iota2/-/tree/develop?ref_type=heads.
- Kirillov, A., E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A.C. Berg, W. Lo, P. Dollár et R. Girshick (2023). « Segment Anything ». In : url : <https://arxiv.org/abs/2304.02643>.
- Körting, T.S., L.M.G. Fonseca et G. Câmara (2013). « GeoDMA—Geographic Data Mining Analyst ». In : *Computers Geosciences* 57, p. 133–145. doi : [10.1016/j.cageo.2013.02.007](https://doi.org/10.1016/j.cageo.2013.02.007).
- Kwan, C., B. Ayhan, B. Budavari, Y. Lu, D. Perez, J. Li, S. Bernabe et A. Plaza (2020). « Deep Learning for Land Cover Classification Using Only a Few Bands ». In : *Remote Sensing*. url : <https://doi.org/10.3390/rs12122000>.
- Lascoumes, Pierre et Patrick Le Galès (2005). « Introduction : L'action publique saisie par ses instruments ». In : *Gouverner par les instruments*. Presses de Sciences Po, p. 11–44.
- Lenormand, A. (2013). « Urbanisme – Projet de loi Alur : une réforme ambitieuse du droit de l'urbanisme ». In : *Banque des Territoires*. url : <https://www.banquedesterritoires.fr/projet-de-loi-alur-une-reforme-ambitieuse-du-droit-de-lurbanisme>.
- « Loi n° 2016-1087 du 8 août 2016 pour la reconquête de la biodiversité, de la nature et des paysages » (s. d.). In : *Légifrance* (). url : <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000033016237>.
- « Loi n° 2021-1104 du 22 août 2021 portant lutte contre le dérèglement climatique et renforcement de la résilience face à ses effets » (2021). In : *Journal officiel de la République française*. url : <https://www.legifrance.gouv.fr/loda/id/JORFTEXT000043956924/>.
- Movement, World Rainforest (2024). Qu'est-ce qu'une forêt ? url : <https://www.wrm.org.uy/sites/default/files/2024-03/WhatIsAForest-FR.pdf>.
- Nigri Happ, Patrick, Rodrigo Ferreira, Cristiana Bentes, Gilson Costa et Raul Feitosa (2010). « Multiresolution segmentation : a parallel approach for high resolution image segmentation in multicore architectures ». In.
- Osco, L.P., Q. Wu, E.L. de Lemos, W.N. Gonçalves, A.P.M. Ramos, J. Li et J. Marcato (2023). « The Segment Anything Model (SAM) for remote sensing applications : From zero to one shot ». In : *International Journal of Applied Earth Observation and Geoinformation* 124, p. 103540. doi : [10.1016/j.jag.2023.103540](https://doi.org/10.1016/j.jag.2023.103540).
- Pelletier, Charlotte (2017). « Cartographie de l'occupation des sols à partir de séries temporelles d'images satellitaires à hautes résolutions Identification et traitement des données mal étiquetées ». Thèse de doct. Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier). url : <https://tel.archives-ouvertes.fr/tel-01665121v2>.
- Rault, Ashwini et al. (2016). « Texture Parameters Extraction of Satellite Image ». In : *International Journal of Science Technology and Engineering*. url : <https://www.ijste.org/articles/IJSTEV2I11015.pdf>.

