

République du Sénégal



Un Peuple - un But - une Foi

MINISTÈRE DE L'ÉCONOMIE, DU PLAN ET DE LA
COOPÉRATION (MEPC)

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



ÉCOLE NATIONALE DE LA STATISTIQUE ET DE L'ANALYSE
ÉCONOMIQUE PIERRE NDIAYE (ENSAE)



PROJET FINAL

STATISTIQUE EXPLORATOIRE SPATIALES : RÉSUMÉ DES TRAVAUX PRATIQUES

Rédigé par :

Kadidja **GUEBEDIANG A NKEN**,

*Élève Ingénieure Statisticienne Économiste en troisième année de
formation du cycle long / ENSAE*

Examiné par :

M. Aboubacar **HEMA**,

Ingénieur Statisticien Économiste,

Research Analyst, Data Scientist ,

Enseignant de Statistique exploratoire spatiale/ ENSAE

Année académique 2025/2026

**RÉSUMÉ DES PROJETS DE STATISTIQUE EXPLORATOIRE SPATIALE DE L'ANNÉE
ACADÉMIQUE 2025-2026**

Document rédigé par :

Kadidja GUEBEDIAN A NKEN

Élève Ingénieure Statisticienne Économiste (ISE1-Cycle Long)

Sous la supervision de :

M Aboubacar HEMA, Ingénieur Statisticien Économiste, Research Analyst, Data Scientist

Année académique 2025/2026

REMERCIEMENTS

Tout d'abord, nous tenons à exprimer notre profonde gratitude à toutes les personnes qui ont contribué, de près ou de loin, à la réalisation de ce travail.

Nous adressons nos sincères remerciements à tout le corps administratif et professoral de l'École nationale de la Statistique et de l'Analyse économique Pierre Ndiaye (ENSAE) pour la qualité de l'enseignement reçu, en particulier à Monsieur **Idrissa DIAGNE**, Directeur de l'ENSAE, ainsi qu'à Monsieur **Mamadou BALDE**, Directeur des études et des stages, Monsieur **Ibrahim BASSE**, chef du bureau des examens et concours, et Monsieur **Thierno BARRY**, responsable de la filière Ingénieur Statisticien Economiste, pour leur accompagnement et leur appui administratif. Nous remercions également tous nos Camarades pour leur collaboration et leur solidarité particulière dont ils ont fait preuve durant la réalisation de ces Travaux pratiques. Enfin, nous remercions M. **Aboubacar HEMA**, enseignant de Statistique Exploratoire Spatiale, pour la richesse du cours dispensé, sa disponibilité ainsi que la clarté de ses explications, qui ont grandement contribué à notre compréhension des outils et méthodes abordés.

SOMMAIRE

REMERCIEMENTS	I
SOMMAIRE	II
LISTE DES IMAGES	III
INTRODUCTION.....	4
PRESENTATION DE GOOGLE EARTH ENGINE	5
ANALYSE SPATIALE DU PALUDISME AU SENEGAL (2000-2024).....	8
ANALYSE DES DISTANCES ENTRE INFRASTRUCTURES ET AIRES PROTEGEES AU SENEGAL	11
ANALYSE DES DONNEES GEOSPATIALES PORTANT SUR L'ACCES AUX INFRASTRUCTURES AU SENEGAL	13
IDENTIFICATION DES TERRES ARABLES EN OUGANDA.....	17
ANALYSE SPATIALE DU DEVELOPPEMENT URBAIN AU RWANDA	22
ANALYSE DES INDICES SPECTRAUX ET FUSION AVEC LES DONNEES EHCVM DU MALI.....	27
MATRICE DE VOISINAGE ET INDICE DE MORAN	32
CONCLUSION	33
TABLE DES MATIERES	34

LISTE DES IMAGES

Image 1 : Interface d'accueil de Google Earth Engine	5
Image 2 : Aperçu du catalogue de Google Earth Engine.....	6
Image 3 : Editeur de Code Google Earth Engine.....	6
Image 4 : comparaison de l'incidence suivant l'année	10
Image 5 : Evolution de l'incidence nationale moyenne annuelle.....	10
Image 6 : Tableau de Bord représentant les infrastructures au Sénégal	15
Image 7 : statistiques obtenues en cliquant sur la carte	16
Image 8 : Pourcentage des terres arables cultivées et non cultivées.....	20
Image 9 : Répartition des terres selon le type.....	21
Image 10 : Présentation de la sortie excel des résultats.....	24
Image 11 : présentation de LCRPGR par région.....	25
Image 12 : interface Dashboard de visualisation de LCRPGR au Rwanda.....	26
Image 13 : Température sur la surface le jour (à gauche) et la nuit (à droite).....	30
Image 14 : représentation spatiale du NDWI (à gauche) et du NDVI (à droite)	30
Image 15 : Sortie excel des résultats.....	31
Image 16 : Représentation des villes de Caroline du Nord selon les différentes matrices de voisinage	32

INTRODUCTION

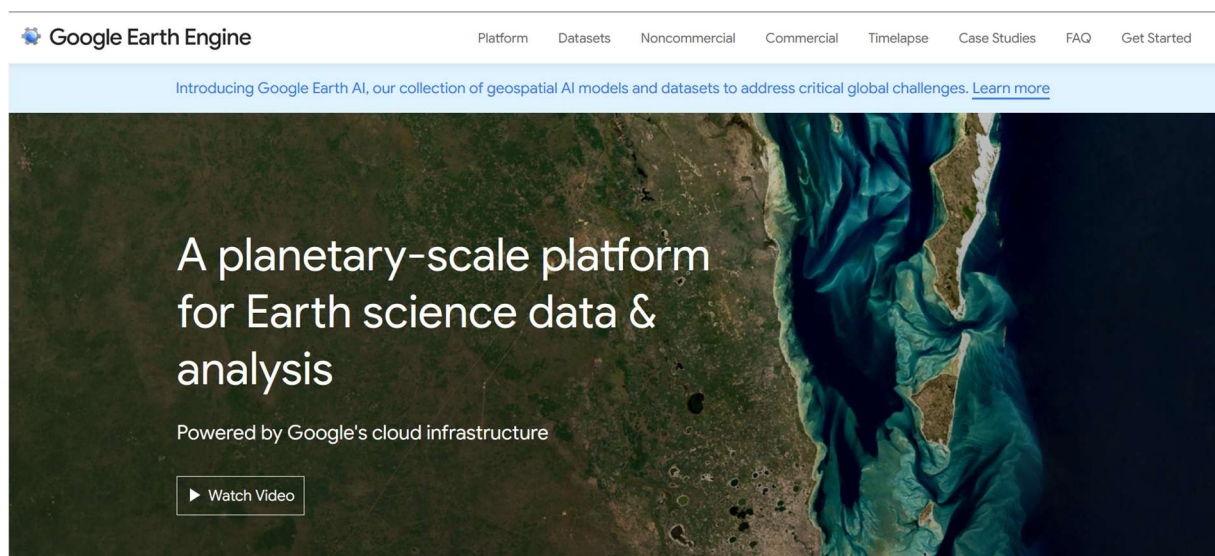
La statistique exploratoire spatiale (SES) est une branche de la statistique ayant pour but de faire des analyses statistiques d'une zone donnée et une représentation cartographique des résultats obtenus. Elle nous a été dispensée comme unité d'enseignement cette année à l'Ecole nationale de la Statistique et de l'Analyse économique Pierre Ndiaye de Dakar. Elle a demandé la mobilisation des connaissances en programmation notamment pour l'utilisation des logiciels R, Python, Google Earth Engine Javascript et Google Earth Engine Python. Le cours s'est étendu sur 30 heures, réparties en séances de 4 heures entre autres. Durant la plupart des séances, un travail pratique (TP) était donné afin de mieux assimiler les connaissances dispensées durant la séance. Les TP se faisaient en groupe, et une séance était à chaque fois dédiée à un groupe devant présenter les résultats de ses recherches. Les membres des groupes et les logiciels utilisés étaient généralement permutés à chaque TP afin de donner l'occasion à chacun de travailler avec les autres camarades ainsi qu'à utiliser les différents logiciels présentés plus haut. De plus, les zones géographiques étudiées par chaque groupe étaient différentes d'un groupe à un autre ce qui permettait non seulement d'approfondir notre culture mais également de faire des comparaisons entre pays. Cette année, sept TP ont été donnés et corrigés durant l'ensemble des 30 heures et l'objectif de ce document est de faire un résumé de chacun de ces travaux pratiques.

PRESENTATION DE GOOGLE EARTH ENGINE

Google Earth Engine (GEE) était le logiciel que nous avons eu à découvrir et explorer durant le cours SES. Etant un nouveau logiciel pour la plupart d'entre nous, il nous a été demandé de faire une brève présentation de son interface ainsi que de ses fonctionnalités. Notre groupe était constitué de Gerald ADDJITA, Michel TEVOEDJRE et Cheikh THIOUB.

L'application du cours de SES demandait l'utilisation de plusieurs langages de programmation et plusieurs logiciels parmi lesquels un logiciel en particulier nous a été présenté : Google Earth Engine (GEE). La particularité de ce logiciel était le fait que ce soit la première fois que nous l'utilisions depuis l'entame de notre parcours académique à l'ENSAE. Les élèves ont été réunis en trois groupe de 04 et un groupe de 05 afin de réaliser chacun cette tâche. Durant ce travail nous avons eu à découvrir ce qu'est GEE, son interface, et les différentes fonctionnalités.

