

Reporte Técnico Final

Análisis de Potencial Solar en Techos de Edificaciones
en Territorios PDET, Colombia

Alejandro Pinzón Fajardo
Juan José Bermúdez Palacios
Juan Manuel Díaz
Victor Peñaranda Florez

Universidad de los Andes
Administración de Bases de Datos

Instructor: Prof. Andrés Oswaldo Calderón Romero, Ph.D.

Noviembre 2025

Resumen

Este reporte presenta el análisis integral del potencial de generación de energía solar mediante instalación de paneles en techos de edificaciones en 146 municipios PDET (Programas de Desarrollo con Enfoque Territorial) de Colombia. Utilizando soluciones NoSQL (MongoDB) y procesando más de 6 millones de edificaciones de Microsoft Buildings y 16 millones de Google Open Buildings, se estimó un potencial de 151.13 km² de área útil instalable en territorios PDET. El proyecto implementó un flujo de trabajo geoespacial reproducible que integra datos abiertos, aplica factores de eficiencia técnicos (47.6 %), y genera recomendaciones estratégicas para la Unidad de Planeación Minero Energética (UPME). Los resultados identifican 5 municipios prioritarios para proyectos piloto y priorizan 3 regiones PDET que concentran el 51 % del potencial solar total.

Índice

1. Introducción	4
1.1. Contexto del Proyecto	4
1.2. Objetivos del Proyecto	4
1.3. Alcance	4
2. Metodología	5
2.1. Stack Tecnológico	5
2.2. Fuentes de Datos	5
2.2.1. Microsoft Building Footprints	5
2.2.2. Google Open Buildings v3	6
2.2.3. DANE Marco Geoestadístico Nacional	6
2.2.4. Lista Oficial Municipios PDET	6
2.3. Factor de Eficiencia Solar	7
3. Diseño de Base de Datos NoSQL	7
3.1. Selección de MongoDB	7
3.2. Esquema de Datos	7
3.2.1. Colección: pdet_municipalities	7
3.2.2. Colección: microsoft_buildings	8
3.2.3. Colección: buildings_by_municipality	8
3.3. Índices Geoespaciales	9
4. Integración de Datos	9
4.1. Municipios PDET	9
4.2. Edificaciones Microsoft	9
4.3. Edificaciones Google	10
4.4. Join Espacial	10
5. Análisis Geoespacial y Cálculo de Área Útil	11
5.1. Scripts de Procesamiento	11
5.2. Pipeline de Agregación MongoDB	11
6. Resultados	12
6.1. Resultados Generales	12
6.2. Top 10 Municipios por Área Útil	12
6.3. Ranking de Regiones PDET	13
6.4. Estadísticas Descriptivas	14
7. Comparación Microsoft vs Google Buildings	15
7.1. Cobertura Geográfica	16
7.2. Cantidad de Edificaciones	16
7.3. Calidad de Datos	16
7.4. Recomendación	16
8. Recomendaciones para UPME	17
8.1. Municipios Prioritarios para Proyectos Piloto	17
8.1.1. Criterios de Selección	17
8.1.2. Top 5 Municipios Recomendados	17
8.2. Estrategia de Implementación por Fases	18
8.2.1. Fase 1: Proyectos Piloto (Meses 1-12)	18
8.2.2. Fase 2: Expansión Regional (Meses 13-24)	18

