

Documentation des Sources de Données

Pipeline SDG 9.1.1 & 11.3.1 – Cameroun
Décembre 2025

M. DIABANG, D. NGUEAJIO, M. SECK, L. NKWA et al.

École Nationale de la Statistique et de l'Analyse Économique (ENSAE) de Dakar

December 7, 2025

Contents

Note méthodologique	3
1 Limites Administratives	4
2 Infrastructure Routière	6
3 Données de Population	8
4 Données de Couverture du Sol – Built-up	13
5 Degré d’Urbanisation (DEGURBA)	19
6 Données Contextuelles	23
7 Synthèse des Datasets	28
8 Limitations Générales et Recommandations	29
9 Références Bibliographiques	31
10 Glossaire	32

Note méthodologique

Ce document constitue la documentation technique exhaustive des sources de données utilisées dans le cadre du projet de calcul des indicateurs ODD 9.1.1 (accessibilité rurale aux routes praticables toute l'année) et 11.3.1 (ratio de consommation des terres sur croissance démographique) pour le Cameroun.

Méthodologie de documentation :

- Recherches documentaires approfondies sur les sources officielles
- Validation des spécifications techniques via les catalogues de données
- Vérification des identifiants Google Earth Engine
- Conformité avec les méthodologies ESCAP (Elena Hristev)

Période d'analyse :

- Données historiques : 2017–2023 (7 années, résolution annuelle)
- Projections : 2030, 2035 (2 époques)

Plateforme : Google Earth Engine (GEE)

1 Limites Administratives

FICHE 1 : GADM v4.1 – Limites Administratives Mondiales

Nom complet du dataset : Database of Global Administrative Areas, version 4.1

Source : GADM Project, University of California, Berkeley

Site officiel : <https://gadm.org/>

Description : Base de données mondiale des limites administratives délimitant 400 276 unités administratives à travers le monde. Pour le Cameroun, fournit trois niveaux : national (ADM0), régional (ADM1), et départemental (ADM2).

Couverture spatiale : Mondiale (247 pays et territoires)

Couverture temporelle : Version 4.1 publiée en 2022

Résolution spatiale : Données vectorielles (polygones haute résolution)

Unité : Sans objet (géométries vectorielles)

Système de référence : EPSG:4326 (WGS84)

Format GEE : ee.FeatureCollection

Asset ID GEE (communautaire) :

projects/sat-io/open-datasets/gadm/gadm_410-levels

Filtrage pour le Cameroun :

```
.filter(ee.Filter.and(  
    ee.Filter.eq('GID_0', 'CMR'),  
    ee.Filter.eq('LEVEL', [0|1|2])  
)
```

Taille approximative :

- ADM0 : jusqu'à 100 KB
- ADM1 : 500 KB – 2 MB
- ADM2 : 2 – 5 MB

Attributs principaux :

- GID_0 : Code ISO3 du pays (CMR)
- NAME_0 : Nom du pays (Cameroon)
- GID_1, NAME_1 : Identifiants et noms des régions
- GID_2, NAME_2 : Identifiants et noms des départements
- LEVEL : Niveau administratif (0, 1, 2)

Méthodologie : Compilation de sources gouvernementales nationales, d'ONG, et de cartes officielles. Les limites sont généralisées à partir de données à haute résolution.

Précision : Résolution supérieure à d'autres bases publiques (Natural Earth, GAUL), mais pas validée par des entités fédérales internationales. Étude comparative (2011) : GADM moins précis que SALB et GAUL mais plus accessible.

Limites connues :

- Possible divergence avec les réformes administratives récentes
- Haute résolution côtière peut être lourde en traitement
- Non validé officiellement (à usage académique/recherche)
- Version 5.0 prévue en janvier 2026

Licence : Gratuit pour usage personnel, académique et recherche. Usage commercial nécessite une licence payante.

Citation recommandée :

University of California, Berkeley, Museum of Vertebrate Zoology, and International Rice Research Institute. GADM database of Global Administrative Areas, version 4.1. 2022. <https://gadm.org>

Usage dans le projet SDG :

- Délimitation de la zone d'étude (Cameroun entier)
- Découpage spatial des analyses par région/département
- Masque pour le clipping des rasters (population, built-up, DEGURBA)
- Agrégation statistique par unités administratives

Alternative envisagée : FAO GAUL (2015) – rejetée car obsolète. GADM 4.1 est le plus récent disponible.

2 Infrastructure Routière

FICHE 2 : HOTOSM – Réseau Routier OpenStreetMap

Nom complet du dataset : Humanitarian OpenStreetMap Team (HOT) – Global OSM Roads

Source : OpenStreetMap contributors via Humanitarian OpenStreetMap Team

Site officiel : <https://www.hotosm.org/>

Description : Réseau routier mondial extrait d'OpenStreetMap (OSM), incluant toutes catégories de routes : autoroutes, routes principales, secondaires, tertiaires, pistes, chemins. Version communautaire disponible sur GEE via sat-io.

Couverture spatiale : Mondiale

Couverture temporelle : Mise à jour continue ; snapshot GEE environ 2023–2024

Résolution spatiale : Vectorielle (lignes)

Unité : Attribut catégoriel `highway` (classification OSM)

Système de référence : EPSG:4326 (WGS84)

Format GEE : ee.FeatureCollection

Asset ID GEE :

`projects/sat-io/open-datasets/OSM/roads`

Taille approximative Cameroun : 20–40 MB (selon complétude régionale)

Attributs principaux :

- `highway` : Type de route (motorway, trunk, primary, secondary, tertiary, unclassified, residential, etc.)
- `surface` : Type de surface (paved, unpaved, asphalt, gravel, etc.) – souvent absent
- `name` : Nom de la route – incomplet
- `ref` : Référence de la route (ex: N1, A1) – incomplet

Méthodologie : Compilation communautaire mondiale via OpenStreetMap. Données contributives, vérifiées par la communauté. Qualité variable selon les régions.

