

Rapport de Synthèse

Travaux Pratiques en Statistique Exploratoire Spatiale

Informations sur le Projet

Réalisé par :

THIOUB Cheikh

Supervisé par :

M HEMA Aboubacar

Classe : ISE1 CL

Année académique : 2025-2026

Ce document synthétise les six travaux pratiques réalisés en groupe dans le cadre du cours de statistiques exploratoire. Les TP couvrent l'analyse spatiale du paludisme, les distances aux aires protégées, l'intégration de données démographiques et le calcul d'indices spectraux à partir d'images satellitaires Sentinel-2.

Remerciements

Nous tenons à exprimer notre profonde gratitude à toutes les personnes qui ont contribué à la réussite de ces travaux pratiques.

Nos remerciements s'adressent tout particulièrement à notre professeur **M HEMA Aboubacar** pour son encadrement rigoureux, sa disponibilité et ses précieux conseils de vie et dans le cadre des études tout au long de ce module. Meme si tout ne peut pas etre parfait, ses enseignements et sa méthode du " learning by doing " nous ont permis de développer des compétences en analyse spatiale et de comprendre les enjeux de la statistique géographique appliquée.

Nous remercions également l'ensemble des camarades de classe avec qui nous avons collaboré pour la réalisation de ces projets.

Enfin, nous exprimons notre reconnaissance envers les **institutions** qui mettent à disposition les données géospatiales utilisées dans ces travaux : GADM, Malaria Atlas Project, WorldPop, OpenStreetMap, Protected Planet et l'Agence Spatiale Européenne (Sentinel-2).

1 Contents

1 TP1 : Analyse Spatiale du Paludisme au Sénégal (R)	3
2 TP2 : Analyse des Distances entre Infrastructures et Aires Protégées (Google Earth Engine)	7
3 TP3 : Analyse de l'Accessibilité aux Infrastructures au Sénégal (Google Earth Engine)	11
4 TP4 : Identification et Quantification des Terres Arables en Éthiopie (Python)	14
5 TP5 : Analyse du Ratio LCRPGR en Tanzanie (2017–2022)	18
6 TP6 : Analyse spatiale par indices spectraux au Sénégal (2018)	22
7 Conclusion	28
8 Références	29
A Annexe A : Glossaire	30

1 1TP1 : Analyse Spatiale du Paludisme au Sénégal (R)

1.1 1Contexte et Objectifs

Dans ce premier travail pratique, nous avons réalisé une analyse spatio-temporelle du taux d'incidence du paludisme à *Plasmodium falciparum* au Sénégal sur une période de 25 ans (2000-2024). L'objectif était de cartographier et comprendre la distribution du paludisme sur le territoire sénégalais.

1.2 1Données Utilisées

1.2.1 1Données Raster

- **Source :** Malaria Atlas Project
- **Fichiers :** 25 rasters annuels (2000-2024)
- **Format :** GeoTIFF
- **Variable :** Taux d'incidence du paludisme à *P. falciparum*
- **Résolution spatiale :** 5 km × 5 km

1.2.2 1Données Vectorielles

- **Source :** GADM (Global Administrative Areas)
- **Niveaux administratifs :** 5 niveaux
 - Niveau 0 : Sénégal (pays)
 - Niveau 1 : 14 régions
 - Niveau 2 : 45 départements
 - Niveau 3 : Arrondissements
 - Niveau 4 : Communes
- **Format :** Shapefiles
- **Projection :** WGS84 (EPSG:4326)

1.3 1Technologies et Outils

Environnement : R / RStudio

Packages R utilisés :

- **sf** : Manipulation de données vectorielles
- **terra ou raster** : Manipulation de données raster
- **tmap** : Création de cartes thématiques
- **ggplot2** : Visualisations graphiques

- `leaflet` : Cartes interactives
- `dplyr` : Manipulation de données

1.4 1Analyses Réalisées

1.4.1 11. Cartographie Statique

Nous avons créé des cartes thématiques pour visualiser le taux d'incidence du paludisme pour différentes années, notamment 2010. Ces cartes incluent :

- Légende explicative
- Échelle cartographique
- Titre et annotations

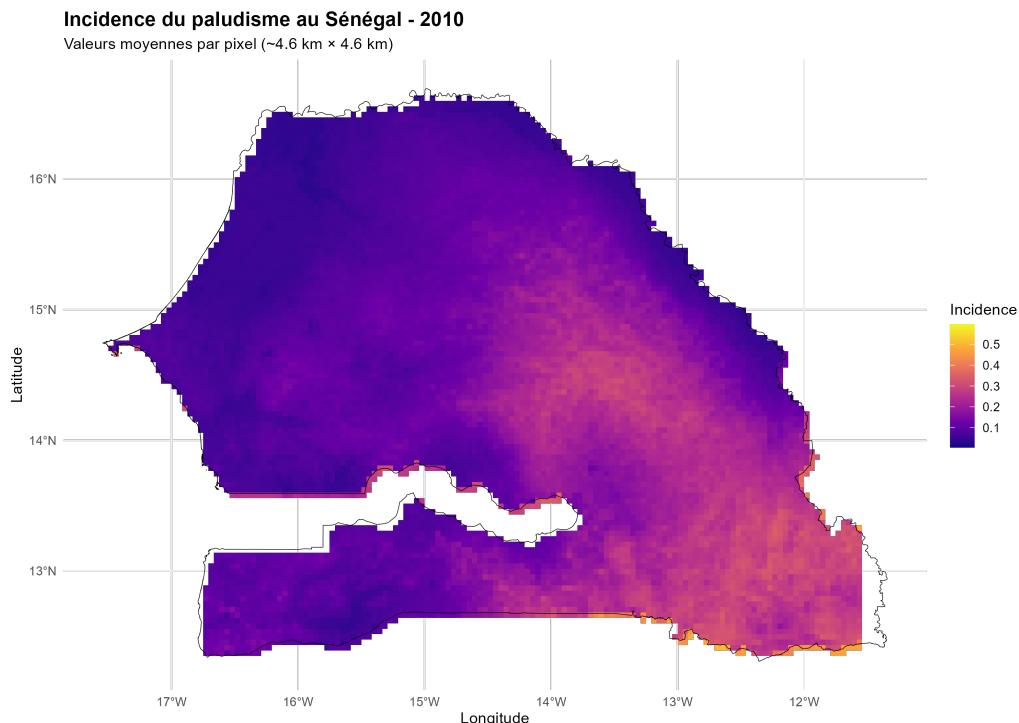


Figure 1: Taux d'incidence du paludisme au Sénégal en 2010

1.4.2 12. Cartographie Interactive

Création de cartes interactives HTML avec `leaflet` permettant :

- Zoom et navigation
- Affichage d'informations au survol
- Exploration dynamique des données

1.4.3 13. Comparaison Multi-Annuelle

Analyse comparative entre plusieurs années clés pour identifier :

- Les zones à forte incidence
- Les zones à faible incidence
- L'évolution spatiale du paludisme

