

Cartographie nationale de données de zones climatiques locales

Exploitation de l'imagerie
satellite SPOT 2022

Rapport technique

Juin 2025

Le Cerema est un établissement public sous la tutelle du ministère de la Transition écologique, présent partout en métropole et dans les Outre-mer grâce à ses 26 implantations et ses 2 400 agents. Détenteur d'une expertise nationale mutualisée, le Cerema accompagne l'État et les collectivités territoriales pour la transition écologique, l'adaptation au changement climatique et la cohésion des territoires par l'élaboration coopérative, le déploiement et l'évaluation de politiques publiques d'aménagement et de transport. Doté d'un fort potentiel d'innovation et de recherche incarné notamment par son institut Carnot Clim'adapt, le Cerema agit dans 6 domaines d'activités : Expertise & ingénierie territoriale, Bâtiment, Mobilités, Infrastructures de transport, Environnement & Risques, Mer & Littoral.

Site web : www.cerema.fr

Cartographie nationale de données de zones climatiques locales

Exploitation de l'imagerie satellite SPOT 2022

Juin 2025

Responsable du rapport :

Benjamin PICCININI – Département Territoires – Observation Satellitaire Et Changement Climatique
Tel. : +33(0)6 58 48 71 25
Mel : benjamin.piccinini@cerema.fr
Direction Territoriale Occitanie – 1 avenue du Colonel Roche – 31400 Toulouse

Suivi du rapport :

Nom	Service	Rôle	Date	Visa
Benjamin PICCININI	DTer Occ / DT / OSECC	Auteur principal		
Julien BOUYER	DTer Est / AN / TEAM	Contributeur		
Loéna TROUVE	DTec TV / DADET / ADCT	Constitutrice		
Arnaud CEYTE	DTer Occ / DT / OSECC	Relecteur		
Florence DROUY	MTEECPR / DGALN / DHUP / AD4	Relectrice		
Yann LANCIEN	MTEECPR / DGALN / DHUP / AD4	Relecteur		
Marine TRANCHANT	DTec TV / DADET / RTC	Relectrice		
Christelle DUC	DTec TV / DADET	Valideur	30/06/2025	CD
Quentin GAUTIER	DTer Occ / DT	Valideur	30/06/2025	QG

Résumé de l'étude

Le présent document a pour objectif de décrire une méthode de diagnostic du comportement climatique des tissus urbains, adaptée par le Cerema et se basant sur le concept des *Local Climate Zones* (LCZ ou zones climatiques locales) défini par Stewart & Oke en 2012.

Cette méthode, qui consiste en une classification de la typo-morphologie d'un territoire à partir d'une série d'indicateurs, a fait l'objet de travaux de R&D durant plusieurs années.

Durant son développement, cette méthode de classification en LCZ est passée par différents stades d'évolution, mobilisant diverses sources de données géographiques, certaines issues d'images satellites, d'autres de bases de données institutionnelles (comme la BD TOPO de l'IGN).

Dans sa version finale, dite opérationnelle, et dans une optique d'application à l'international, elle est conçue pour n'utiliser que des données dérivées de l'imagerie satellite, même si l'exploitation qui peut en être faite sur des territoires en France inclut l'usage de la BD TOPO. Elle a cependant la capacité de prendre en entrée toute donnée, pour peu qu'elle soit correctement formatée.

Ce document accompagne et décrit spécifiquement la production d'une carte nationale des LCZ sur les aires urbaines de plus de 50 000 habitants de France métropolitaine, basée sur la couverture annuelle SPOT 2022, la BD TOPO v3.3 de décembre 2022 et le découpage Urban Atlas 2018.

5 à 10 mots clés à retenir de l'étude

Surchauffe urbaine	ICU
LCZ	satellite
SPOT	2022
BD TOPO	Urban Atlas
Îlot de Chaleur Urbain	Zone Climatique Locale

Statut de communication de l'étude

- ☒ Accès libre : document accessible au public sur internet
- ☐ Accès restreint : document accessible uniquement aux agents du Cerema
- ☐ Accès confidentiel : document non accessible

Cette étude est capitalisée sur la plateforme documentaire [CeremaDoc](https://doc.cerema.fr/depot-rapport.aspx), via le dépôt de document : <https://doc.cerema.fr/depot-rapport.aspx>

SOMMAIRE

INTRODUCTION.....	6
Le phénomène d'Îlot de Chaleur Urbain.....	6
La classification en Zones Climatiques Locales.....	8
1 HISTORIQUE DE LA MÉTHODE CEREMA.....	10
1.1 L'algorithme analytique Stewart & Oke.....	10
1.2 Les algorithmes intermédiaires.....	11
1.3 L'algorithme opérationnel.....	11
2 LES DONNÉES D'ENTRÉE.....	12
2.1 Segmentation morphologique.....	12
2.1.1 Traitement de territoires complémentaires.....	13
2.2 Vecteur bâti 3D.....	15
2.3 Raster OCS 6 classes.....	16
2.3.1 Focus sur la caractérisation de la végétation.....	17
3 LES INDICATEURS.....	18
3.1 Morphologie urbaine.....	18
3.2 Occupation du sol.....	19
4 OBTENTION DE LA CARTE LCZ.....	20
4.1 Préparation des données d'entrée.....	20
4.1.1 Segmentation morphologique.....	21
4.1.2 Bâti 3D.....	21
4.1.3 OCS 6 classes.....	21
4.2 Calcul des indicateurs et attribution de la typologie LCZ.....	22
5 LIMITES DE LA DONNÉE LCZ.....	24
6 OUTILS DE PRODUCTION.....	26
6.1 Scripts Python.....	26
6.2 Modeleurs QGIS.....	26
ANNEXES.....	27
Bibliographie.....	27
Glossaire.....	27
Index des figures.....	27

Crédits, sauf mention contraire :

- images d'illustrations : © Cerema

- images satellites : SPOT © Airbus DS

INTRODUCTION

Le phénomène d'Îlot de Chaleur Urbain

Les données de Météo-France montrent que la fréquence et l'intensité des vagues de chaleur ont augmenté au cours de ces trente dernières années, et que cette tendance va se poursuivre sur les décennies à venir (Figure 1). La canicule d'août 2003 fait toujours référence, car c'est l'événement le plus sévère que la France ait connu depuis au moins 1947 en termes de vagues de chaleur.

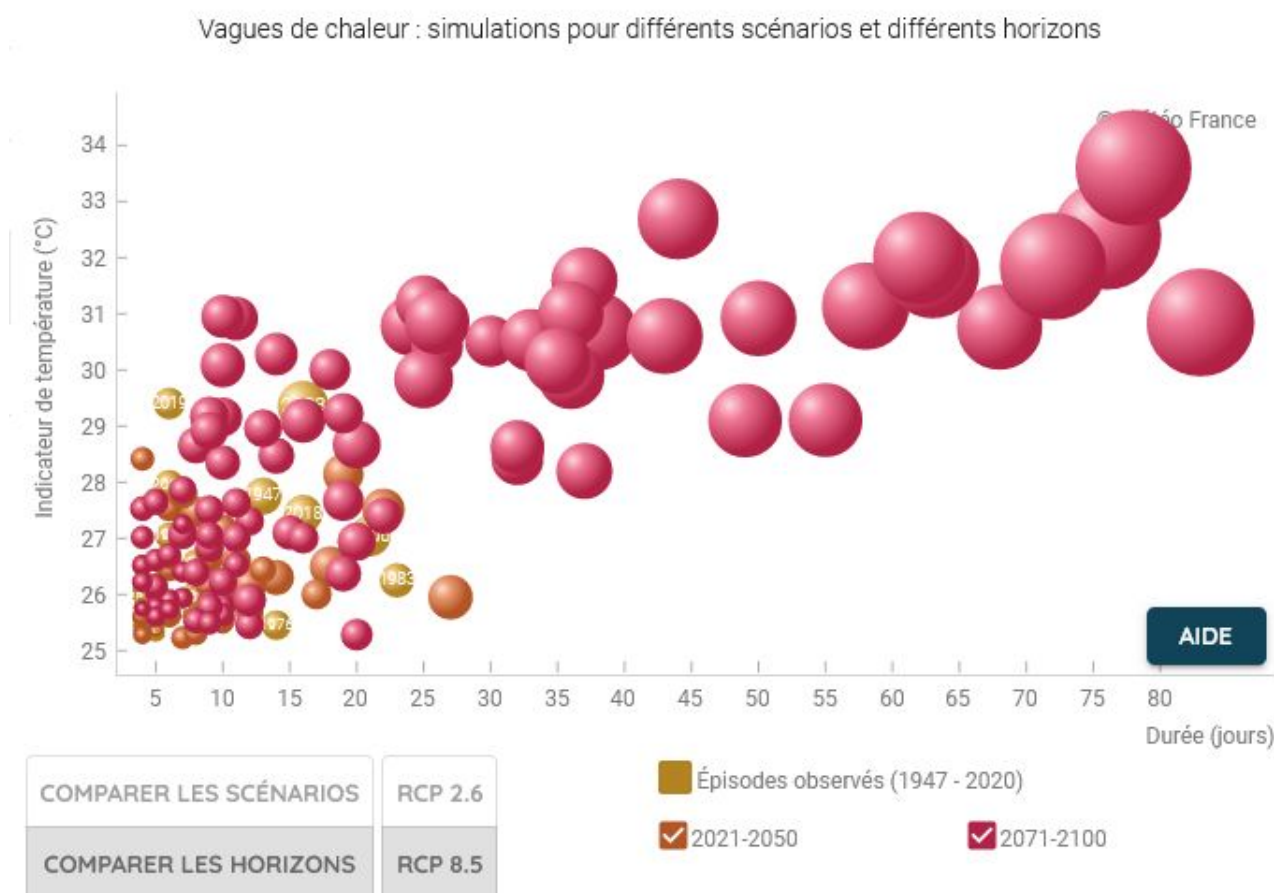


Figure 1 : Vagues de chaleur observées et projetées en France métropolitaine.
(© Météo-France / CLIMAT HD)

Dans ce contexte, le phénomène d'Îlot de Chaleur Urbain (ICU) devrait augmenter en fréquence et en intensité dans de nombreuses villes.