Complétude estimée Cameroun :

- Zones urbaines (Yaoundé, Douala) : 70–80%
- Zones rurales : 30–50%
- Zones reculées : 10–30%

Limites connues :

- Couverture hétérogène selon les régions (biais urbain)
- Pas d'information explicite sur la praticabilité saisonnière
- Absence d'attribut `all-season` direct
- Qualité dépendante des contributeurs locaux

- Données de surface souvent manquantes

Proxy all-season utilisé :

Classification basée sur le type de route (méthodologie ESCAP) :

```
[‘motorway’, ‘motorway_link’,
‘trunk’, ‘trunk_link’,
‘primary’, ‘primary_link’,
‘secondary’, ‘secondary_link’]
```

Hypothèse : Routes principales et autoroutes = praticables toute l'année.

Validation envisagée :

- Comparaison visuelle avec Sentinel-2
- Analyse NDVI pour identifier végétation sur routes abandonnées
- Croisement avec données nationales si disponibles

Licence : Open Data Commons Open Database License (ODbL)

Citation recommandée :

OpenStreetMap contributors. *Planet dump*. Retrieved from <https://planet.osm.org>. © OpenStreetMap contributors. Data available under the ODbL.

Usage dans le projet SDG :

- Calcul SDG 9.1.1 : Analyse d'accessibilité rurale
- Buffer 2 km autour des routes all-season
- Calcul de la population rurale accessible

Note ESCAP : Elena Hristev reconnaît dans sa méthodologie que national road network data may sometimes lack explicit information on seasonality , justifiant l'usage de proxies.

3 Données de Population

FICHE 3 : WorldPop – Population Historique (2017–2023)

Nom complet du dataset : WorldPop Global Project Population Data – Unconstrained Estimates

Source : WorldPop, University of Southampton / University of Louisville / Université de Namur

Site officiel : <https://www.worldpop.org/>

Financement : Bill & Melinda Gates Foundation

Description : Estimations démographiques désagrégées spatialement à 100 m via modèles Random Forest. Version unconstrained : population répartie sur toutes les cellules terrestres sans contrainte de zones bâties.

Couverture spatiale : Mondiale (242 pays)

Couverture temporelle : 2000–2020 annuel (Global1) ; 2015–2030 annuel (Global2/R2025A)

Période utilisée projet : 2017–2023 (7 années annuelles)

Résolution spatiale : 100 m (3 arc-secondes)

Unité : Nombre de personnes par pixel

Système de référence : EPSG:4326 (WGS84)

Format GEE : ee.ImageCollection

Asset ID GEE :

WorldPop/GP/100m/pop

Filtrage Cameroun :

```
.filter(ee.Filter.eq('country', 'CMR'))  
.filter(ee.Filter.eq('year', YYYY))
```

Taille approximative par année : 150–250 MB (avant clipping)

Bandes disponibles :

- population : Estimation du nombre de personnes

Méthodologie : Modélisation dasymétrique par Random Forest. Désagrégation de données de recensement administratif vers grilles 100 m en utilisant :

- Imagerie satellite (Landsat, Sentinel-2)
- Données de luminosité nocturne (VIIRS)
- Couverture du sol
- Données d'infrastructure (routes, bâtiments)
- Variables environnementales (élévation, pente, accès à l'eau)

Source de calibration : Recensements nationaux + projections UN World Population Prospects.

Différence Unconstrained vs Constrained :

Unconstrained (utilisé ici) :

- Population répartie sur toutes cellules terrestres
- Peut allouer population dans zones non habitables
- Utile si limites de peuplement incertaines
- Disponible 2000–2020 annuel

Constrained :

- Population limitée aux zones avec bâtiments/infrastructures
- Utilise masques de peuplement satellite
- Plus réaliste pour zones urbaines denses
- Disponible uniquement 2020 (snapshot)

Précision : Corrélation avec recensements ≥ 0.85 en moyenne. Erreurs plus importantes en zones urbaines denses (slums) et zones rurales isolées. MAE (Mean Absolute Error) variable selon contexte.

Limites connues :

- Erreurs cellulaires importantes (jusqu'à 50% en zones urbaines)
- Sous-estimation en slums urbains
- Sur-estimation possible en zones rurales isolées
- Dépendance à la qualité des recensements nationaux
- Biais si recensements obsolètes (≥ 10 ans)
- Pas de distinction urbain/rural native

Validation : Comparaison avec recensements nationaux à l'échelle administrative. Validation croisée avec GHS-POP et Meta HRSL.

Licence : Creative Commons Attribution 4.0 International (CC BY 4.0)

Citation recommandée :

WorldPop (www.worldpop.org) and Center for International Earth Science Information Network (CIESIN), Columbia University (2018). Global High Resolution Population Denominators Project. Funded by The Bill and Melinda Gates Foundation (OPP1134076).

Bondarenko M., Kerr D., Sorichetta A., Tatem A. Census/projection-disaggregated gridded population datasets, adjusted to match the corresponding UNPD 2020 estimates, for 183 countries in 2020 using Built-Settlement Growth Model (BSGM) outputs. WorldPop, University of Southampton, UK, 2020. doi:10.5258/SOTON/WP00685

Usage dans le projet SDG :

- SDG 9.1.1 : Calcul de la population rurale totale et accessible
- SDG 11.3.1 : Calcul du Population Growth Rate (PGR) 2017–2023
- Combinaison avec DEGURBA pour identifier population rurale

FICHE 4 : GHS-POP – Population Projections (2030, 2035)

Nom complet du dataset : Global Human Settlement Layer – Population Grid R2023A

Source : European Commission, Joint Research Centre (JRC)

Programme : Copernicus Emergency Management Service

Site officiel : <https://human-settlement.emergency.copernicus.eu/>

Description : Grille de population mondiale multitemporelle (1975–2030) produite par désagrégation spatiale de données de recensement via modèles basés sur les surfaces bâties GHS-BUILT-S. Projections 2025–2030 basées sur tendances 1975–2020.

Couverture spatiale : Mondiale

Couverture temporelle : 1975–2030 (intervalles de 5 ans) ; 12 époques historiques + 2 projections

Période utilisée projet : 2030, 2035 (projections)

Note : 2035 n'existe pas officiellement dans GHS-POP R2023A. Utiliser 2030 comme proxy ou interpler 2030 avec taux de croissance UN.

