

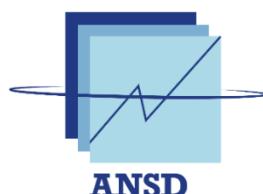
République du Sénégal



Un Peuple - un But - une Foi

MINISTERE DE L'ECONOMIE, DU PLAN ET DE LA COOPERATION (MEPC)

AGENCE NATIONALE DE LA STATISTIQUE ET DE LA DEMOGRAPHIE (ANSD)



Agence Nationale de la
Statistique et de la Démographie

ECOLE NATIONALE DE LA STATISTIQUE ET DE L'ANALYSE ECONOMIQUE
PIERRE NDIAYE (ENSAE)



Statistique exploratoire spatiale

Résumé des travaux pratiques

Rédigé par :

Herman Parfait NGAKE YAMAHA

Etudiant en ISE1-CL à l'ENSAE

Sous la supervision de :

M. Aboubacar HEMA

Research Scientist

Année académique 2025-2026

Sommaire

Sommaire	2
Liste des figures	3
Introduction.....	4
TP 0 : Présentation de Google Earth Engine.....	6
TP 1 : Analyse spatiale du paludisme et de la population au Cameroun	8
TP 2 : Analyse des données géospatiales portant sur les infrastructures, les aires protégées et la population du Cameroun.....	10
TP 3 : Analyse des données géospatiales portant sur les infrastructures, les aires protégées et la population du Cameroun (suite du TP2)	12
TP 4 : Identification des terres arables au Burundi	14
TP 5 : Analyse du ratio LCRPGR en Tanzanie entre 2017 et 2022.....	16
TP 6 : Analyse des dynamiques environnementales et territoriales au Sénégal à partir des indices spectraux de Sentinel-2	18
TP 7 (individuel) : Analyse des matrices de voisinage et de l'autocorrélation spatiale : Application de l'indice de Moran global avec R	23
Conclusion	26

Liste des figures

Figure 1 : Interface de la carte interactive des subdivisions administratives du Cameroun.....	9
Figure 2 : Interface de la carte interactive de l'évolution du taux d'incidence du paludisme au Cameroun	9
Figure 3 : Vue d'ensemble de quelques infrastructures du Cameroun.....	10
Figure 4 : Interface de la carte interactive des aires protégées du Cameroun	11
Figure 5 : Interface de la carte interactive de l'analyse spatiale de l'accès à l'éducation et à la santé au Cameroun	11
Figure 6 : Distribution des distances villages-hôpitaux au Cameroun	12
Figure 7 : Interface de la carte interactive de couverture sanitaire au Cameroun en 2022.	13
Figure 8 : Visualisation des chemins de fer du Cameroun avec buffer.....	13
Figure 9 : Interface de la carte interactive des surfaces arabes au Burundi.....	15
Figure 10 : Interface de la carte interactive de l'étalement urbain en Tanzanie.....	17
Figure 11 : Résultats issus de l'analyse du NDBI au Sénégal en 2018	21
Figure 12 : Top 10 des départements avec les indices de nudité (BSI) les plus faibles au Sénégal en 2018	21
Figure 13 : Interface de la carte interactive de visualisation du NDVI dans les départements du Sénégal en 2018	22
Figure 14 : Classification des quartiers de Columbus suivant le revenu annuel moyen des ménages (en milliers de dollars US)	24
.....	24
Figure 15 : Structure de voisinage avec contiguïté Bishop (3*3).....	24
.....	24
Figure 16 : Illustration d'un réseau de voisinage KNN (K=5).....	25
Figure 17 : Illustration d'un réseau de voisinage par contiguïté Queen	25
Figure 18 : Illustration d'un réseau de voisinage par distance seuil (seuil = 0.42 dégré)	25

Introduction

Le cours de **statistiques exploratoires spatiales** vise à développer des compétences en cartographie thématique et en analyse spatiale, en privilégiant une approche pratique et interactive. Les travaux pratiques réalisés en équipe ont constitué des espaces privilégiés pour renforcer le travail collaboratif, affiner la capacité de synthèse des résultats et développer les aptitudes à la communication scientifique à travers des présentations orales.

Pour mener à bien ces exercices, plusieurs environnements de programmation et plateformes de traitement géospatial ont été mobilisés :

- **Python**, aussi bien en tant en programmation classique qu'à travers son API intégrée à Google Earth Engine ;
- **R**, pour l'analyse statistique et la manipulation avancée des données spatiales ;
- **JavaScript**, dans l'environnement natif de Google Earth Engine.

Un premier travail pratique introductif (**TP0**) a été entièrement consacré à la présentation de Google Earth Engine, permettant de se familiariser avec l'interface du Code Editor, d'explorer le vaste catalogue de données disponibles et de maîtriser les principales fonctionnalités de cette plateforme cloud de calcul géospatial. Les travaux suivants ont abordé des thématiques variées, principalement ancrées dans le contexte africain, à savoir :

- l'analyse spatiale et temporelle du paludisme en relation avec la distribution de la population au Cameroun ;
- l'étude de l'accessibilité aux infrastructures sociales (santé et éducation), la cartographie des aires protégées et l'analyse des réseaux de transport au Cameroun ;
- l'identification et la quantification des terres arables au Burundi ;
- l'évaluation de l'indicateur LCRPGR (ratio de consommation des terres par rapport à la croissance démographique) en Tanzanie, pour mesurer l'efficacité de la gestion de l'étalement urbain ;
- l'analyse des dynamiques environnementales et territoriales du Sénégal à partir d'indices spectraux multiples calculés sur des images Sentinel-2 ;

- et enfin, un travail pratique individuel consacré à la construction de matrices de voisinage et à l'analyse de l'autocorrélation spatiale via l'indice de Moran global.

Le présent document propose une synthèse de l'ensemble de ces travaux, en mettant en évidence entre autres les objectifs poursuivis, les sources de données, quelques résultats obtenus et les difficultés rencontrées.

TP O : Présentation de Google Earth Engine

Membres du groupe :

- AGNANGMA SANAM David Landry
- DIOP Astou
- DIOP Marème
- NGAKE YAMAHA Herman Parfait

Il était question dans ce premier TP, à partir d'un document PowerPoint, de présenter Google Earth Engine (GEE), une plateforme de calcul dédiée à l'analyse géospatiale à grande échelle. L'introduction a rappelé le contexte du Big Data géospatial, marqué par des volumes massifs de données et la nécessité de traitements parallèles rapides. GEE, lancé en 2010 lors de la COP16 à Cancún, répond à ces défis en combinant un vaste catalogue de données, une grande capacité de calcul et des interfaces de programmation (JavaScript et Python).

La première partie de la présentation a consisté à définir GEE et à mettre en avant son utilité : réduction drastique des temps de traitement, analyse de séries temporelles longues, exploration de l'hétérogénéité spatiale et détection d'anomalies. Les principaux domaines d'application incluent l'environnement et le climat (déforestation, sécheresses), l'agriculture (santé des cultures, gestion de l'eau), l'urbanisme (suivi de l'étalement urbain), et la gestion des catastrophes (inondations, incendies).

La deuxième partie a décrit les étapes de création d'un compte GEE et d'accès au catalogue de données (plus de 800 jeux publics incluant Landsat, Sentinel, MODIS, ERA5, données socio-économiques, etc.), ainsi que l'interface du Code Editor. Les composants essentiels de ladite interface sont l'éditeur de script, la console de sortie, le visualiseur cartographique, l'inspecteur de carte et le gestionnaire de tâches.

La troisième partie a présenté les langages de programmation intégrés :

- JavaScript, utilisé dans l'éditeur web pour le prototypage rapide et la visualisation interactive.
- Python API, permettant une intégration avec des environnements scientifiques (Jupyter, Google Colab) et des analyses avancées.

Enfin, un cas pratique a illustré l'utilisation de GEE pour analyser l'évolution des températures au Sénégal entre 2014 et 2024, à partir des données provenant de ERA5-Land. L'étude a mis en évidence les tendances annuelles, les contrastes régionaux et les variations saisonnières, grâce à des scripts JavaScript exécutés dans GEE.

La conclusion a souligné le rôle de GEE comme outil puissant pour la recherche environnementale et territoriale, en offrant un accès rapide et global aux données géospatiales et en facilitant leur traitement à grande échelle.

TP 1 : Analyse spatiale du paludisme et de la population au Cameroun

Ce projet a constitué un travail pratique d'initiation à Google Earth Engine (GEE) avec JavaScript. Il a visé à importer, manipuler et visualiser des données rasters et vectorielles pour analyser la distribution spatiale et temporelle de phénomènes (paludisme et population). L'étude a été réalisée sur le Cameroun.