Il s'agit d'un phénomène qui consiste en un écart positif de quelques degrés de la température de l'air entre le centre des villes et les zones rurales ou naturelles périphériques. Son existence dépend de différents paramètres, comme la densité et la forme urbaine, l'artificialisation des sols, les propriétés d'absorption et de stockage des matériaux de constructions, les activités anthropiques productrices de chaleur (industrie, transport, équipements domestiques), la présence/absence d'espaces végétalisés ou en eau, ou encore le climat régional.

L'intensité et la répartition spatiale de l'ICU fluctuent selon la saison et le moment de la journée, la force, la direction des vents dominants, et la couverture nuageuse. L'intensité de l'ICU est :

- plus importante l'été : lors de longues périodes sans précipitations, dans des conditions de journées ensoleillées et faiblement ventées ;
- maximale la nuit : le refroidissement de l'air est moins rapide en milieu urbain dense qu'en milieu naturel ou agricole, à cause de la restitution de la chaleur piégée par les surfaces des bâtiments compacts et denses et par les chaussées (Figure 2). Certaines villes voient ainsi des différences de températures entre centre-ville et périphérie qui peuvent atteindre 10 °C à ce moment pendant les périodes de vagues de chaleur. Or, la nuit est un moment crucial pour permettre aux organismes de récupérer.

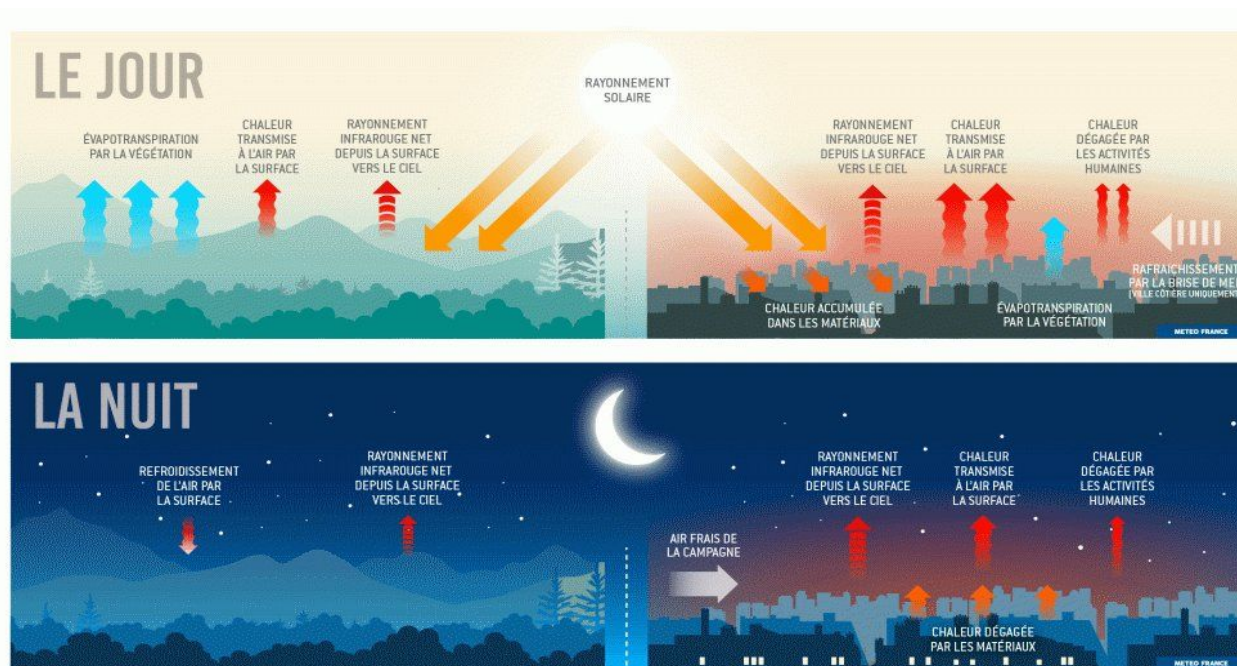


Figure 2 : Le mécanisme d'Îlot de Chaleur Urbain.
(© Météo-France)

En période de canicule, le phénomène d'ICU accentue donc l'impact des vagues de chaleur sur les populations, les infrastructures des villes mais également la biodiversité. Certaines personnes sont davantage vulnérables face aux ICU, telles que les personnes âgées, les jeunes enfants et les personnes en situation de pauvreté. Dans un contexte d'urbanisation croissante, l'effet d'ICU constitue un véritable risque pour les territoires.

La classification en Zones Climatiques Locales

Les LCZ, pour *Local Climate Zones* ou Zones Climatiques Locales, sont un référentiel de classification géo-climatique, développé par deux chercheurs canadiens, Stewart & Oke, en 2012¹. Elles décrivent un territoire en caractérisant chaque îlot/quartier selon ses formes urbaines, son organisation spatiale, l'occupation du sol dominante, et en le classant dans une des 17 classes du référentiel : 10 classes bâties (LCZ 1 à 10) et 7 LCZ non-bâties (LCZ A à G) ([Figure 3](#)), en tenant compte d'une série d'indicateurs.

Les 10 indicateurs proposés par Stewart & Oke peuvent être classés en trois familles distinctes, qui sont à l'origine des facteurs explicatifs du phénomène d'ICU. Ces trois familles englobent respectivement :

- les indicateurs morphologiques : « *Sky view factor* »², « *Aspect ratio* »³, « *Height of roughness elements* »⁴ et « *Terrain roughness class* »⁵. Cette première famille rend compte de l'aérodynamisme des tissus urbains, et donc de leur propension à être plus ou moins rafraîchis grâce aux vents. En particulier, le « *Sky view factor* » représente la capacité d'un tissu urbain à piéger le rayonnement solaire ;
- les indicateurs d'occupation du sol : « *Building surface fraction* »⁶, « *Impervious surface fraction* »⁷ et « *Pervious surface fraction* »⁸. Ils vont rendre compte de la capacité des sols à refléter ou emmagasiner l'énergie solaire, et la capacité de rafraîchissement grâce à l'évapotranspiration ou la présence d'humidité ;
- les indicateurs thermo-physiques : « *Surface admittance* »⁹, « *Surface albedo* »¹⁰ et « *Anthropogenic heat output* »¹¹. Ces derniers vont rendre compte des flux et stocks en fonction de la capacité des matériaux à stocker de l'énergie, ou des émissions d'énergie dues aux activités humaines telles que les climatisations, les industries ou les véhicules.

Chaque classe LCZ est ainsi définie suivant ce panel d'indicateurs sous-jacents, dont les valeurs sont bornées en fonction de la classe LCZ.

Dans le diagnostic de l'ICU, la classification LCZ est utilisée pour localiser les typologies urbaines favorables à son apparition (forte densité de bâtiments hauts, forte imperméabilité, faible végétation...), et estimer l'intensité de cet ICU dans les différents îlots/quartiers en fonction des valeurs des différents indicateurs sous-jacents.

1 <https://doi.org/10.1175/BAMS-D-11-00019.1>

2 Facteur de vue du ciel

3 Rapport d'aspect

4 Hauteur des éléments de rugosité

5 Classe de rugosité du terrain

6 Taux de surface bâtie

7 Taux de surface imperméable

8 Taux de surface perméable

9 Conductivité thermique de surface

10 Albédo de surface

11 Flux de chaleur anthropique

Les zones climatiques locales (LCZ)