- Ronneberger, O., P. Fischer et T. Brox (2015). « U-Net : Convolutional Networks for Biomedical Image Segmentation ». In : url : <https://arxiv.org/pdf/1505.04597>.
- Smith, B. (1995). « On drawing lines on a map ». In : *Spatial Information Theory : A Theoretical Basis for G/S*. Sous la dir. d'A.U. Frank et W. Kuhn. T. 988. Lecture Notes in Computer Science. Springer, p. 475-484. doi : [10.1007/3-540-60392-1_31](https://doi.org/10.1007/3-540-60392-1_31).
- Transition écologique, Ministère de la (sept. 2022). *Loi Solidarité et renouvellement urbain (SRU)*. url : <https://www.ecologie.gouv.fr/politiques-publiques/loi-solidarite-renouvellement-urbain-sru>.
- Uhart, T. (2023). « CARTE. Votre commune respecte-t-elle la loi en matière de logement social ? » In : Franceinfo. url : https://www.francetvinfo.fr/societe/carte-votre-commune-respecte-t-elle-la-loi-en-matiere-de-logement-social_6195918.html.
- Wikipedia (2025a). *Classification*. url : <https://fr.wikipedia.org/wiki/Classification>.
- (2025b). *LeNet*. url : https://en.wikipedia.org/wiki/LeNet#1989_LeNet.
- Wu, Q., L.P. Osco, E.L. de Lemos, W.N. Gonçalves, A.P.M. Ramos, J. Li et J. Marcato (2023). « samgeo : A Python package for segmenting geospatial data with the Segment Anything Model (SAM) ». In : *Journal of Open Source Software* 8.89, p. 5663. doi : [10.21105/joss.05663](https://doi.org/10.21105/joss.05663).

Table des figures

1	Carte de localisation du Grand Avignon et ses communes. N. Massot, 2025	16
2	Territorialisation des objectifs ZAN. Source : La lutte contre l'artificialisation des sols, Notre affaire à tous, 2023, modifié par N. Massot, 2025	21
3	Évolution du ZAN N. Massot, 2025	24
4	Matrice de l'artificialisation selon l'occupation et l'usage des sols de l'OCS GE. Source : IGN	32
5	Classes de la couverture et de l'usage du sol. N. Massot, 2024. Source : IGN	33
6	Nombre de départements cartographiés par l'OCS GE chaque année. N. Massot, 2024. Source : Institut National de l'Information Géographique et Forestière, OCS GE - Occupation du sol à grande échelle, Consulté le 29 décembre 2024, https://geoservices.ign.fr/ocsge	34
7	Comparatif de l'OCS GE et CoSIA. Légende OCS GE : cf. figure 5; Légende CoSIA : cf. figure 9. N. Massot, 2024. Source : IGN	35
8	Pyramides de segmentation. Source : Guigues, L., Cocquerez, J. P., & Le Men, H. (2006). Scale-sets Image Analysis. <i>International Journal of Computer Vision</i> , 68(3), 289–317. https://doi.org/10.1007/s11263-005-6299-0	37
9	Classes de CoSIA. N. Massot, 2024. Source : Institut National de l'Information Géographique et Forestière (IGN). (2023). Cosia - Documentation technique. Consulté le 29 décembre 2024, sur https://cosia.ign.fr/pdf/Cosia_Documentation_Technique_IGN_2023.pdf	38
10	Chaîne de traitements de CoSIA. N. Massot, 2024. Source : Source : IGN. (2023). Cosia - Documentation technique. Consulté le 29 décembre 2024, sur https://cosia.ign.fr/pdf/Cosia_Documentation_Technique_IGN_2023.pdf	39
11	Le Corine Land Cover sur Avignon : des zones homogènes de 25 ha. N. Massot, 2024. Source : Copernicus	40
12	Classes du CLC niv. 3. N. Massot, 2024. Source : Copernicus	41
13	Classes de l'ESA World Cover. N. Massot, 2024. Source : ESA	42
14	Classes du produit OSO. N. Massot, 2024. Source : Pôle Theia	44
15	Classes de ESRI Land Cover. N. Massot, 2024. Source : ESRI	45
16	Part de la superficie des tags de la clé 'landuse' sur le Grand Avignon. N. Massot, 2024. Source : OpenStreetMap	46