Membres du groupe :

- AGNANGMA SANAM David Landry
- DIOP Astou
- DIOP Marème
- NGAKE YAMAHA Herman Parfait

Les données utilisées pour ce premier TP sont les suivants :

- **Les shapefiles des limites administratives du Cameroun** : téléchargés via le site GADM, en sélectionnant le pays souhaité et en récupérant le fichier zippé au format *shapefile*.
- **Les données sur le paludisme** : données chronologiques (2000-2021) disponibles sur le site du MAP (Malaria Atlas Project). L'interface propose une carte interactive avec un panneau *Layers* permettant de sélectionner un indicateur, une période, une zone géographique (pays ou subdivisions), et de télécharger les données au format souhaité (GeoTIFF, CSV, etc.).

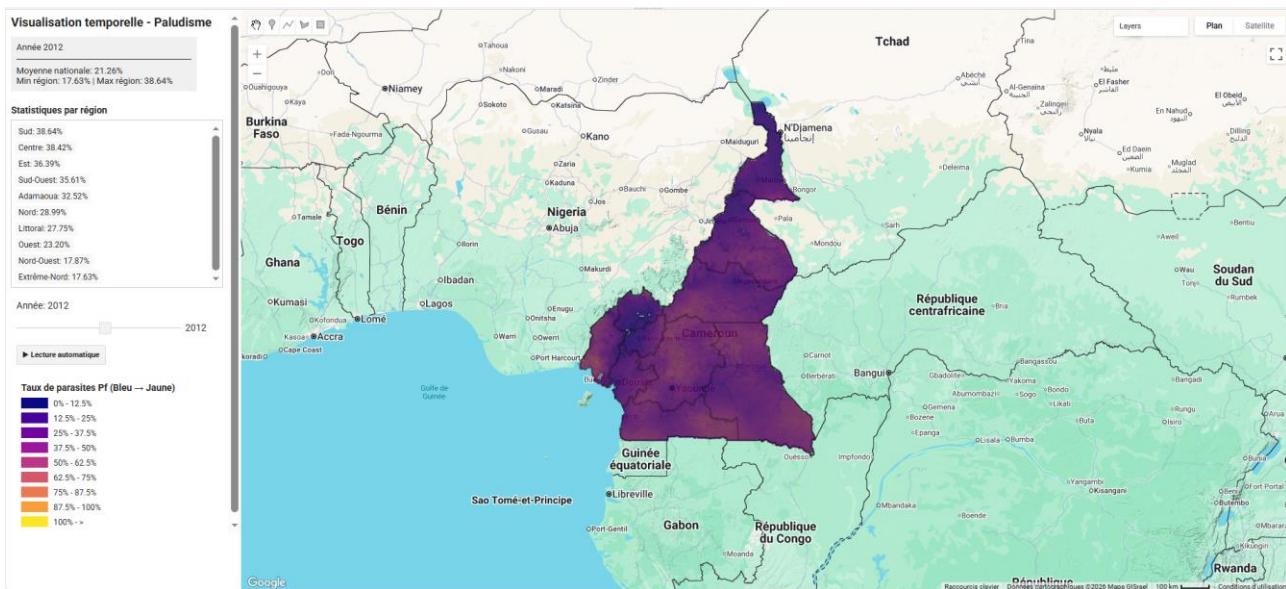
Pour ce TP, nous avons utilisé les bibliothèques *Rasterio*, *Geopandas*, *Matplotlib* et *Folium* afin de charger, manipuler et visualiser les données géospatiales (shapefiles *GADM*, rasters *WorldPop* et *Malaria Atlas Project*). Après prévisualisation, des graphiques et cartes interactives ont été générés pour illustrer la distribution spatiale des indicateurs.

Figure 1 : Interface de la carte interactive des subdivisions administratives du Cameroun



Source : GADM.

Figure 2 : Interface de la carte interactive de l'évolution du taux d'incidence du paludisme au Cameroun



Source : GADM, MAP.

Difficultés rencontrées : La principale difficulté rencontrée a été relative à l'export des données sous les différents formats, notamment HTML pour les cartes interactives depuis GEE. C'est ce qui a expliqué d'ailleurs le fait que le dossier "outputs" de l'arborescence du projet soit vide. Il est également important de préciser ici qu'il n'est pas possible avec GEE de rendre les chemins d'accès dynamiques pour des utilisations ultérieures depuis d'autres appareils. L'utilisateur doit donc importer les données disponibles dans le dossier "Data" sur GEE puis modifier les chemins d'accès manuellement.

TP 2 : Analyse des données géospatiales portant sur les infrastructures, les aires protégées et la population du Cameroun

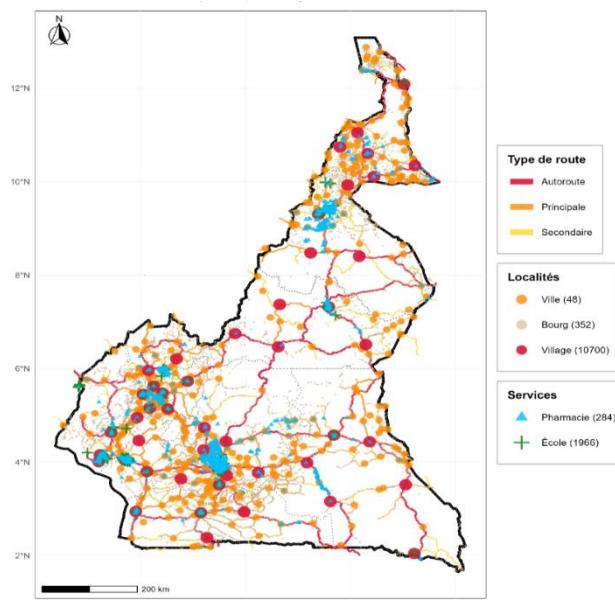
L'objectif principal de ce TP était de générer des visualisations thématiques, statiques et interactives sur les infrastructures sociales, les localités, les aires protégées, la population, les réseaux d'eau et les réseaux de transport. Les données proviennent essentiellement de *OpenStreetMap (OSM)*, *World Database on Protected Areas (WDPA)* et de *Protected Planet*, avec une utilisation complémentaire des shapefiles *GADM* pour les limites administratives.

Membres du groupe :

- AGNANGMA SANAM David Landry
- DIOP Astou
- DIOP Marème
- NGAKE YAMAHA Herman Parfait

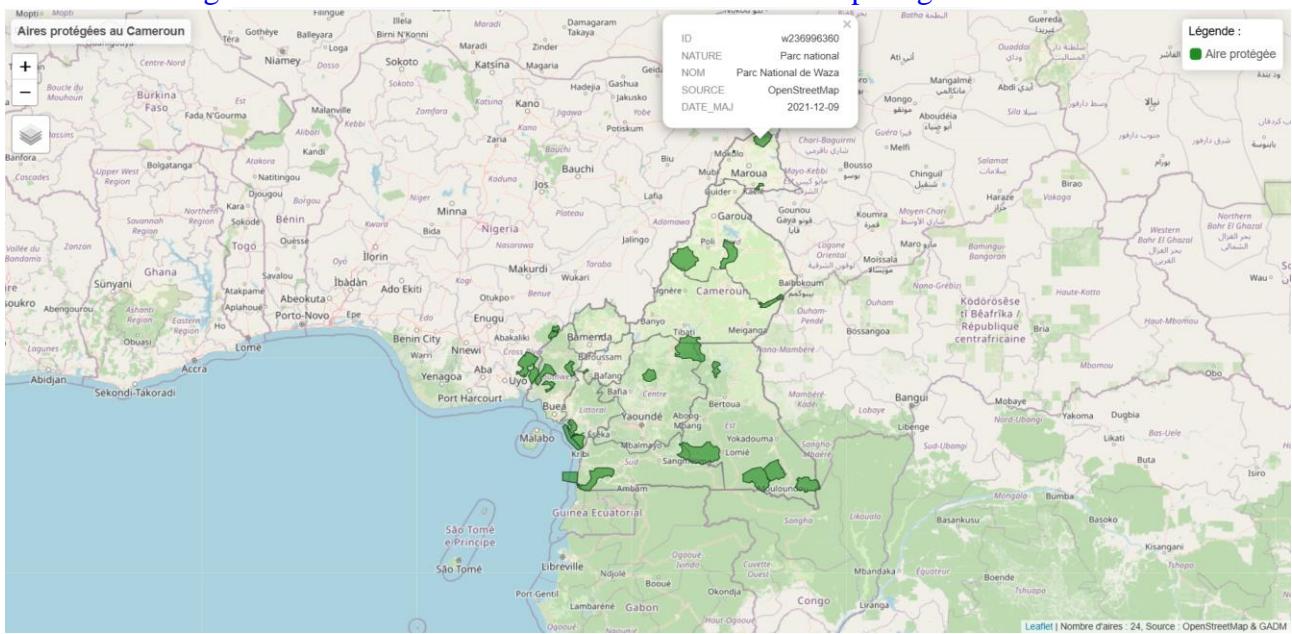
Le projet comprend a été réalisé à partie du logiciel R. Ainsi, il comprend plusieurs scripts R pour le traitement et la visualisation des données ainsi qu'une application interactive Shiny et un tableau de bord. Les analyses ont permis de visualiser la distribution des écoles, hôpitaux, grandes villes, etc. et de produire des cartes stratégiques pour la planification et la prise de décision.