Stewart et Oke (2012) — Traduction Cerema







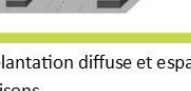


Classes bâties	Définitions	Classes non bâties	Définitions
1. Ensemble compact de tours 	<p>Mélange dense de grands immeubles avec des dizaines d'étages. Peu ou pas d'arbres. Sol principalement revêtu. Matériaux de construction : béton, acier, pierre, verre.</p>	A. Espace densément arboré 	<p>Paysage fortement boisé composé d'arbres à feuilles caduques et/ou à feuilles persistantes. Couverture du sol principalement perméable (plantes basses). Fonction de la zone : forêt, arboriculture, parc urbain.</p>
2. Ensemble compact d'immeubles 	<p>Mélange dense de bâtiments de hauteur moyenne (3 à 9 étages). Peu ou pas d'arbres. Sol principalement revêtu. Matériaux de construction : pierre, brique, tuile, béton.</p>	B. Espace arboré clairsemé 	<p>Paysage légèrement boisé composé d'arbres à feuilles caduques et/ou à feuilles persistantes. Couverture du sol principalement perméable (plantes basses). Fonction de la zone : forêt, arboriculture, parc urbain.</p>
3. Ensemble compact de maisons 	<p>Mélange dense de bâtiments de faible hauteur (1 à 3 étages). Peu ou pas d'arbres. Sol principalement revêtu. Matériaux de construction : pierre, brique, tuile, béton.</p>	C. Espace végétalisé hétérogène 	<p>Buissons, arbustes et ligneux espacés. Couverture du sol principalement perméable (sol nu ou sable). Fonction de la zone : maquis, agriculture.</p>
4. Ensemble de tours espacées 	<p>Tours espacées de plus de 10 étages. Sol perméable végétalisé en abondance (plantes basses, arbres épars). Matériaux de construction : béton, acier, pierre, verre.</p>	D. Végétation basse 	<p>Paysage plat composé d'herbe ou plantes herbacées, de cultures. Peu ou pas d'arbres. Fonction de la zone : prairie, agriculture, parc urbain.</p>
5. Ensemble d'immeubles espacés 	<p>Bâtiments de hauteur moyenne espacés (3 à 9 étages). Sol perméable végétalisé en abondance (plantes basses, arbres épars). Matériaux de construction : béton, acier, pierre, verre.</p>	E. Sol imperméable naturel ou artificiel 	<p>Paysage plat rocheux ou revêtu. Peu ou pas d'arbres ou de plantes. Fonction de la zone : désert naturel (roche) ou route.</p>
6. Ensemble de maisons espacées 	<p>Bâtiments de faible hauteur espacés (1 à 3 étages). Sol perméable végétalisé en abondance (plantes basses, arbres épars). Matériaux de construction : bois, brique, pierre, tuile, béton.</p>	F. Sol nu 	<p>Paysage plat composé de sol nu ou de sable. Peu ou pas d'arbres ou de plantes. Fonction de la zone : désert naturel ou agriculture.</p>
7. Ensemble dense de maisons ou constructions légères 	<p>Mélange dense de bâtiments à un étage. Peu ou pas d'arbres. Sol principalement compacté. Matériaux de construction légers : bois, chaume, tôle ondulée, etc.</p>	G. Eau 	<p>Plans d'eau (mers, lacs, rivières, bassins, lagon).</p>
8. Bâtiments bas de grande emprise 	<p>Bâtiments bas (1 à 3 étages) de grande emprise, et espacés. Peu ou pas d'arbres. Sol principalement revêtu. Matériaux de construction en acier, béton, métal et pierre.</p>	10. Industrie lourde 	<p>Structure industrielle basse ou de hauteur moyenne (tours, réservoirs, cheminées). Peu ou pas d'arbres. Sol principalement revêtu ou compacté. Matériaux de construction : métal, acier, béton.</p>
9. Implantation diffuse et espacée de maisons 	<p>Bâtiments de petite ou moyenne hauteur et éloignés dans un cadre naturel ou végétalisé. Sol perméable végétalisé en abondance (plantes basses, arbres épars).</p>		

Figure 3 : Définition des typologies de LCZ.

1 HISTORIQUE DE LA MÉTHODE CEREMA

1.1 L'algorithme analytique Stewart & Oke

Cette méthode dite analytique s'appuie sur une implémentation la plus fidèle possible des fiches LCZ proposées dans le référentiel de Stewart & Oke [2]. Elle a été développée dans le cadre du projet DiaClimAP (2015-2017), retenu lors de l'APR MODEVAL URBA 2014 de l'ADEME.

L'objectif premier de la méthode est d'être reproductible facilement, à des coûts raisonnables, avec des données facilement accessibles. Or, concernant les indicateurs thermo-physiques (conductivité thermique de surface, albédo de surface, flux de chaleur anthropique), aucune base de données sur étagère ou facile à construire à l'échelle d'une agglomération n'est disponible à ce jour.

On observe par ailleurs que les indicateurs proposés par Stewart & Oke sont parfois redondants ou très similaires pour décrire un même phénomène. Ils sont également parfois peu discriminants, ce qui est particulièrement le cas pour les indicateurs thermo-physiques. Ils sont donc abandonnés sans avoir a priori d'impact majeur sur la classification.

La mise en œuvre stricte des seuils et combinatoire des indicateurs, décrite dans le corpus de fiches du référentiel LCZ, ne permet pas de classer de manière exhaustive tous les cas de figure.

Les premiers travaux ont donc consisté, sur la base d'un arbre de décisions composé des branches principales respectant les seuils de ce référentiel, à proposer des branches complémentaires permettant d'attribuer une classe LCZ aux îlots premièrement non-classés.

Stewart & Oke ne proposent pas de méthode d'implémentation pratique, ni de hiérarchisation des indicateurs. Les choix propres au Cerema sont :

- une mobilisation de données disponibles gratuitement en France
- la hiérarchisation des indicateurs entre eux
- le respect des seuils proposés dans le référentiel
- une sélection des indicateurs à mettre en œuvre, en fonction des données disponibles en France : BD TOPO, Urban Atlas, occupation du sol (OCS) dérivée d'images satellite optique à très haute résolution spatiale
- une dépendance de l'algorithme à la nomenclature de l'Urban Atlas
- pas de mobilisation des indicateurs thermo-physiques
- une écriture de branches dites « expertes » pour combler les surfaces non-classées, avec une introduction d'indicateurs d'OCS complémentaires : sol nu vs. végétation basse vs. végétation moyenne vs. végétation haute vs. autre occupation du sol

Les premières cartes LCZ ont montré des résultats satisfaisants, mais avec des limites. Certaines classes présentaient des confusions, et l'algorithme ne traitant pas tous les cas de figure, des trous de classification restent présents. La traduction des fiches au travers d'un arbre de décision conduisait à une structure particulièrement complexe, et à de nombreux cas particuliers pour palier les limites de la méthode proposée par Stewart & Oke. Enfin, la maintenance de l'algorithme et son évolution sont peu ergonomiques et chronophages.

1.2 Les algorithmes intermédiaires

Sur la base des limites évoquées précédemment, plusieurs évolutions de l'algorithme ont été conduites. Ces étapes nous ont conduit à nous écarter du cadre strict de la nomenclature proposée par Stewart & Oke, et à développer un algorithme original. Nous ne décrivons pas ici dans le détail ces évolutions, mais elles visaient les objectifs suivants :

- s'affranchir de la dépendance de l'algorithme aux nomenclatures de l'Urban Atlas
- s'affranchir du calcul de certains indicateurs particulièrement gourmands en ressources de calcul
- mobilisation plus forte des indicateurs OCS issus des images avec notamment la végétation
- s'affranchir des indicateurs redondants et n'apportant pas de plus-value significative pour la qualification fiable des LCZ : parfois certains indicateurs pouvaient apporter de la confusion si les données d'entrée n'étaient pas d'une qualité suffisante
- fusion des classes LCZ 8 et LCZ 10, différenciables quasi-exclusivement avec l'indicateur de flux de chaleur anthropique (dont la donnée n'est généralement pas disponible)
- simplifier la structure de l'algorithme pour le faire évoluer et s'affranchir du cadre strict proposé par Stewart & Oke

Plusieurs versions successives ont été réalisées, et ont permis de s'approcher des objectifs visés.

1.3 L'algorithme opérationnel

Cet algorithme est la dernière étape du développement. Les dernières évolutions de celui-ci ont été réalisées dans le cadre du projet SatLCZ¹², labellisé par le *Space Climate Observatory* en 2020 et subventionné par le CNES, et conduit en partenariat avec les sociétés Airbus Defence and Space et TerraNIS.

Ce projet a permis de rendre l'algorithme opérationnel et applicable sur n'importe quel territoire (y compris à l'étranger) [1].

Le processus de production, incluant les données d'entrée nécessaires, les indicateurs dérivés, et les outils développés pour produire la carte LCZ sont décrites plus en détail dans la suite du document.

¹² <https://www.spaceclimateobservatory.org/fr/satlcz-lille>

2 LES DONNÉES D'ENTRÉE

2.1 Segmentation morphologique

Cette première donnée constitue nos îlots d'études : les polygones dans lesquels sont calculés les différents indicateurs et où est attribuée la classe LCZ correspondante aux valeurs de ces indicateurs. Elle doit respecter au maximum la définition liée à la production de données LCZ, à savoir une homogénéité de la typo-morphologie au sein d'un même îlot et une taille minimale de 200 m de rayon (soit une superficie d'environ 10 ha).

Cette donnée peut avoir trois origines possibles, soit :

- une donnée déjà existante : il peut s'agir de données produites par les collectivités et qui décrivent leur territoire, ou des données produites à plus large échelle, comme l'Urban Atlas du programme européen Copernicus ;
- une donnée nouvelle produite de façon automatique, semi-automatique ou manuelle : sur de petits territoires, il est possible de la produire en digitalisant à la main sous QGIS, tout en s'aidant d'autres données pour mieux caractériser le milieu urbain (images aériennes historiques, infrastructures de transport, bâti...). La production (semi-)automatique peut être réalisée à partir de différents algorithmes, tels que l'outil GeoCLIMATE¹³ qui permet de générer des *Reference Spatial Unit* (RSU) ;
- une donnée nouvelle correspondant à un carroyage régulier (carré ou hexagone) : elle présente l'avantage de pouvoir être produite sur n'importe quel territoire, mais avec comme principale limite le non-respect de l'homogénéité typo-morphologique, compte-tenu du fait qu'une telle grille ne suit pas les limites des îlots urbains ou des champs par exemple, ce qui peut conduire à mélanger des typologies urbaines et/ou des occupations du sol.

Cas de la production nationale LCZ :

Pour notre production nationale LCZ 2022, la donnée de segmentation morphologique utilisée est l'Urban Atlas de 2018¹⁴, une donnée déjà existante en large échelle. En effet, l'Urban Atlas est produit pour les aires urbaines européennes de plus de 50 000 habitants.

Même si celle-ci est à l'origine une cartographie de l'occupation et de l'usage des sols (obtenue à partir de données Sentinel-2), c'est uniquement sa géométrie vectorielle (qui respecte la définition d'homogénéité typo-morphologique des îlots) que nous récupérons et utilisons.