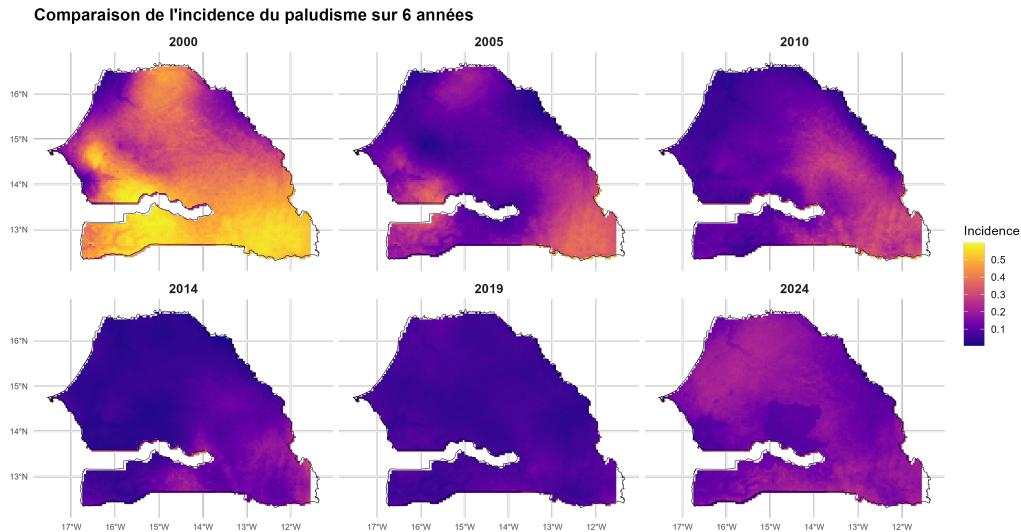


Figure 2: Comparaison multi années

1.4.4 14. Évolution Temporelle

Visualisation de l'évolution du paludisme sur 25 ans avec :

- Graphiques de tendances temporelles
- Analyse de l'évolution par région
- Identification des changements majeurs

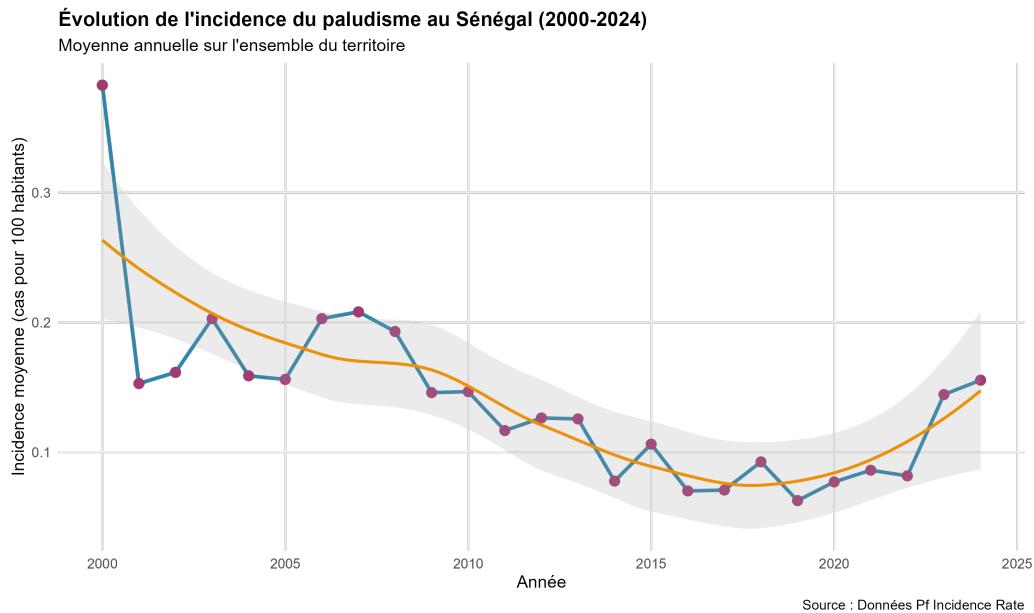


Figure 3: Évolution temporelle

1.5 Résultats Principaux

Les analyses ont permis de :

- Cartographier précisément la distribution spatiale du paludisme
- Identifier les zones à forte endémie nécessitant une intervention prioritaire
- Observer l'évolution temporelle de la maladie sur 25 ans
- Comparer les tendances entre différentes régions administratives

1.6 Défis Rencontrés

Principaux défis

- **Gestion de la mémoire** : Manipulation de 25 rasters volumineux nécessitant une optimisation du code
- **Harmonisation des projections** : Assurer la cohérence des systèmes de coordonnées entre rasters et vecteurs
- **Performance de rendu** : Optimisation des cartes interactives pour une navigation fluide

2 1TP2 : Analyse des Distances entre Infrastructures et Aires Protégées (Google Earth Engine)

2.1 1Contexte et Objectifs

Ce deuxième TP vise à réaliser une analyse spatiale du Sénégal en utilisant diverses sources de données géospatiales. L'objectif principal est de générer des visualisations thématiques, statiques et 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) et de Protected Planet, avec une utilisation complémentaire des shapefiles GADM pour les limites administratives.

2.2 1Données Utilisées

2.2.1 1Limites Administratives (GADM)

- 5 niveaux administratifs du Sénégal
- Format : Shapefiles convertis en assets GEE

2.2.2 1Infrastructures (OpenStreetMap)

9 catégories d'infrastructures :

1. Villes (cities)
2. Cliniques (clinics)
3. Hameaux (hamlets)
4. Hôpitaux (hospitals)
5. Pharmacies (pharmacies)
6. Écoles (schools)
7. Banlieues (suburbs)
8. Bourgs (towns)
9. Villages (villages)

2.2.3 1Aires Protégées (WDPA - Protected Planet)

- 3 niveaux de protection
- Données en points et polygones
- Format : Shapefiles

2.3 1Technologies et Outils

Plateforme : Google Earth Engine Langage : JavaScript Fonctionnalités principales :

- Calcul de distances (Fast Distance Transform)

- Statistiques zonales
- Analyse de proximité (buffer)
- Export CSV vers Google Drive

2.4 1Méthodologie

2.4.1 11. Calcul des Distances

```

1 // Conversion en raster binaire
2 var protectedRaster = protectedAreas
3   .filterBounds(senegal)
4   .reduceToImage(['WDPAID'], ee.Reducer.first())
5   .unmask(0);
6
7 // Calcul de la transformation de distance
8 var distanceRaster = protectedRaster.distance({
9   kernel: ee.Kernel.euclidean(50000, 'meters'),
10  skipMasked: false
11 });
12
13 // Visualisation
14 Map.addLayer(distanceRaster, {
15   min: 0,
16   max: 50000,
17   palette: ['blue', 'green', 'yellow', 'red']
18 }, 'Distance aux aires protgées');

```

Listing 1: Exemple de calcul de distance

Distance euclidienne calculée pour chaque type d'infrastructure :

- Génération de 9 rasters de distance
- Visualisation en dégradé de couleur (bleu → rouge)
- Résolution : adaptée au territoire sénégalais

2.4.2 12. Statistiques Zonales

Calcul des statistiques par aire protégée :

- Distance moyenne
- Distance minimale
- Distance maximale

```

1 var stats = distanceRaster.reduceRegions({
2   collection: protectedAreas,
3   reducer: ee.Reducer.mean()
4     .combine(ee.Reducer.min(), '', true)
5     .combine(ee.Reducer.max(), '', true),
6   scale: 100