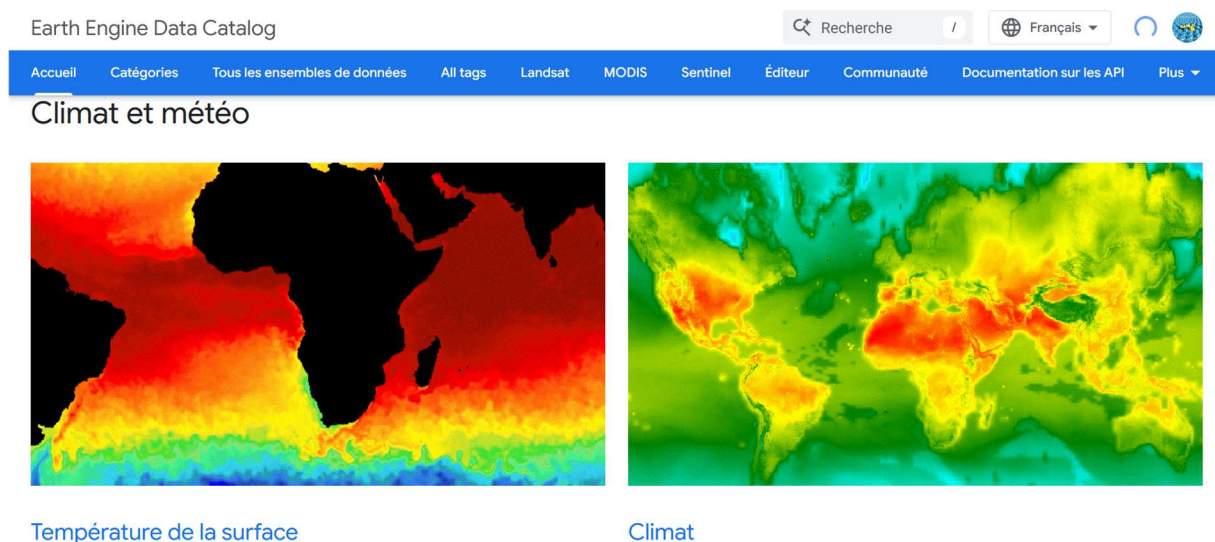
Image 1 : Interface d'accueil de Google Earth Engine



Source : Site de Google Earth Engine

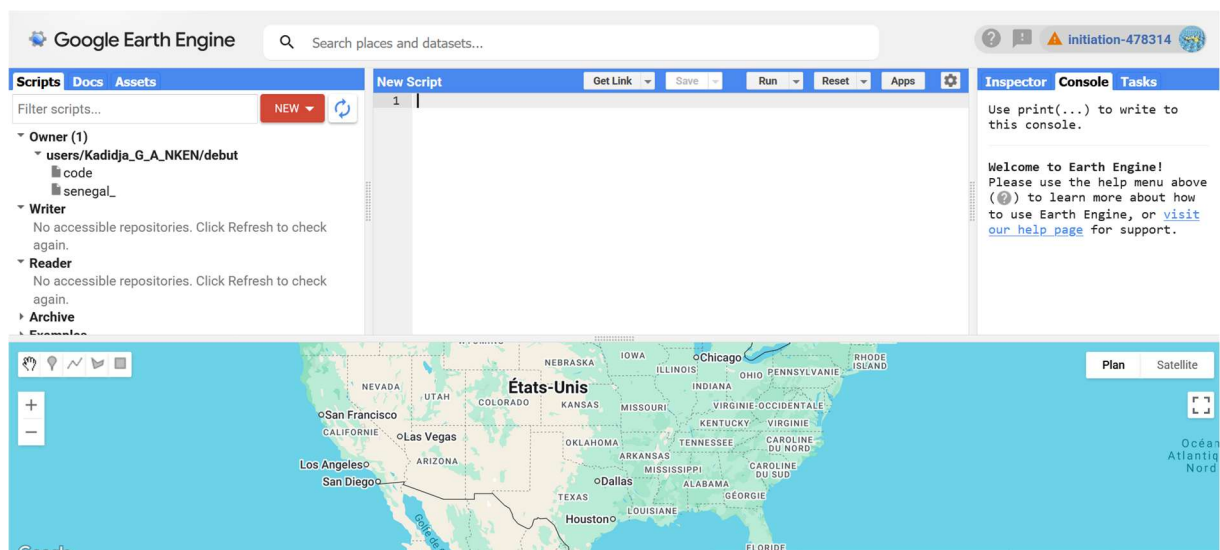
En termes d'outil, nous avons découvert que GEE permettait de faire des représentations planétaires facilitant les analyses spatiales sur certains aspects de la terre. Avec son catalogue de données mis à jour au quotidien, il facilite l'accès à certaines données notamment ceux sur la température de MODIS et Landsat, les données sur le climat de GridMET, ou encore les données atmosphériques de la NASA. De plus, GEE peut être utilisée avec le langage Javascript (au niveau de l'éditeur de code de GEE) ou avec Python en liant votre compte GEE à l'éditeur de code que vous utilisez grâce aux API de GEE.

Image 2 : Aperçu du catalogue de Google Earth Engine



Source : Site de Google Earth Engine

Image 3 : Editeur de Code Google Earth Engine



Source : Site de Google Earth Engine

A partir d'un PowerPoint/ Canva, chaque groupe a donc eu la possibilité de faire un document présentant ces différents aspects de GEE.

Grâce à cette expérience, nous avons eu l'opportunité de manier pour la première l'Outil GEE, découvrir son interface et découvrir les analyses qu'il nous permet de réaliser.

ANALYSE SPATIALE DU PALUDISME AU SENEGAL (2000-2024)

L'analyse spatiale du paludisme ici était restreinte à l'évolution du taux d'incidence sur 25 ans. Le pays étudié dans notre groupe était le Sénégal et le logiciel était R. Les membres du groupe n'avaient pas été permutés pour ce travail.

I. OBJECTIF DU TRAVAIL

L'objectif de ce travail était de faire une analyse spatiale de l'évolution spatio-temporelle du taux d'incidence du paludisme à *Plasmodium falciparum* au Sénégal sur une période de 25 ans (2000-2024). Il s'agissait spécifiquement de :

- Faire une analyse descriptive spatiale de l'évolution du paludisme au Sénégal ;
- Générer des cartes afin de présenter des résultats exploitables ;
- Manipuler les données de Malaria Atlas Project et de GADM.

II. METHODOLOGIE DU TRAVAIL

i) La source de données

Concernant le shapefile du Sénégal, il a été obtenu à partir des données du site GADM (<https://gadm.org/>) et les données sur l'incidence du paludisme ont été obtenues à partir du site de Malaria Atlas Project (<https://data.malariaatlas.org/>).

ii) Le logiciel et le langage de programmation utilisés

Le logiciel utilisé dans notre groupe était R et son utilisation dans l'atteinte de nos objectifs a demandé les packages suivants :

- sf : Pour la manipulation des shapefiles ;
- terra ou raster : Pour la manipulation de données raster ;

- tmap : Pour la création de cartes thématiques ;
- ggplot2 : Pour la création des graphes ;
- leaflet : Pour la création des cartes interactives ;
- dplyr : Pour la manipulation de données.

iii) Traitement de données

Nous sommes d'abord passés par l'inspection des données avec quelques vérifications essentielles notamment :

- Le système de projection de notre shapefile (st_crs)
- L'emprise géographique (st_bbox)
- La résolution de nos rasters (res)
- Les valeurs minimales et maximales de l'incidence du paludisme mise en exergue par nos raster (global (r, fun=c(« min », « max »), na.rm =TRUE)

Nous avons 25 rasters qui constituaient chacune un raster par année (de 2000 à 2024) comportant l'incidence annuelle du paludisme au Sénégal. A chaque représentation d'un raster annuel sur le Shapefile du Sénégal, une reprojection du shapefile par rapport aux coordonnées du Raster était faite.

Concernant les calculs des statistiques exploratoires, une analyse descriptive globale a été menée en calculant la moyenne annuelle de l'incidence du paludisme sur l'ensemble du territoire Sénégal.