Cependant, l'Urban Atlas ayant ses propres caractéristiques techniques, il ne respecte pas toujours la taille minimale de 200 m de rayon des îlots. Il faudra donc bien garder en tête cette limite au moment de l'analyse de la carte LCZ pour le diagnostic de l'ICU (et raisonner en moyennant des îlots voisins de typologie différente).

Pour la production de territoires complémentaires non-couverts par cette donnée Urban Atlas 2018, le Cerema a développé un algorithme semi-automatique de découpage morphologique, dont le déroulé technique est expliqué [ci-après](#).

¹³ <https://github.com/orbisgis/geoclimate>

¹⁴ <https://land.copernicus.eu/en/products/urban-atlas/urban-atlas-2018>

2.1.1 Traitement de territoires complémentaires

Pour cartographier des territoires complémentaires non-couverts par la donnée Urban Atlas 2018 (Arles, Cholet, La Roche-sur-Yon, Montauban, Narbonne), et pour pallier aux limites de cette donnée (notamment la superficie de certaines entités bien en dessous des préconisations d'étude de la méthode LCZ), le Cerema a développé une méthode afin de construire un maillage continu du territoire en entités plus cohérentes avec la méthode LCZ et répliquable sur tout le territoire.

Cette méthode mobilise des bases de données géomatiques ainsi que d'occupation du sol (OCS). Elle vise à assembler des zones en fonction du tissu routier et de l'occupation du sol. En plus de la BD TOPO de l'IGN, la donnée OSO¹⁵ du millésime 2022 est utilisée, un produit d'occupation des sols en 24 classes issu de données satellite Sentinel-2, d'une résolution spatiale de 10 m.

Sans rentrer dans le détail technique (un article scientifique est en cours de rédaction), les grandes phases de production de cette segmentation morphologique sont les suivantes ([Figure 4](#)) :

- 1 Préparation des données :
 - 1.1 Préparation de l'OSO en 3 bandes et optimisation de la palette de couleurs de l'occupation du sol pour l'algorithme de superpixels
 - 1.2 Extraction de la BD TOPO pour conserver le tissu routier structurant (squelette)
 - 1.3 Suppression des voies parallèles et ronds points par extraction de squelette et fusion avec le réseau ferroviaire
- 2 Création de la segmentation brute du territoire :
 - 2.1 Découpage du territoire à partir des réseaux structurants pré-traités
 - 2.2 Fusion des très petits polygones générés
 - 2.3 Segmentation du territoire avec un algorithme de superpixels :
 - 2.3.1 Utilisation d'une méthode d'optimisation paramétrique MCMC (*Markov Chain Monte Carlo*) de la librairie EEMC avec une fonction de coût basée sur des indicateurs de superficie, de compacité et d'uniformité d'occupation du sol des polygones. La cible de cette fonction de coût est commune à tous les territoires
 - 2.3.2 Découpage du territoire avec la configuration paramétrique optimale
- 3 Fusions des polygones :
 - 3.1 Suppression des zones en eau et fusion des très petits polygones
 - 3.2 Fusion des petits polygones avec leur voisin partageant la plus grande frontière commune
 - 3.3 Application d'un tag « route » aux polygones voisins d'une route pour éviter les fusions qui « traversent » une route par la suite
 - 3.4 Calcul de statistiques pour chaque polygone : hauteur moyenne de bâti, classe majoritaire et valeur de la classe majoritaire de l'OCS OSO, compacité
 - 3.5 Application d'un algorithme de fusion pour les polygones de tailles moyenne et grande : une minimisation de distance euclidienne calculée avec les statistiques du point précédent permet de fusionner un polygone avec son « meilleur » voisin

15 <https://www.theia-land.fr/ces-occupation-des-terres/occupation-des-sols/>

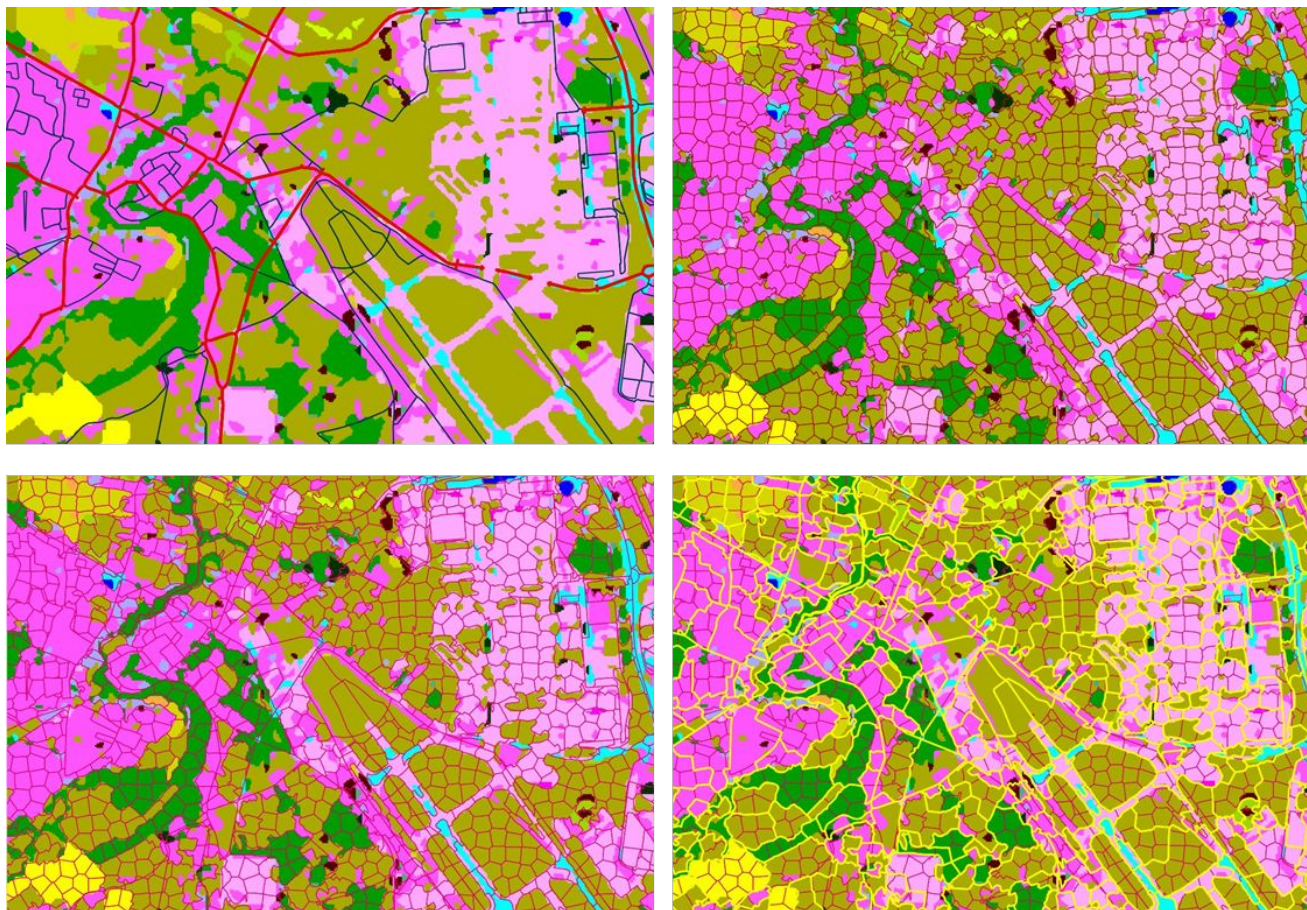


Figure 4 : Grandes étapes de la production de la segmentation morphologique Cerema.

- (a) extraction du squelette des routes principales en rouge*
- (b) découpage SLIC sur tout le territoire*
- (c) assemblage du découpage « routes » avec le découpage SLIC*
- (d) découpage avant/après fusions en rouge/jaune*

S'agissant d'une nouvelle méthode expérimentale développée selon une approche semi-automatique, et non encore validée scientifiquement, des erreurs de découpage peuvent apparaître. L'analyse visuelle en 1^{ère} approche du résultat de la classification en LCZ rend assez bien compte des tendances globales des territoires.

De futures versions pourront être livrées, en fonction des améliorations apportées à la méthode.

2.2 Vecteur bâti 3D

Cette deuxième donnée décrit l'emprise au sol de chaque bâtiment du territoire d'étude, sous forme de vecteur. A chacun de ces bâtiments est rattachée une valeur d'attribut correspondant à la hauteur de celui-ci.

Le cas le plus classique est de disposer d'une donnée déjà existante et bien formatée (au format vecteur polygone, avec un attribut hauteur), comme la BD TOPO de l'IGN (cf. ci-dessous), mais il est également possible de la produire, en croisant deux données distinctes :

- une donnée de la répartition des bâtiments (au format vecteur, ou raster qu'on pourra vectoriser), issue d'une couche d'occupation du sol ou d'un recensement terrain par exemple ;
- une donnée de la hauteur des bâtiments, issue par exemple d'un Modèle Numérique de Hauteur (MNH) (cf. [2.3.1](#)).

Cas de la production nationale LCZ :

Pour la production nationale LCZ 2022, c'est la BD TOPO¹⁶ v3.3 de décembre 2022 de l'IGN qui est utilisée.

La BD TOPO est une description et une modélisation vectorielle 3D du territoire et de ses infrastructures d'une précision métrique. Elle présente l'avantage de couvrir de façon cohérente l'ensemble du territoire national français. Elle est divisée en 8 thèmes et 60 classes.