Résolution spatiale : 100 m (native), également disponible 1 km

Unité : Nombre de personnes par cellule (décimales Float)

Système de référence : EPSG:54009 (World Mollweide)

Format GEE : ee.ImageCollection

Asset ID GEE :

JRC/GHSL/P2023A/GHS_POP

Filtrage par année :

```
.filter(ee.Filter.eq('year', 2030))
```

Taille approximative par époque : 150–250 MB (clipped Cameroun)

Bandes disponibles :

- `population_count` : Estimation du nombre d'habitants

Méthodologie : Désagrégation de données de recensement et projections CIESIN GPWv4.11 vers grilles 100 m, informée par :

- Distribution spatiale des surfaces bâties (GHS-BUILT-S)
- Volume bâti (GHS-BUILT-V)
- Classification résidentiel/non-résidentiel (GHS-BUILT-C)
- Modèles de croissance démographique UN

Méthode : Symbolic Machine Learning + regression spatiale.

Projections 2025–2030 : Basées sur :

- Tendances spatiales 1975–2020
- Projections démographiques UN World Population Prospects 2022

- Modèles de croissance urbaine
- Expansion prévue des surfaces bâties

Précision historique (1975–2020) :

- Précision totale d'allocation : 83% à 100 m
- MAE surfaces bâties : 6% de l'aire cellulaire
- Corrélation avec recensements : $\rho \approx 0.80$

Incertitude projections (2025–2030) :

- Incertitude croissante avec l'horizon temporel
- Projections 2030 : incertitude modérée
- Projections hypothétiques 2035 : incertitude élevée
- Dépendance aux scénarios de croissance UN
- Sensible aux changements de politiques démographiques

Limites connues :

- Projections uniquement jusqu'à 2030 officiellement
- 2035 nécessite extrapolation manuelle (non recommandé)
- Incertitude élevée hors zones métropolitaines
- Hypothèses fortes sur continuité des tendances
- Pas de prise en compte d'événements exceptionnels (pandémies, guerres, migrations massives)

Avantages GHS-POP :

- Cohérence méthodologique avec GHS-BUILT-S et GHS-SMOD
- Profondeur historique (1975–2020)
- Projection Mollweide (préserve aires)
- Intégration native avec DEGURBA
- Validation extensive (≈ 20 publications scientifiques)

Licence : Open and free. European Commission Reuse and Copyright Notice. Attribution requise.

Citation recommandée :

Schiavina, M.; Freire, S.; Carioli, A.; MacManus, K. (2023): GHS-POP R2023A – GHS population grid multitemporal (1975–2030). European Commission, Joint Research Centre (JRC). PID: <http://data.europa.eu/89h/2ff68a52-5b5b-4a22-8f40-c41da8332cfe>
doi:10.2905/2FF68A52-5B5B-4A22-8F40-C41DA8332CFE

Pesaresi, M., Schiavina, M., Politis, P., et al. (2024). Advances on the Global Human Settlement Layer by joint assessment of Earth Observation and Population Survey data. *International Journal of Digital Earth*, 17(1), 2390454. doi:10.1080/17538947.2024.2390454

Usage dans le projet SDG :

- SDG 11.3.1 : Calcul du Population Growth Rate (PGR) pour horizons 2030
- Scénarios prospectifs d'urbanisation
- Comparaison avec tendances historiques WorldPop

Note importante : Pour 2035, deux options :

1. Utiliser 2030 comme proxy (conservateur)
2. Interpoler linéairement 2030 → 2040 avec taux UN (risqué)

Recommandation : Limiter l'analyse prospective à 2030 et documenter l'incertitude.

4 Données de Couverture du Sol – Built-up

FICHE 5 : Esri Land Cover – Surfaces Bâties Historiques (2017–2023)

Nom complet du dataset : ESRI Global Land Use/Land Cover 10-Meter Time Series

Source : Esri, Microsoft, Impact Observatory

Partenariat : National Geographic Society (données d'entraînement)

Site officiel : <https://livingatlas.arcgis.com/landcover/>

Description : Série temporelle annuelle mondiale de classification d'occupation du sol à 10 m dérivée de Sentinel-2 par deep learning. 9 classes LULC incluant classe 7 Built Area (surfaces bâties).

Couverture spatiale : Mondiale

Couverture temporelle : 2017–2024 (8 années annuelles à ce jour)

Période utilisée projet : 2017–2023 (7 années)

Résolution spatiale : 10 m

Unité : Classes catégorielles (1–9)

Système de référence : EPSG:4326 (WGS84) ; traitement en UTM zones multiples

Format GEE : ee.ImageCollection

Asset ID GEE (communautaire) :

projects/sat-io/open-datasets/landcover/ESRI_Global-LULC_10m_TS

Filtrage par année :

```
.filter(ee.Filter.eq('year', YYYY))
```

Taille approximative par année : 50–150 MB (après clipping Cameroun)

Classes LULC (9 classes) :

1. Water
2. Trees
3. Flooded vegetation
4. Crops
5. **Built Area (classe 7)** ← Utilisé pour SDG 11.3.1
6. Bare ground
7. Snow/Ice
8. Clouds
9. Rangeland (combinaison Grass + Scrub depuis v3)

Méthodologie :

Données d'entrée :

- Sentinel-2 Level-2A (Surface Reflectance)
- 6 bandes : Blue, Green, Red, NIR, SWIR1, SWIR2
- Composite annuel médian (réduction nuages)
- $\approx 400\,000$ observations Sentinel-2 par an

Entraînement du modèle :

- Modèle : Deep Learning (Convolutional Neural Networks)
- Données d'entraînement : ≈ 5 milliards de pixels étiquetés manuellement
- Source labels : National Geographic Society Dynamic World Training Dataset
- Échantillonnage : $\approx 20\,000$ sites répartis sur tous biomes mondiaux
- Plateforme : Microsoft Planetary Computer + Azure Batch

Processus de classification :

1. Classification par dates multiples sur l'année entière
2. Composite des prédictions pour snapshot représentatif
3. Post-traitement pour cohérence spatiale
4. Filtrage des nuages et ombres