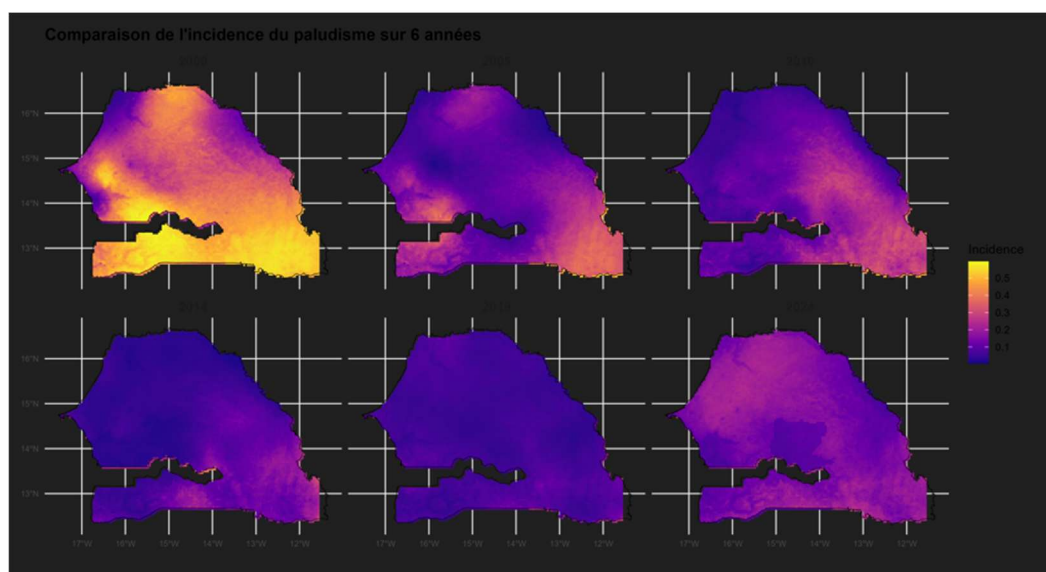
III. PRINCIPAUX RESULTATS ET OUTPUTS

Les principaux résultats du TP sont :

La sortie des données avec statistiques_annuelles.csv : nous avons pu faire une base de données recensant l'incidence moyenne annuelle du paludisme.

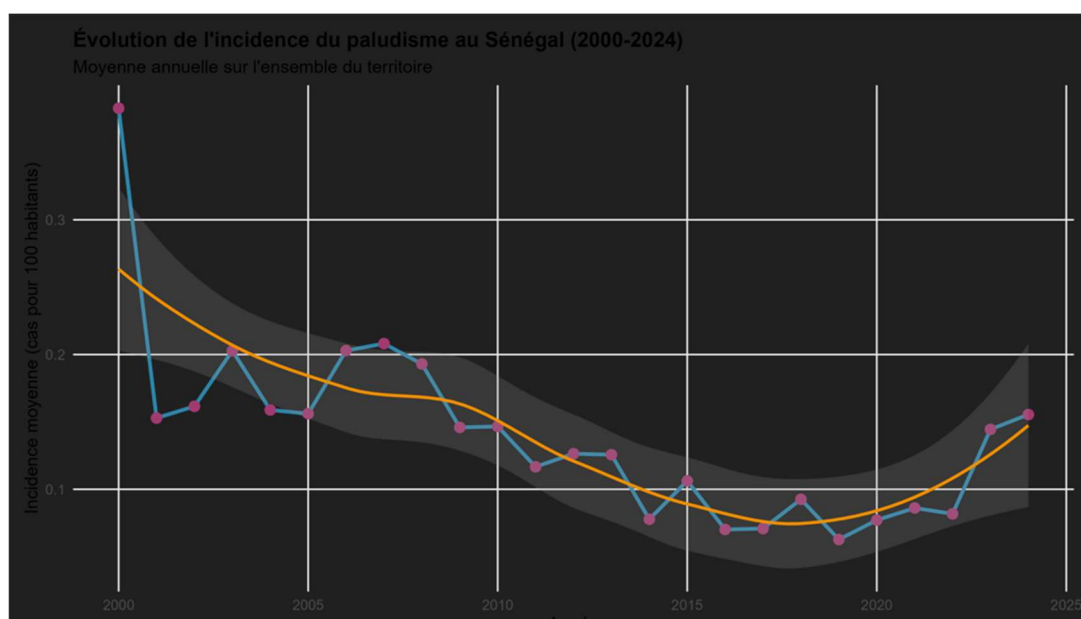
Une comparaison multi-année nous a également permis de conclure quant à la diminution progression de l'incidence du paludisme au Sénégal, ce qui pourrait soulever un fort impact de l'action gouvernementale quant à la santé de sa population. Cette tendance à la baisse est plus visible avec le nuage de point, bienqu'elle met également en évidence une tendance haussière après l'année 2020, les années post-Covid.

Image 4 : comparaison de l'incidence suivant l'année



Source : travaux du groupe

Image 5 : Evolution de l'incidence nationale moyenne annuelle



Source : travaux du groupe

ANALYSE DES DISTANCES ENTRE INFRASTRUCTURES ET AIRES PROTEGEES AU SENEGAL

L'analyse faite ici demandait dans un premier temps de repérer les aires protégées du pays étudié, ainsi que les infrastructures, et d'évaluer la distance entre les deux. Les groupes n'avaient pas été changé ainsi que le pays, mais le logiciel utilisé avait été permuté. Pour ce travail, notre groupe avait utilisé GEE Javascript.

I. OBJECTIF DU TRAVAIL

L'objectif principal de travail était de mettre en évidence les zones protégées exposées, c'est-à-dire les zones protégées qui sont très proche des infrastructures et de la population. Il s'agissait spécifiquement de :

- Repérer et cartographier les zones protégées ;
- Repérer et cartographier les infrastructures ;
- Faire des Buffers pour faire une analyse de proximité (savoir sur quel rayon les infrastructures recouvrent les zones protégées).

II. METHODOLOGIE DE TRAVAIL

i) La source de données

Concernant les sources de données, nous avons utilisé trois principalement. Il s'agissait premièrement de GADM ([GADM - Database of Global Administrative Areas](#)) pour obtenir le shapefile du Sénégal, OpenStreetMap ([OpenStreetMap \(OSM\)](#)) pour obtenir les différentes infrastructures étudiées, et enfin Protected Planet ([Protected Planet - World Database on Protected Areas \(WDPA\)](#)) qui permet d'obtenir le shapefile sur les aires protégées.

ii) Logiciel et langage de Programmation

Pour réaliser ce travail, notre groupe a eu à utiliser Google Earth Engine, avec pour langage de programmation Javascript.

iii) Traitement de données

Le premier traitement s'est fait au niveau de rasters des aires protégées. Le site de planet protected ne nous permet pas d'avoir un raster unique, mais plutôt des fragments de Raster. Avec la fonction « merge » il a fallu merger ces différents fragments de Rasters.

Concernant les infrastructures, pour chaque pixel ayant des infrastructures, on lui attribue la valeur de 1 `features.map(function(f) {...})` et les autres on attribue les valeurs nulles. Cette opération permet la rasterisation des infrastructures en image binaire de présence, étape indispensable pour le calcul des distances avec les aires protégées.

Le code suivant a été utilisé pour calculer les distances :

```
var distance = image.fastDistanceTransform({  
  neighborhood: 256,  
  units: 'pixels',  
  metric: 'squared_euclidean'  
}).sqrt().multiply(ee.Image.pixelArea().sqrt());
```

La distance utilisée était la distance euclidienne ('squared_euclidean') et le calcul se faisait pour chaque pixel. Puis ces distances seront utilisées pour chaque zones protégées pour savoir quelles les infrastructures les plus proches ou les oins proches des aires protégées.

Au terme de ce TP, nous avons appris à faire des buffers et ainsi à mettre en évidence les notions d'accès ou encore de proximité, selon un rayon donné.

ANALYSE DES DONNEES GEOSPATIALES PORTANT SUR L'ACCES AUX INFRASTRUCTURES AU SENEGAL

Cette analyse portait sur localisation des infrastructures par rapport à une zone donnée afin de connaître les zones dans lesquelles les habitants ont un accès difficile voire impossible aux infrastructures. Le logiciel utilisé ici était GEE javascript, le pays était le Sénégal et le groupe n'avait pas changé.

I. OBJECTIF DE TRAVAIL

L'objectif n'est pas très différent de celui précédent, mais ici il s'agit de la distance de la population aux infrastructures. Il était question de faire un état des lieux de l'accès des populations aux différentes infrastructures (hôpitaux, écoles, pharmacies,...). Les objectifs spécifiques étaient :

- Faire des Buffers et repérer les zones ayant peu d'infrastructures à proximité
- Mettre en évidence la situation socio-démographiques des populations par rapport à l'accès aux infrastructures.

II.METHODOLOGIE DE TRAVAIL

i) La source de données

Les données ont pour source GADM ([GADM - Database of Global Administrative Areas](#)) pour obtenir le shapefile du Sénégal, OpenStreetMap ([OpenStreetMap \(OSM\)](#)) pour obtenir les différentes infrastructures étudiées, et enfin WorldPop ([Protected Planet - World Database on Protected Areas \(WDPA\)](#)) pour mes données sur la démographie.

ii) Logiciel et langage de programmation

Le travail a été réalisé sur Google Earth Engine (GEE), en JavaScript, avec développement :

- d'une carte interactive (Map et couches),

- d'une interface UI (panneaux, boutons, légende),
- de traitements géospatiaux (buffers, comptages, agrégations, choroplèthes, statistiques).

iii) Traitement de données

La plupart des données utilisées ne nécessitaient pas de traitement particulier. L'étape la plus cruciale dans notre analyse était celle avec la fonction `createBuffers` qui nous permettait de varier le rayon du buffers et également la méthode « onclick » que nous avons découvert et grâce à laquelle nous avons pu créer des fonctions qui s'active en cliquant sur la carte de GEE ; comme la fonction qui nous permet de compter le nombre d'infrastructure sur un rayon de 10km en prenant pour centre le point sur lequel on a cliqué.

III. PRINCIPAUX RESULTATS ET OUPUTS

L'implémentation du script a permis d'obtenir plusieurs résultats analytiques et cartographiques.