Dans notre cas, c'est la classe « Bâtiment » du thème « Bâti » qui est utilisée, et qui contient un attribut « Hauteur » (qui correspond à la hauteur du bâtiment mesurée entre le sol et le point haut de la gouttière, exprimée en mètres), attribut nécessaire au calcul de l'indicateur de hauteur moyenne du bâti.

Avant traitement, un nettoyage de cette donnée est réalisée à partir des différents attributs, par exemple pour supprimer les bâtiments en « État » = 'En projet' (il n'est pas encore construit et n'a donc pas de géométrie ou une géométrie incorrecte, mais il peut fausser le calcul de l'indicateur de surface moyenne du bâti) ou en « Hauteur » = 0 (la hauteur n'est pas connue, mais il peut fausser le calcul de l'indicateur de hauteur moyenne du bâti).

¹⁶ <https://geoservices.ign.fr/bdtopo>

2.3 Raster OCS 6 classes

La dernière donnée d'entrée nécessaire pour produire la carte LCZ est un raster représentant l'occupation du sol, suivant six classes bien définies :

- bâti ;
- route et surface imperméable ;
- sol nu perméable ;
- surface en eau ;
- végétation arborée ;
- végétation herbacée.

Dans nos productions classiques, cette couche OCS est obtenue par classification supervisée orientée pixel d'une ou plusieurs images satellites, et on nomme le résultat « OCS sat' ». Les algorithmes utilisés pour classer ces images sont généralement issus du *Machine learning* (*Random Forest* ou *SVM*) ou du *Deep learning* (*UNet* ou *ResUNet*)¹⁷.

Cas de la production nationale LCZ :

Pour notre production nationale LCZ 2022, cette OCS sat' est en partie dérivée de la couverture nationale SPOT 2022 acquise dans le cadre de DINAMIS¹⁸. Comme pour la BD TOPO de l'IGN, cette couverture SPOT nous permet de disposer d'une information à très large échelle, et de façon homogène sur tout le territoire.

L'usage d'imagerie satellite permet d'avoir une information objective, puisque ces images sont des photographies du territoire. L'usage de cette couverture annuelle garantit un couvert nuageux quasi-absent, nécessaire pour cartographier les objets et surfaces au sol.

Cependant, cette information est obtenue à un instant t , à la date de la prise de vue du satellite. Ainsi, il peut y avoir des différences entre l'OCS produite et la réalité d'aujourd'hui (de nouveaux quartiers résidentiels ou des réaménagements de centres urbains par exemple).

La plage d'acquisition de cette couverture nationale (entre le 5 avril et le 27 juillet 2022) induit également des différences de couvert végétal entre deux images adjacentes (il n'y a pas de logique d'acquisition est-ouest ou nord-sud, la programmation du satellite dépend notamment des conditions météorologiques pour avoir le minimum de couvert nuageux) : des parcelles agricoles à des stades de développements différents ou des sols à des niveaux de sécheresse différents peuvent influencer la détection de la végétation, pouvant conduire à plutôt détecter du sol nu perméable.

Compte tenu du volume de données et des temps de traitement, la couverture nationale SPOT 2022 n'a été exploitée que pour extraire la végétation par seuillage NDVI avant que ne soit faite la distinction végétation arborée/végétation herbacée par l'usage de RGE ALTI 1M et MNS-Correl (cf. [2.3.1](#)).

Pour compléter et ajuster cette « OCS végétation », nous utilisons des bases de données exogènes : BD TOPO de l'IGN, HRL Imperviousness de Copernicus (cf. [4.1.3](#)).

¹⁷ <https://github.com/CEREMA/dterocc.chaineTraitement.traitementImageSatelliteEtIndicateursDerives>

¹⁸ <https://dinamis.data-terra.org/spot-6-7-couverture-france-metropolitaine-2022/>

2.3.1 Focus sur la caractérisation de la végétation

Comme expliqué au début de la présentation de la donnée raster OCS 6 classes, il est nécessaire de bien connaître la répartition et la stratification de la végétation pour générer la carte LCZ.

Ainsi, la répartition de cette végétation peut être obtenue :

- par classification supervisée : cas classique, elle est cartographiée en même temps que les autres classes de l'OCS, avec les mêmes méthodes et algorithmes de classification ;
- par seuillage NDVI : le NDVI (indice de végétation par différence normalisée) est un indice radiométrique obtenu à partir des bandes spectrales rouge (R) et proche infrarouge (PIR). Cet indice normalisé $[-1;1]$ permet de facilement caractériser la végétation, dont la valeur est proche de 1. Ainsi, une fois cet indice calculé, il suffit de trouver le seuil de bascule entre la végétation et les autres classes, généralement autour de 0,33 pour des images SPOT-6/7 ou Pléiades, et de construire un masque binaire végétation/non-végétation.

Une fois la carte de végétation obtenue, la stratification de celle-ci peut être obtenue :

- par étude de l'homogénéité radiométrique (texture) : on étudie, pour une bande spectrale définie (généralement le panchromatique ou le rouge), la radiométrie des pixels environnants un pixel central. La végétation herbacée va présenter une certaine homogénéité, contrairement à la végétation arborée qui va donc exprimer une faible valeur texturale. Cette forte hétérogénéité radiométrique de la végétation arborée s'explique notamment par les branches et les feuilles qui, de part leur orientation, réfléchissent différemment la lumière et peuvent créer des micro-zones d'ombres ;
- directement par étude de la hauteur : il est nécessaire ici de disposer, comme son nom l'indique, d'une donnée rendant compte de la hauteur de la végétation, un Modèle Numérique de Hauteur (MNH). Celui-ci est lui-même obtenu par soustraction d'un Modèle Numérique de Terrain (MNT) et d'un Modèle Numérique de Surface (MNS), deux représentations 3D d'un territoire, la première rendant compte de l'altitude du sol et la seconde de l'altitude des éléments de sursol (bâtiments, ponts, arbres...).

Cas de la production nationale LCZ :

Pour notre production nationale LCZ 2022, la distinction herbacée/arborée a été obtenue par utilisation d'un MNH, issue de deux sources de données disponibles à grande échelle :

- côté MNT, c'est le RGE ALTI 1M¹⁹ qui a été exploité : donnée mise à jour en continu, elle peut être produite à partir de relevés LiDAR, RaDAR ou par méthode de corrélation de photographies aériennes ;
- côté MNS, c'est le MNS-Correl²⁰ qui a été exploité : donnée récente et mise à jour tous les 3-4 ans, en fonction des acquisitions réalisées pour la BD ORTHO dont elle est dérivée.

Afin d'être cohérent avec les images SPOT-6/7 et l'OCS dérivées, le MNH a été produit à une résolution spatiale de 1,5 m, soit en sous-échantillonnant le RGE ALTI 1M et le MNS-Correl, respectivement à 1 m et 0,5 m de résolution spatiale originellement.

¹⁹ <https://geoservices.ign.fr/rgealti>

²⁰ <https://geoservices.ign.fr/modeles-numeriques-de-surfaces-correles>

3 LES INDICATEURS

3.1 Morphologie urbaine

Les indicateurs de morphologie urbaine de notre méthode sont au nombre de deux :

- **HRE** = hauteur moyenne du bâti
- **ARE** = surface moyenne du bâti

Ils reposent tous les deux sur le croisement de la donnée de segmentation morphologique et de la donnée vecteur de bâti 3D.

Les calculs de ces indicateurs sont réalisés en SQL sous PostgreSQL/PostGIS, qui permet de traiter de façon optimisée de grands volumes de données, notamment via des requêtes spatiales.

La première étape consiste à faire un croisement pour récupérer les îlots de notre donnée de segmentation morphologique intersectant au moins un bâtiment de notre donnée vecteur de bâti 3D.

Sur ces îlots intersectés, résultant en une table PostgreSQL/PostGIS, on calcule ainsi la moyenne de la hauteur et de la surface des bâtiments à l'intérieur de chaque îlot.

Les îlots n'intersectant pas de bâtiment sont récupérés par ailleurs dans une nouvelle table, et une valeur par défaut de 0 est donnée aux deux indicateurs.

Une fois les valeurs d'indicateurs attribuées à tous les îlots, les deux tables précédentes (celle des îlots intersectant du bâti et celle des îlots n'intersectant pas de bâti) sont unies pour reconstituer notre donnée de segmentation morphologique complète.

3.2 Occupation du sol

Les indicateurs d'occupations du sol de notre méthode sont au nombre de six :

- **BuR** = taux d'occupation de la classe « bâti »
- **RoR** = taux d'occupation de la classe « route et surface imperméable »
- **BsR** = taux d'occupation de la classe « sol nu perméable »
- **WaR** = taux d'occupation de la classe « surface en eau »
- **VeR** = taux d'occupation des classes « végétation »
- **VHR** = part de la végétation arborée sur la végétation globale

Ils reposent tous les six sur le croisement de la donnée de segmentation morphologique et de la donnée raster d'OCS 6 classes.

Les cinq premiers (BuR, RoR, BsR, WaR, VeR) sont calculés de la même façon, à savoir en effectuant un croisement statistique entre notre raster d'OCS 6 classes et notre donnée de segmentation morphologique, pour récupérer, pour chaque îlot de notre donnée de segmentation morphologique, le taux d'occupation de chaque classe OCS. La particularité de l'indicateur VeR est de regrouper les deux classes végétation arborée et végétation herbacée.