17	Nombre de classes par couche d'occupation du sol. N. Massot, 2024. Source : IGN, Copernicus, ESRI, OSM, ESA, Pôle Théia	47
18	Reclassification des couches d'occupation du sol. N. Massot, 2024. Source : IGN, Copernicus, ESRI, OSM, ESA, Pôle Théia	48
19	Échantillon pour la création de la vérité terrain de 2024. N. Massot, 2024	50
20	Mesure de la précision : Occupation du sol à Grande Échelle (OCS GE). N. Massot, 2024	51
21	Mesure de la précision : CoSIA. N. Massot, 2024	52
22	Mesure de la précision : Corine Land Cover. N. Massot, 2024	53
23	Mesure de la précision : ESA World Cover. N. Massot, 2024	54
24	Mesure de la précision : Occupation des Sols Opérationnelle (OSO). N. Massot, 2024	55
25	Mesure de la précision : ESRI Land Cover. N. Massot, 2024	56
26	Mesure de la précision : OpenStreetMap. N. Massot, 2024	57
27	Classes du LiDAR HD. N. Massot, 2024. Source : IGN	58
28	Incohérences de la BD Topo en comparaison avec la BD Ortho 2024 sur Avignon. N. Massot, 2024. Source : IGN	60
29	Incohérences dans le Registre Parcellaire Graphique. N. Massot, 2024. Source : IGN	62
30	Incohérences de la BD Forêt en comparaison avec la BD Ortho 2021 pour le Vaucluse et 2018 pour le Gard sur le Grand Avignon. N. Massot, 2024. Source : IGN	64
31	Typologie non-exhaustive des classificateurs pour la cartographie de l'occupation du sol. N. Massot, 2025	71
32	Architecture U-net. Chaque boîte bleue correspond à une image multi- canaux. Le nombre de canaux est marqué au-dessus de chaque boîte. Les boîtes grises représentent les images concaténées. La dimension des images est marquée en bas à gauche de chaque boîte. N. Massot, 2025	74
33	Convolution multi-canaux. Source : FIDLE, 2024. <i>Programme FIDLE</i> 2024/25. CNRS. https://fidle.cnrs.fr/w3/programme.html	75
34	Max-pooling. Source : FIDLE, 2024. <i>Programme FIDLE 2024/25</i> . CNRS. https://fidle.cnrs.fr/w3/programme.html	76
35	Sous-apprentissage, apprentissage approprié et sur-apprentissage en apprentissage profond. N. Massot, 2025	77
36	Pondération des pixels selon la fenêtre de Hann. Source : N. Pielawski, and C. Wählby (2019), Introducing Hann windows for reducing edge- effects in patchbased image segmentation, https://arxiv.org/pdf/1910.07831	78

37	Chaîne de traitements mobilisée. N. Massot, 2025	81
38	Tentative de segmentation des bâtiments et des arbres avec samgeo. N. Massot, 2025. Source : BD Ortho, Google satellite	86
39	Limites de l'occupation du sol multi-sources. N. Massot, 2025	88
40	Effet du seuil de densité de point minimal sur le détourage des éléments du sursol. N. Massot, 2025	89
41	Chaîne de traitements pour la réalisation du masque d'entraînement avec FME. N. Massot, 2025	91
42	Matrice de confusion du masque d'entraînement. N. Massot, 2025 . . .	93
43	Extraits du masque d'entraînement, et comparaison avec la vérité terrain et l'OCS GE. N. Massot, 2025	94
44	Fréquence d'apparition des classes dans le masque d'entraînement avant et après découpe en patchs. N. Massot, 2025	96
45	Éléments spectraux et morphologiques utilisés pour les images d'en- traînement. N. Massot, 2025	99
46	Techniques d'augmentation. Source : Tian, Y., & Zhang, Y. (2022). A com- prehensive survey on regularization strategies in machine learning. In- formation Fusion, 80, 146–166. https://doi.org/10.1016/j.inffus.2021.11.005	100
47	Signatures spectrales des classes selon la bande spectrale (valeurs normalisées). N. Massot, 2025	101
48	Relation conceptuelle nombre de bandes spectrales – nombre de classes avec la distance de Jeffries-Matusita. N. Massot, 2025	102
49	Séparabilité des classes selon la distance de Jeffries-Matusita pour les canaux spectraux et morphologiques utilisés pour les images d'entraî- nement. N. Massot, 2025	103
50	Matrice de corrélation des canaux spectraux et morphologiques utilisés pour les images d'entraînement. N. Massot, 2025	104
51	Indicateurs de performance du modèle après entraînement. N. Massot, 2025	108
52	Cartographie de la segmentation (b) en comparaison de l'image satel- litaire (a) N. Massot, 2025	112
53	Carte des probabilités (c). N. Massot, 2025	113
54	Fréquence cumulée des probabilités maximales. N. Massot, 2025	113
55	Mesure de Précision du modèle U-net. N. Massot, 2025	116
56	Mesure du rappel du modèle U-net. N. Massot, 2025	117
57	IoU global, par classe et moyen obtenue par l'architecture U-Net basé sur le masque d'entraînement. N. Massot, 2025	118