Figure 3 : Vue d'ensemble de quelques infrastructures du Cameroun



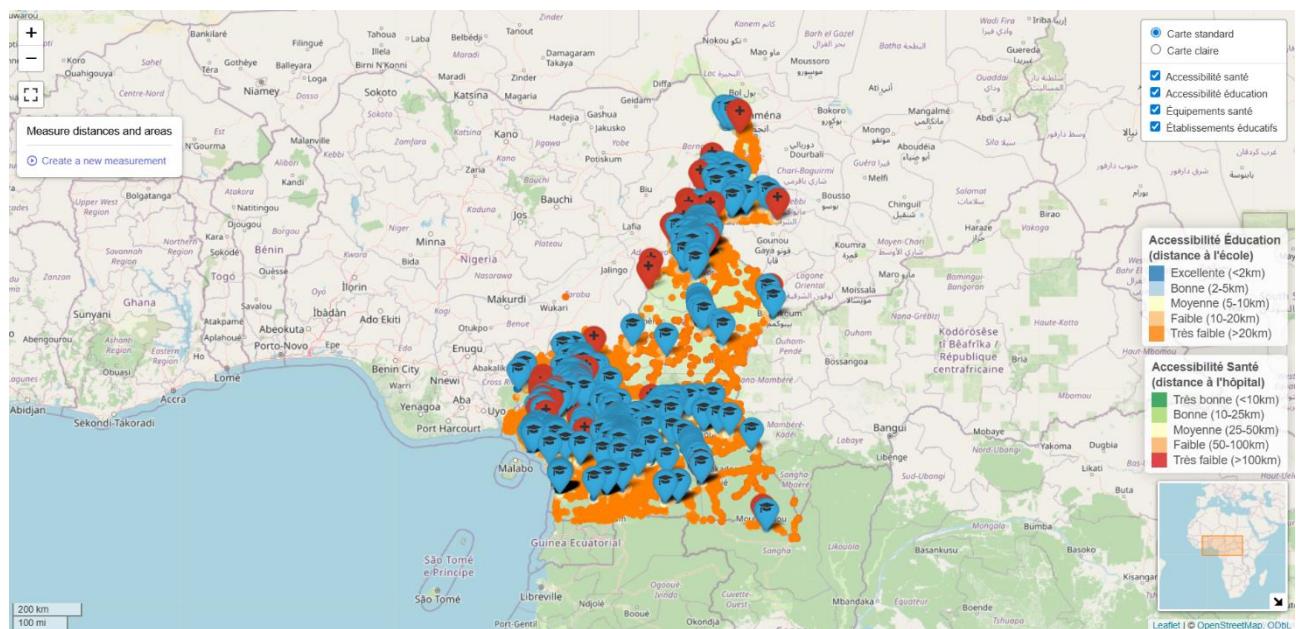
Source : GADM, OSM.

Figure 4 : Interface de la carte interactive des aires protégées du Cameroun



Source : GADM, OSM.

Figure 5 : Interface de la carte interactive de l'analyse spatiale de l'accès à l'éducation et à la santé au Cameroun



Source : GADM, OSM.

TP 3 : Analyse des données géospatiales portant sur les infrastructures, les aires protégées et la population du Cameroun (suite du TP2)

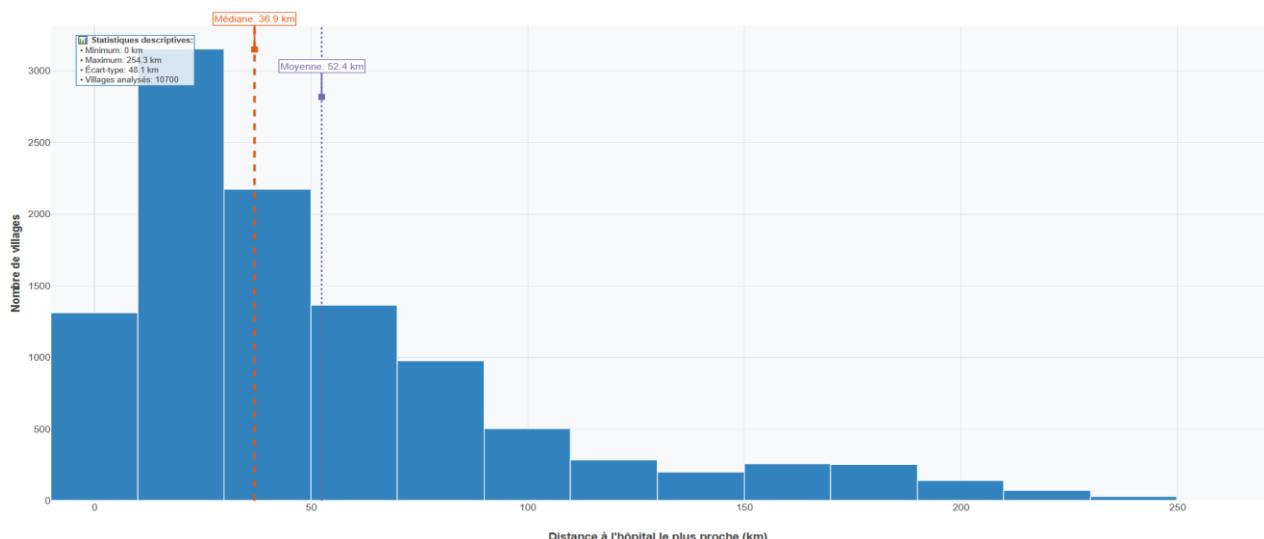
L'objectif principal était de générer des cartes thématiques et interactives pour créer des zones tampon autour des infrastructures sociales, des localités, du réseau ferroviaire et de visualiser la répartition de la population à l'intérieur des buffers et des aires protégées. Les données proviennent essentiellement de OpenStreetMap (OSM) pour les infrastructures, de Worldpop pour la population et de Protected Planet pour les aires protégées, avec une utilisation complémentaire des shapefiles GADM pour les limites administratives.

Membres du groupe :

- AGNANGMA SANAM David Landry
- DIOP Astou
- DIOP Marème
- NGAKE YAMAHA Herman Parfait

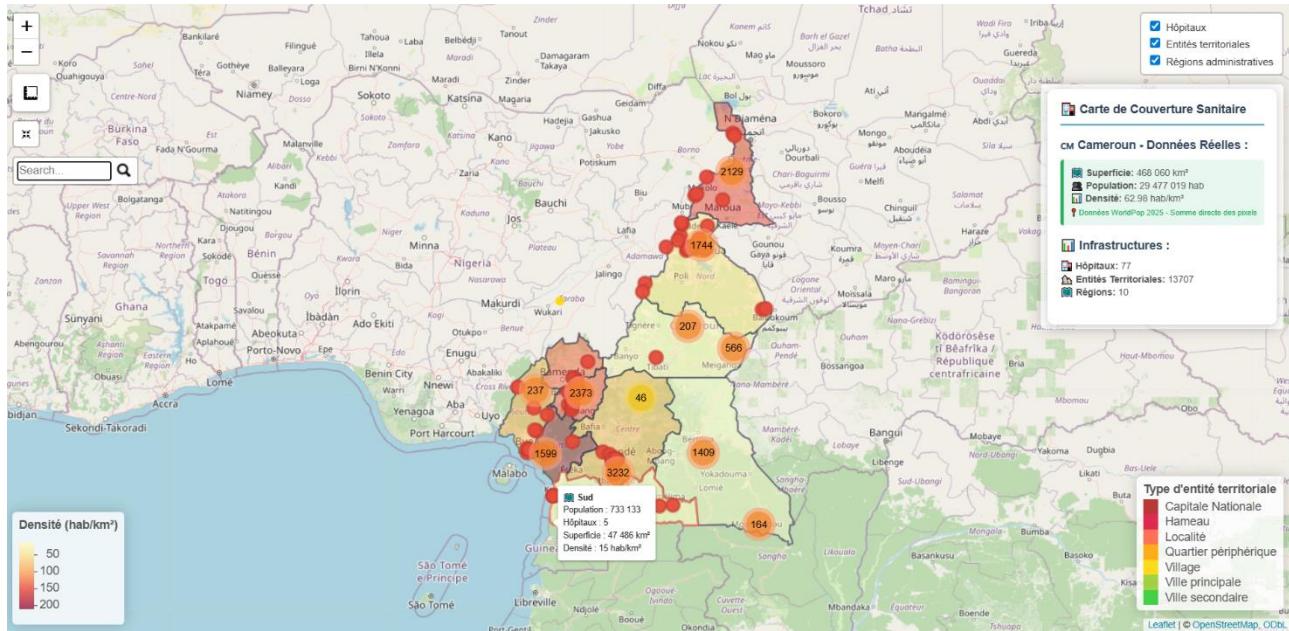
Le projet comprend plusieurs scripts R pour le traitement et la visualisation des données ainsi qu'une application interactive Shiny et un tableau de bord pour produire des cartes stratégiques pour la planification et la prise de décision.