Ainsi, la formule mathématique permettant d'aboutir à ces indicateurs est la suivante :

$\frac{S_{classe}}{S_{ilot}} \times 100$ où S_{classe} est la surface occupée par la classe analysée (ou la somme des surfaces des classes végétation arborée et végétation herbacée dans le cas de l'indicateur VeR) et S_{ilot} est la surface de l'îlot.

Le dernier indicateur (VHR) se concentre sur la distinction de la strate végétale : on exclut toutes les autres classes de notre analyse, et on regarde le taux d'occupation de la végétation arborée sur la somme de la végétation globale :

$\frac{S_{veg_arborée}}{S_{veg_arborée} + S_{veg_herbacée}} \times 100$ où $S_{veg_arborée}$ et $S_{veg_herbacée}$ sont respectivement les surfaces occupées par les classes végétation arborée et végétation herbacée de l'OCS 6 classes.

Tous ces indicateurs servent principalement à distinguer les classes LCZ non-bâties A à G.

L'indicateur VHR sert exclusivement à différencier les classes LCZ A à D, différenciables par leur taux de végétation arborée (plus VHR est élevé, plus on est proche de la classe LCZ A).

4 OBTENTION DE LA CARTE LCZ

4.1 Préparation des données d'entrée

Pour produire la carte LCZ de chacun des 88 territoires retenus pour cette production nationale, un grand nombre de données ont été nécessaires (certaines ayant déjà été évoquées en section [2](#)).

Pour cette production massive, nous nous sommes basées sur des données de 4 principaux fournisseurs de données :

- Airbus Defence and Space : imagerie spatiale SPOT-6/SPOT-7 millésime 2022²¹, mise à disposition via DINAMIS (cf. [2.3](#))
- IGN : BD TOPO v3.3 12/2022²², RGE ALTI 1M²³, MNS-Correl²⁴
- Copernicus : Urban Atlas 2018²⁵, HRL Imperviousness 2018²⁶
- Theia : OSO 2022²⁷

Au total, ce sont près de 6 To de données qui ont été téléchargées et traitées.

Le déroulé global de la production (incluant la préparation des données d'entrée) est donné dans la figure ci-dessous :

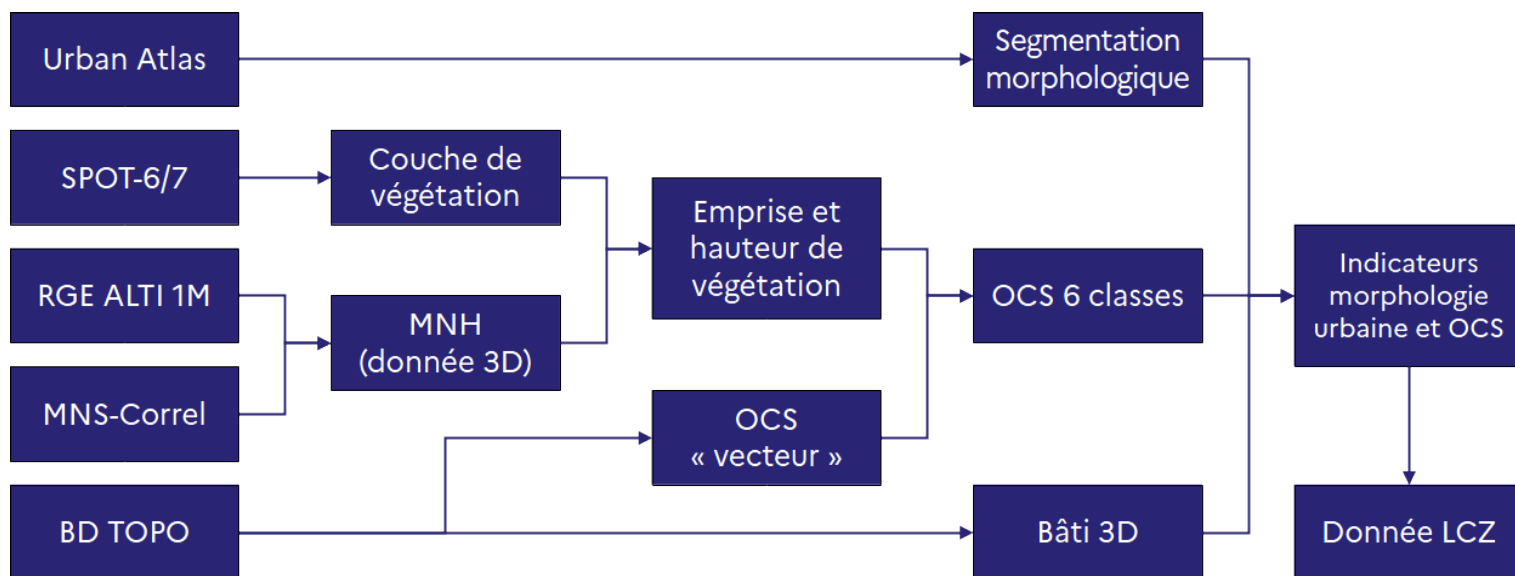


Figure 5 : Grandes étapes de la production de la donnée LCZ.

21 <https://intelligence.airbus.com/imagery/our-optical-and-radar-satellite-imagery/spot/>

22 <https://geoservices.ign.fr/bdtopo>

23 <https://geoservices.ign.fr/rgealti>

24 <https://geoservices.ign.fr/modeles-numeriques-de-surfaces-correles>

25 <https://land.copernicus.eu/en/products/urban-atlas/urban-atlas-2018>

26 <https://land.copernicus.eu/en/products/high-resolution-layer-imperviousness/imperviousness-density-2018>

27 <https://www.theia-land.fr/ces-occupation-des-terres/occupation-des-sols/>

4.1.1 Segmentation morphologique

Cette donnée d'entrée ne nécessite que l'Urban Atlas 2018, qui a simplement été dézippée et re-projetée en *RGF93 v1 / Lambert-93* (EPSG:2154), s'agissant d'une donnée européenne en *ETRS89-extended / LAEA Europe* (EPSG:3035) à l'origine.

Pour les territoires complémentaires ne disposant pas de cette donnée, un traitement spécifique a été réalisé, sur la base de la donnée OSO et de la BD TOPO (cf. [2.1.1](#)).

4.1.2 Bâti 3D

Cette donnée d'entrée n'a nécessité que la BD TOPO, et uniquement la classe « bâtiment » du thème « bâti ». Pour nettoyer la donnée et ne pas introduire trop d'erreurs, un filtrage SQL a été réalisé, pour ne conserver que les bâtiments dont les informations sont suffisamment précises :

```
hauteur IS NOT NULL AND hauteur >= 3 AND construction_legere IS false AND  
etat_de_l_objet = 'En service' AND precision_altimetrique != 9999
```

4.1.3 OCS 6 classes

C'est ici la partie la plus complexe et la plus longue du traitement, faisant appel à l'ensemble des données citées au début de ce chapitre (à l'exception de l'Urban Atlas 2018).

La brique « hauteur de végétation » est issue du croisement de la végétation issue de l'imagerie satellite avec une donnée 3D, elle-même construite à partir de deux données de l'IGN (cf. [2.3.1](#)) : le RGE ALTI 1M qui représente l'altitude du sol et le MNS-Correl qui représente l'altitude du sur-sol. La différence de ces deux données permet d'aboutir à la hauteur des éléments du sur-sol, et spécifiquement de la végétation dans notre cas.

Ensuite, un certain nombre de classes de la BD TOPO sont exportées de la base Postgres dans laquelle celle-ci avait été montée. Celles-ci viendront enrichir l'OCS dérivée de l'imagerie SPOT. De même, l'usage de la donnée HRL Imperviousness de Copernicus viendra ajouter de l'information. Ainsi, sur les différentes sources de données et leur apport :

- BD TOPO « bâtiment » : ajout à la classe « bâti »
- BD TOPO « tronçon de route » : ajout à la classe « route et surface imperméable »
- BD TOPO « tronçon de voie ferrée » : ajout à la classe « route et surface imperméable »
- BD TOPO « équipement de transport » : ajout à la classe « route et surface imperméable »
- BD TOPO « piste d'aérodrome » : ajout à la classe « route et surface imperméable »
- BD TOPO « terrain de sport » : ajout à la classe « route et surface imperméable »
- BD TOPO « tronçon hydrographique » : ajout à la classe « surface en eau »
- BD TOPO « surface hydrographique » : ajout à la classe « surface en eau »
- BD TOPO « limite terre-mer » : distinction des classes « sol nu perméable » / « surface en eau » (pour les territoires côtiers)
- Copernicus HRL Imperviousness : distinction des classes « route et surface imperméable » / « sol nu perméable » (dernière étape, sur les pixels non-encore classés précédemment)

4.2 Calcul des indicateurs et attribution de la typologie LCZ

Une fois l'ensemble des données d'entrées correctement formatées, le calcul des indicateurs a pu être réalisé. Ce calcul a été expliqué en section 3, et la production nationale LCZ n'a pas nécessité d'adaptation particulière, malgré le volume de données conséquent.

Sitôt les huit indicateurs obtenus, on peut attribuer la classe de LCZ correspondante aux valeurs de ces différents indicateurs. Notre méthode se base sur un arbre de décision (Figure 6), dont les seuils ont été définis suite aux projets de recherche DiaCliMAP²⁸ et SatLCZ²⁹.

La première analyse dans ce logigramme consiste à regarder l'indicateur BuR (taux de surface bâtie) : il permet de distinguer les classes LCZ bâties (1 à 9) des LCZ non-bâties (A à G) suivant si la valeur de cet indicateur est, respectivement, supérieur ou inférieur à 10 %.