58	Flux d'occupation du sol entre la sortie du modèle U-Net et CoSIA. N. Massot, 2025	139
----	--	-----

Liste des tableaux

2	Objectif de consommation d'espaces des SRADDET N. Massot, 2025. Source : Doré, G. (2023). <i>Les SRADDET et l'objectif ZAN en 2023</i>	22
3	Définitions de la consommation d'espaces et de l'artificialisation. N. Massot, 2025. Source : Décret n° 2023-1096 du 27 novembre 2023 relatif à l'évaluation et au suivi de l'artificialisation des sols	25
4	Catégories de surfaces et seuils de référence. N. Massot, 2025 (*) Les infrastructures linéaires sont qualifiées à partir d'une largeur minimale de cinq mètres. (**) Une surface végétalisée est qualifiée d'herbacée dès lors que moins de vingt-cinq pour cent du couvert végétal est arboré. Source : Article R. 101-1 du Code de l'Urbanisme	26
5	Précision globale des différentes données de couverture du sol. N. Massot, 2025.	65
6	Comparaison des concepts d'outils nationaux et territorialisés. N. Massot, 2025	68
7	Composition des classes du modèle. N. Massot, 2025	82
8	Agrégation des classes d'occupation du sol les plus précises. Les chiffres sont issus des matrices de confusion de la partie 3.1.2.2. N. Massot, 2025	84
9	Les trois manières d'utiliser samgeo. N. Massot, 2025	85
10	Structure de la variable <code>Classe_NDVI_Hauteur</code> . N. Massot, 2025	87
11	Interprétation de <code>Classe_NDVI_Hauteur</code> N. Massot, 2025	87
12	Sources des classes d'occupation du sol et ordre de priorité. Ce tableau synthétise le diagramme d'activité de la figure 41 . N. Massot, 2025	90
13	Liste des indices spectraux et leur définition. N. Massot, 2025	97
14	Tests de performance unitaires en enlevant les bandes unes par unes. N. Massot, 2025	106
15	Effectifs des échantillons pour chaque classe. N. Massot, 2025	114
16	Résumé des données, des usages et fréquences de mise à jour. N. Massot, 2025	122
17	Synthèse des dépôts et des licences. N. Massot, 2025	126
18	Règles de classification hiérarchique pour l'occupation du sol	135

19	Hyperparamètres utilisés pour l'entraînement du modèle U-Net. N. Massot, 2025	137
20	Configuration matérielle utilisée pour l'entraînement et l'inférence du modèle. N. Massot, 2025	137
21	Correspondance entre les classes U-Net et les classes Cosia. N. Massot, 2025	138

Annexes

.1 Règles de classification hiérarchique pour l'occupation du sol (méthode rejetée)