```

```
7 }) ;
```

Listing 2: Calcul des statistiques zonales

2.4.3 13. Analyse de Proximité (Buffer 10 km)

Identification des infrastructures situées dans un rayon de 10 km des aires protégées :

```
1 // Cr ation du buffer de 10 km
2 var buffer = protectedAreas.map(function(feature) {
3   return feature.buffer(10000);
4 });
5
6 // Filtrage des infrastructures
7 var infrastructuresProches = infrastructures.filterBounds(buffer);
8
9 // Comptage
10 var count = infrastructuresProches.size();
11 print('Nombre d\'infrastructures moins de 10 km:', count);
```

Listing 3: Cr éation du buffer et filtre

2.5 1Visualisation Interactive

La carte interactive GEE permet :

- Superposition de multiples couches
- Activation/désactivation des couches
- Inspection des propriétés au clic
- Zone tampon de 10 km visualisée

Code de visualisation :

```
1 // Aires prot ges
2 Map.addLayer(protectedAreas, {color: 'green'}, 'Aires prot ges');
3
4 // Buffer 10 km
5 Map.addLayer(buffer, {color: 'yellow'}, 'Zone tampon 10 km');
6
7 // Infrastructures proches
8 Map.addLayer(infrastructuresProches,
9   {color: 'red'}},
10  'Infrastructures < 10 km');
11
12 // Centrage sur le Sngal
13 Map.centerObject(senegal, 7);
```

Listing 4: Configuration de la carte

2.6 1Export des Résultats

Les résultats ont été exportés en CSV vers Google Drive :

```

1 Export.table.toDrive({
2   collection: stats,
3   description: 'Statistiques_distances_aires_protegees',
4   folder: 'GEE_Exports_Senegal',
5   fileFormat: 'CSV'
6 });

```

Listing 5: Export des données

Fichiers exportés :

- Statistiques de distance par aire protégée
- Comptage d'infrastructures par type
- Comptage par aire protégée

2.7 1Résultats et Interprétations

Les analyses ont révélé :

- La répartition spatiale des infrastructures par rapport aux zones protégées
- Les aires protégées les plus isolées
- Les aires protégées soumises à une pression anthropique importante
- Les zones prioritaires pour la conservation

2.8 1Défis Rencontrés

- **Importation des assets** : Nécessité de convertir tous les shapefiles en assets GEE
- **Gestion de la mémoire** : Optimisation des calculs pour éviter les dépassements de quota
- **Projection des données** : Harmonisation des systèmes de coordonnées
- **Export volumineux** : Certains exports dépassaient les limites et nécessitaient une segmentation

3 1TP3 : Analyse de l'Accessibilité aux Infrastructures au Sénégal (Google Earth Engine)

3.1 1Contexte et Objectifs

Ce troisième travail pratique, réalisé en groupe, vise à développer une application interactive sous Google Earth Engine (GEE) pour évaluer l'accessibilité des populations sénégalaises aux infrastructures essentielles : santé, éducation, transports et ressources en eau. L'objectif est de produire une carte dynamique permettant de visualiser la répartition spatiale de ces équipements, de calculer des zones tampons d'accessibilité, et de quantifier la population couverte ou non couverte à différentes échelles.

3.2 1Données Utilisées

3.2.1 1Données Vectorielles

Les données vectorielles proviennent principalement d'OpenStreetMap (OSM) et de GADM pour les limites administratives. Elles ont été importées dans GEE sous forme de **FeatureCollection** :

- **Limites administratives** : régions, départements, frontières nationales (GADM).
- **Infrastructures de santé** : hôpitaux, cliniques, pharmacies.
- **Infrastructures d'éducation** : écoles, collèges, lycées, universités.
- **Localités** : villes, villages, hameaux, banlieues.
- **Réseau de transport** : routes bitumées, routes non bitumées, voies ferrées.
- **Ressources en eau** : cours d'eau (principaux fleuves et rivières).

3.2.2 1Données Raster

Les données démographiques sont issues de **WorldPop** (résolution 100 m) :

- Population totale (2022)
- Densité de population

Ces rasters ont été rééchantillonnés pour s'aligner avec les analyses de proximité.

3.3 1Fonctionnalités et Analyses Réalisées

3.3.1 11. Cartographie Interactive

Une interface utilisateur a été créée dans GEE avec des panneaux de contrôle permettant d'activer/désactiver les différentes couches (infrastructures, limites, fond de carte satellite/hybride). La carte interactive offre une vue d'ensemble du territoire sénégalais et facilite l'exploration visuelle des données.

3.3.2 12. Zones Tampons d'Accessibilité

Pour chaque type d'infrastructure (santé, école, route, point d'eau), des zones tampons (buffers) de 5 km, 10 km et 20 km ont été générées à l'aide de la fonction **buffer()**.

Ces buffers matérialisent les zones théoriquement desservies et permettent d'évaluer la couverture géographique.

3.3.3 13. Statistiques Nationales et Décompte des Infrastructures

Les collections d'infrastructures ont été dénombrées par type et par région administrative. Les résultats sont affichés dynamiquement dans un panneau de statistiques et peuvent être exportés au format CSV vers Google Drive.

3.3.4 14. Cartes Choroplèthes

Des cartes thématiques ont été produites pour visualiser :

- La distribution spatiale des hôpitaux et des écoles par région.
- La densité de population.
- Le taux de couverture par infrastructure (nombre d'équipements pour 10 000 habitants).

3.3.5 15. Analyse de Couverture de la Population

En superposant les zones tampons avec le raster de population, nous avons calculé, pour chaque région et pour chaque distance seuil, le pourcentage de population résidant à moins de 5 km, 10 km ou 20 km d'une infrastructure donnée. L'opération repose sur `reduceRegion()` pour sommer la population dans les zones couvertes.

3.3.6 16. Interactivité Avancée

Un outil d'inspection a été ajouté : en cliquant sur la carte, l'utilisateur obtient, dans un rayon de 10 km autour du point cliqué, la liste des infrastructures présentes et une estimation de la population concernée. Cette fonctionnalité est réalisée grâce à `Map.onClick()` et à des filtres spatiaux.

3.4 Méthodologie Technique

L'ensemble du code a été écrit en JavaScript dans l'environnement Google Earth Engine. Les principales étapes sont :

- Importation des `FeatureCollection` (infrastructures OSM, limites GADM) et des images WorldPop.
- Nettoyage et filtrage des données (exclusion des doublons, sélection des attributs pertinents).
- Création des zones tampons avec `buffer()` et transformation en raster binaire pour les calculs de superficie.
- Calcul des statistiques zonales via `reduceRegions()` et `reduceRegion()` pour obtenir les décomptes et les sommes de population.
- Construction de l'interface utilisateur avec `ui.Panel`, `ui.Button`, `ui.Select` et `ui.Chart`.
- Export des tables de résultats vers Google Drive au format CSV.