Figure 6 : Distribution des distances villages-hôpitaux au Cameroun



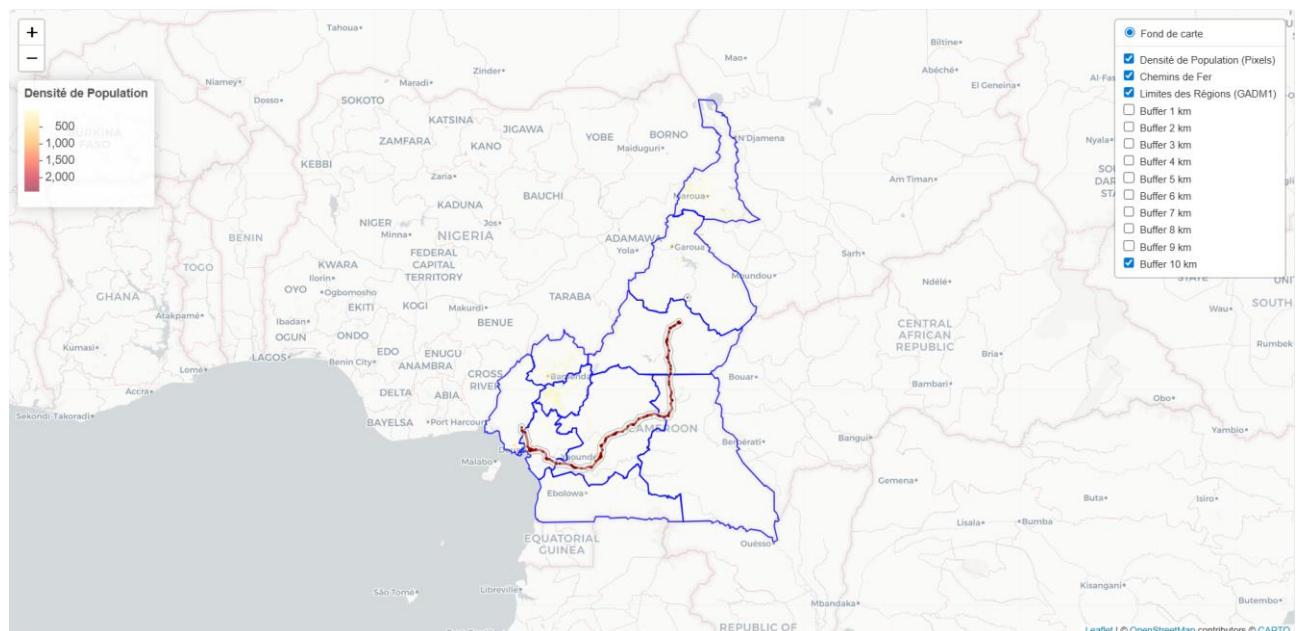
Source : GADM, OSM, calculs des auteurs.

Figure 7 : Interface de la carte interactive de couverture sanitaire au Cameroun en 2022.



Source : GADM, OSM, calculs des auteurs.

Figure 8 : Visualisation des chemins de fer du Cameroun avec buffer.



Source : GADM, OSM, calculs des auteurs.

TP 4 : Identification des terres arables au Burundi

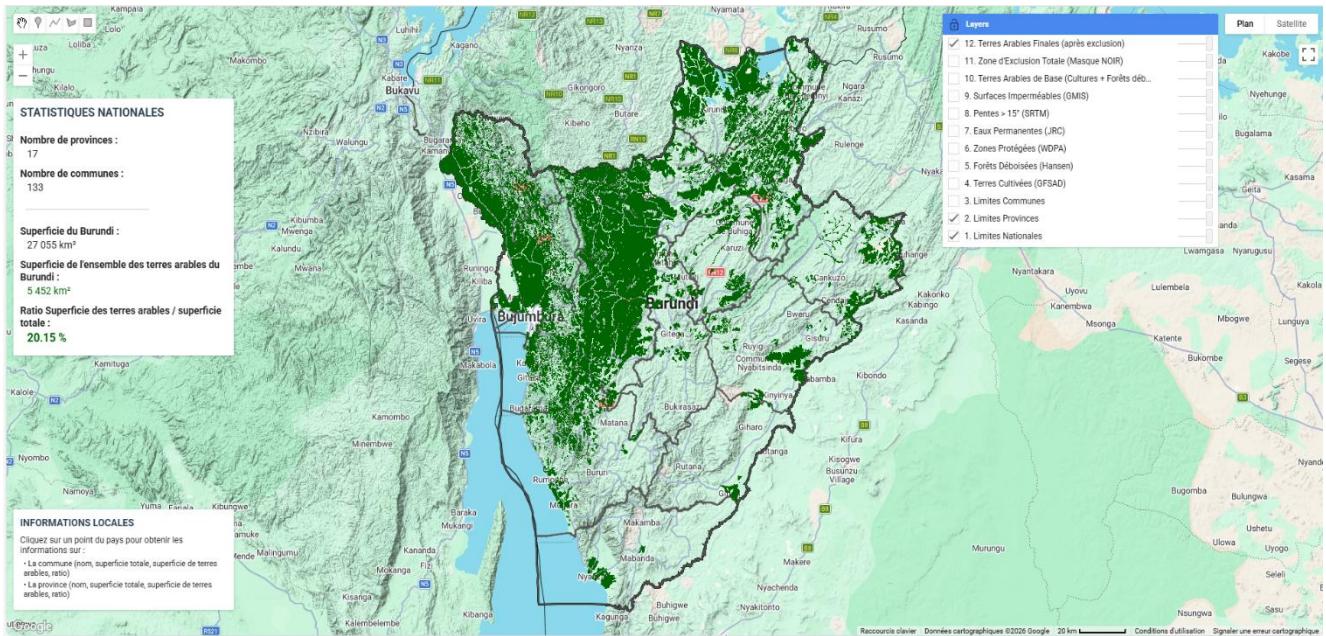
Ce projet a visé à réaliser une **analyse spatiale des terres arables au Burundi** afin d'estimer la superficie totale des zones réellement aptes à l'agriculture. L'approche consistait à identifier et cartographier les terres cultivables en excluant les espaces non exploitables ou soumis à des restrictions légales. L'ensemble du travail a été réalisé sur la plateforme **Google Earth Engine (GEE)** à l'aide du langage **JavaScript**, en mobilisant à la fois des données satellitaires (terres cultivées, couverture forestière, eaux permanentes, surfaces imperméables et pentes raides) et des données vectorielles (limites administratives et zones protégées). Les résultats produits ont permis de générer des statistiques précises par niveau administratif, aussi bien à l'échelle des **provinces (ADMIN1)** qu'à celle des **communes (ADMIN2)**.

Membres du groupe :

- Herman Parfait NGAKE YAMAHA
- Joe Young Véridique Gabriel DIOP
- David Landry AGNANGMA SANAM
- Cheikh Ahmadou Bamba FALL
- Gérald Guerngue ADDJITA

Les données mobilisées pour ce projet proviennent de plusieurs sources complémentaires. Les **shapefiles GADM** définissent les limites administratives du Burundi et servent de base pour l'agrégation statistique. Les terres cultivées ont été identifiées grâce au jeu de données **GFSAD** (2015), tandis que les pertes de forêts entre 2000 et 2015 ont été issues du produit **Hansen Global Forest Change (HGFC)**, permettant de caractériser la base arable. Les **zones protégées** ont été intégrées via la base **WDPA**, constituant un masque légal d'exclusion. Les **eaux permanentes** ont été extraites du jeu de données **JRC Global Surface Water**, utilisé comme masque physique (occurrence > 75 %). La **topographie** a été prise en compte grâce au modèle **SRTM**, afin d'exclure les zones de forte pente (> 15°). Enfin, les **surfaces imperméables** ont été identifiées par l'asset privé **GMIS Burundi**, permettant d'écartier les zones bâties et les routes (imperméabilité > 10 %).

Figure 9 : Interface de la carte interactive des surfaces arabes au Burundi



Source : GADM, GFSAD, HGFC, WPDA, JRC, SRTM, GMIS, calculs des auteurs.