Table 18. Règles de classification hiérarchique pour l'occupation du sol

Priorité	Classe_NDVI_Hauteur	Champ(s) déclencheur(s) (OCS GE)	Classe attribuée
1	301	—	Végétation haute
2.1	301, 201	<code>nature = 'Serre'</code> , <code>hauteur ≤ 3</code>	Serre
2.2	101	<code>nature = 'Serre'</code>	Végétation basse
2.3	100	<code>nature = 'Serre'</code>	Sol nu
2.4	—	<code>nature = 'Serre'</code>	Serre
3	—	<code>Libellé Code CS = 'Bâtiment'</code>	Bâtiment
4.1	101	—	Végétation basse
4.2	201	—	Végétation moyenne
5	—	<code>Libellé Code CS = 'Surfaces d'eau'</code>	Surfaces d'eau
6.1	100	<code>Libellé Usage = 'Réseaux routhiers'</code> , <code>CODE_CS = 'Formations herbacées'</code>	Sol nu
6.2	101	<code>Libellé Usage = 'Réseaux routhiers'</code> , <code>CODE_CS = 'Formations herbacées'</code>	Végétation basse

Suite page suivante

Suite de la Table 18

Priorité	Classe_NDVI_Hauteur	Champ(s) déclencheur(s) (OCS GE)	Classe attribuée
6.3	—	Libellé Usage = 'Réseaux rou- tiers'	Réseaux routiers
7.1	—	Libellé Usage = 'Réseaux fer- rés', CODE_CS = 'Zones à maté- riaux minéraux'	Réseaux ferrés
7.2	100	Libellé Usage = 'Réseaux fer- rés', CODE_CS = 'Formations her- bacées'	Sol nu
7.3	100	Libellé Usage = 'Réseaux fer- rés', CODE_CS = 'Zones bâties'	Imperméable
7.4	—	Libellé Usage = 'Réseaux fer- rés'	Réseaux ferrés
8	NULL	Libellé Usage = 'Production se- condeaire', CODE_CS ∈ {'Forma- tions herbacées', 'Zones bâties'}	Photovoltaïque au sol
9.1	100, NULL	CODE_CS = 'Zones bâties'	Imperméable
9.2	100	CODE_CS ≠ 'Zones bâties', usage_1 ∈ {'Résidentiel', 'Mixte', 'Tertiaire', 'Aérien'}	Imperméable
9.3	100	CODE_CS = 'Zones à matériaux minéraux'	Imperméable
10	100, NULL	Cas restants	Sol nu

.2 Hyperparamètres

Hyperparamètres	Valeur
Epochs	80
Learning Rate	0.0001
Neuron	64
Kernel (convolution)	3
Kernel (conv. transposée)	2
Dropout	0.1
Strides	2
Pool	2
Padding	Same

Table 19. Hyperparamètres utilisés pour l'entraînement du modèle U-Net.
N. Massot, 2025

.3 Configuration

Élément	Configuration
Processeur	AMD Ryzen 7 5800X
Fréquence CPU	3.80 GHz
Cœurs	8
Mémoire RAM	64,0 Go
Carte graphique	NVIDIA GeForce RTX 3060
Mémoire GPU	12 Go

Table 20. Configuration matérielle utilisée pour l'entraînement et l'inférence du modèle.
N. Massot, 2025

4 Comparatif avec CoSIA

Il est pertinent de comparer le résultat de ce mémoire avec une autre source de données afin d'évaluer leur cohérence mutuelle. Dans ce travail, nous avons choisi de confronter notre cartographie issue du modèle U-Net au produit CoSIA. L'objectif est de quantifier, à l'échelle du pixel, les changements d'occupation du sol entre ces deux jeux de données. Pour cela, CoSIA a d'abord été reclassé selon les six mêmes catégories (cf. tableau 21). Ensuite, nous avons construit un diagramme de Sankey (cf. figure 58) pour visualiser les flux proportionnels des transitions d'occupation du sol entre U-Net et CoSIA.

Ce diagramme révèle qu'aucun flux supérieur à 2 % n'apparaît pour les catégories "bâtiments", "bitume" et "eau". En revanche, des transitions plus significatives sont observées pour les classes végétales : notamment du "arbre" vers "herbe" et de "herbe" vers "sol nu". Ces écarts peuvent s'expliquer par un diachronisme des prises de vue, qui ne sont pas nécessairement effectuées durant la même année, ni à la même saison. Malgré ces variations localisées, la structure générale des flux reste cohérente, ce qui souligne la fiabilité globale de la comparaison.