8.2.3. Fase 3: Implementación Completa (Meses 25-36)	19
8.3. Análisis Complementario Recomendado	19
8.3.1. Datos de Irradiación Solar Real	19
8.3.2. Demanda Energética Actual	19
8.3.3. Análisis Socio-económico	19
8.3.4. Validación de Campo	19
8.4. Recomendaciones Técnicas Adicionales	20
8.4.1. Refinamiento de Datos	20
8.4.2. Diferenciación por Tipo de Edificación	20
9. Limitaciones y Trabajo Futuro	20
9.1. Limitaciones del Estudio	20
9.1.1. Supuestos del Análisis	20
9.1.2. Limitaciones de Datos	21
9.1.3. Alcance de Resultados	21
9.2. Trabajo Futuro Recomendado	22
9.2.1. Mejoras Técnicas Inmediatas	22
9.2.2. Análisis Complementarios	22
9.2.3. Escalabilidad	23
10. Conclusiones	23
10.1. Cumplimiento de Objetivos	23
10.1.1. Objetivos Técnicos	23
10.1.2. Alineación con Objetivos UPME	24
10.2. Hallazgos Principales	24
10.3. Contribuciones del Proyecto	25
10.3.1. Contribuciones Técnicas	25
10.3.2. Contribuciones Estratégicas	25
10.4. Impacto Potencial	25
10.4.1. Impacto Energético	25
10.4.2. Impacto Social	25
10.5. Reflexión Final	26
11. Archivos Entregables	26
11.1. Repositorio GitHub	26
11.2. Estructura del Repositorio	26
11.3. Documentos Principales	26
11.4. Scripts de Análisis	27
11.5. Outputs de Datos	27
11.6. Visualizaciones	27
12. Referencias	27

1. Introducción

1.1. Contexto del Proyecto

La transformación rápida de la infraestructura energética en Colombia requiere enfoques innovadores para identificar ubicaciones adecuadas para el despliegue de soluciones de energía renovable. Como parte de una estrategia nacional más amplia para promover el desarrollo sostenible y mejorar el acceso a la energía en regiones desatendidas, la Oficina de Investigación e Información de la Unidad de Planeación Minero Energética (UPME) ha lanzado una iniciativa estratégica centrada en evaluar la viabilidad de la energía solar en territorios seleccionados.

Los territorios PDET (Programas de Desarrollo con Enfoque Territorial) representan áreas clave de enfoque para el desarrollo posconflicto y la mejora de infraestructura en Colombia. Al aprovechar conjuntos de datos geoespaciales de acceso abierto que contienen miles de millones de contornos de edificaciones derivados de imágenes satelitales de alta resolución, este proyecto tiene como objetivo cuantificar las superficies potenciales para la captación de energía en contextos urbanos y rurales.

1.2. Objetivos del Proyecto

Los objetivos principales de este proyecto son:

1. **Contar edificaciones** dentro de cada municipio PDET
2. **Estimar el área total de techos** adecuada para instalación de paneles solares
3. **Comparar resultados** de diferentes conjuntos de datos de edificaciones abiertas
4. **Implementar soluciones NoSQL** para almacenamiento escalable y operaciones espaciales eficientes
5. **Proporcionar recomendaciones estratégicas** para ubicaciones de granjas solares como prueba de concepto

1.3. Alcance

Este proyecto cubre:

- **Territorio:** 146 municipios PDET (de 170 oficiales) en 14 regiones
- **Área territorial:** 358,181 km²
- **Edificaciones procesadas:** 6.08M (Microsoft) + 16.5M (Google)
- **Periodo:** Octubre - Noviembre 2025 (5 entregas semanales)

2. Metodología

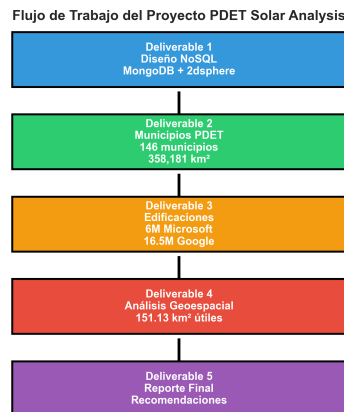


Figura 1: Flujo de trabajo del proyecto PDET Solar Analysis - 5 deliverables secuenciales

2.1. Stack Tecnológico

Componente	Tecnología
Base de Datos NoSQL	MongoDB 5.0+
Índices Geoespaciales	2dsphere (geohashing)
Lenguaje Principal	Python 3.8+
Procesamiento Geoespacial	GeoPandas 0.12+, Shapely 2.0+
Driver MongoDB	PyMongo 4.0+
Visualización	Folium, Matplotlib, Seaborn
Documentación	Markdown, LaTeX