Dans le 1^{er} cas, on regarde ensuite l'indicateur de morphologie urbaine de hauteur moyenne de bâti, qui permet de faire trois grands groupes (dont les seuils sont à 9 m et 30 m, soit à 3 étages et 10 étages) : les îlots de maisons (**LCZ 3** / **LCZ 6** / **LCZ 7** / **LCZ 9**), les îlots d'immeubles (**LCZ 2** / **LCZ 5**) et les îlots de tours (**LCZ 1** / **LCZ 4**).

La classe **LCZ 8**, correspondant aux zones d'activités ou équipements publics avec de grands bâtiments relativement bas, est à part et est attribué directement à partir de l'indicateur de surface moyenne du bâti > 700 m² (pour une hauteur moyenne de bâti inférieure à 21 m).

S'ensuit la réutilisation de l'indicateur de taux de surface bâtie, qui permet de faire la distinction entre les classes de bâti dense (**LCZ 1** / **LCZ 2** / **LCZ 3**) et de bâti peu dense (**LCZ 4** / **LCZ 5** / **LCZ 6** / **LCZ 9**), avec un taux, respectivement, supérieur et inférieur à 40 %.

La **LCZ 7** est là-aussi un cas particulier au sein des LCZ « îlots de maisons », correspondant aux zones de bidonville ou aux zones agricoles avec de grandes serres (il faut donc que le sol nu perméable soit majoritaire dans l'îlot, en plus d'avoir un taux de bâti supérieur à 60 %).

Dans le 2nd cas, on regarde la classe d'OCS majoritaire de l'îlot, en comparant les quatre indicateurs RoR/BsR/WaR/VeR :

- si c'est l'eau (WaR), l'îlot est classé **LCZ G**
- si c'est le sol nu perméable (BsR), l'îlot est classé **LCZ F**
- si c'est le sol imperméable (RoR), l'îlot est classé **LCZ E**
- si c'est la végétation (VeR), l'îlot est ensuite différencié en fonction de la valeur de l'indicateur de part de végétation arborée sur la végétation globale (VHR) : **LCZ A** si VHR > 50 %, **LCZ B** si VHR > 25 %, **LCZ C** si VHR > 10 % et **LCZ D** si VHR < 10%

Pour finir, un post-traitement automatique sur les îlots « réseau de transport » et « surfaces en eau » de l'Urban Atlas a été nécessaire. Ceci s'explique par la géométrie de l'Urban Atlas qui ne suit pas parfaitement le découpage réel du territoire : certains bâtiments se retrouvent en partie dans ces îlots, faisant artificiellement monter le taux de surface bâtie, ce qui peut aboutir à une fausse classification en LCZ bâtie. Ces îlots sont donc respectivement reclassés LCZ E et LCZ G.

D'autres post-traitements ont pu être réalisés ponctuellement pour corriger des erreurs liées là aussi majoritairement à la géométrie de l'Urban Atlas.

²⁸ <https://www.cerema.fr/fr/actualites/cerema-concoit-outil-aider-villes-s-adapter-aux-ilots>

²⁹ <https://www.spaceclimateobservatory.org/fr/satlcz-lille>

HRE = hauteur moyenne du bâti
 ARE = superficie moyenne du bâti
 BuR = taux de bâti
 RoR = taux de route/surface minérale imperméable
 BsR = taux de sol nu perméable
 WaR = taux de surface en eau
 VeR = taux de végétation
 VHR = part de la végétation haute sur la végétation globale

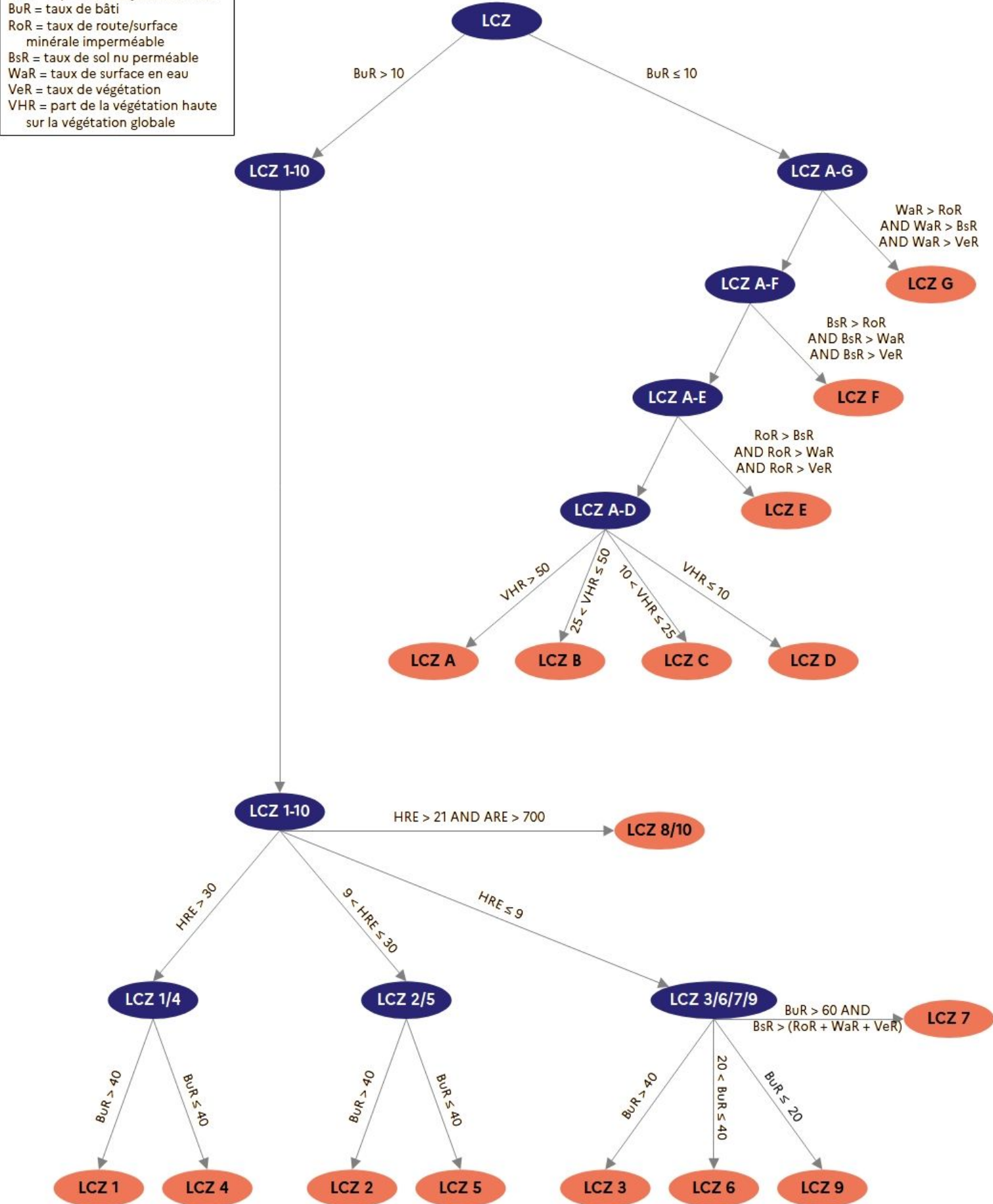


Figure 6 : Logigramme pour l'attribution de la classe de LCZ en fonction de la valeur des indicateurs.

5 LIMITES DE LA DONNÉE LCZ

Comme toute donnée, celle-ci présente des limites dont certaines sont expliquées ici.

Validation partielle de la donnée :

La validation, au moment de la rédaction de ce rapport et de la livraison des données, a été réalisée sur la base d'une vérification visuelle de polygones sur l'ensemble des 83 aires urbaines (en comparant à l'imagerie satellite SPOT 2022 ayant servi à la production), et sur la cohérence de la répartition des typologies LCZ à l'échelle d'une même aire urbaine : vérifier qu'une zone urbaine ressorte bien en « typologie urbaine » (majorité de LCZ 1 à LCZ 6), une zone rurale en « typologie rurale » (majorité de LCZ 6, LCZ 9, LCZ D, LCZ F), une zone naturelle en « typologie naturelle » (majorité de LCZ A à LCZ D et de LCZ F et LCZ G)...

Un rapport de qualité devrait être produit à posteriori, avec une validation statistique plus précise.

La donnée LCZ n'est pas une donnée d'occupation du sol :

Même si la symbologie peut y faire penser, la donnée LCZ n'a pas vocation à se substituer à une donnée d'occupation du sol. Elle est elle-même le résultat du croisement entre une couche OCS (dérivée de la couverture nationale SPOT 2022, cf. 2.3) et une donnée de bâti 3D, introduisant une notion de morphologie urbaine.

De ce fait, il peut arriver que des îlots, pourtant majoritairement végétalisés, ne ressortent pas en végétation, notamment certains parcs urbains : ceci s'explique par la présence de quelques bâtiments, qui font basculer ces îlots en LCZ « bâties », différentes des LCZ A à D.

La végétation n'apparaît pas toujours :

L'utilisation de données satellite SPOT issues de la couverture annuelle, acquise en l'occurrence entre début avril et fin juillet, pour dériver une couche d'occupation du sol peut conduire à des différences sur la détection de la végétation : en fonction de la date d'acquisition de l'image, le stade phénologique de la végétation va varier, et la végétation peut ne pas être correctement détectée.

Cela est particulièrement visible au niveau des zones agricoles, qui peuvent engendrer une « coupure » sur le territoire si celui-ci est couvert par plusieurs images acquises à des dates différentes.