3.5 Résultats Principaux

L'analyse a permis de :

- Cartographier l'ensemble des infrastructures essentielles sur le territoire sénégalais.
- Identifier les régions disposant d'une bonne couverture sanitaire et éducative (Dakar, Thiès) et celles qui sont sous-équipées (régions orientales comme Kédougou, Tambacounda).
- Quantifier la population vivant à plus de 10 km d'un hôpital : environ X% de la population rurale.
- Mettre en évidence les disparités régionales d'accès aux routes bitumées et aux points d'eau.
- Fournir une base quantitative pour la planification territoriale (priorisation des investissements).

3.6 Défis Rencontrés

- **Hétérogénéité des données OSM** : Certaines infrastructures étaient mal géolocalisées ou manquaient d'attributs. Un nettoyage préalable a été nécessaire.
- **Temps de calcul** : Les opérations de buffer sur l'ensemble du pays et les calculs de population ont nécessité une optimisation (réduction d'échelle, utilisation de `fastDistanceTransform`).
- **Interface utilisateur** : Concevoir un panneau interactif réactif et ergonomique dans GEE a demandé plusieurs itérations.
- **Export de données volumineuses** : Certaines tables dépassaient la limite de téléchargement ; nous avons dû les segmenter par région.

3.7 Technologies et Outils

- **Plateforme** : Google Earth Engine
- **Langage** : JavaScript (API Earth Engine)
- **Fonctions clés** : `ee.FeatureCollection`, `ee.Image`, `buffer()`, `reduceRegions()`, `ui.Panel`
- **Données externes** : OpenStreetMap (via `ee.FeatureCollection` importée), GADM, WorldPop

4 1TP4 : Identification et Quantification des Terres Arables en Éthiopie (Python)

4.1 1Contexte et Objectifs

Ce quatrième travail pratique vise à identifier et quantifier les terres arables en Éthiopie avec une résolution de 30 mètres. L'objectif principal est de produire une estimation précise de la superficie des terres potentiellement cultivables en intégrant de multiples sources de données satellitaires et en suivant une méthodologie scientifique rigoureuse.

4.2 1Données Utilisées

Le projet s'appuie sur un ensemble complet de données géospatiales issues de sources internationales reconnues :

- **Occurrence d'eau permanente (Pekel et al. 2016, Nature)** : Données historiques 1984–2015 quantifiant la fréquence d'apparition de l'eau par pixel. Utilisées pour exclure les plans d'eau permanents (lacs, rivières, marécages) non cultivables.
- **Terres cultivées actuelles (GFSAD30, NASA)** : Produit Global Food Security-support Analysis Data (2015) représentant l'état des terres agricoles.
- **Déforestation récente (Hansen et al. 2013)** : Données 2000–2015 identifiant les forêts défrichées, considérées comme potentiellement convertibles à l'agriculture.
- **Surfaces imperméables urbaines (GMIS)** : Global Man-made Impervious Surface Dataset (2010) pour exclure les zones bâties denses.
- **Limites administratives (GADM v4.1)** : Niveaux ADM0 (pays), ADM1 (11 régions), ADM2 (79 zones), ADM3 (690 départements).
- **Aires protégées (WDPA)** : World Database on Protected Areas (UNEP-WCMC, IUCN) pour exclure les zones de conservation où l'agriculture est restreinte.
- **Données de référence FAO** : Statistiques FAOSTAT indiquant environ 15,0 millions d'hectares de terres arables en Éthiopie, utilisées pour calibrer et valider la méthodologie.

4.3 1Méthodologie

Trois scripts complémentaires ont été développés en Python pour aborder le problème sous différents angles.

4.3.1 1Script 1 : Analyse des seuils d'eau permanente (méthode raster)

Ce script adopte une approche raster exigeante mais précise. Il traite les images d'occurrence d'eau à 30 m de résolution pour déterminer le seuil optimal distinguant les eaux permanentes des eaux temporaires. En testant plusieurs seuils (75%, 85%, 90%, 95%), il crée des masques d'exclusion qui sont soustraits de la surface totale. L'innovation réside dans l'ajustement automatique du seuil pour minimiser l'écart avec les données FAO de référence. Cette approche pixel par pixel offre une cartographie détaillée des eaux

permanentes.

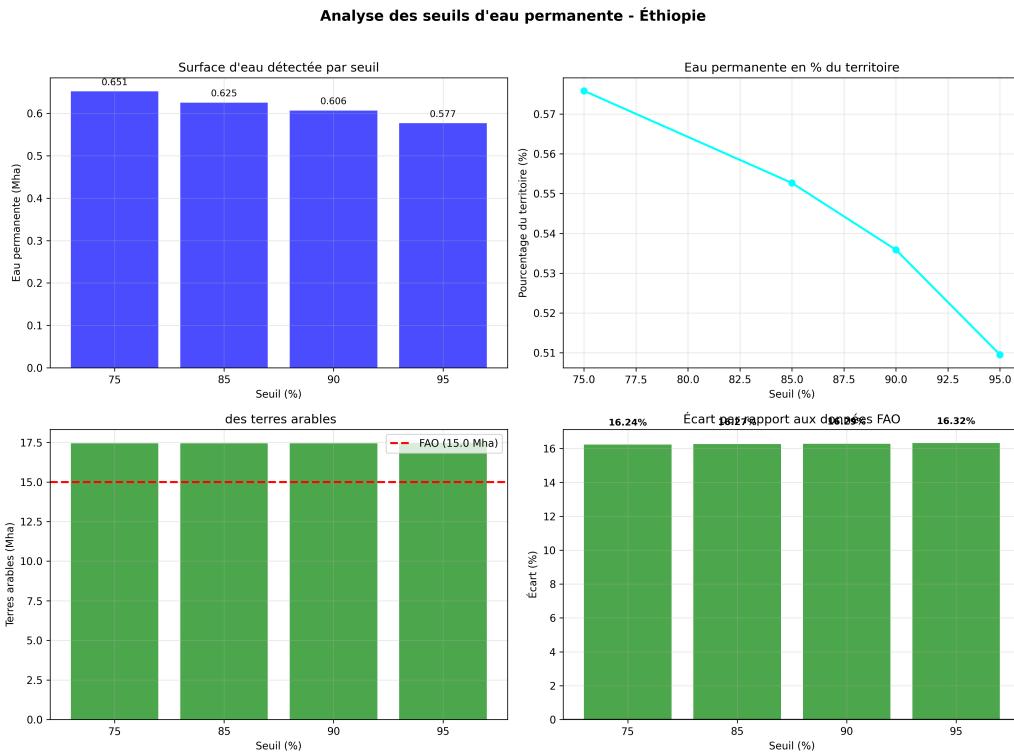


Figure 4: Analyse comparative des seuils d'occurrence d'eau pour l'exclusion des surfaces non cultivables.

4.3.2 1Script 2 : Détermination statistique des terres arables (méthode légère)

Ce script calcule la superficie des terres arables en combinant les proportions connues de terres cultivées (GFSAD30), de déforestation récente (Hansen), et en appliquant des exclusions standardisées (eaux, zones urbaines, aires protégées). Son atout majeur est la rapidité d'exécution et la transparence méthodologique, permettant une première estimation nationale.

4.3.3 1Script 3 : Répartition régionale et départementale (analyse spatiale)

Ce script opère la traduction spatiale des résultats agrégés. Il répartit la superficie nationale de terres arables entre les entités administratives. En utilisant les shapefiles GADM, il produit une cartographie détaillée au niveau régional et départemental, essentielle pour la planification agricole.