Précision : Précision moyenne évaluée à 75% globalement. Précision variable selon classes et biomes :

- Water, Trees, Built : $\approx 80\%$
- Crops, Rangeland : 70–80%
- Flooded vegetation, Bare ground : 60–75%

Précision classe Built (classe 7) : Estimée 80–85% en zones urbaines claires. Confusion possible avec :

- Bare ground (sols nus/chantiers)
- Rangeland (zones végétalisées avec structures dispersées)
- Ambiguïtés en zones périurbaines mixtes

Limites connues :

- Ambiguïtés zones mixtes urbain-végétation (périurbain)
- Confusion possible avec sols nus en zones arides
- Biais en zones de végétation sèche/saisonnière
- Résolution 10 m : structures < 10 m non détectées
- Bâtiments isolés en zones rurales parfois manqués
- Pas de distinction résidentiel/non-résidentiel

Avantages :

- Résolution temporelle annuelle (idéal pour tendances)

- Haute résolution spatiale (10 m)
- Méthodologie cohérente 2017–2024
- Conformité méthodologie Elena Hristev ESCAP
- Mise à jour annuelle continue
- Validation extensive (publications IEEE)

Compatibilité méthodologie ESCAP : Elena Hristev utilise exactement Esri LULC classe 7 pour le calcul SDG 11.3.1 (étude de cas Kirghizistan). Compatible 100%.

Licence : Creative Commons BY-4.0. Attribution requise à Esri, Impact Observatory, Microsoft, National Geographic Society.

Citation recommandée :

Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J.C., Mathis, M., Brumby, S.P. (2021). Global land use/land cover with Sentinel-2 and deep learning. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pp. 4704–4707. IEEE. doi:10.1109/IGARSS47720.2021.9553499

This dataset was produced by Impact Observatory for Esri. © 2021–2024 Esri. Dataset based on the Dynamic World training dataset by National Geographic Society in partnership with Google and World Resources Institute.

Usage dans le projet SDG :

- SDG 11.3.1 : Calcul du Land Consumption Rate (LCR) 2017–2023
- Extraction surfaces bâties annuelles : `lulc.eq(7)`
- Calcul de l'expansion urbaine temporelle
- Validation croisée avec GHS-BUILT-S

Validation envisagée :

- Comparaison avec GHS-BUILT-S 2020
- Contrôle visuel Sentinel-2 RGB
- Analyse de cohérence temporelle (pas de décroissance)

FICHE 6 : GHS-BUILT-S – Surfaces Bâties Projections (2030, 2035)

Nom complet du dataset : Global Human Settlement Layer – Built-up Surface Grid R2023A

Source : European Commission, Joint Research Centre (JRC)

Programme : Copernicus Emergency Management Service

Site officiel : <https://human-settlement.emergency.copernicus.eu/>

Description : Grille multitemporelle mondiale (1975–2030) de distribution des surfaces bâties exprimées en m² par cellule de 10 m. Produite par analyse d'imagerie Landsat et Sentinel-2 via Symbolic Machine Learning.

Couverture spatiale : Mondiale

Couverture temporelle : 1975–2030 (intervalles de 5 ans) ; 12 époques historiques + 2 projections

Période utilisée projet : 2030 (projection) ; 2035 par extrapolation

Note : 2035 non disponible officiellement. Utiliser 2030 ou interpoler avec tendances.

Résolution spatiale : 10 m (haute résolution, disponible pour 2018) ; 100 m (série temporelle 1975–2030)

Unité : Surface bâtie en mètres carrés (m²) par pixel

Système de référence : EPSG:54009 (World Mollweide)

Format GEE : ee.ImageCollection

Asset ID GEE :

JRC/GHSL/P2023A/GHS_BUILT_S (100m, 1975–2030)

JRC/GHSL/P2023A/GHS_BUILT_S_10m (10m, 2018 seulement)

Filtrage par année :

.filter(ee.Filter.eq('year', 2030))

Taille approximative par époque : 200–300 MB (clipped Cameroun, 10 m)

Bandes disponibles :

- built_surface : Surface bâtie totale (m²)
- built_surface_nres : Surface bâtie non-résidentielle (m²)

Méthodologie :

Données d'entrée :

- Landsat Collection 2 (MSS, TM, ETM+, OLI) : 1975–2020
- Sentinel-2 L2A composite 2018 : haute résolution
- Classification via Symbolic Machine Learning

Processus :

1. Extraction surfaces bâties par époque (1975–2020)
2. Classification résidentiel/non-résidentiel via :

- GHS-BUILT-H (hauteur bâtiments)
- GHS-BUILT-V (volume bâti)
- Indices spectraux (NDVI, NDBI)
- Patterns spatiaux

3. Interpolation/extrapolation spatiotemporelle 1975–2030

4. Projections 2025–2030 basées sur :

- Tendances d'expansion urbaine 1975–2020
- Modèles de croissance socio-économique
- Projections démographiques UN
- Politiques d'aménagement territorial

Précision historique (1975–2020) :

- IoU (Intersection over Union) classe built-up : 0.92 à 10 m
- MAE surface bâtie : 6% de l'aire cellulaire à 100 m
- Précision classe built-up : ↗ 85%
- Précision classe non-résidentiel : 80%

Incertitude projections (2025–2030) :

- Incertitude élevée hors zones métropolitaines
- Projections basées sur extrapolation linéaire des tendances
- Sensible aux changements de politiques urbaines
- Ne capture pas événements exceptionnels (catastrophes, politiques brutales)
- Incertitude croissante avec horizon temporel

Limites connues :

- Projections officielles seulement jusqu'à 2030
- 2035 nécessite extrapolation manuelle (non recommandé)
- Résolution temporelle 5 ans (pas annuelle)
- Incertitude importante en zones rurales isolées
- Hypothèses fortes sur continuité tendances d'urbanisation
- Pas de prise en compte politiques urbaines futures

Avantages GHS-BUILT-S :

- Cohérence méthodologique avec GHS-POP et GHS-SMOD
- Profondeur historique exceptionnelle (1975–2020)
- Résolution spatiale fine (10–100 m)
- Distinction résidentiel/non-résidentiel
- Projection Mollweide (préserve aires → calculs surfaces exacts)

- Validation extensive (> 30 publications JRC)

Masque binaire built-up :

```
var builtup = built_surface.gt(0); // Pixel bâti si surface > 0 m2
```

Licence : Open and free. European Commission Reuse and Copyright Notice. Attribution required.