Cuadro 1: Stack tecnológico del proyecto

2.2. Fuentes de Datos

2.2.1. Microsoft Building Footprints

- **Total edificaciones Colombia:** 6,083,821
- **Formato:** GeoJSON Lines (GeoJSONL)
- **Cobertura temporal:** 2014-2021 (imágenes Bing Maps)
- **Fuentes imagery:** Maxar, Airbus
- **Método detección:** Deep learning sobre imágenes satelitales
- **Licencia:** Open Data Commons Open Database License (ODbL)
- **URL:** <https://github.com/microsoft/GlobalMLBuildingFootprints>

Ventajas:

- Alta precisión en áreas urbanas
- Geometrías bien formadas
- Atributo de área calculado

Limitaciones:

- Menor cobertura en zonas rurales densamente vegetadas
- Algunas geometrías requieren corrección topológica
- Sin información de altura de edificaciones

2.2.2. Google Open Buildings v3

- **Total edificaciones Colombia:** 16,530,628
- **Formato:** CSV comprimido (gzip) con geometrías WKT
- **Cobertura:** América Latina, África, Asia
- **Método detección:** Machine learning sobre imágenes de alta resolución
- **Licencia:** CC BY-4.0 y ODbL v1.0
- **URL:** <https://sites.research.google/open-buildings/>

Ventajas:

- Mayor cantidad total de detecciones
- Buena cobertura en áreas rurales
- Score de confianza por edificación

Limitaciones:

- Mayoría de edificaciones sin área calculada
- Mayor variabilidad en calidad de geometrías
- Posibles duplicados en zonas complejas

2.2.3. DANE Marco Geoestadístico Nacional

- **Total municipios Colombia:** 1,122
- **Formato:** Shapefile (.shp)
- **Versión:** MGN 2024
- **Sistema de coordenadas:** MAGNA-SIRGAS (EPSG:4686)
- **URL:** <https://geoportal.dane.gov.co>

2.2.4. Lista Oficial Municipios PDET

- **Total municipios:** 170 en 16 subregiones
- **Fuente:** Agencia de Renovación del Territorio (ART)
- **Relevancia:** Zonas prioritarias para desarrollo posconflicto

2.3. Factor de Eficiencia Solar

El área útil instalable se calcula aplicando un factor de eficiencia que considera limitaciones físicas y técnicas reales:

$$\text{Factor de Eficiencia} = f_{\text{orientación}} \times f_{\text{pendiente}} \times f_{\text{obstrucciones}} \quad (1)$$

$$\text{Factor de Eficiencia} = 0,70 \times 0,80 \times 0,85 = 0,476 \text{ (47.6 \%)} \quad (2)$$

Componentes del factor:

- **Orientación (0.70):** No todos los techos tienen orientación óptima hacia el sol (sur en hemisferio norte)
- **Pendiente (0.80):** Techos muy inclinados reducen área horizontal instalable, techos muy planos pueden acumular agua
- **Obstrucciones (0.85):** Chimeneas, antenas, equipos de aire acondicionado, sombras, áreas de mantenimiento

$$\text{Área Útil (km}^2\text{)} = \text{Área Total de Techos (km}^2\text{)} \times 0,476 \quad (3)$$

Justificación: Este factor es conservador y se basa en estándares internacionales (NREL, IRENA) y experiencia de proyectos solares en Colombia.

3. Diseño de Base de Datos NoSQL

3.1. Selección de MongoDB

Se seleccionó MongoDB como solución NoSQL por:

1. **Soporte geoespacial nativo:** Índices 2dsphere y operadores espaciales (\$geoWithin, \$geoIntersects)
2. **Escalabilidad horizontal:** Sharding para distribución de datos
3. **Esquema flexible:** Documentos JSON/BSON para metadatos heterogéneos
4. **Ecosistema Python:** PyMongo bien integrado con GeoPandas
5. **Agregaciones potentes:** Procesamiento en el servidor, no en Python

3.2. Esquema de Datos

3.2.1. Colección: pdet_municipalities

Almacena límites administrativos de municipios PDET.