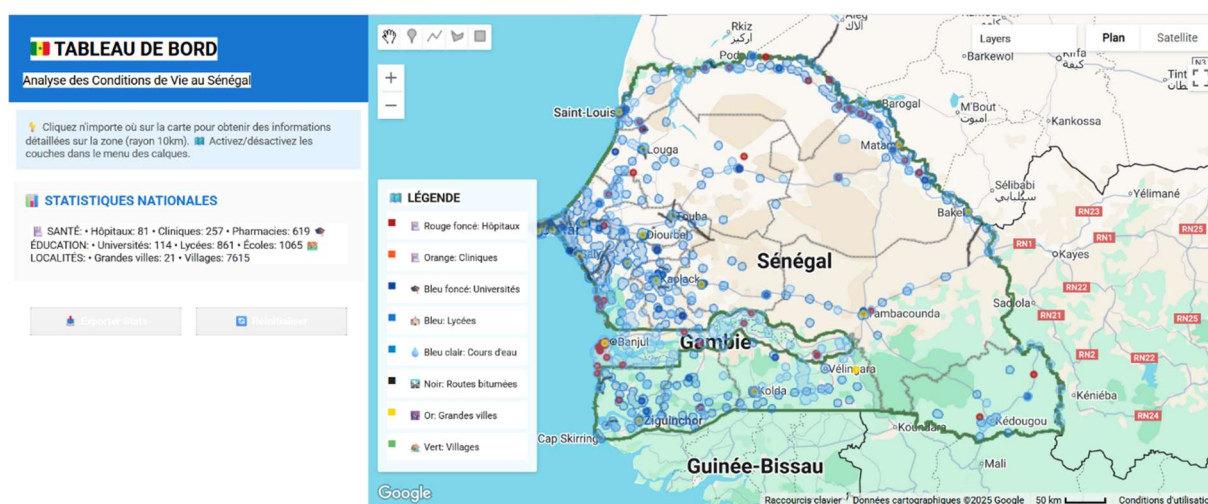
- Visualisation spatiale des infrastructures

Une carte interactive du Sénégal a été produite, intégrant :

- les infrastructures de santé (hôpitaux, cliniques, pharmacies) ;
- les infrastructures éducatives (écoles, collèges, lycées, universités) ;
- les localités et réseaux de transport ;
- les données de population.

Cette visualisation permet une première lecture des disparités territoriales.

Image 6 : Tableau de Bord représentant les infrastructures au Sénégal



Source : travaux du groupe

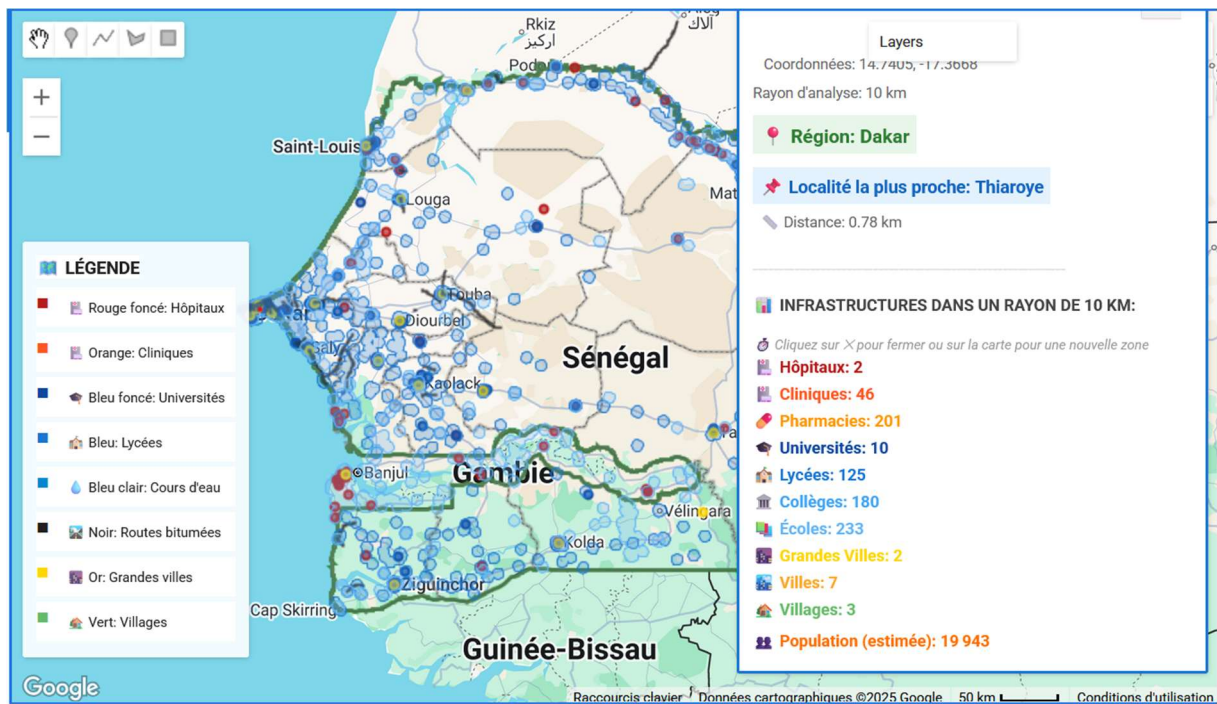
- Analyse interactive locale

Une fonctionnalité de clic sur la carte permet désormais :

- d'identifier la région du point sélectionné ;
- de trouver la localité la plus proche ;
- de compter les infrastructures dans un rayon de 10 km ;
- d'estimer la population locale.

Cette approche offre un outil d'exploration spatiale dynamique utile pour le diagnostic territorial.

Image 7 : statistiques obtenues en cliquant sur la carte



Source : travaux du groupe

Grâce à ce TP nous avons pu approfondir la création des buffers et cette fois-ci avec des analyses multivariées et en incluant des analyses statistiques au sein des buffers.

IDENTIFICATION DES TERRES ARABLES EN OUGANDA

L'analyse de ce TP reposait sur l'identification des terres arables en Ouganda. Ce travail marqua le premier changement de membre des groupes. Notre groupe comptait cette fois Astou DIOP, Mamadou Lamine DIABANG, et Cheikh M. Moustapha NDIAYE. Le logiciel utilisé ici était Python.

I. OBJECTIF DU TRAVAIL

L'objectif de ce travail était d'analyser, à l'aide de données géospatiales, l'occupation du sol et les contraintes environnementales en Ouganda, afin d'évaluer le potentiel agricole et la répartition spatiale de certaines composantes du territoire.

Plus précisément, il s'agissait de produire un diagnostic spatial combinant plusieurs variables biophysiques et anthropiques (forêt, imperméabilisation, pente, eau, aires protégées, terres cultivées).

Les objectifs spécifiques étaient :

- Identifier et cartographier différentes composantes de l'occupation du sol (forêt, surfaces imperméables, zones en pente, eaux permanentes, etc.) ;
- Produire des rasters binaires permettant de distinguer les zones favorables ou contraintes;
- Analyser la répartition des terres arables et cultivées ;
- Calculer des superficies et produire des graphiques statistiques par région ;
- Mettre en place une visualisation interactive pour faciliter l'exploration spatiale.

L'enjeu principal était de fournir une lecture spatiale intégrée du territoire ougandais en lien avec le potentiel agricole et les contraintes physiques.

II. METHODOLOGIE DE TRAVAIL

i) La source de données

Plusieurs bases de données géospatiales ont été mobilisées :

- eaux permanentes le site JRC_GSW1_4_Global sur disponible sur GEE : https://developers.google.com/earthengine/datasets/catalog/JRC_GSW1_4_GlobalSurfaceWater?hl=fr
- Pour la déforestation il s'agissait de hansen_global_forest_change disponible également sur GEE: https://developers.google.com/earthengine/datasets/catalog/UMD_hansen_global_forest_change_2024_v1_12?hl=fr
- Les données des terres cultivées ont été trouvées sur Copernicus : <https://essd.copernicus.org/articles/17/3777/2025/>
- Google Earth Engine nous permettait également d'obtenir les données de USGS_SRTMGL pour les terres cultivées : https://developers.google.com/earthengine/datasets/catalog/USGS_SRTMGL1_003?hl=fr ainsi que les aires protégées : https://developers.google.com/earthengine/datasets/catalog/WCMC_WDPA_current_polygons?hl=fr
- Les limites administratives du Rwanda ont été retrouvées dans le catalogue de GEE recueillant les données de la FAO a : https://developers.google.com/earthengine/datasets/catalog/FAO_GAUL_2015_level0?hl=fr
- Les terres imperméables sont disponibles sur le site <https://gee-community-catalog.org/projects/gisa/>

ii) Logiciel et langage de programmation

Le traitement a été réalisé avec Google Earth Engine (GEE) en utilisant le langage Python via l'API Earth Engine, couplé à :

- Folium pour la cartographie interactive ;
- Matplotlib, ggplot pour la production de graphiques ;
- Pandas et NumPy pour le traitement tabulaire.

iii) Principales étapes de traitement

Nous avons commencé par un prétraitement des données comprenant le clip des données à l'emprise de l'Ouganda, l'harmonisation des projections et la conversion de certaines couches en rasters binaires (0 ou 1) afin d'identifier les surfaces arables.