Citation recommandée :

Pesaresi, M.; Politis, P. (2023): GHS-BUILT-S R2023A – GHS built-up surface grid, derived from Sentinel2 composite and Landsat, multitemporal (1975–2030). European Commission, Joint Research Centre (JRC). PID: <http://data.europa.eu/89h/9f06f36f-4b11-47ec-abb0-4f8b7b1d72ea>. doi:10.2905/9F06F36F-4B11-47EC-ABB0-4F8B7B1D72EA

Pesaresi, M., Schiavina, M., Politis, P., et al. (2024). Advances on the Global Human Settlement Layer by joint assessment of Earth Observation and Population Survey data. *International Journal of Digital Earth*, 17(1), 2390454. doi:10.1080/17538947.2024.2390454

Usage dans le projet SDG :

- SDG 11.3.1 : Scénarios prospectifs Land Consumption Rate (LCR) 2030
- Comparaison avec tendances historiques Esri LULC
- Validation croisée périodes de chevauchement (2020)

Recommandation 2035 :

1. **Option conservatrice** : Utiliser 2030 comme proxy maximum
2. **Option interpolation** : Appliquer taux de croissance 2020–2030 pour estimer 2035 (documenter incertitude élevée)
3. **Option recommandée** : Limiter l'analyse prospective à 2030

5 Degré d'Urbanisation (DEGURBA)

FICHE 7 : GHS-SMOD – Settlement Model Grid (2020, 2030, 2035)

Nom complet du dataset : Global Human Settlement Layer – Settlement Model Grid R2023A

Source : European Commission, Joint Research Centre (JRC)

Programme : Copernicus Emergency Management Service

Site officiel : <https://human-settlement.emergency.copernicus.eu/>

Description : Classification mondiale multitemporelle du degré d'urbanisation (DEGURBA) appliquant la méthodologie Degree of Urbanisation Stage I recommandée par la Commission Statistique des Nations Unies. Basée sur grilles de population GHS-POP et surfaces bâties GHS-BUILT-S.

Couverture spatiale : Mondiale

Couverture temporelle : 1975–2030 (intervalles de 5 ans) ; 12 époques

Période utilisée projet : 2020 (référence), 2030 (projection) ; 2035 par extrapolation

Note : 2035 non disponible. Utiliser 2030 ou extrapoler avec tendances.

Résolution spatiale : 1 km (basse résolution) ; version expérimentale 100 m disponible pour 2020 uniquement

Unité : Code SMOD (classification DEGURBA)

Système de référence : EPSG:54009 (World Mollweide)

Format GEE : ee.ImageCollection

Asset ID GEE :

JRC/GHSL/P2023A/GHS_SMOD (1km, 1975–2030)

JRC/GHSL/P2023A/GHS_SMOD/2020_100 (100m, 2020 uniquement)

Filtrage par année :

```
.filter(ee.Filter.eq('year', 2020))
```

Taille approximative par époque : 10–15 MB (1 km) ; 100–200 MB (100 m)

Bandes disponibles :

- smod_code : Code de classification DEGURBA (10–30)

Classification DEGURBA (8 classes détaillées – Niveau 2) :

Code	Classe	Définition
10	Water	Surface en eau
11	Very low density rural	Rural densité très faible
12	Low density rural	Rural densité faible
13	Rural cluster	Cluster rural
21	Suburban	Suburbain / périurbain
22	Semi-dense urban cluster	Cluster urbain semi-dense
23	Dense urban cluster	Cluster urbain dense
30	Urban centre	Centre urbain ← Utilisé pour SDG 11.3.1

Classification simplifiée (Niveau 1) :

- **Urban** : 21–30 (agglomération $\geq 50\ 000$ hab, densité ≥ 1500 hab/km 2)
- **Semi-urban/Towns** : 21–23 (clusters ≥ 5000 hab)
- **Rural** : 11–13 (reste)

Méthodologie DEGURBA Stage I :

Inputs :

- GHS-POP R2023A (grille population 100 m)
- GHS-BUILT-S R2023A (grille surfaces bâties 10 m)

Processus (méthodologie UN) :

1. Identification cellules urbaines :

- Densité population ≥ 1500 hab/km 2
- + Surface bâtie \geq seuil biome-spécifique

2. Clustering spatial :

- Agglomération cellules contigües
- Formation clusters urbains

3. Seuils de population :

- Urban centre (30) : Cluster $\geq 50\ 000$ habitants
- Dense urban cluster (23) : Cluster 20 000–50 000 hab
- Semi-dense urban cluster (22) : Cluster 5000–20 000 hab
- Suburban (21) : Zones périurbaines ; 5000 hab mais haute densité

4. Classification rurale :

- Rural cluster (13) : Clusters 500–5000 hab
- Low density rural (12) : Densité ≤ 300 hab/km 2
- Very low density rural (11) : Densité ≤ 50 hab/km 2

Résolution 1 km vs 100 m :

1 km (standard, 1975–2030) :

- Résolution grossière
- Perte détails en zones périurbaines
- Adapté analyses macro-régionales
- Disponible toutes époques

100 m (expérimental, 2020 uniquement) :

- Haute résolution
- Détails zones périurbaines
- Meilleure délimitation frontières urbain/rural
- Non disponible pour séries temporelles

Recommandation projet : Utiliser 1 km pour cohérence temporelle 2020–2030.

Précision : Validation extensive avec recensements nationaux et données urbaines. Accord $\geq 85\%$ avec classifications nationales. Sous-estimation possible petites villes en zones rurales isolées.