Enfin, ces résultats illustrent l'importance de produire des cartes d'occupation du sol à haute résolution temporelle. En effet, en multipliant les millésimes, on limite les biais liés à la saisonnalité et on améliore la robustesse des analyses diachroniques, tout en s'assurant que les évolutions de l'occupation du sol ne sont pas liées à d'autres facteurs exogènes.

Classes U-Net	Classes Cosia
Bâtiment	Bâtiment, Serre
Arbre	Feuillu, Conifère
Bitume	Zone imperméable
Eau	Surface d'eau, Piscine
Herbe	Culture, Broussaille, Pelouse
Sol nu	Zone perméable, Terre labourée, Sol nu, Vignes

Table 21. Correspondance entre les classes U-Net et les classes Cosia.
N. Massot, 2025

Apport des réseaux de neurones convolutifs à la cartographie de l'occupation du sol :
Cas d'usage sur le Grand Avignon

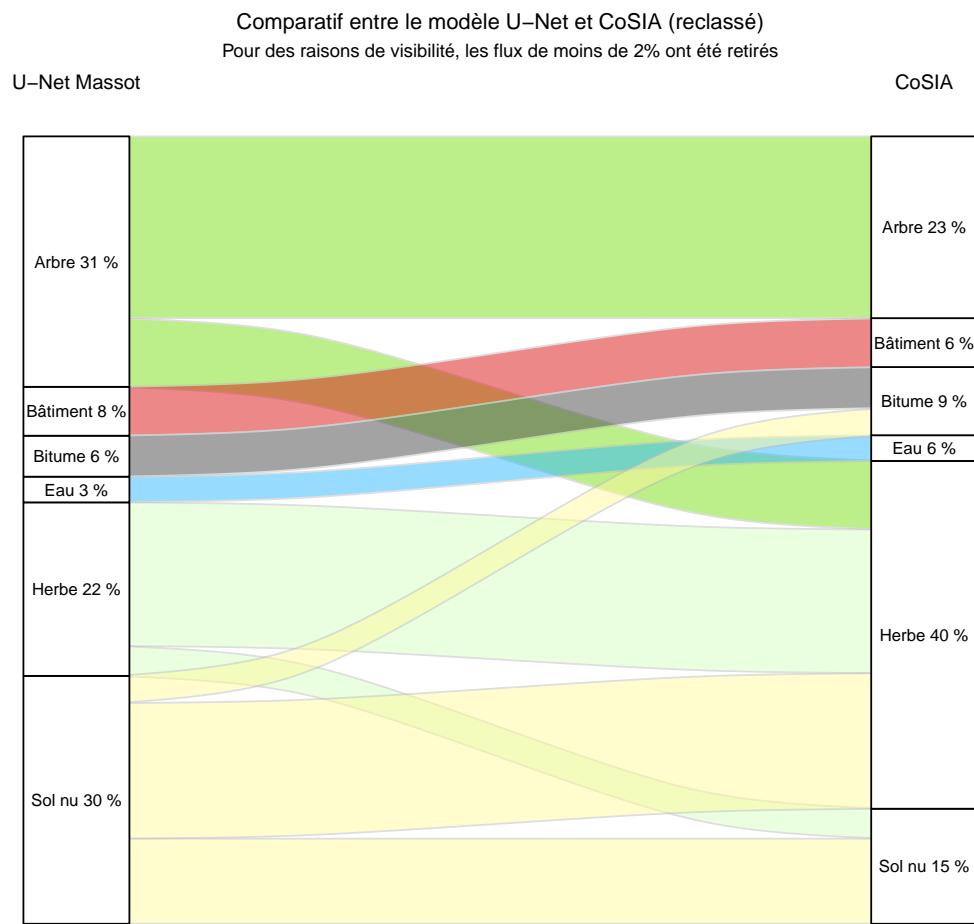


Figure 58. Flux d'occupation du sol entre la sortie du modèle U-Net et CoSIA.
N. Massot, 2025