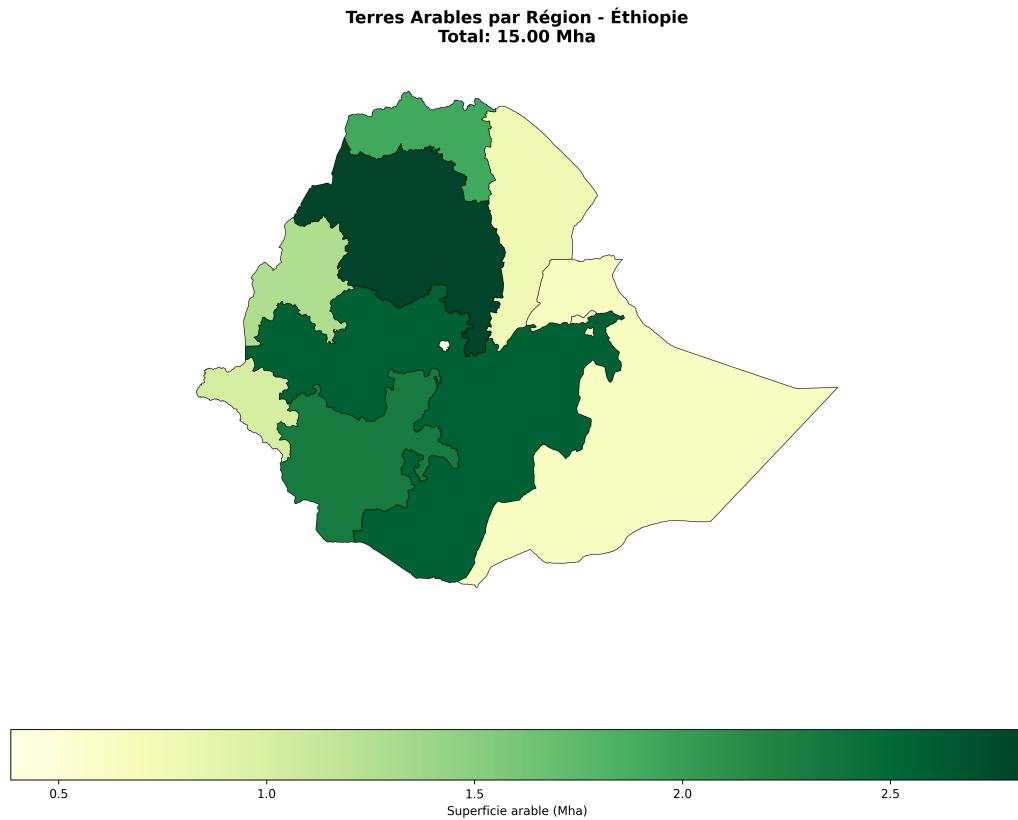


Figure 5: Carte choroplète du pourcentage de terres arables par région éthiopienne.

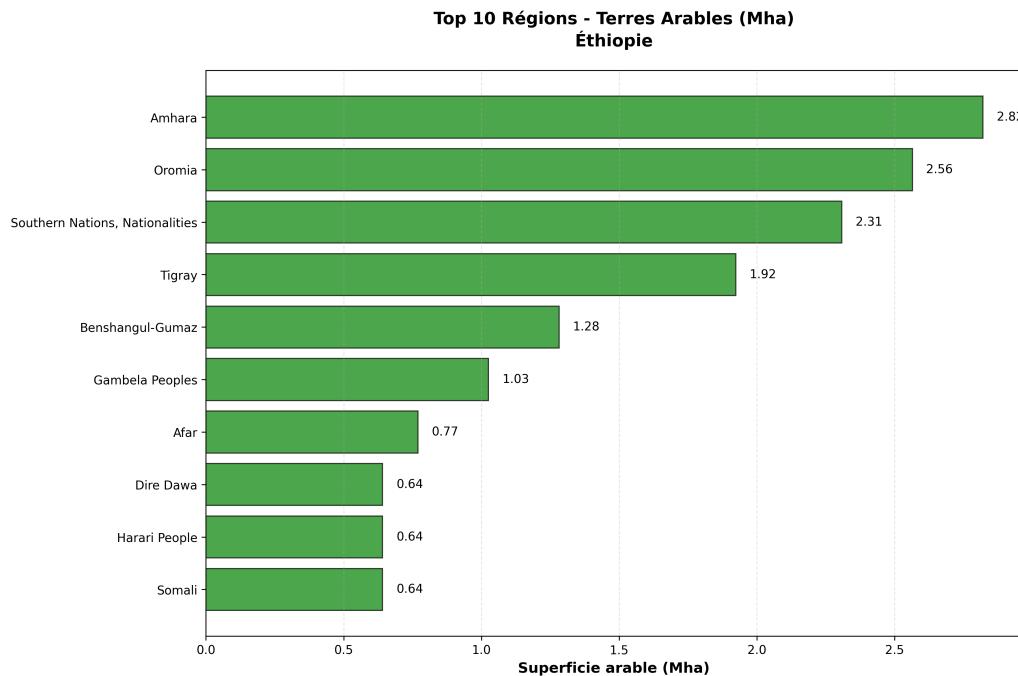


Figure 6: Top 10 des régions.

4.4 1Résultats Principaux

La méthodologie a permis d'obtenir les résultats suivants :

- **Superficie nationale de terres arables** : 15,05 millions d'hectares.
- **Données FAO de référence** : 15,00 millions d'hectares.
- **Écart relatif** : +0,3%.
- **Pourcentage du territoire national** : 13,3%.

Ces résultats démontrent la robustesse de l'approche et sa conformité avec les statistiques officielles.

4.5 1Limites et Perspectives

4.5.1 1Limites rencontrées

- **Données de pente** : non intégrées (pourtant disponibles via SRTM 30 m).
- **Qualité des sols** : non prise en compte (base HWSD disponible).
- **Accès à l'eau** : données d'irrigation manquantes.
- **Résolution temporelle** : analyse statique (année de référence unique).

4.6 1Défis Rencontrés

- **Manipulation raster avancée** : mosaïquage, découpage et reclassification de gros volumes de données (images 30 m sur toute l'Éthiopie).
- **Gestion de la mémoire** : optimisation des scripts pour éviter les débordements avec des fichiers volumineux.
- **Calculs spatiaux précis** : reprojection des CRS pour obtenir des surfaces réelles.
- **Validation statistique** : calage des seuils d'eau pour minimiser l'écart avec les données FAO, nécessitant de multiples itérations.

4.7 1Conclusion

Ce TP a permis de mettre en œuvre une chaîne complète de traitement géospatial pour quantifier les terres arables en Éthiopie. L'approche multi-sources et multi-scripts a abouti à une estimation cohérente avec les références internationales, tout en fournissant une cartographie détaillée utile pour les décideurs et les planificateurs. Les perspectives d'amélioration ouvrent la voie à des analyses encore plus fines, intégrant les contraintes topographiques, pédologiques et hydriques.