TP 5 : Analyse du ratio LCRPGR en Tanzanie entre 2017 et 2022

Ce projet a visé à analyser l'indicateur ODD 11.3.1, également connu sous le nom de **Ratio de la Consommation des Terres par rapport au Taux de Croissance Démographique (LCRPGR)**. L'objectif principal fût d'évaluer l'efficacité de la gestion de l'étalement urbain en Tanzanie en comparant la vitesse à laquelle les surfaces bâties augmentent par rapport à la croissance démographique sur la période 2017-2022.

Membres du groupe :

- Cheikh THIOUB
- Marème DIOP
- Gerald ADDJITA
- David Landry AGNANGMA SANAM
- Herman Parfait NGAKE YAMAHA

Trois sources de données principales sont intégrées :

- **ESRI Global Land Cover** : Données d'occupation du sol à 10m de résolution pour identifier les zones bâties (classe 7).
- **WorldPop** : Rasters de population à 100m pour les années 2017 et 2022.
- **GADM** : Limites administratives des 31 régions de Tanzanie.

Le projet a été réalisé avec l'API Python GEE. Chaque source subit un prétraitement spécifique : mosaïquage des tuiles pour l'occupation du sol, correction des facteurs d'échelle pour la population, et harmonisation des systèmes de coordonnées (reprojection en UTM Zone 35S).

Trois indicateurs principaux ont été calculés pour chaque région et au niveau national :

- **LCR (Land Consumption Rate)** : taux annuel de croissance des surfaces bâties.
- **PGR (Population Growth Rate)** : taux annuel de croissance démographique.
- **LCRPGR (ODD 11.3.1)** : ratio entre LCR et PGR, utilisé pour évaluer la relation entre expansion spatiale et croissance démographique.

Interprétation :

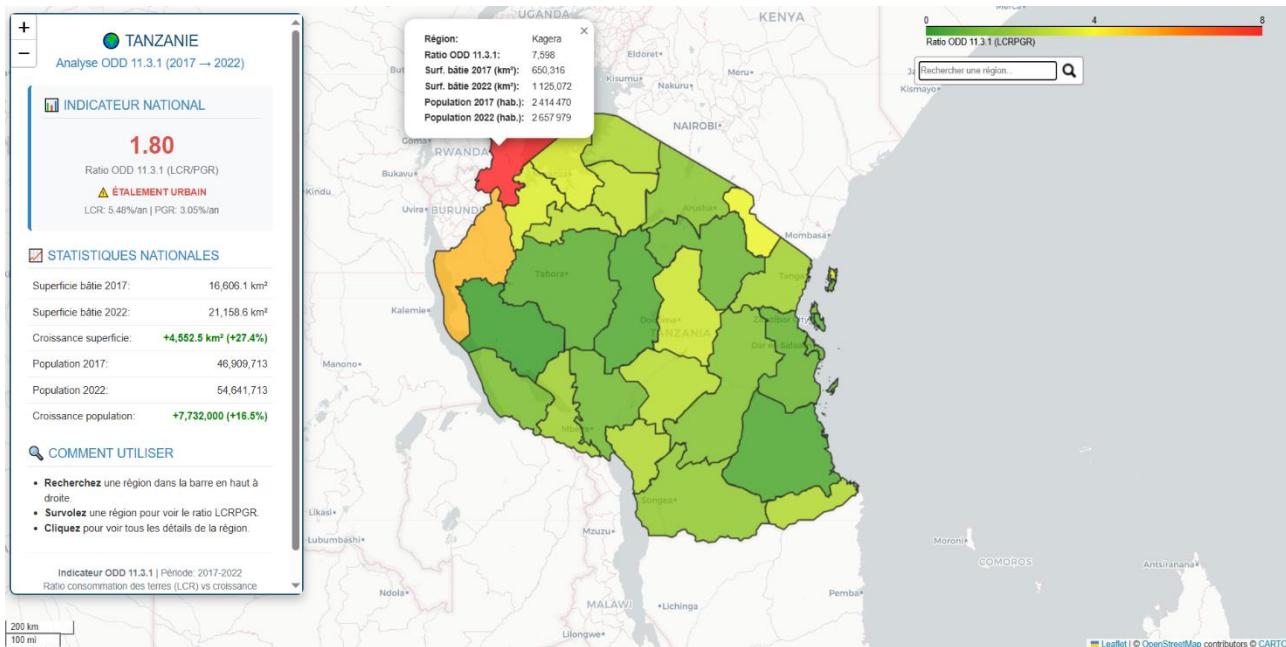
- **LCRPGR > 1** : étalement urbain (les terres sont consommées plus vite que la population ne croît).
 - **LCRPGR < 1** : densification (la population croît plus vite que l'expansion spatiale).
 - **LCRPGR = 1** : équilibre relatif entre croissance spatiale et démographique.

Des indicateurs secondaires ont été calculés pour compléter l'analyse : densité urbaine (hab/km²), superficie bâtie par habitant (m²/personne), et variation absolue des surfaces bâties.

Les aspects techniques :

- **Optimisations** : calculs réalisés à une résolution adaptée (50 m pour les surfaces bâties), gestion des erreurs (valeurs nulles, divisions par zéro), et agrégation parallèle région par région pour réduire la charge mémoire.
 - **Validation** : contrôles de qualité intégrés, incluant la vérification de la plausibilité démographique (comparaison avec les estimations officielles) et l'analyse de cohérence spatiale (déttection visuelle des valeurs aberrantes).

Figure 10 : Interface de la carte interactive de l'étalement urbain en Tanzanie



Source : GADM, WorldPop, ESRI Global Land Cover, calculs des auteurs.

TP 6 : Analyse des dynamiques environnementales et territoriales au Sénégal à partir des indices spectraux de Sentinel-2

Ce projet a visé à analyser plusieurs dimensions environnementales et territoriales du Sénégal à partir d'indices spectraux calculés à partir des images Sentinel-2, traitées avec Google Earth Engine via l'API Python. L'étude a été réalisée à l'échelle des 45 départements du Sénégal, pour l'année 2018. Les indicateurs produits ont vocation à être intégrés et combinés avec les données socio-économiques issues de l'enquête EHCVM, dans une perspective d'analyse conjointe environnement – conditions de vie.

Membres du groupe :

- Cheikh THIOUB
- Marème DIOP
- Gerald ADDJITA
- David Landry AGNANGMA SANAM
- Herman Parfait NGAKE YAMAHA

Objectifs spécifiques de l'étude :

- Exploiter la télédétection pour produire des indicateurs environnementaux robustes ;
- Analyser les disparités spatiales à l'échelle départementale ;
- Produire des cartes et indicateurs spatialisés exploitables ;
- Enrichir les bases EHCVM 2018 avec des variables environnementales.

Données communes au projet :

Imagerie satellitaire :

- Satellite : Sentinel-2 (Copernicus)
- Collection : COPERNICUS/S2_SR_HARMONIZED
- Niveau : réflectance de surface (SR)
- Période : 1er janvier – 31 décembre 2018
- Résolution spatiale : 10 à 20 mètres

- Prétraitement : correction atmosphérique et masquage des nuages

Limites administratives :

- Source : GADM
- Niveaux :
 - Niveau 0 : Sénégal
 - Niveau 2 : 45 départements
- Projection : WGS84 (EPSG:4326)

Données socio-économiques :

- Enquête EHCVM 2018
 - Base individu
 - Base welfare
- Variables clés : grappe, ménage, département

Le projet mobilise plusieurs indices spectraux issus des images Sentinel-2 pour analyser différentes dimensions environnementales et territoriales du Sénégal en 2018 à savoir :

- Le **NDTI** (Normalized Difference Turbidity Index) : qui permet d'évaluer la turbidité des eaux de surface en distinguant les zones à forte charge sédimentaire des eaux plus claires. Sa formule est :

$$\text{NDTI} = \frac{\text{Rouge}-\text{Vert}}{\text{Rouge}+\text{Vert}}$$

Où la bande rouge correspond à B4 (665 nm) et la bande verte à B3 (560 nm). Les eaux turbides présentent une réflectance plus élevée dans le rouge, ce qui permet de distinguer les zones à forte charge sédimentaire.

Les valeurs supérieures à 0,3 indiquent une eau très turbide, entre 0,1 et 0,3 une eau turbide, entre -0,1 et 0,1 une eau modérément claire, et entre -0,3 et -0,1 une eau claire. Les valeurs inférieures à -0,3 correspondent généralement à des surfaces non aquatiques ou végétalisées.