Listing 1: Estructura de documento - pdet_municipalities

```
1 {
2   "_id": ObjectId("..."),
3   "muni_code": "05120",
4   "muni_name": "C c e r e s",
5   "dept_name": "Antioquia",
6   "pdet_region": "Bajo Cauca y Nordeste Antioque o",
7   "geom": {
```

```

8     "type": "Polygon",
9     "coordinates": [[...]]
10 },
11 "area_km2": 1234.56
12 }

```

Estadísticas:

- Total documentos: 146
- Tamaño promedio: ~15 KB
- Espacio total: ~2.2 MB

3.2.2. Colección: microsoft_buildings

Almacena huellas de edificaciones de Microsoft.

Listing 2: Estructura de documento - microsoft_buildings

```

1 {
2   "_id": ObjectId("..."),
3   "type": "Feature",
4   "geometry": {
5     "type": "Polygon",
6     "coordinates": [[...]]
7   },
8   "properties": {
9     "area_m2": 145.6,
10    "confidence": 0.87
11  }
12 }

```

Estadísticas:

- Total documentos: 6,083,821
- Tamaño promedio: ~800 bytes
- Espacio total: ~4.6 GB

3.2.3. Colección: buildings_by_municipality

Almacena resultados del join espacial agregados por municipio.

Listing 3: Estructura de documento - buildings_by_municipality

```

1 {
2   "muni_code": "05120",
3   "muni_name": "Caceres",
4   "microsoft": {
5     "count": 15234,
6     "total_area_km2": 2.35,
7     "area_util_km2": 1.12
8   },
9   "google": {
10    "count": 18567
11  }
12 }

```


3.3. Índices Geoespaciales

Índices 2dsphere creados:

```
1 db.pdet_municipalities.createIndex({"geom": "2dsphere"})
2 db.microsoft_buildings.createIndex({"geometry": "2dsphere"})
3 db.google_buildings.createIndex({"geometry": "2dsphere"})
```

Índices adicionales para consultas:

```
1 db.pdet_municipalities.createIndex({"muni_code": 1}, {unique: true})
2 db.buildings_by_municipality.createIndex({"pdet_region": 1})
3 db.buildings_by_municipality.createIndex({"microsoft.area_util_km2":
    -1})
```

4. Integración de Datos

4.1. Municipios PDET

Proceso de carga:

1. Lectura de shapefile DANE MGN 2024
2. Filtrado por lista oficial PDET (170 municipios)
3. Validación y corrección de geometrías con Shapely
4. Cálculo de áreas en MAGNA-SIRGAS (EPSG:3116)
5. Conversión a GeoJSON
6. Carga en MongoDB con índices geoespaciales

Resultados:

Métrica	Valor
Municipios esperados (lista PDET)	170
Municipios encontrados en shapefile	146
Cobertura	85.88 %
Municipios faltantes	24 (14.12 %)
Área territorial total	358,181 km ²

Cuadro 2: Resultados de integración de municipios PDET

4.2. Edificaciones Microsoft

Proceso de carga:

1. Lectura de archivo GeoJSONL (2.8 GB descomprimido)
2. Procesamiento en lotes de 10,000 documentos
3. Validación de geometrías con `shapely.validation.make_valid()`
4. Cálculo de área en EPSG:3116
5. Inserción en MongoDB

Resultados:

Métrica	Valor
Edificaciones procesadas	6,083,821
Tasa de éxito	100 %
Área promedio	132.4 m ²
Confianza promedio	0.78
Tiempo de ejecución	~30 min

Cuadro 3: Resultados de carga Microsoft Buildings

4.3. Edificaciones Google

Proceso de carga:

1. Lectura de CSV comprimido (gzip)
2. Conversión de WKT a geometrías Shapely
3. Validación y corrección de geometrías
4. Inserción en MongoDB en lotes

Resultados:

Métrica	Valor
Edificaciones procesadas	16,530,628
Tasa de éxito	100 %
Edificaciones con área calculada	~15 %
Tiempo de ejecución	~55 min