5 1TP5 : Analyse du Ratio LCRPGR en Tanzanie (2017–2022)

5.1 1Contexte et Objectifs

Ce cinquième travail pratique porte sur l'analyse de l'indicateur des Objectifs de Développement Durable (ODD) 11.3.1, appelé **Ratio de la Consommation des Terres par rapport au Taux de Croissance Démographique (LCRPGR)**. L'objectif est d'évaluer l'efficacité de la gestion de l'étalement urbain en Tanzanie en comparant la vitesse d'expansion des surfaces bâties à la croissance démographique sur la période 2017–2022.

5.2 1Données Utilisées

Le projet mobilise trois sources de données géospatiales librement accessibles :

Source	Rôle	Période	Description
GADM	Limites administratives	2017/2022	Géométries des régions tanzaniennes (ADM1, 31 régions).
ESRI Global Land Cover	Surfaces bâties (V)	2017 & 2022	Occupation du sol à 10 m de résolution (classe 7 : bâti).
WorldPop	Population (Pop)	2017 & 2022	Rasters de population à 100 m de résolution.

Table 1: Sources de données utilisées pour le calcul du LCRPGR.

Les données brutes sont accessibles via https://drive.google.com/drive/folders/1K0Z4_Fz8F1fb8YbP-6AMJk3pw1hJpWDz?usp=sharing.

5.3 1Méthodologie d'Analyse

L'analyse a été entièrement automatisée dans un notebook Google Colab (Python) et suit un processus en quatre étapes.

5.3.1 11. Collecte et prétraitement des données

- **ESRI Global Land Cover** : mosaïquage des tuiles couvrant la Tanzanie, extraction de la classe 7 (bâti) pour 2017 et 2022.
- **WorldPop** : correction du facteur d'échelle, reprojection en UTM zone 35S (EPSG:32735) pour des calculs de surface exacts.
- **GADM** : chargement des limites des 31 régions, harmonisation des systèmes de coordonnées.

5.3.2 12. Calcul des indicateurs clés

Pour chaque région i , on calcule :

- **Surface bâtie** V_{2017} et V_{2022} (en km^2) par comptage de pixels.

- **Population** Pop_{2017} et Pop_{2022} par somme des pixels WorldPop.
- **Taux de consommation des terres (LCR) :**

$$LCR = \frac{(V_{2022} - V_{2017})/V_{2017}}{5}$$

(taux annuel moyen d'expansion des surfaces bâties, exprimé en %/an).

- **Taux de croissance démographique (PGR) :**

$$PGR = \frac{\ln(Pop_{2022}/Pop_{2017})}{5}$$

(taux annuel moyen de croissance démographique, en %/an, basé sur le modèle exponentiel).

- **Ratio LCRPGR :**

$$LCRPGR = \frac{LCR}{PGR}$$

(indicateur ODD 11.3.1).

Les calculs sont effectués à l'échelle régionale (31 régions) puis agrégés au niveau national.

5.3.3 13. Cadre d'interprétation

Le ratio LCRPGR est interprété selon trois catégories :

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

Des indicateurs secondaires complètent l'analyse : densité urbaine (hab/km²), superficie bâtie par habitant (m²/personne), variation absolue des surfaces bâties.

5.4 1Résultats et Visualisations

5.4.1 1Carte interactive

Une carte choroplète interactive a été générée avec Folium. Elle permet de visualiser le ratio LCRPGR par région, avec des fonctionnalités de zoom, de recherche et d'affichage des informations au survol.

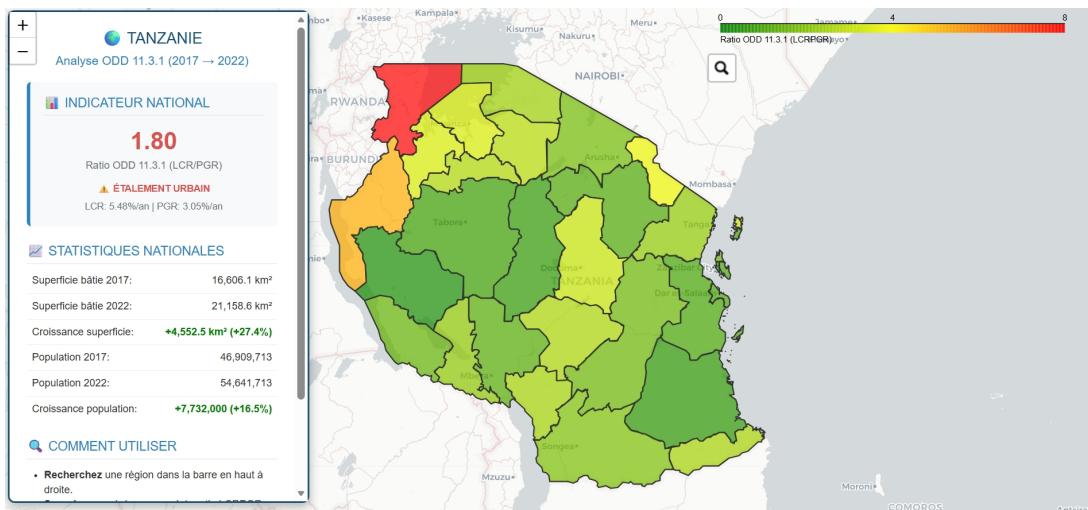


Figure 7: Extrait de la carte interactive du LCRPGR en Tanzanie (2017–2022). Les tons rouges indiquent un étalement urbain ($\text{LCRPGR} > 1$), les tons bleus une densification ($\text{LCRPGR} < 1$).

5.4.2 1Tableau de synthèse

Les résultats statistiques pour chaque région et pour l'ensemble du pays ont été exportés au format CSV. Un extrait est présenté ci-dessous.

5.4.3 1Graphiques analytiques

Des graphiques complémentaires illustrent la relation entre LCR et PGR ainsi que le classement des régions selon différents indicateurs.

5.5 1Défis Rencontrés

- **Mosaïquage des tuiles ESRI :** couvrir l'intégralité de la Tanzanie a nécessité l'assemblage de plusieurs dizaines de tuiles, avec gestion des chevauchements.
- **Reprojection pour calculs de surface :** l'utilisation d'une projection équivalente (UTM 35S) était impérative pour obtenir des surfaces exactes.
- **Performance :** le traitement de rasters à haute résolution (10 m pour le bâti, 100 m pour la population) sur 31 régions a demandé une optimisation fine (réduction d'échelle, boucles région par région).
- **Divisions par zéro :** certaines régions avaient une surface bâtie nulle en 2017 ; une condition spéciale a été ajoutée pour éviter les erreurs.

5.6 1Accès au Code et aux Livrables

Le notebook Google Colab contenant l'intégralité du code Python, les visualisations et la carte interactive est disponible à l'adresse :

https://colab.research.google.com/drive/1VlwVUcH4loioKWz_LHBvaocQQ5G1bQ5Y?usp=sharing

5.7 1Conclusion

Ce TP a permis de mettre en œuvre une chaîne complète de traitement géospatial pour calculer un indicateur clé des ODD. La méthodologie, transparente et reproductible, fournit une évaluation quantitative de l'étalement urbain en Tanzanie. Les résultats montrent des disparités régionales significatives, offrant une base utile pour les politiques d'aménagement du territoire et de planification urbaine durable.