Limites connues :

- Résolution 1 km très faible pour analyses locales
- Perte détails en zones périurbaines complexes
- Classification binaire urbain/rural simpliste
- Seuils de densité fixes (pas d'adaptation contextuelle)
- Projections 2030 basées sur extrapolation linéaire
- 2035 non disponible
- Pas de distinction types d'urbanisation (industrielle, résidentielle)

Avantages :

- Méthodologie standardisée UN (reproductibilité internationale)
- Cohérence totale avec GHS-POP et GHS-BUILT-S
- Série temporelle 1975–2030 (analyse évolution urbaine)
- Classification officielle adoptée par UN Statistical Commission (2020)
- Validation extensive (≥ 200 pays)

Extraction zones urbaines (SDG 11.3.1) :

```
var urban_centres = smod.eq(30); // Uniquement centres urbains
```

Ou version inclusive :

```
var urban_all = smod.gte(21); // Tout urbain (21-30)
```

Définition population rurale (SDG 9.1.1) :

```
var rural_mask = smod.lt(21); // Zones rurales (10-13)
var rural_pop = population.updateMask(rural_mask);
```

Licence : Open and free. European Commission Reuse and Copyright Notice. Attribution requise.

Citation recommandée :

Schiavina, M.; Melchiorri, M.; Pesaresi, M. (2023): GHS-SMOD R2023A – GHS settlement layers, application of the Degree of Urbanisation methodology (stage I) to GHS-POP R2023A and GHS-BUILT-S R2023A, multitemporal (1975–2030). European Commission, Joint Research Centre (JRC). PID: <http://data.europa.eu/89h/a0df7a6f-49de-46ea-9bde-563437a6e2ba>. doi:10.2905/A0DF7A6F-49DE-46EA-9BDE-563437A6E2BA

Dijkstra, L., Florczyk, A., Freire, S., et al. (2021). Applying the Degree of Urbanisation to the globe: A new harmonised definition reveals a different picture of global urbanisation. *Journal of Urban Economics*, 125, 103312. doi:10.1016/j.jue.2020.103312

Usage dans le projet SDG :

- SDG 9.1.1 : Identification population rurale (masque DEGURBA | 30)
- SDG 11.3.1 : Identification zones urbaines (DEGURBA = 30)
- Calcul du ratio de consommation de terres urbaines

Note méthodologique ESCAP : Elena Hristev utilise classification similaire mais génère DEGURBA custom via outil GHS-DUG (MATLAB). Dans GEE, GHS-SMOD est l'équivalent direct et validé.

6 Données Contextuelles

FICHE 8 : SRTM – Modèle Numérique d’Élévation

Nom complet du dataset : Shuttle Radar Topography Mission (SRTM) – Global 1 Arc-Second V3

Source : NASA / USGS (United States Geological Survey)

Mission : Space Shuttle Endeavour, février 2000

Site officiel : <https://www2.jpl.nasa.gov/srtm/>

Description : Modèle numérique d’élévation (DEM) mondial quasi-global acquis par radar interférométrique en bande C. Fournit élévation du terrain en mètres au-dessus du niveau de la mer (ellipsoïde WGS84).

Couverture spatiale : 60°N – 56°S (couvre 80% des terres émergées)

Couverture temporelle : Février 2000 (acquisition unique, mission 11 jours)

Résolution spatiale : 1 arc-seconde (30 m à l’équateur) ; version 3 arc-secondes (90 m) également disponible

Unité : Mètres (élévation)

Système de référence : EPSG:4326 (WGS84)

Format GEE : ee.Image

Asset ID GEE :

USGS/SRTMGL1_003 (30m, 1 arc-second)

USGS/SRTM90_V4 (90m, 3 arc-seconds)

Taille approximative Cameroun : 100 MB (30 m)

Bandes disponibles :

- elevation : Élévation en mètres

Méthodologie : Radar interférométrique en bande C (5.6 cm longueur d’onde). Deux antennes sur bras de 60 m pour créer interférogrammes. Génération DEM par traitement phase radar.

Précision verticale :

- Précision absolue : ±16 m (90% confiance)
- Précision relative : ±10 m (pente ± 20%)
- Zones montagneuses : précision dégradée (±30 m)

Précision horizontale : ±20 m (90% confiance)

Traitement version 3 : Amélioration 2013 avec :

- Comblement vides (océans, lacs, zones d’ombre radar)
- Filtrage artefacts
- Correction aberrations
- Amélioration cohérence

Limites connues :

- Artefacts possibles en zones montagneuses abruptes
- Radar pénètre végétation : élévation = sommet canopée (biais +5 à +15 m en forêts denses)
- Zones d'ombre radar en reliefs complexes
- Pas de mise à jour (données 2000, 24 ans d'âge)
- Changements topographiques récents (mines, déforestation, urbanisation) non capturés

Usage dans le projet SDG :

- Contexte physique du territoire
- Analyse de l'accessibilité routière (pente, relief)
- Identification zones difficiles d'accès
- Calcul pente et rugosité terrain
- Validation cohérence données (ex: bâti en zones à fort dénivelé)

Produits dérivés calculables :

- Pente (slope) : ee.Terrain.slope(elevation)
- Aspect (orientation) : ee.Terrain.aspect(elevation)
- Rugosité (TRI, Terrain Ruggedness Index)
- Ombrage (hillshade)

Licence : Domaine public (données NASA/USGS). Libre utilisation sans restriction.

Citation recommandée :

NASA Shuttle Radar Topography Mission (SRTM) (2013). Shuttle Radar Topography Mission (SRTM) Global. Distributed by OpenTopography. <https://doi.org/10.5069/G9445JDF>

Farr, T.G., Rosen, P.A., Caro, E., et al. (2007). The Shuttle Radar Topography Mission. *Reviews of Geophysics*, 45, RG2004. doi:10.1029/2005RG000183

Note : SRTM est un standard de facto pour DEM global. Utilisé comme référence dans de nombreux projets (GHS-BUILT-H, analyses hydrologiques, etc.).