- Le **NDVI** (Normalized Difference Vegetation Index) sert à mesurer la vigueur de la végétation et la biomasse, apportant des informations essentielles sur les contrastes agroécologiques. Il est calculé à partir de la formule suivante :

$$\text{NDVI} = \frac{\text{B8}-\text{B4}}{\text{B8}+\text{B4}}$$

Où B8 est la bande proche infrarouge (NIR) et B4 la bande rouge. Les valeurs élevées (proche de 1) traduisent une végétation dense et vigoureuse, tandis que les valeurs faibles (proche de -1) indiquent des zones dégradées ou peu végétalisées. Cet indice est particulièrement pertinent dans le contexte sénégalais, où l'agriculture et le pastoralisme jouent un rôle important. Il permet de mettre en évidence les contrastes agroécologiques entre régions plus humides et zones plus arides.

- Le **NDBI** (Normalized Difference Built-up Index) : qui est utilisé pour identifier et quantifier l'urbanisation et les surfaces artificialisées. Il est calculé à partir de la formule suivante :

$$\text{NDBI} = \frac{B11-B8}{B11+B8}$$

Où B11 correspond au SWIR (infrarouge à ondes courtes) et B8 au NIR (proche infrarouge). Les valeurs positives indiquent généralement des zones bâties ou imperméabilisées, tandis que les valeurs négatives correspondent à des surfaces naturelles (végétation, eau). Cet indice permet de cartographier l'expansion urbaine et d'analyser les dynamiques territoriales liées à l'urbanisation.

- Le **BAI** (Burned Area Index) : qui permet de détecter et caractériser les zones brûlées. Il est donné par :

$$\text{BAI} = \frac{1}{(0.1-\text{RED})^2+(0.06-\text{NIR})^2}$$

Il repose sur la réflectance des bandes rouge et NIR, exploitant les contrastes spectraux entre végétation saine et sols nus brûlés. Les valeurs du BAI sont interprétées ainsi : inférieures à 10, absence de brûlure ; entre 10 et 50, végétation sèche ; entre 50 et 150, zones probablement brûlées ; supérieures à 150, zones brûlées confirmées. Cet indice est particulièrement utile pour analyser l'impact des feux de brousse sur les écosystèmes et les activités agricoles.

- Le **BSI** (Bare Soil Index) : qui sert à repérer les sols nus et les surfaces minérales, en les distinguant de la végétation et des plans d'eau. Il est donné par :

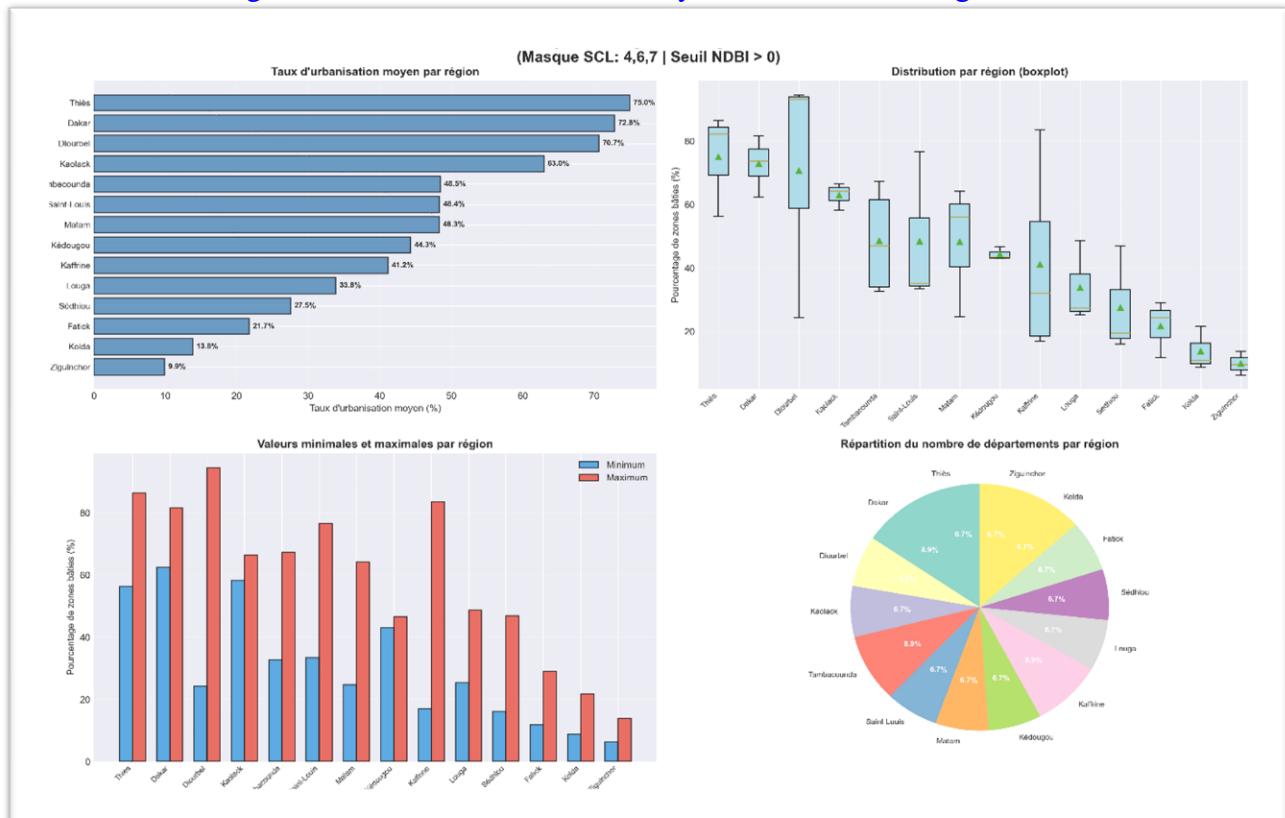
$$\text{BSI} = \frac{(SWIR+\text{Rouge})-(NIR+\text{Bleu})}{(SWIR+\text{Rouge})+(NIR+\text{Bleu})}$$

Où le SWIR correspond à la bande B11, le rouge à B4, le NIR à B8 et le bleu à B2. Les sols nus présentent une réflectance élevée dans le rouge et le SWIR, et plus faible dans le bleu et le NIR. Les valeurs élevées du BSI (proche de 1) traduisent des zones sèches et sableuses, tandis que les valeurs

faibles (proche de -1) correspondent à des zones plus humides et végétalisées. Cet indice est particulièrement pertinent pour analyser la répartition spatiale des sols nus au Sénégal, notamment dans les régions nord et est du pays.

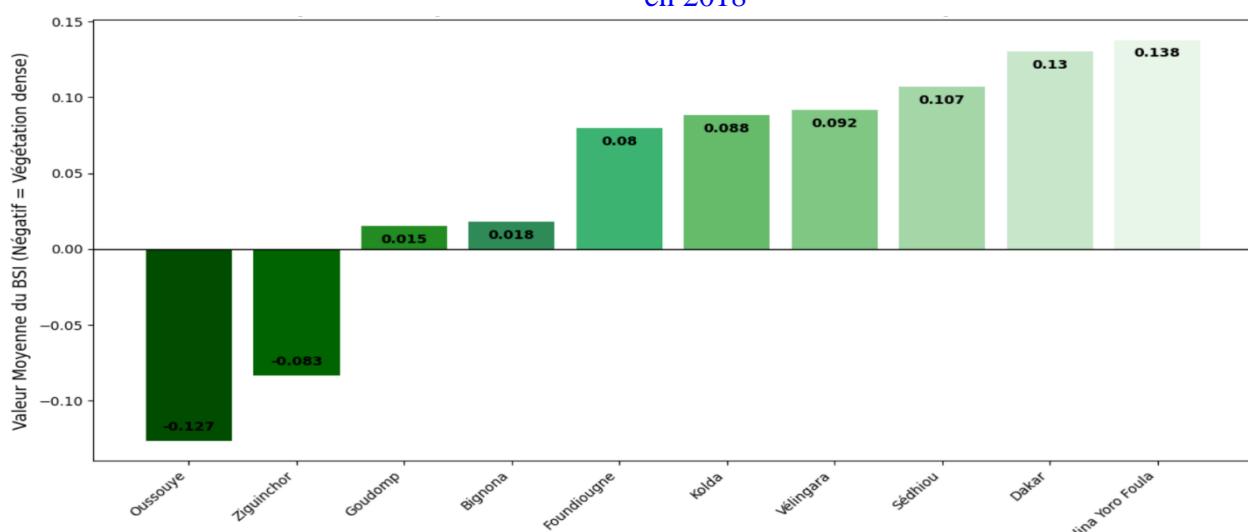
Quelques outputs obtenus :