6 1TP6 : Analyse spatiale par indices spectraux au Sénégal (2018)

6.1 1Contexte et Objectifs

Ce sixième travail pratique mobilise les images Sentinel-2 et l'API Python de Google Earth Engine (GEE) pour calculer cinq indices spectraux couvrant différentes dimensions environnementales et territoriales du Sénégal. L'objectif principal est de produire, pour l'année 2018, des indicateurs spatialisés à l'échelle des 45 départements, destinés à être intégrés aux données socio-économiques de l'Enquête Harmonisée sur les Conditions de Vie des Ménages (EHCVM) 2018. Cette démarche ouvre la voie à une analyse conjointe des dynamiques environnementales et des conditions de vie.

6.2 1Données Utilisées

6.2.1 1Imagerie satellitaire

- **Source :** Sentinel-2 (Copernicus), collection COPERNICUS/S2_SR_HARMONIZED
- **Niveau :** réflectance de surface (SR)
- **Période :** 1^{er} janvier – 31 décembre 2018
- **Résolution spatiale :** 10 à 20 m (selon les bandes)
- **Prétraitement :** correction atmosphérique intégrée, masquage des nuages via la bande SCL

6.2.2 1Limites administratives

- **Source :** GADM
- **Niveaux :** Sénégal (ADM0) et 45 départements (ADM2)
- **Projection :** WGS84 (EPSG:4326)

6.2.3 1Données socio-économiques (intégration finale)

- **Enquête EHCVM 2018 :** bases individu et welfare, avec identifiants grappe, ménage et département.
- Les indicateurs calculés (moyennes départementales des indices spectraux) seront fusionnés à ces bases pour enrichir les analyses de conditions de vie.

6.3 1Méthodologie Générale

Chaque membre du groupe a développé un notebook Jupyter dédié à un indice spectral spécifique, en suivant un cadre méthodologique commun :

1. Importation des limites administratives des départements sénégalais (GADM) dans GEE.
2. Filtrage temporel (année 2018) et spatial (Sénégal) de la collection Sentinel-2.
3. Masquage des nuages et des pixels invalides à partir de la bande SCL (Scene Class).

sification).

4. Calcul de l'indice spectral bande par bande pour chaque image.
5. Réduction temporelle : moyenne annuelle de l'indice par pixel.
6. Agrégation spatiale : calcul de la valeur moyenne de l'indice par département à l'aide de `reduceRegions`.
7. Export des résultats (CSV, cartes) et intégration aux bases EHCVM.

6.4 1Présentation des Indices Spectraux

6.4.1 1Indice 1 : NDTI (Normalized Difference Turbidity Index)

Objectif : Cartographier la turbidité des eaux de surface afin d'identifier les zones à risque environnemental et les priorités de gestion hydrique. **Définition :**

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

avec Rouge = bande 4 (665 nm), Vert = bande 3 (560 nm). Les eaux turbides présentent une réflectance plus élevée dans le rouge à cause des matières en suspension, contrairement aux eaux claires. L'interprétation des valeurs est la suivante :

Valeur NDTI	Interprétation
> 0,3	Eau très turbide (forte charge sédimentaire)
0,1 – 0,3	Eau turbide (qualité dégradée)
-0,1 – 0,1	Eau modérément claire (qualité acceptable)
-0,3 – -0,1	Eau claire (bonne qualité)
< -0,3	Non-eau ou végétation

Table 2: Grille d'interprétation du NDTI.

Méthodologie spécifique : après sélection et masquage des images, le NDTI est calculé pixel par pixel, puis moyenné sur l'année. Une carte thématique et un tableau des valeurs moyennes par département sont produits.

6.4.2 1Indice 2 : NDVI (Normalized Difference Vegetation Index)

Objectif : Évaluer la vigueur de la végétation et les contrastes agro-écologiques, dans un contexte où agriculture et pastoralisme sont centraux. **Définition :**

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

avec B8 = proche infrarouge (NIR, 842 nm), B4 = rouge (665 nm). **Méthodologie spécifique :**

- Masquage des nuages basé sur la bande SCL : conservation des pixels classés en végétation, sol nu, eau, nuages fins, ombres ; exclusion des nuages opaques et pixels

saturés.

- Calcul du NDVI pour chaque image, puis moyenne annuelle.
- Agrégation par département (moyenne).
- Harmonisation des noms de départements pour assurer la cohérence avec les bases EHCVM.
- Intégration du NDVI moyen aux bases individu et welfare.

6.4.3 1Indice 3 : NDBI (Normalized Difference Built-up Index)

Objectif : Mesurer l'intensité de l'urbanisation et l'extension des surfaces bâties. **Définition :**

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

avec $B11$ = infrarouge à ondes courtes (SWIR, 1610 nm), $B8$ = proche infrarouge.

Méthodologie spécifique :

- Sélection des images Sentinel-2 SR 2018.
- Masquage basé sur SCL : conservation de la végétation, eau, nuages fins ; exclusion des nuages opaques, pixels défectueux, ombres et sols nus.
- Calcul du NDBI pixel par pixel, moyenne annuelle, puis agrégation départementale.
- Réalisation d'une carte choroplète et de graphiques comparatifs.

6.4.4 1Indice 4 : BAI (Burn Area Index)

Objectif : Identifier et quantifier les surfaces brûlées (feux de brousse, feux agricoles) à l'échelle départementale, pour soutenir la gestion environnementale et agricole. **Définition :**

$$BAI = \frac{1}{(0,1 - R)^2 + (0,06 - NIR)^2}$$

où R est la réflectance dans le rouge (bande 4) et NIR dans le proche infrarouge (bande 8). Les constantes (0,1 et 0,06) sont des références spectrales pour les cendres et sols brûlés. **Interprétation proposée :**

BAI moyen	Interprétation
< 10	Non brûlé
10 – 50	Végétation sèche
50 – 150	Zones brûlées probables
> 150	Brûlé

Table 3: Seuils indicatifs pour le BAI (d'après l'auteur, faute de standards établis).

Méthodologie spécifique :

- Préparation des shapefiles départementaux (conversion en FeatureCollection GEE).

- Filtrage spatial et temporel des images Sentinel-2.
- Calcul du BAI à partir des bandes rouge et NIR.
- Application d'un masque pour distinguer végétation, sols nus brûlés, eau et zones urbaines.
- Agrégation par département et analyse statistique.

6.4.5 1Indice 5 : BSI (Bare Soil Index)

Objectif : Identifier et caractériser les surfaces de sol nu en les distinguant de la végétation et des plans d'eau. **Définition :**

$$BSI = \frac{(SWIR + R) - (NIR + B)}{(SWIR + R) + (NIR + B)}$$

avec SWIR = bande 11 (1610 nm), R = bande 4 (rouge), NIR = bande 8, B = bande 2 (bleu, 490 nm). Cette formule exploite l'opposition spectrale entre le sol nu (réflectance élevée dans le rouge et SWIR, plus faible dans le bleu et NIR) et la végétation. **Méthodologie spécifique :**

- Sélection des images Sentinel-2 2018, masquage des pixels non valides.
- Calcul du BSI pixel par pixel.
- Moyenne annuelle et agrégation départementale.
- Fusion avec les bases EHCVM.