FICHE 9 : Sentinel-2 – Imagerie Optique Haute Résolution

Nom complet du dataset : Sentinel-2 MSI: MultiSpectral Instrument, Level-2A (Surface Reflectance)

Source : European Space Agency (ESA) / Copernicus Programme

Satellites : Sentinel-2A (lancé 2015), Sentinel-2B (lancé 2017)

Site officiel : <https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-2>

Description : Constellation de satellites d'observation optique haute résolution fournissant imagerie multispectrale pour surveillance des terres émergées, zones côtières et eaux intérieures. 13 bandes spectrales du visible à l'infrarouge moyen.

Couverture spatiale : Mondiale (terres émergées entre 84°N et 56°S)

Couverture temporelle : 2015–présent (continue)

Période utilisée projet : 2023 (composite annuel médian)

Résolution spatiale :

- 10 m : B2 (Blue), B3 (Green), B4 (Red), B8 (NIR)
- 20 m : B5, B6, B7 (Red Edge), B8A (NIR Narrow), B11, B12 (SWIR)
- 60 m : B1 (Coastal Aerosol), B9 (Water Vapour), B10 (Cirrus)

Fréquence temporelle : 5 jours (constellation 2 satellites) ; 10 jours (satellite individuel)

Unité : Réflectance de surface (Surface Reflectance, sans unité 0–1)

Système de référence : UTM (zones multiples) ; reprojection WGS84 dans GEE

Format GEE : ee.ImageCollection

Asset ID GEE :

COPERNICUS/S2_SR_HARMONIZED (Level-2A, Surface Reflectance)

COPERNICUS/S2_HARMONIZED (Level-1C, Top-of-Atmosphere)

Note : Utiliser version HARMONIZED (correction radiométrique Sentinel-2A/2B).

Taille approximative composite Cameroun : 500 MB – 1 GB (avant clipping et sélection bandes)

Bandes utilisées projet :

- B2, B3, B4 : RGB (visualisation)
- B8 : NIR (végétation, NDVI)
- B11, B12 : SWIR (humidité, urbain)

Méthodologie de composite :

Objectif : Créer image représentative annuelle sans nuages.

Processus :

1. Filtrage spatial : .filterBounds(cameroon)
2. Filtrage temporel : .filterDate('2023-01-01', '2023-12-31')

3. Filtrage nuages : `.filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE', 20))`
4. Agrégation : `.median()` (réduction médiane par pixel)
5. Clipping : `.clip(cameroon)`

Indices spectraux calculables :

NDVI (Normalized Difference Vegetation Index) :

$$\text{NDVI} = (\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red}) = (\text{B8} - \text{B4}) / (\text{B8} + \text{B4})$$

Utilité : Identifier végétation active, valider routes abandonnées.

NDBI (Normalized Difference Built-up Index) :

$$\text{NDBI} = (\text{SWIR} - \text{NIR}) / (\text{SWIR} + \text{NIR}) = (\text{B11} - \text{B8}) / (\text{B11} + \text{B8})$$

Utilité : Identifier zones bâties, valider Esri LULC.

NDWI (Normalized Difference Water Index) :

$$\text{NDWI} = (\text{Green} - \text{NIR}) / (\text{Green} + \text{NIR}) = (\text{B3} - \text{B8}) / (\text{B3} + \text{B8})$$

Utilité : Identifier surfaces en eau.

Précision radiométrique : Surface Reflectance corrigée atmosphère, géométrie, radiométrie.
Précision absolue 5% sur surfaces homogènes.

Limites connues :

- Couverture nuageuse importante en zones tropicales (Cameroun centre-sud : 50–70% nuages)
- Nécessité de composites temporels (médiane, percentiles)
- Saison sèche (décembre–février) : meilleure qualité
- Saison des pluies (juin–septembre) : forte couverture nuageuse
- Résolution temporelle limitée en zones très nuageuses

Usage dans le projet SDG :

- Validation visuelle des données OSM (routes réelles ?)
- Validation Esri LULC classe built-up (zones urbaines cohérentes ?)
- Contexte géographique pour rapports et cartes
- Identification zones de changement (déforestation, urbanisation)
- Calcul NDVI pour identifier routes abandonnées/végétalisées

Recommandation composite :

- Période saison sèche : novembre–février (moins de nuages)
- Filtrage CLOUDY_PIXEL_PERCENTAGE < 20%
- Agrégation médiane (robuste aux outliers)
- Sélection bandes : B2, B3, B4, B8 (10 m) pour réduire taille

Licence : Copernicus Sentinel Data – Full, open and free access. Attribution ESA requise.

Citation recommandée :

European Space Agency (ESA). Copernicus Sentinel-2 data. Retrieved from <https://scihub.copernicus.eu/>

Drusch, M., Del Bello, U., Carlier, S., et al. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, 120, 25–36. doi:10.1016/j.rse.2011.11.026

Note : Sentinel-2 est la référence mondiale pour imagerie optique haute résolution gratuite. Base de nombreux produits dérivés (Esri LULC, Dynamic World, etc.).

7 Synthèse des Datasets

Tableau récapitulatif complet

Dataset	Variable	Résolution	Période	Projection
LIMITES ADMINISTRATIVES				
GADM v4.1	Frontières	Vectoriel	2022	WGS84
INFRASTRUCTURE				
HOTOSM OSM	Routes	Vectoriel	2023-24	WGS84
POPULATION				
WorldPop	Population	100 m	2017-2023	WGS84
GHS-POP	Population (proj)	100 m	2030	Mollweide
BUILT-UP / LULC				
Esri LULC	Classe built-up	10 m	2017-2023	WGS84
GHS-BUILT-S	Surface bâtie	10-100 m	2030	Mollweide
DEGURBA				
GHS-SMOD	Classification	1 km	2020, 2030	Mollweide
CONTEXTE				
SRTM	Élévation	30 m	2000	WGS84
Sentinel-2	Imagerie RGB+NIR	10 m	2023	WGS84/UTM

Cohérence méthodologique

Suite GHS (JRC) :

- GHS-POP, GHS-BUILT-S, GHS-SMOD : Méthodologie cohérente
- Projection Mollweide commune (préserve aires)
- Intervalles 5 ans synchronisés
- Validation croisée extensive

Données annuelles (historique) :

- WorldPop + Esri LULC : Résolution temporelle fine
- Projection WGS84 commune
- Périodes 2017–2023 alignées

Conformité ESCAP : Méthodologie Elena Hristev respectée avec adaptations documentées (proxies OSM, grilles population).