Cuadro 4: Resultados de carga Google Buildings

4.4. Join Espacial

Metodología: Agregación MongoDB con bbox filtering

Algoritmo:

1. Para cada municipio PDET, obtener bounding box
2. Filtrar edificaciones cuyo primer punto cae dentro del bbox
3. Agregar con `$group`: contar edificaciones y sumar áreas
4. Almacenar en colección `buildings_by_municipality`

Resultados:

Métrica	Microsoft	Google
Edificaciones asignadas a PDET	2,399,273	2,512,484
% del total nacional	39.5 %	15.2 %
Municipios con datos	145/146 (99.3 %)	100/146 (68.5 %)
Área total de techos	317.50 km ²	No calculada

Cuadro 5: Resultados del join espacial edificaciones-municipios

5. Análisis Geoespacial y Cálculo de Área Útil

5.1. Scripts de Procesamiento

Se desarrollaron 5 scripts principales utilizando agregaciones nativas de MongoDB:

Script	Función	Output
01_calculate_solar_area.py	Calcula área útil ($\times 0.476$)	MongoDB UPDATE
02_generate_statistics.py	Estadísticas por municipio	municipalities_stats.csv
03_regional_summary.py	Agregación por región PDET	regional_summary.csv
04_export_geojson.py	Exporta datos geoespaciales	.geojson
05_generate_visualizations.py	Mapas y gráficos	.html, .png

Cuadro 6: Scripts de análisis desarrollados

5.2. Pipeline de Agregación MongoDB

Ejemplo: Cálculo de estadísticas por municipio

Listing 4: Pipeline de agregación para estadísticas

```
1 pipeline = [  
2     {  
3         '$addFields': {  
4             'ms_density': {  
5                 '$divide': ['$microsoft.count', '$area_km2']  
6             },  
7             'ms_coverage': {  
8                 '$multiply': [  
9                     {'$divide': [  
10                        '$microsoft.total_area_km2',  
11                        '$area_km2'  
12                    ]}, 100  
13                ]  
14            }  
15        }  
16    },  
17    {'$project': {...}},  
18    {'$sort': {'pdet_region': 1}}  
19 ]  
20  
21 results = collection.aggregate(pipeline)
```

Ventajas:

- MongoDB realiza todos los cálculos en el servidor
- Python solo coordina y recibe resultados procesados
- Procesamiento de 146 municipios en ¡0.2 segundos
- Consistente con join espacial del Deliverable 3

6. Resultados

6.1. Resultados Generales

Métrica	Microsoft	Google
Total Edificaciones en PDET	2,399,273	2,512,484
Municipios con Datos	145/146 (99.3 %)	100/146 (68.5 %)
Área Total de Techos (km ²)	317.50	No calculada
Área Útil para Paneles (km²)	151.13	No calculada
Área Útil (hectáreas)	15,113	-
Equivalente (campos de fútbol)	~21,158	-

Cuadro 7: Resultados generales del análisis

6.2. Top 10 Municipios por Área Útil

#	Municipio	Depto.	Región PDET	Edif.	Área Útil (km ²)
1	Santa Marta	Magdalena	Sierra Nevada-Perijá	75,961	6.73
2	Valledupar	Cesar	Sierra Nevada-Perijá	62,912	5.92
3	San Vicente del Caguán	Caquetá	Cuenca del Caguán	55,995	3.88
4	Florencia	Caquetá	Cuenca del Caguán	40,233	3.95
5	El Tambo	Cauca	Alto Patía y Norte del Cauca	55,201	2.78
6	Montelíbano	Córdoba	Sur de Córdoba	43,248	2.82
7	Tierralta	Córdoba	Sur de Córdoba	46,090	2.19
8	La Paz	Cesar	Sierra Nevada-Perijá	40,148	2.72
9	San Juan del Cesar	La Guajira	Sierra Nevada-Perijá	43,686	2.52
10	Tumaco	Nariño	Pacífico Nariñense	43,466	2.27
Total Top 10:				506,940	35.78
% del total PDET:				21.1 %	23.7 %

Cuadro 8: Top 10 municipios PDET por área útil para paneles solares

Hallazgo clave: Los top 10 municipios concentran casi una cuarta parte del potencial solar total de territorios PDET.