Figure 11 : Résultats issus de l'analyse du NDBI au Sénégal en 2018



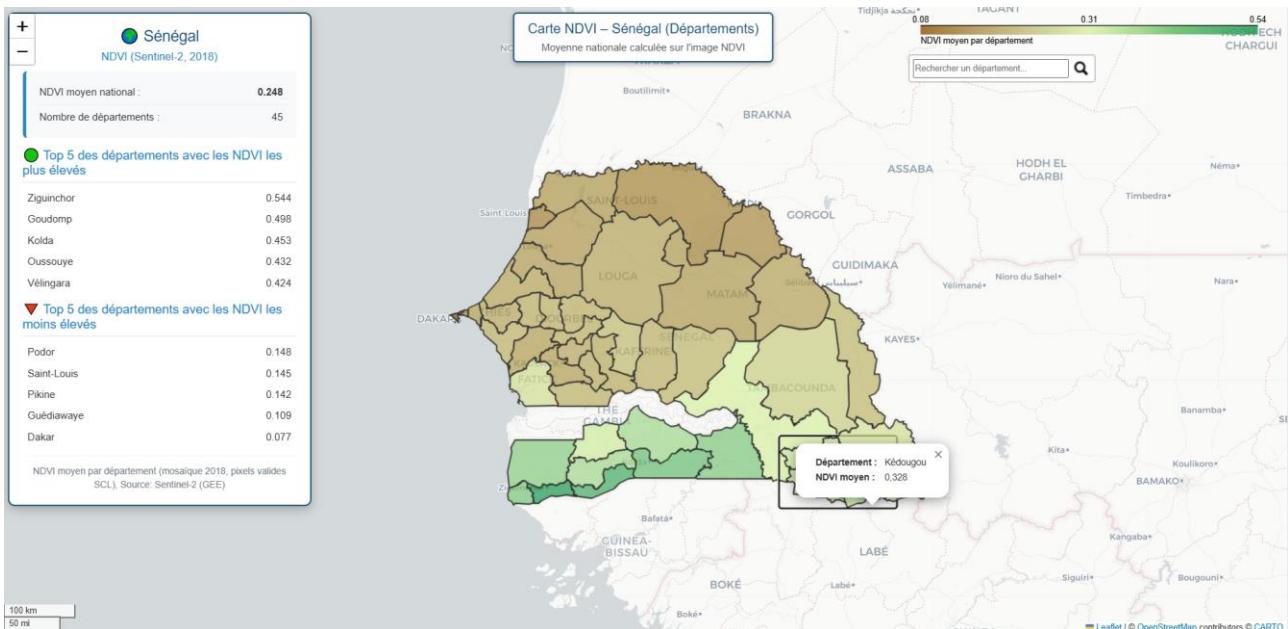
Source : COPERNICUS/S2_SR_HARMONIZED, calculs des auteurs.

Figure 12 : Top 10 des départements avec les indices de nudité (BSI) les plus faibles au Sénégal en 2018



Source : COPERNICUS/S2_SR_HARMONIZED, calculs des auteurs.

Figure 13 : Interface de la carte interactive de visualisation du NDVI dans les départements du Sénégal en 2018



Source : COPERNICUS/S2_SR_HARMONIZED, calculs des auteurs.

Limites : Malgré l'intérêt des résultats obtenus, ce travail présente plusieurs limites qu'il convient de souligner pour une interprétation rigoureuse des indicateurs. Dans un premier temps, les images Sentinel-2, bien que de haute résolution spatiale, restent sensibles à la couverture nuageuse, particulièrement durant la saison des pluies. Malgré les procédures de masquage, certains pixels résiduels peuvent affecter les moyennes annuelles des indices, en particulier pour les zones humides et forestières. Dans un second temps, les indices spectraux utilisés reposent sur des signatures spectrales simplifiées et peuvent présenter des confusions entre certaines classes. Par exemple, les surfaces urbaines peuvent être confondues avec des sols nus pour le BSI, ou certaines zones sèches avec des zones brûlées pour le BAI. L'utilisation d'un indice isolé ne permet donc pas une discrimination parfaite des types de surfaces.

TP 7 (individuel) : Analyse des matrices de voisinage et de l'autocorrélation spatiale : Application de l'indice de Moran global avec R

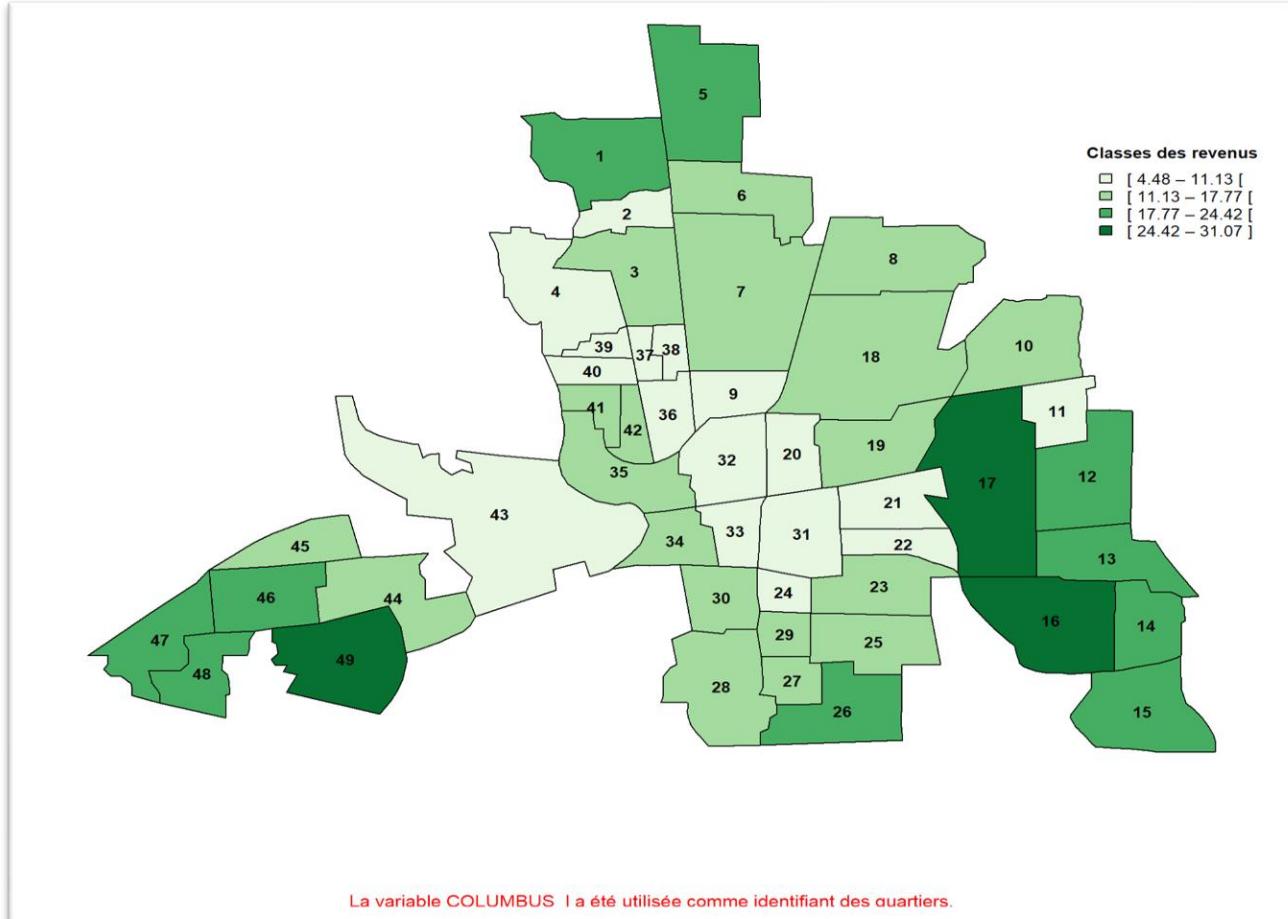
Ce travail pratique a visé à analyser les relations de voisinage entre des entités géographiques ainsi que l'autocorrélation spatiale des données. Réalisé en langage **R**, il s'est articulé autour de trois volets principaux :

- Dans une première étape, les données brutes sont importées et transformées en objets spatiaux afin de produire des cartes thématiques, notamment sur les revenus moyens des ménages de Columbus (aux USA) en 1980.
- La deuxième partie a été consacrée à la construction et à la comparaison de différentes structures de voisinage (contiguïté, k plus proches voisins et seuil de distance), permettant d'explorer ainsi la densité et la cohérence des différents réseaux spatiaux.
- Enfin, la troisième partie aborde l'indice de **Moran global**, utilisé pour mesurer l'autocorrélation spatiale. L'exercice a mis en évidence l'équivalence entre la pente de la régression de WX sur X et l'indice de Moran global lorsque la matrice de poids est normalisée.

Ce TP a mobilisé plusieurs packages R spécialisés (*spData*, *spDataLarge*, *sp*, *sf*, *spdep*, *classInt*, *here*) et illustre de manière pratique les concepts fondamentaux de la statistique spatiale, en combinant visualisation cartographique, modélisation des voisinages et analyse mathématique de l'autocorrélation.