Ceci peut se traduire par une typologie LCZ F (sol nu perméable) d'un côté et LCZ D ou C (végétation herbacée ou hétérogène) de l'autre, mais sans que cela n'impacte le diagnostic de l'ICU.

Des imprécisions de classifications :

Du fait de l'utilisation d'un arbre de décision fixe qui attribue une classe LCZ à un îlot en fonction de la valeur des différents indicateurs (cf. 4.2), il peut y avoir une différence de classes LCZ entre deux îlots de typo-morphologie visiblement très proches.

Par exemple, deux îlots « urbains », constitués d'immeubles, dont la seule différence concerne l'indicateur « taux de surface bâtie », l'un à 39 % l'autre à 41 %, vont, respectivement, être classés « Ensemble d'immeubles espacés » et « Ensemble compact d'immeubles ».

De même, un îlot « bâti » mais présentant moins de 10 % de surface bâtie se verra automatiquement classé en typologie LCZ « non-bâtie » (LCZ A à G).

On peut examiner les indicateurs sous-jacents, en plus de la classe LCZ en elle-même, pour comprendre pourquoi tel îlot est classé en telle classe LCZ (impact de la hauteur du bâti ? Impact de la présence de végétation ?...).

La donnée LCZ peut avoir des polygones plus petits que le seuil minimal de 200 m de rayon :

Dans les spécifications de production de la donnée Urban Atlas, on trouve une MMU (*Minimum Mapping Unit*, unité minimale de cartographie) de 0,25 ha ou 1 ha en fonction des classes, ce qui est bien en dessous des 200 m de rayon théoriques nécessaires à l'exploitation de la cartographie LCZ (~ 10 ha).

Il faut donc faire preuve de prudence au moment de l'analyse de cette donnée LCZ, qui peut faire ressortir des micro-typologies de LCZ qui n'ont pas lieu d'être si on fait la même analyse avec des polygones de taille plus importante. Plutôt que de s'attacher à la classe LCZ de tel ou tel îlot, il convient donc de regarder la typologie majoritaire autour de celui-ci.

Le découpage de la donnée LCZ ne suit pas le découpage administratif communal :

L'unité administrative utilisée pour produire la donnée Urban Atlas, sous-jacente à la donnée LCZ, se base sur la définition d'aire urbaine³⁰. Ainsi, un îlot peut se retrouver à cheval sur plusieurs communes, ce qui peut provoquer des erreurs topologiques si l'on découpe la donnée LCZ produite avec l'emprise d'une commune ou d'un EPCI.

Il vaut mieux réaliser une sélection par localisation (récupérer toutes les entités LCZ qui intersectent l'emprise souhaitée) plutôt qu'un découpage, même si l'on déborde un peu sur les territoires voisins (en affichage dans un SIG ou pour la production de cartes, on pourra toujours superposer une emprise négative de celle souhaitée pour cacher ces fragments hors territoire).

Une autre possibilité consiste à travailler avec la donnée LCZ au format raster, mais avec l'inconvénient que l'information attributaire des indicateurs n'est plus disponible.

Une autre conséquence de ce découpage particulier de l'Urban Atlas en fonction de l'aire urbaine (ou aire d'attraction d'une ville aujourd'hui) est le fait qu'une commune (ou qu'un EPCI) peut être partiellement couverte par la donnée LCZ.

La géométrie de la donnée LCZ peut présenter des erreurs :

Dans le processus de production des îlots Urban Atlas, pouvant utiliser des données exogènes, des décalages géométriques peuvent être observés par rapport à la vérité terrain, ce qui peut fausser le classement en LCZ.

Par exemple, un bâti qui intersecte un îlot de forêt va artificiellement faire augmenter le taux de surface bâtie de cet îlot, au risque de le faire basculer dans une typologie LCZ « non-bâties ».

Ceci est particulièrement impactant sur les îlots des routes, qui résultent d'un tampon appliqué à un linéaire routier issu d'OSM, en fonction de « l'importance » de cette route (autoroute, route départementale, route communale, piste cyclable...). C'est pour cette raison qu'un post-traitement a été appliqué sur ces îlots, en récupérant les classes « réseau de transport » de l'Urban Atlas, pour les reclasser sans tenir compte de la valeur des indicateurs (cf. 4.2).

³⁰ <https://www.insee.fr/fr/information/3642744>

6 OUTILS DE PRODUCTION

La production de données LCZ nécessite des outils robustes (grand volume de données en entrée) et automatisés (répétabilité sur plusieurs territoires), qui ont été développés par le Pôle satellite du Cerema. Suivant les cas, on peut utiliser des scripts Python (comme ça a été le cas pour cette production nationale) ou des modeleurs QGIS (comme c'est le cas pour des productions ponctuelles).

6.1 Scripts Python

Depuis plusieurs années, le Pôle satellite du Cerema développe une chaîne de traitement d'images satellites pour la production d'OCS et d'indicateurs dérivés. Celle-ci permet une exploitation de données satellitaires sur des thématiques telles que l'aménagement du territoire, la gestion des risques ou le suivi du littoral. Elle est basée sur du Python, et n'utilise que des outils libres : GDAL/OGR, OTB, PostGIS... Elle est accessible en Open Source sur GitHub³¹.

Dans cette chaîne, une branche dite LCZ est présente. Celle-ci se basant à l'origine sur la méthode analytique de Stewart & Oke, on y retrouve un script Python par indicateur (à l'exception des indicateurs de propriétés thermo-physiques), et un dernier pour la mise en place de la cartographie LCZ, à partir du croisement de ces indicateurs, calculés par polygone de notre fichier de segmentation morphologique du territoire d'étude.

De plus, un script calculant les nouveaux indicateurs dérivés de l'OCS est ajouté à cette branche, ainsi qu'un script final d'attribution des classes LCZ spécifique à la méthode opérationnelle, avec le dernier logigramme (en version SQL et non Python comme au début).

6.2 Modeleurs QGIS

Dans un souci de diffusion de la méthode et de prise en main facilitée, des modeleurs QGIS ont été déployés dans le cadre du projet SCO SatLCZ. À la différence des scripts Python précédents, seuls les indicateurs retenus dans la version opérationnelle de la méthode ont été considérés. Il y a donc 3 modeleurs :

- un modeleur « HRE » pour les indicateurs de morphologie urbaine : hauteur moyenne du bâti et surface moyenne du bâti;
- un modeleur « OCS » pour les indicateurs d'occupation du sol : taux d'occupation des différentes classes OCS (bâti, route, sol nu, eau, végétation) + part de végétation arborée sur la végétation globale ;
- un modeleur final « LCZ » pour la cartographie des LCZ par croisement des indicateurs ;

accompagnés d'un quatrième modeleur qui est simplement l'enchaînement de ces trois.

Le modeleur « OCS », basé sur un croisement statistique raster-vecteur, existe en deux versions : une version qui utilise les outils intégrés à QGIS (Histogramme zonal), et une version qui utilise un script issu de la chaîne de traitement du Pôle satellite et adapté à QGIS (plus rapide, mais ne fonctionnant que sous Windows).

Il est nécessaire d'avoir en plus PostgreSQL/PostGIS d'installer et en lien avec QGIS pour les utiliser. Et comme pour la chaîne de traitements, ces modeleurs sont disponibles sur GitHub³².

31 <https://github.com/CEREMA/dterocc.chaineTraitement.traitementImageSatelliteEtIndicateursDerives>

32 <https://github.com/CEREMA/dterocc.modeleursQgis3.classificationLCZ>

ANNEXES

BIBLIOGRAPHIE

- [1] Hebrard D. & Piccinini B., Rapport d'étude – SCO SatLCZ – Cartographie LCZ par satellite, Cerema, 2021, 46 p., https://www.spaceclimateobservatory.org/sites/default/files/2021-09/rapport_SCO_SatLCZ_VF_2021-07-01.pdf
- [2] Stewart I.D. & Oke T.R., Local Climate Zones for Urban Temperature Studies, *Bulletin of the American Meteorological Society*, 2012/93/12, p.1879-1900, <https://doi.org/10.1175/BAMS-D-11-00019.1>

GLOSSAIRE

Cerema	Centre d'études et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement
DGALN	Direction Générale de l'Aménagement, du Logement et de la Nature
DINAMIS	Dispositif Institutionnel National d'Accès Mutualisé en Imagerie Satellitaire
EPSG	<i>European Petroleum Survey Group</i>
HRL	<i>High Resolution Layer</i>
ICU	Îlot de Chaleur Urbain
IGN	Institut national de l'information géographique et forestière
LCZ	<i>Local Climate Zone</i>
OCS	Occupation du sol
MNT	Modèle Numérique de Terrain
MNS	Modèle Numérique de Surface
MNH	Modèle Numérique de Hauteur
NDVI	<i>Normalized Difference Vegetation Index</i>
OSM	<i>Open Street Map</i>
SCO	<i>Space Climate Observatory</i>
SQL	<i>Structured Query Language</i>

INDEX DES FIGURES

Figure 1 : Vagues de chaleur observées et projetées en France métropolitaine.....	6
Figure 2 : Le mécanisme d'Îlot de Chaleur Urbain.....	7
Figure 3 : Définition des typologies de LCZ.....	9
Figure 4 : Grandes étapes de la production de la segmentation morphologique Cerema.....	14
Figure 5 : Grandes étapes de la production de la donnée LCZ.....	20
Figure 6 : Logigramme pour l'attribution de la classe de LCZ en fonction de la valeur des indicateurs.	23



**RÉPUBLIQUE
FRANÇAISE**

*Liberté
Égalité
Fraternité*



Cerema

CLIMAT & TERRITOIRES DE DEMAIN