8 Limitations Générales et Recommandations

Limitations transversales

Projections 2030–2035 :

- 2035 non disponible officiellement (GHS-POP, GHS-BUILT-S, GHS-SMOD)
- Incertitude croissante avec horizon temporel
- Hypothèses fortes sur continuité tendances
- Recommandation : Limiter analyses prospectives à 2030

Proxy OSM all-season :

- Pas d'attribut praticabilité saisonnière direct
- Classification par type de route (approximation)
- Validation recommandée avec Sentinel-2 + NDVI

Grilles population vs bâtiments ponctuels :

- Elena Hristev utilise bâtiments géocodés (vecteur)
- Projet utilise grilles 100 m (raster)
- Différence méthodologique acceptable (désagrégation spatiale équivalente)

Résolution DEGURBA 1 km :

- Très grossière pour analyses locales
- Perte détails zones périurbaines
- Alternative : GHS-SMOD 100 m (2020 uniquement)

Recommandations qualité

Validation croisée :

1. Comparer WorldPop 2020 GHS-POP 2020
2. Comparer Esri LULC 2020 GHS-BUILT-S 2020
3. Valider OSM avec Sentinel-2 RGB
4. Analyser cohérence temporelle (pas de décroissance anormale)

Traitements des incertitudes :

1. Documenter toutes hypothèses et proxies
2. Quantifier incertitudes (intervalles de confiance si possible)
3. Analyse de sensibilité (ex: seuils DEGURBA, buffer routes)
4. Scénarios alternatifs pour projections 2030

Reprojections :

- Privilégier projections équivalentes (Mollweide pour calculs d'aires)

- Documenter transformations de projection
- Vérifier cohérence échelles après reprojection

9 Références Bibliographiques

Méthodologie SDG

Hristev, E. (2021). *Calculating SDG Indicators 9.1.1 and 11.3.1 using Earth Observation Data: A Step-by-Step Methodology*. United Nations Economic and Social Commission for Asia and the Pacific (ESCAP). Bangkok.

United Nations. (2022). *SDG Indicators Metadata Repository*. <https://unstats.un.org/sdgs/metadata/>

GADM

University of California, Berkeley. (2022). *GADM database of Global Administrative Areas, version 4.1*. <https://gadm.org>

WorldPop

Bondarenko, M., Kerr, D., Sorichetta, A., Tatem, A. (2020). Census/projection-disaggregated gridded population datasets for 183 countries in 2020 using BSGM outputs. University of Southampton. doi:10.5258/SOTON/WP00685

Stevens, F.R., Gaughan, A.E., Linard, C., Tatem, A.J. (2015). Disaggregating census data for population mapping using Random Forests with remotely-sensed and ancillary data. *PLOS ONE*, 10(2), e0107042. doi:10.1371/journal.pone.0107042

GHS (JRC)

Pesaresi, M., Schiavina, M., Politis, P., et al. (2024). Advances on the Global Human Settlement Layer by joint assessment of Earth Observation and Population Survey data. *International Journal of Digital Earth*, 17(1), 2390454. doi:10.1080/17538947.2024.2390454

Schiavina, M., Melchiorri, M., Pesaresi, M., et al. (2023). *GHSL Data Package 2023*. Publications Office of the European Union, Luxembourg. doi:10.2760/098587

Dijkstra, L., Florczyk, A., Freire, S., et al. (2021). Applying the Degree of Urbanisation to the globe: A new harmonised definition. *Journal of Urban Economics*, 125, 103312. doi:10.1016/j.jue.2020.103312

Esri Land Cover

Karra, K., Kontgis, C., Statman-Weil, Z., et al. (2021). Global land use/land cover with Sentinel-2 and deep learning. *2021 IEEE IGARSS*, 4704–4707. doi:10.1109/IGARSS47720.2021.9553499

Sentinel-2

Drusch, M., Del Bello, U., Carlier, S., et al. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, 120, 25–36. doi:10.1016/j.rse.2012.02.018

SRTM

Farr, T.G., Rosen, P.A., Caro, E., et al. (2007). The Shuttle Radar Topography Mission. *Reviews of Geophysics*, 45, RG2004. doi:10.1029/2005RG000183

10 Glossaire

All-season road : Route praticable toute l'année, indépendamment des conditions météorologiques.

Built-up area : Zone construite, surface occupée par bâtiments et infrastructures.

DEGURBA : Degree of Urbanisation, classification du degré d'urbanisation (méthodologie UN).

DEM : Digital Elevation Model, modèle numérique d'élévation.

GEE : Google Earth Engine, plateforme d'analyse géospatiale cloud.

GHS : Global Human Settlement Layer, couche mondiale de peuplement humain (JRC).

IoU : Intersection over Union, métrique de précision classification (0–1).

JRC : Joint Research Centre, centre de recherche de la Commission Européenne.

LCR : Land Consumption Rate, taux de consommation des terres (SDG 11.3.1).

LULC : Land Use/Land Cover, occupation et utilisation du sol.

MAE : Mean Absolute Error, erreur absolue moyenne.

NDVI : Normalized Difference Vegetation Index, indice de végétation normalisé.

OSM : OpenStreetMap, projet cartographique collaboratif mondial.

PGR : Population Growth Rate, taux de croissance démographique (SDG 11.3.1).

Proxy : Variable de substitution utilisée en l'absence de données directes.

SDG : Sustainable Development Goals, Objectifs de Développement Durable (ODD).

SMOD : Settlement Model Grid, grille de modèle de peuplement (DEGURBA).

SRTM : Shuttle Radar Topography Mission, mission topographique radar spatiale.

UN : United Nations, Nations Unies.

WGS84 : World Geodetic System 1984, système géodésique de référence mondial (EPSG:4326).

Document généré dans le cadre du TP de statistique exploratoire spatial portant sur le calcul des indicateurs SDG 9.1.1 & 11.3.1

École Nationale de la Statistique et de l'Analyse Économique (ENSAE) Dakar

Décembre 2025