Pour la création de rasters binaires, plusieurs variables ont été discrétisées, notamment la distinction forêt versus non-forêt avec un seuil de 20 %, les surfaces imperméables, les zones à forte pente supérieures ou égales à 15°, ainsi que les eaux permanentes. Cette étape permet de transformer des variables continues en indicateurs exploitables.

Le script réalise également l'identification des terres arables, l'intersection avec les terres cultivées et le calcul des superficies correspondantes. Cela permet d'estimer les terres arables exploitées ainsi que les terres arables non exploitées. Les superficies sont agrégées par région administrative afin de comparer les territoires, calculer les parts relatives et produire des indicateurs synthétiques. Plusieurs formes de sortie ont été produites, comprenant des cartes interactives Folium, des cartes thématiques, des diagrammes en camembert et des graphiques en barres empilées.

III. RESULTATS ET OUTPUTS

Le script a permis de générer plusieurs cartes thématiques de l'Ouganda montrant :

- la couverture forestière,
- les surfaces imperméables,
- les zones de forte pente,
- les eaux permanentes,
- les aires protégées.

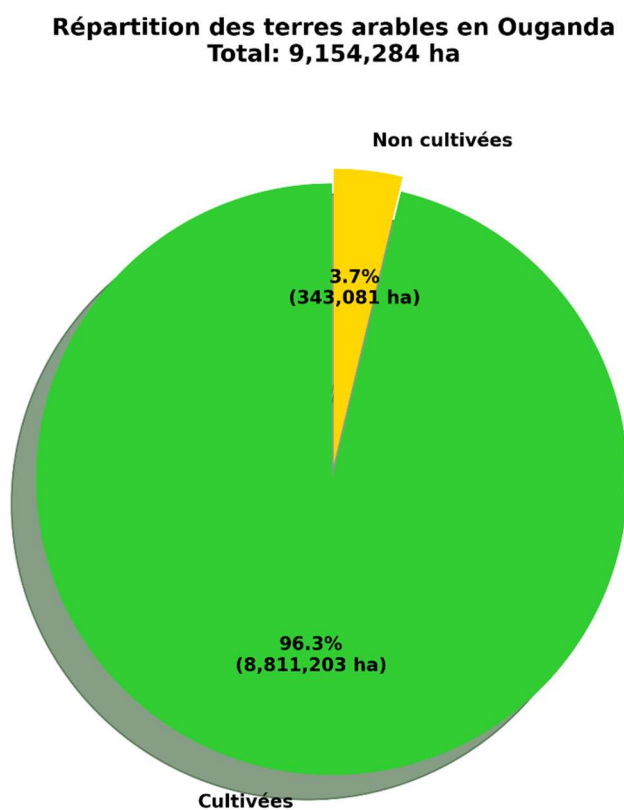
Ces cartes offrent une lecture spatiale des contraintes physiques du territoire.

Concernant l'identification des terres arables, l'analyse a permis de distinguer :

- les terres arables cultivées,
- les terres arables non cultivées.

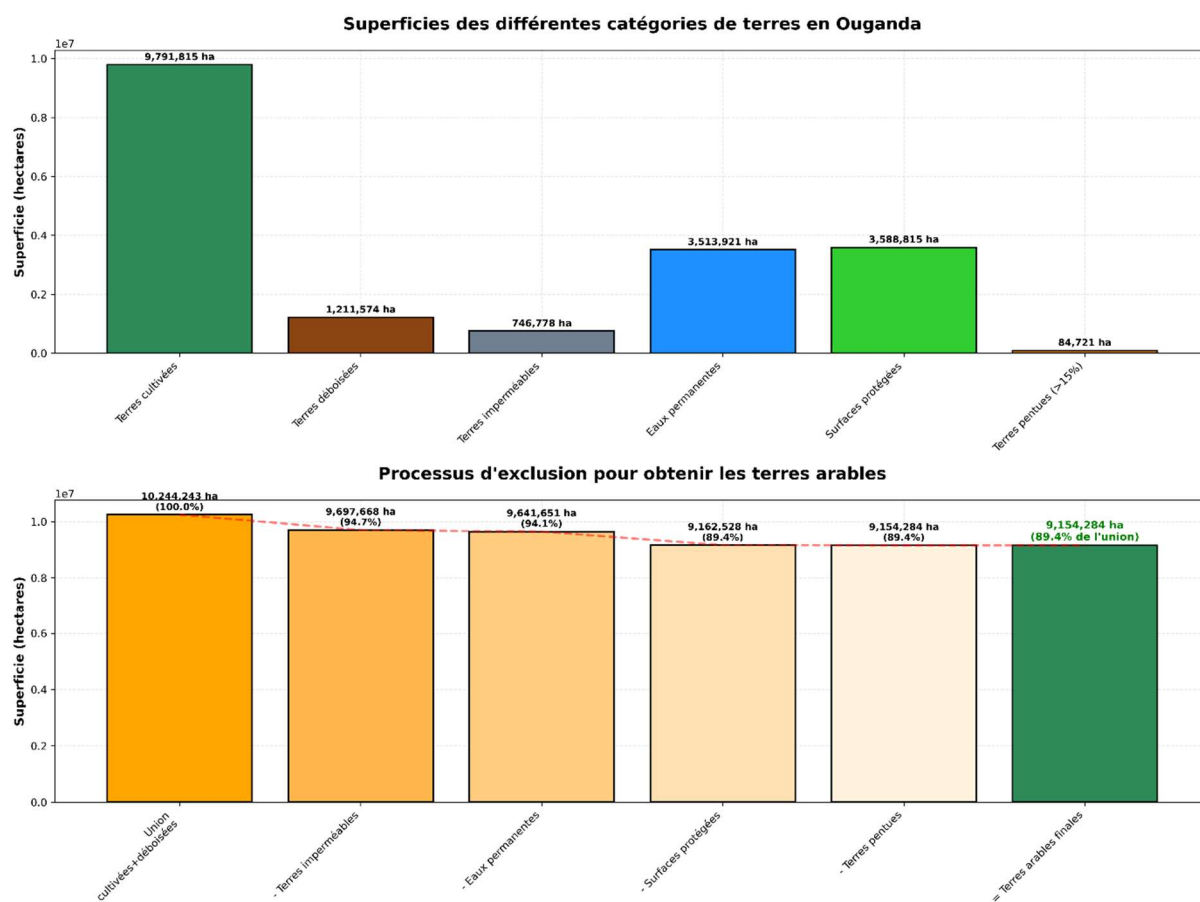
Un diagramme circulaire (camembert) met en évidence la part relative de ces deux catégories, fournissant un indicateur du niveau d'exploitation agricole.

Image 8 : Pourcentage des terres arables cultivées et non cultivées



Source : travaux du groupe

Image 9 : Répartition des terres selon le type



Source : travaux du groupe

Ce TP nous a permis de découvrir la situation agricole de l'Ouganda et également de découvrir plusieurs méthodes permettant de filtrer des couches. De nouvelles sources de données ont également été découvertes, à l'instar de Copernicus et de WCMC.

ANALYSE SPATIALE DU DEVELOPPEMENT URBAIN AU RWANDA

A partir du ratio LCRPGR, ce travail demandait de mettre en évidence l'objectif de développement durable 11.3.1 portant sur l'étude du taux d'utilisation des terres par rapport au taux de croissance démographique. Notre groupe était constitué de Astou DIOP, Trésor RIRADJIM, et Michel TEVOEDJRE. Le logiciel utilisé ici était GEE Python et le pays étudié était le Rwanda.

I. OBJECTIF DU TRAVAIL

L'objectif principal du script est d'évaluer la dynamique d'urbanisation au Rwanda entre 2015 et 2025 en comparant :

- l'évolution des surfaces bâties (Built-up) ;
- l'évolution de la population (WorldPop) ;

puis en calculant un indicateur synthétique : LCRPGR (*Land Consumption Rate to Population Growth Rate*), afin d'identifier si l'urbanisation correspond plutôt à :

- une densification (croissance spatiale plus lente que la population),
- ou un étalement/expansion urbaine (croissance spatiale plus rapide que la population).

II. MÉTHODOLOGIE DE TRAVAIL

i) Sources de données

Le script mobilise des données raster et vectorielles :

- Surfaces bâties : Rwanda_BuiltUp_2015.tif et Rwanda_BuiltUp_2025.tif
- Les données de la population avec Worldpop

- **Limites administratives Rwanda (GADM) :**

- niveau pays : gadm41_RWA_0.shp
- niveau régions/provinces : gadm41_RWA_1.shp

ii) Logiciel et langage de programmation

Le logiciel utilisé était GEE avec pour langage Python (notebook Jupyter). Les principales librairies utilisées étaient :

- rasterio,
- geopandas,
- numpy,
- matplotlib, + openpyxl (export Excel).

iii) Traitements de données

Au niveau du raster de la population, les pixels sans valeur avaient été imputés par (-99999). Pour les besoins de calcul, ces valeurs ont été systématiquement remplacées par NaN pour éviter les biais dans les calculs statistiques.

Ensuite, le script lit chaque raster et calcule des statistiques de base :

- min, max, moyenne, somme, nombre de pixels valides,
- résolution, CRS, dimensions.

Ce pour harmoniser les projections et nettoyer les données.