Quelques outputs :

Figure 14 : Classification des quartiers de Columbus suivant le revenu annuel moyen des ménages (en milliers de dollars US)



Source : GeoDa Center for Geospatial Analysis and Computation (Recensement de 1980, ville de Columbus, Etat de l'OHIO, USA), calculs des auteurs.

Figure 15 : Structure de voisinage avec contiguïté Bishop (3*3)

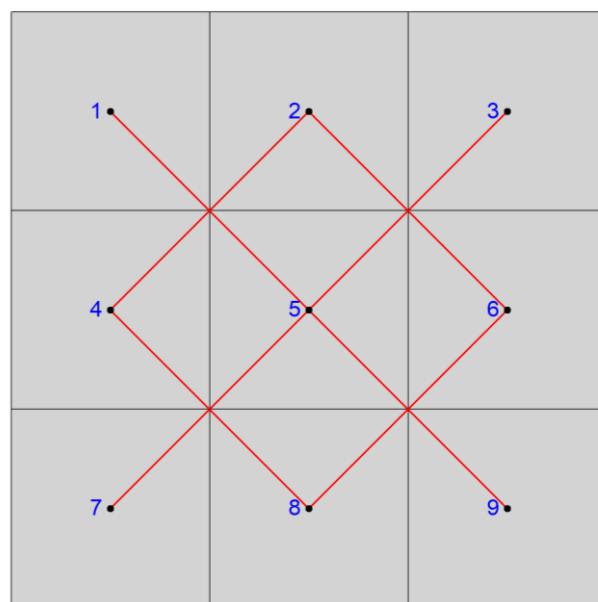
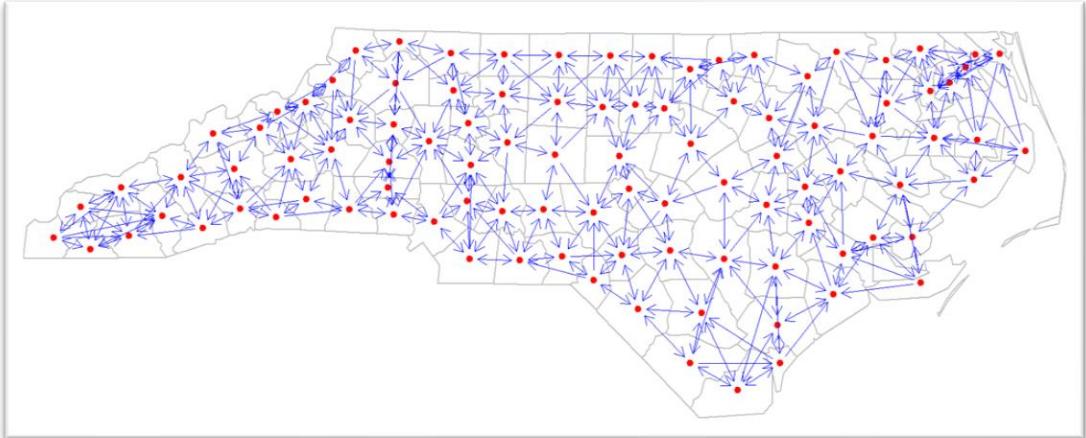
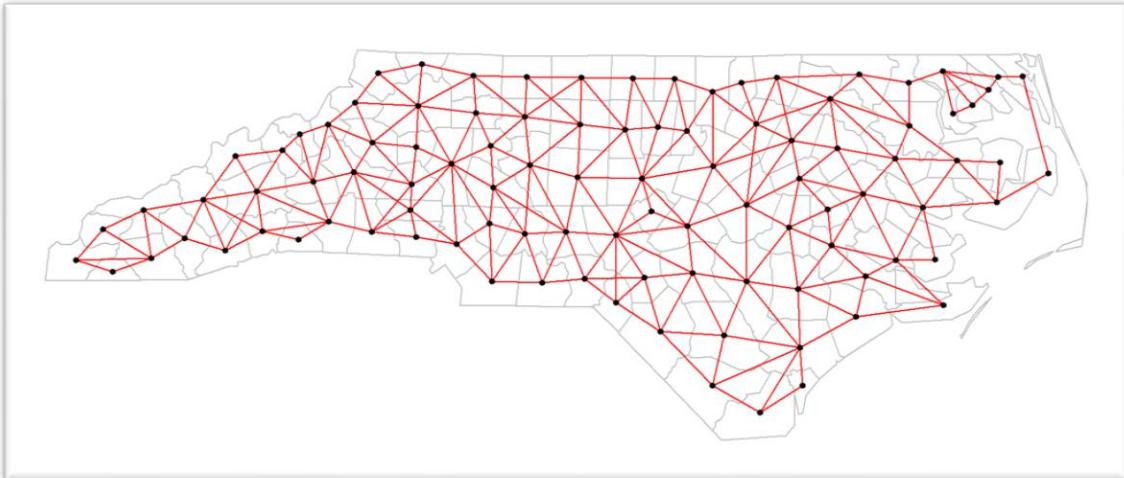


Figure 16 : Illustration d'un réseau de voisinage KNN (K=5)



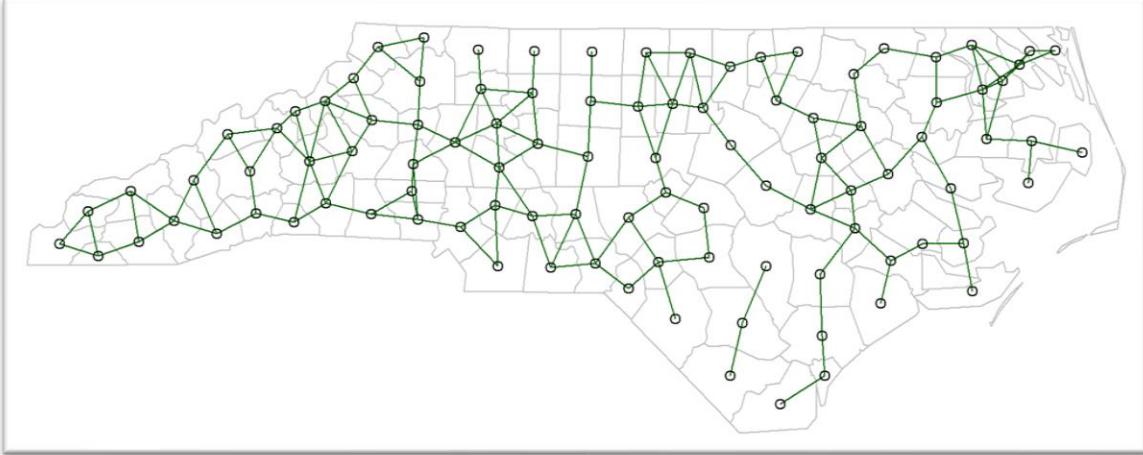
Source : GeoDa Center for Geospatial Analysis and Computation (Shapefile SIDS, Caroline du Nord, 1974-1984).

Figure 17 : Illustration d'un réseau de voisinage par contiguïté Queen



Source : GeoDa Center for Geospatial Analysis and Computation (Shapefile SIDS, Caroline du Nord, 1974-1984).

Figure 18 : Illustration d'un réseau de voisinage par distance seuil (seuil = 0.42 degré)



Source : GeoDa Center for Geospatial Analysis and Computation (Shapefile SIDS, Caroline du Nord, 1974-1984).

Conclusion

Au terme de ces 30 heures de formation et de sept travaux pratiques riches et variés, ce parcours en statistiques exploratoires spatiales a permis de développer des compétences solides dans la manipulation et l'analyse de données géospatiales. L'utilisation de multiples environnements de programmation (Google Earth Engine via JavaScript et l'API Python, R et Python classique) constitue un acquis technique précieux, tandis que l'exploration de problématiques concrètes – du paludisme au Cameroun à l'étalement urbain en Tanzanie, en passant par l'identification des terres arables au Burundi et l'analyse environnementale du Sénégal – a révélé toute la portée opérationnelle de ces méthodes pour éclairer les enjeux de santé publique, d'aménagement du territoire ou encore de gestion environnementale en Afrique. Les défis techniques rencontrés (gestion des projections, prétraitement d'images satellitaires, optimisation des calculs, création d'applications interactives, etc.) ont constitué autant d'opportunités d'apprentissage approfondi, renforcées par le travail collaboratif au sein des groupes. Nous tenons ainsi à exprimer notre profonde gratitude à Monsieur Aboubacar HEMA pour sa pédagogie et ses précieux conseils, ainsi qu'à nos camarades de promotion pour leur engagement et leur esprit d'équipe. Ces compétences, ancrées dans des problématiques réelles du continent africain, constituent désormais une base solide pour contribuer, dans nos futures carrières professionnelles ou académiques, à la résolution des grands défis territoriaux, environnementaux et sanitaires auxquels notre continent est confronté.