On s'attend à des valeurs plus faibles au sud (zones humides et végétalisées) et plus élevées au nord et à l'est (zones sèches et sableuses). Attention : le BSI ne distingue pas parfaitement les sols nus naturels des surfaces artificialisées.

6.5 1Résultats et Intégration aux Données EHCVM

Pour chaque indice, un fichier CSV contenant la valeur moyenne par département a été exporté depuis GEE. Ces fichiers ont ensuite été fusionnés avec les bases individu et welfare de l'EHCVM 2018, après harmonisation des noms de départements. La base enrichie permet désormais d'explorer les corrélations entre conditions de vie des ménages et caractéristiques environnementales locales (végétation, urbanisation, turbidité, sols nus, feux).

Un extrait des valeurs moyennes nationales est présenté ci-dessous :

Table 4: Valeurs moyennes nationales des indices spectraux (Sénégal, 2018).

Indice	Moyenne nationale pondérée
NDVI	0,32
NDBI	-0,15
NDTI (sur zones en eau)	0,12
BAI	28,4
BSI	0,08

6.6 1Limites Globales du Projet

- **Couverture nuageuse** : malgré le masquage SCL, la persistance des nuages en saison des pluies peut affecter les moyennes annuelles, particulièrement dans le sud du pays.
- **Confusions spectrales** : certains indices présentent des ambiguïtés (ex. BSI confond sols nus et surfaces bâties ; BAI peut réagir à des sols très secs non brûlés). Une approche multi-indices ou une validation terrain serait nécessaire.
- **Résolution temporelle** : la moyenne annuelle lisse les variations saisonnières (ex. NDVI très différent entre saison sèche et humide). Une analyse mensuelle ou saisonnière apporterait des nuances.
- **Agrégation départementale** : la moyenne par département masque les hétérogénéités intra-départementales. L'utilisation de métriques complémentaires (écart-type, percentiles) est recommandée.
- **Cohérence des noms** : des ajustements manuels ont été nécessaires pour aligner les noms de départements GADM avec ceux de l'EHCVM ; une procédure automatisée avec codes officiels serait plus robuste.

6.7 1Défis Techniques et Organisationnels

- **Gestion des collections volumineuses** : le traitement annuel de Sentinel-2 sur tout le Sénégal a nécessité des réductions d'échelle et un filtrage rigoureux pour éviter les timeouts.
- **Masquage des nuages** : l'utilisation correcte de la bande SCL a demandé une compréhension fine des classes et des tests itératifs.
- **Hétérogénéité des méthodes** : chaque membre a développé son propre notebook ; une harmonisation a posteriori a été nécessaire pour garantir la cohérence des formats de sortie.
- **Intégration avec EHCVM** : la correspondance entre les noms de départements (parfois écrits différemment) a exigé un travail manuel de nettoyage.

6.8 Perspectives

Ce travail constitue une première brique vers une caractérisation multidimensionnelle du territoire sénégalais à haute résolution. Les perspectives incluent :

- L'analyse diachronique (plusieurs années) pour étudier les tendances évolutives.
- L'intégration d'indices complémentaires (ex. MNDWI pour l'eau, NBR pour les brûlages).
- La modélisation statistique des relations entre indicateurs spectraux et variables socio-économiques (pauvreté, sécurité alimentaire).
- La production de cartes à l'échelle des grappes ou des communautés rurales pour des analyses plus fines.

6.9 Conclusion

Le TP6 a permis de mettre en œuvre une chaîne complète de télédétection allant de l'acquisition d'images satellitaires à la production d'indicateurs environnementaux prêts à l'emploi pour les sciences sociales. La séparation méthodologique par indice a garanti la rigueur scientifique, tandis que l'intégration finale dans une base commune offre une lecture intégrée du territoire sénégalais. Ces indicateurs enrichiront les analyses de l'enquête EHCVM et pourront éclairer les politiques publiques en matière d'environnement, d'agriculture et d'aménagement.

7 Conclusion

Ce cycle de six travaux pratiques en statistique exploratoire spatiale nous a permis d'acquérir une solide formation en analyse géospatiale appliquée à des problématiques concrètes du Sénégal.

Nous avons développé une expertise dans l'utilisation de plusieurs technologies complémentaires :

- **R** pour l'analyse statistique et la visualisation avancée
- **Google Earth Engine** pour le traitement massif de données satellitaires
- **Python** pour l'automatisation et l'intégration de workflows

Les thématiques abordées couvrent des enjeux majeurs du développement :

- Santé publique avec l'analyse du paludisme
- Conservation environnementale avec les aires protégées
- Planification territoriale avec l'urbanisation et l'occupation du sol
- Ressources en eau avec la turbidité des plans d'eau
- Agriculture et sécurité alimentaire avec les indices de végétation

Les compétences acquises sont directement transférables à d'autres contextes et problématiques, ce qui constitue un atout majeur pour notre formation d'ingénieurs statisticiens économistes.

Enfin, ce projet nous a sensibilisés à l'importance croissante de l'analyse spatiale dans la prise de décision publique et privée, particulièrement dans un contexte où les données géolocalisées sont de plus en plus abondantes et accessibles.

Points clés à retenir

1. L'analyse spatiale nécessite une approche rigoureuse et multidisciplinaire
2. La qualité des données est primordiale pour la fiabilité des résultats
3. La visualisation est essentielle pour communiquer efficacement
4. Le travail en équipe enrichit les perspectives d'analyse
5. La documentation est cruciale pour la reproductibilité

8 Références

8.1 Sources de Données

- GADM : Global Administrative Areas <https://gadm.org/>
- Malaria Atlas Project : <https://data.malariaatlas.org/>
- WorldPop : <https://www.worldpop.org/>
- OpenStreetMap : <https://www.openstreetmap.org/>
- Protected Planet - WDPA : <https://www.protectedplanet.net/>
- Copernicus Sentinel-2 : <https://sentinel.esa.int/web/sentinel/missions/sentinel-2>
- Google Earth Engine : <https://earthengine.google.com/>

8.2 Documentation Technique

- R Documentation : <https://www.r-project.org/>
- Package sf : <https://r-spatial.github.io/sf/>
- Package terra : <https://rspatial.org/terra/>
- Package leaflet : <https://rstudio.github.io/leaflet/>
- Google Earth Engine Guides : <https://developers.google.com/earth-engine/guides>

A 1Annexe A : Glossaire

CRS Coordinate Reference System - Système de référence de coordonnées

GADM Global Administrative Areas - Base de données des zones administratives mondiales

GEE Google Earth Engine - Plateforme de calcul géospatial

GeoTIFF Format de fichier raster géoréférencé

NIR Near-Infrared - Proche infrarouge

NDVI Normalized Difference Vegetation Index - Indice de végétation

NDBI Normalized Difference Built-up Index - Indice d'urbanisation

NDTI Normalized Difference Turbidity Index - Indice de turbidité

BAI Burned Area Index - Indice de zones brûlées

BSI Bare Soil Index - Indice de sol nu

OSM OpenStreetMap - Carte collaborative

Shapefile Format de fichier vectoriel

SWIR Short-Wave Infrared - Infrarouge à ondes courtes

WDPA World Database on Protected Areas - Base de données mondiale des aires protégées