Il produit des cartes comparables entre années en utilisant une échelle commune (mêmes bornes min/max) calculée par :

- `get_common_vmin_vmax([...])`
- et une fonction de rendu cartographique : `plot_raster_with_border(...)`

À partir des sommes nationales faite en sommant les pixels après nettoyage :

- V_{2015}, V_{2025} = surfaces bâties totales (converties en km²)
- P_{2015}, P_{2025} = populations totales
- $Y = 10$ ans

Nous avons eu à faire les calculs suivants :

- taux de croissance spatiale : $(V_{2025} - V_{2015}) / (V_{2015} \times Y)$
- taux de croissance démographique : $\ln(P_{2025} / P_{2015}) / Y$

Ce qui nous a permis de calculer le LCRPGR

$$\text{LCRPGR} = (\text{taux spatial}) / (\text{taux démographique})$$

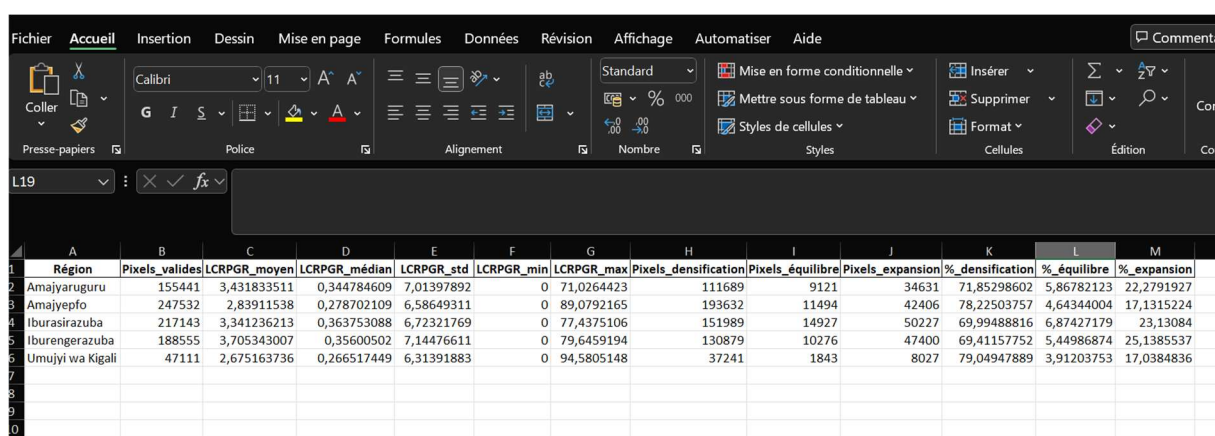
Notons de plus que le script permet de projeter les régions dans le CRS des rasters. Il extrait ensuite les pixels LCRPGR appartenant à chaque région (masque spatial), ce qui permet de calculer le LCRPGR par région (moyenne, médiane, std, min, max) puis à partir du résultat, nous pourrions classer les pixels par catégories (densification (inférieur à 1), équilibre (LCRGPGR=1) ou expansion (LCRPGR supérieur à 1)).

Ainsi, nous pouvons produire ensuite :

- un classement des régions (barres horizontales),
- une répartition en % par catégories (barres empilées),
- une carte choroplèthe du LCRPGR moyen,
- une carte « dominante » (expansion , densification et mixte).

Nous avons également eu à exporter nos résultats dans un fichier Excel : LCRPGR_Analyse_Regionale.xlsx

Image 10 : Présentation de la sortie excel des résultats



	A	B	C	D	E	F	G	H	I	J	K	L	M
	Région	Pixels_valides	LCRPGR_moyen	LCRPGR_médian	LCRPGR_std	LCRPGR_min	LCRPGR_max	Pixels_densification	Pixels_équilibre	Pixels_expansion	%_densification	%_équilibre	%_expansion
2	Amajyaruguru	155441	3,431833511	0,344784609	7,01397892	0	71,0264423	111689	9121	34631	71,85298602	5,86782123	22,2791927
3	Amajyepfo	247532	2,83911538	0,278702109	6,58649311	0	89,0792165	193632	11494	42406	78,22503757	4,64344004	17,1315224
4	Iburasirazuba	217143	3,341236213	0,363753088	6,72321769	0	77,4375106	151989	14927	50227	69,99488816	6,87427179	23,13084
5	Iburengerazuba	188555	3,705343007	0,35600502	7,14476611	0	79,6459194	130879	10276	47400	69,41157752	5,44986874	25,1385537
6	Umujiyi wa Kigali	47111	2,675163736	0,266517449	6,31391883	0	94,5805148	37241	1843	8027	79,04947889	3,91203753	17,0384836
7													
8													
9													
10													

Source : travaux du groupe

III. PRINCIPAUX RÉSULTATS ET OUTPUTS

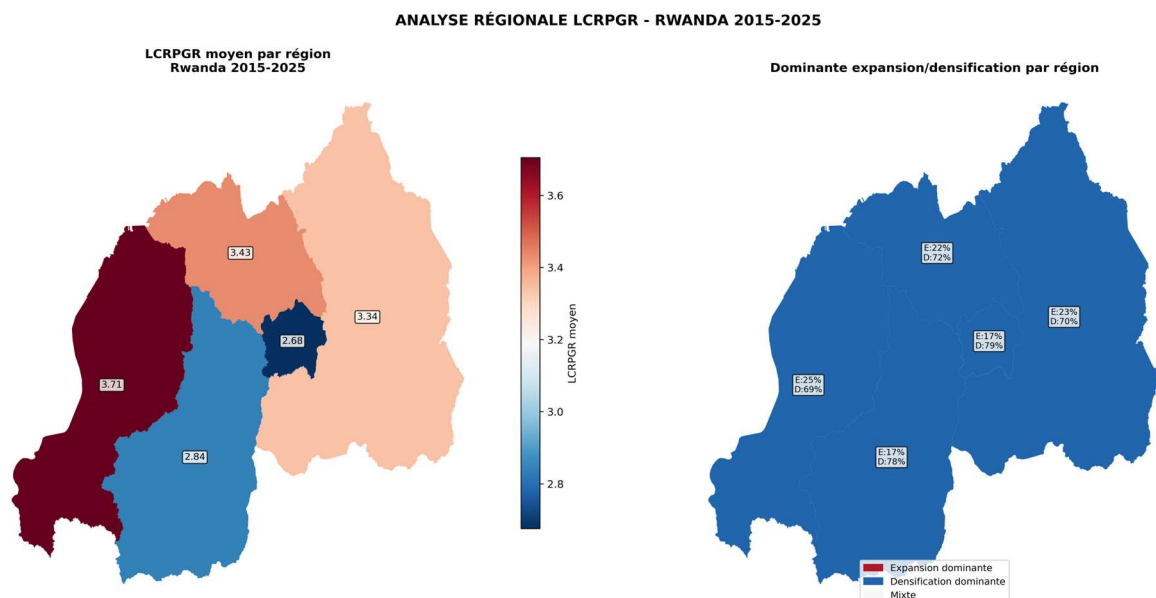
Les statistiques descriptives complètes sur les rasters (surface bâti et population) pour 2015 et 2025, ainsi que la valeur LCRPGR globale et son interprétation (densification vs expansion).

Cartographie de :

- surfaces bâties 2015/2025,
 - population 2015/2025,
 - LCRPGR par pixel,
 - LCRPGR par région (choroplèthes).
- Analyse régionale détaillée : classement des provinces + structure % des catégories.

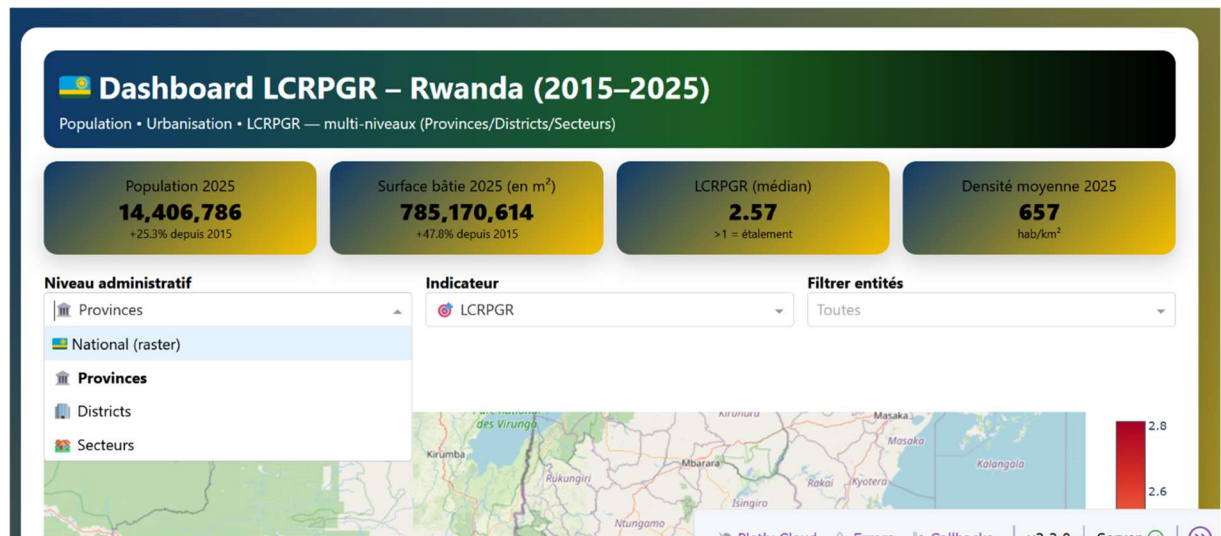
Le script enregistre plusieurs figures PNG (cartes + graphiques) et un export Excel, notamment le LCRPGR_Analyse_Rwanda_2015-2025.png, le LCRPGR_Classement_Regions.png le LCRPGR_Carte_Regions.png et LCRPGR_Analyse_Régionale.xlsx.