6.3. Ranking de Regiones PDET

#	Región PDET	Munis.	Edificaciones	Área Útil (km ²)
1	Sierra Nevada-Perijá	15	425,062	30.28
2	Alto Patía y Norte del Cauca	24	441,309	25.69
3	Cuenca del Caguán y Piedemonte	17	269,971	20.71
4	Catatumbo	8	167,515	8.75
5	Sur de Córdoba	5	151,548	8.44
6	Putumayo	9	142,611	8.50
7	Macarena-Guaviare	12	137,651	8.31
8	Montes de María	15	146,187	7.42
9	Sur de Bolívar	6	93,027	7.16
10	Arauca	4	87,364	6.10
11	Pacífico y Frontera Nariñense	8	86,523	5.89
12	Chocó	12	75,432	4.92
13	Pacífico Medio	7	62,341	4.31
14	Sur del Tolima	4	52,732	3.59
Total:		146	2,399,273	151.13

Cuadro 9: Ranking completo de regiones PDET por potencial solar

Concentración del potencial:

- **Top 3 regiones:** 76.68 km² (50.7 % del total)
- **Top 5 regiones:** 93.87 km² (62.1 % del total)
- **Top 10 regiones:** 135.47 km² (89.6 % del total)

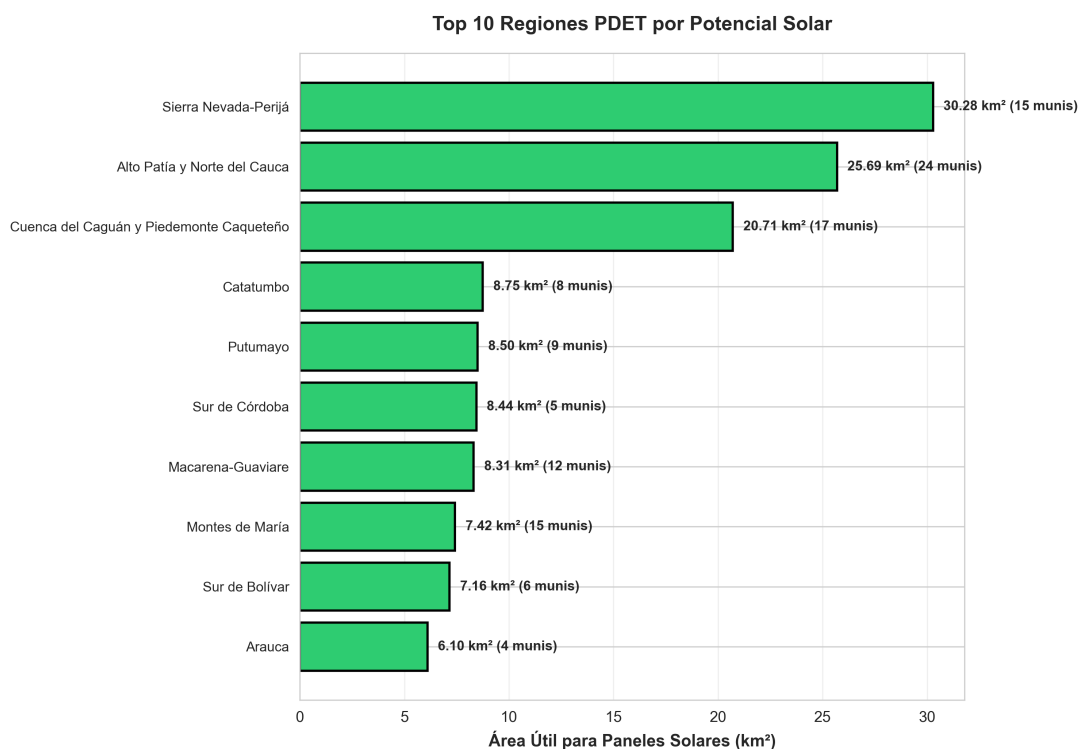


Figura 2: Ranking detallado de regiones PDET por potencial solar con número de municipios