Image 11 : présentation de LCRPGR par région



Source : travaux du groupe

Image 12 : interface Dashboard de visualisation de LCRPGR au Rwanda



Source : travaux du groupe

Grâce à ce TP, nous avons appris un nouvel indicateur qu'est le Land Consumption Rate to Population Growth Rate ainsi que sa méthodologie. De plus

ANALYSE DES INDICES SPECTRAUX ET FUSION AVEC LES DONNEES EHCVM DU MALI

Ce dernier travail portait sur une analyse spatiale de quelques indicateurs spectraux de certains pays de l'UEMOA. Les indicateurs pouvaient aller de la température, à la végétation, en passant par la pollution. Les groupes n'avaient pas été changés mais le pays étudié pour ce travail était le Mali. Le logiciel utilisé était Python.

I. OBJECTIF DU TRAVAIL

L'objectif principal de ce travail était d'exploiter les indices spectraux issus de l'imagerie satellitaire afin de caractériser l'état de surface du territoire et d'analyser les conditions environnementales en lien avec les dynamiques territoriales.

Il s'agissait plus spécifiquement de :

- Calculer et cartographier plusieurs indicateurs spectraux;
- Caractériser les différentes composantes du paysage (végétation, eau, bâti, température, pollution) ;
- Fusionner les résultats obtenus avec les données de l'Enquête Harmonisée des Conditions de Vie des Ménages.

II. METHODOLOGIE DE TRAVAIL

i) La source de données

Ce projet d'analyse territoriale s'appuie sur plusieurs sources de données complémentaires.

- Images Sentinel-2 : pour le calcul des indices spectraux ;
- MODIS Terra LST pour les données concernant la température ;

- Limites administratives : GADM ;
- Données EHCVM du Mali : Programme d'Harmonisation et de Modernisation des Enquêtes sur les Conditions de Vie des ménages dans les États membres de l'UEMOA.

Ces données combinent imagerie raster multispectrale et couches vectorielles administratives.

ii) Logiciel et langage de programmation

Le traitement a été réalisé sous Google Earth Engine et Python (notebooks).

Les outils principaux utilisés sont :

- Earth Engine API pour le calcul des indices ;
- Python (NumPy, Pandas) pour le traitement statistique ;
- Matplotlib pour les graphiques ;
- Folium / visualisation pour le dashboard interactif.

iii) Traitement de données

Prétraitement des images Sentinel-2

Le prétraitement des images Sentinel-2 a suivi plusieurs étapes essentielles : filtrage temporel des acquisitions, nettoyage des nuages, mosaïquage lorsque nécessaire, et découpage à l'emprise de la zone d'étude. Cette chaîne de traitement garantit la qualité radiométrique nécessaire au calcul des indices.

Les indicateurs spectraux

Pour ce projet, Quatre groupes d'indices spectraux principaux ont été calculés : un groupe avec des indicateurs de végétation, un autre avec des indices de température, le troisième correspondait aux indicateurs de Pollution, le Quatrième à l'eau.

Dans le groupe des indicateurs de végétation nous avons cinq indicateurs parmi lesquels, le NDVI (Normalized Difference Vegetation Index). Le NDVI mesure la vigueur de la végétation selon la formule. Les valeurs élevées indiquent une végétation dense, tandis que les valeurs faibles signalent des sols nus ou des zones urbaines. Cet indice a permis de cartographier le couvert végétal national, de produire des histogrammes de distribution et d'agréger les résultats par région et département. Cet indicateur a pour formule

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

NIR= Near InfraRed

Concernant les indicateurs des activités aquatiques nous avons eu à calculer Le NDWI (Normalized Difference Water Index). Le NDWI permet d'identifier les surfaces en eau et les zones humides. Les valeurs positives indiquent une présence probable d'eau, tandis que les valeurs négatives correspondent aux surfaces terrestres. Il sert à détecter les plans d'eau et à analyser l'humidité de surface. Il a pour formule :

$$NDWI = \frac{Green - NIR}{Green + NIR}$$

La température de surface (Land Surface Température abrégé LST) a été estimée à partir des bandes thermiques satellitaires. Ici une fonction centrale de lecture des données MODIS a permis de traiter les fichiers HDF contenant la température de surface. Elle ouvre chaque fichier, extrait la couche spécifiée (LST_Day_1km ou LST_Night_1km), et applique la conversion d'échelle MODIS standard : la température est calculée selon la formule

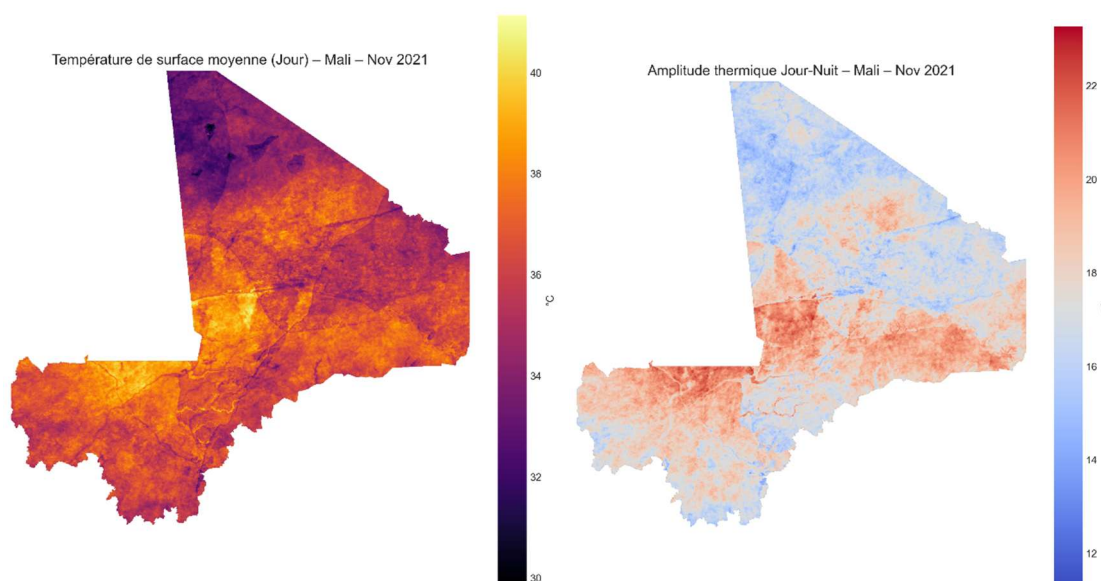
$$temp = lst \times 0.02 - 273.15,$$

permettant le passage des valeurs brutes aux degrés Celsius. Les valeurs nulles sont transformées en NaN (Not a Number) et un filtrage élimine les valeurs aberrantes, notamment celles inférieures à -20°C ou supérieures à 60°C. La date d'acquisition est extraite du nom du fichier selon le format MODIS, par exemple A2021329 indique l'année 2021 et le jour julien 329. Cette étape transforme les pixels MODIS bruts en températures exploitables accompagnées de leur référence temporelle.

III. RESULTATS PRINCIPAUX ET OUTPUTS

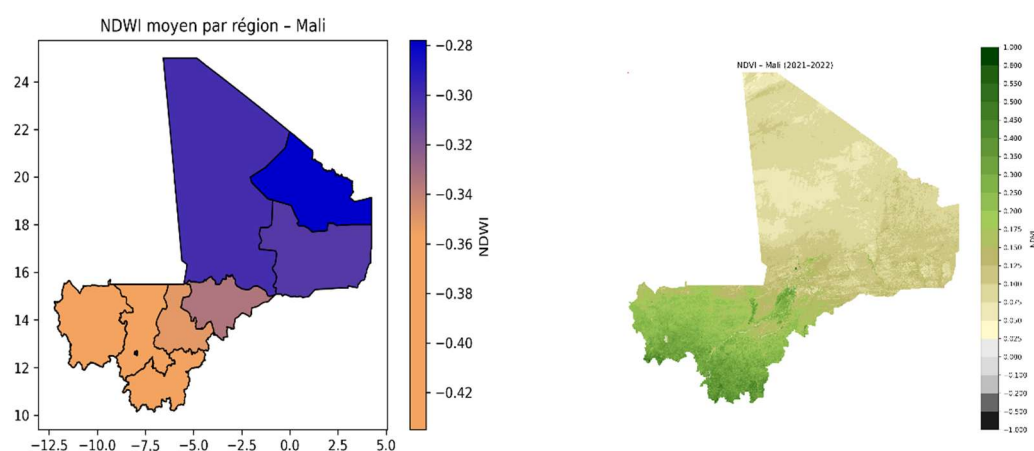
Des cartes nationales ont été générées pour six indicateurs clés permettant une caractérisation complète de l'état de surface du territoire : le NDVI pour la végétation, le NDWI pour les surfaces en eau, le NDMI pour l'humidité de la végétation, le NDBI pour les zones bâties, le BSI pour les sols nus, et la température de surface (LST). Ces cartes offrent une lecture synoptique et complémentaire de l'état de surface, permettant d'appréhender simultanément les différentes composantes du territoire.

Image 13 : Température sur la surface le jour (à gauche) et la nuit (à droite)



Source : travaux du groupe

Image 14 : représentation spatiale du NDWI (à gauche) et du NDVI (à droite)



Source : travaux du groupe

L'analyse statistique multi-échelle révèle une forte variabilité spatiale des indices, tant au niveau régional que départemental. Cette hétérogénéité spatiale reflète la diversité des conditions biophysiques et anthropiques du territoire étudié. Les histogrammes nationaux produits pour chaque indice permettent d'analyser la distribution globale des valeurs et d'identifier les seuils caractéristiques ainsi que les tendances centrales de chaque indicateur à l'échelle du pays.

Pour faciliter les analyses complémentaires et l'intégration dans d'autres systèmes, des fichiers excel de synthèse ont été produits pour chaque indice.

Image 15 : Sortie excel des résultats

BASE EHCVM		INDEX VEGETATION			INDEX BUILT-UP	INDEX WATER	INDEX TEMPERATURE					
grappe	menage	vNDVI	BSI	NDMI	NDBI	NDWI	T_Jour_moy	T_Nuit_moy	IVT_par_régio	IVT_categorie	Score_risque	Alerte_chaleur
1	1.0	10.19759421	0.21684463	-0.14775878	0.14775878	-0.36332738	36.78	18.34	0.53	Modérée	68.15	Non
1	2.0	10.19759421	0.21684463	-0.14775878	0.14775878	-0.36332738	36.78	18.34	0.53	Modérée	68.15	Non
1	3.0	10.19759421	0.21684463	-0.14775878	0.14775878	-0.36332738	36.78	18.34	0.53	Modérée	68.15	Non
1	4.0	10.19759421	0.21684463	-0.14775878	0.14775878	-0.36332738	36.78	18.34	0.53	Modérée	68.15	Non
1	5.0	10.19759421	0.21684463	-0.14775878	0.14775878	-0.36332738	36.78	18.34	0.53	Modérée	68.15	Non
1	6.0	10.19759421	0.21684463	-0.14775878	0.14775878	-0.36332738	36.78	18.34	0.53	Modérée	68.15	Non
1	7.0	10.19759421	0.21684463	-0.14775878	0.14775878	-0.36332738	36.78	18.34	0.53	Modérée	68.15	Non
1	8.0	10.19759421	0.21684463	-0.14775878	0.14775878	-0.36332738	36.78	18.34	0.53	Modérée	68.15	Non
1	9.0	10.19759421	0.21684463	-0.14775878	0.14775878	-0.36332738	36.78	18.34	0.53	Modérée	68.15	Non
1	10.0	10.19759421	0.21684463	-0.14775878	0.14775878	-0.36332738	36.78	18.34	0.53	Modérée	68.15	Non
1	11.0	10.19759421	0.21684463	-0.14775878	0.14775878	-0.36332738	36.78	18.34	0.53	Modérée	68.15	Non
1	12.0	10.19759421	0.21684463	-0.14775878	0.14775878	-0.36332738	36.78	18.34	0.53	Modérée	68.15	Non
2	1.0	10.19759421	0.21684463	-0.14775878	0.14775878	-0.36332738	36.78	18.34	0.53	Modérée	68.15	Non
2	2.0	10.19759421	0.21684463	-0.14775878	0.14775878	-0.36332738	36.78	18.34	0.53	Modérée	68.15	Non

Source : travaux du groupe

L'analyse par indices spectraux met en évidence la diversité biophysique du Mali, avec des contrastes marqués entre zones végétalisées, surfaces en eau, espaces urbanisés, sols nus, température, pollution. Ce TP nous a permis d'enrichir notre connaissance par rapport aux indicateurs spatiaux et par la même occasion de découvrir la situation environnementale du Mali

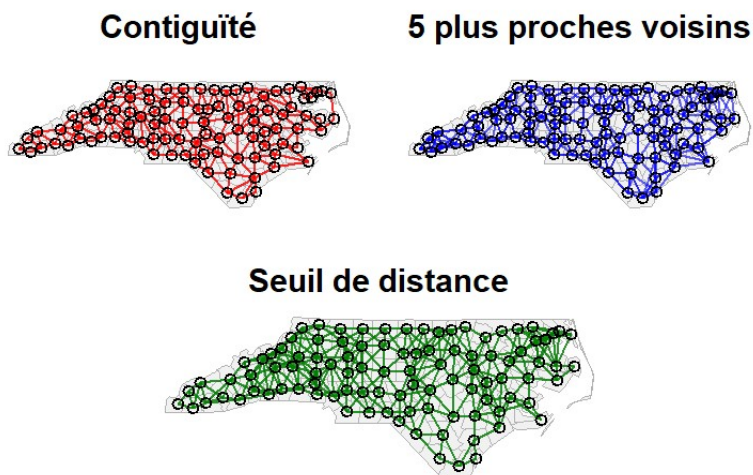
MATRICE DE VOISINAGE ET INDICE DE MORAN

Ce travail était individuel et portait sur l'application des notions de matrice de voisinage et d'indice de Moran vu durant le cours.

Il s'agissait principalement ici d'appliquer ce qui avait été vu en cours notamment les notions de Matrice de voisinage basée sur la distance, les matrices de voisinage basée sur les plus proches voisins et les matrices de contiguïté de type Queen, Bishop et Hook. Ce travail devait être fait dans le logiciel R.

Ce travail a pu être fait grâce aux données de Caroline du Nord qu'on a pu retrouver grâce au code suivant : « `nc <- st_read("https://geodacenter.github.io/data-and-lab/data/sids.geojson")` »

Image 16 : Représentation des villes de Caroline du Nord selon les différentes matrices de voisinage



Source : travaux du groupe

Il était également question de démontrer l'assertion selon laquelle « la pente de la droite de régression linéaire simple de WX contre X est égale à l'indice de Moran si W est normalisée »

Le présent TP nous a permis de mieux nous familiariser avec ces concepts et à les manipuler grâce au Logiciel R.

CONCLUSION

Durant ces 30 heures de cours de statistique exploratoire spatiale, nous avons appris à manier diverses données qui étaient méconnues par la plupart d'entre nous, des rasters, aux Geopackage, en passant par les shapefiles. La découverte de Google Earth Engine a de plus enrichi notre connaissance sur les logiciels de programmation et d'analyse spatiale. L'utilisation de certaines données spécifiques à nos TP, notamment les rasters de la malaria, des terres arables ou encore les données démographiques nous ont permis de découvrir plusieurs sources de données parmi lesquelles MODIS, World Pop ou encore Malaria Atlas project. En outre, spécifiquement les TP5 et TP6 nous ont permis de découvrir plusieurs indicateurs en particulier le LCRPGR ainsi que sa méthodologie, et également le NDVI. Le travail Pratique sur les matrices de voisinage nous a permis de mieux comprendre de manière les notions d'indices de Moran et de matrice de contiguïté sur R qui était encore jusqu'ici très théorique. Cette unité d'enseignement a ainsi constitué une occasion pour nous de découvrir de nouvelles manières d'appliquer la statistique et de faire parler les résultats, avec comme méthode le « learning by doing ».

TABLE DES MATIERES

REMERCIEMENTS	I
SOMMAIRE	II
LISTE DES IMAGES	III
INTRODUCTION.....	4
PRESENTATION DE GOOGLE EARTH ENGINE	5
ANALYSE SPATIALE DU PALUDISME AU SENEGAL (2000-2024).....	8
I. OBJECTIF DU TRAVAIL.....	8
II.METHODOLOGIE DU TRAVAIL	8
III.PRINCIPAUX RESULTATS ET OUTPUTS	9
ANALYSE DES DISTANCES ENTRE INFRASTRUCTURES ET AIRES PROTEGEES AU SENEGAL.....	11
I. OBJECTIF DU TRAVAIL.....	11
II. METHODOLOGIE DE TRAVAIL.....	11
ANALYSE DES DONNEES GEOSPATIALES PORTANT SUR L'ACCES AUX INFRASTRUCTURES AU SENEGAL.....	13
I. OBJECTIF DE TRAVAIL	13
II. METHODOLOGIE DE TRAVAIL.....	13
III. PRINCIPAUX RESULTATS ET OUPUTS.....	14
IDENTIFICATION DES TERRES ARABLES EN OUGANDA	17
I. OBJECTF DU TRAVAIL	17
II. METHODOLOGIE DE TRAVAIL.....	18
III. RESULTATS ET OUTPUTS	19
ANALYSE SPATIALE DU DEVELOPPEMENT URBAIN AU RWANDA.....	22
I. OBJECTIF DU TRAVAIL.....	22
II. MÉTHODOLOGIE DE TRAVAIL.....	22
III. PRINCIPAUX RÉSULTATS ET OUTPUTS	25
ANALYSE DES INDICES SPECTRAUX ET FUSION AVEC LES DONNEES EHCVM DU MALI	27

I. OBJECTIF DU TRAVAIL.....	27
II. METHODOLOGIE DE TRAVAIL.....	27
III. RESULTATS PRINCIPAUX ET OUTPUTS	29
MATRICE DE VOISINAGE ET INDICE DE MORAN	32
CONCLUSION	33
TABLE DES MATIERES.....	34