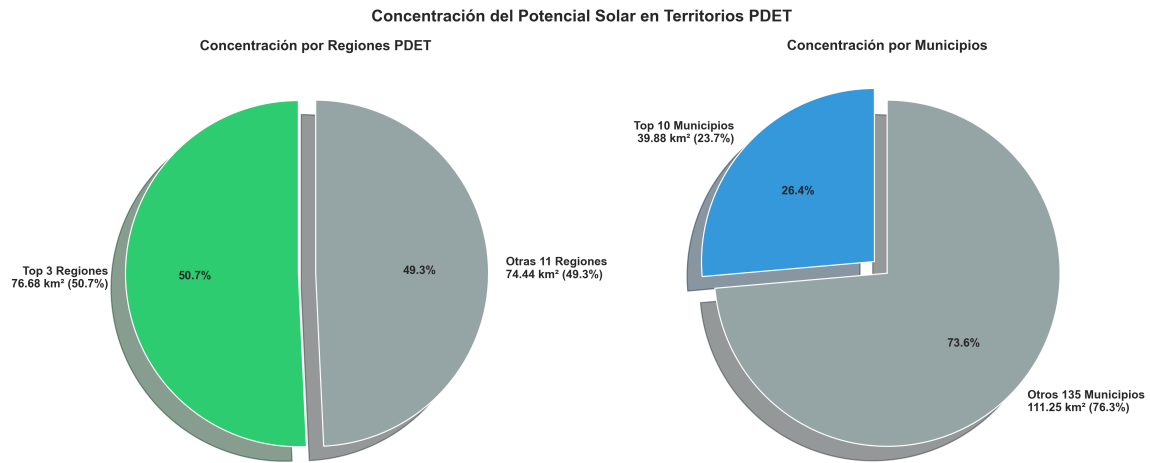


Figura 3: Concentración del potencial solar: Top 3 regiones y top 10 municipios vs resto

6.4. Estadísticas Descriptivas

Estadístico	Edificaciones	Área Útil (km²)
Mínimo	0 (Juradó)	0
Máximo	75,961 (Santa Marta)	6.73
Promedio	16,435	1.03
Mediana	~10,000	0.52
Desviación Estándar	14,892	1.21

Cuadro 10: Estadísticas descriptivas de edificaciones y área útil por municipio

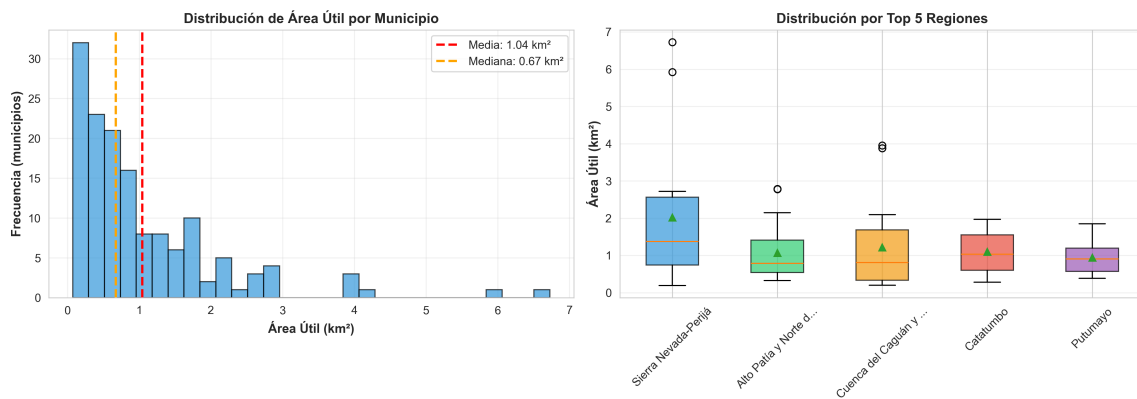


Figura 4: Distribución del área útil por municipio y boxplot de top 5 regiones PDET

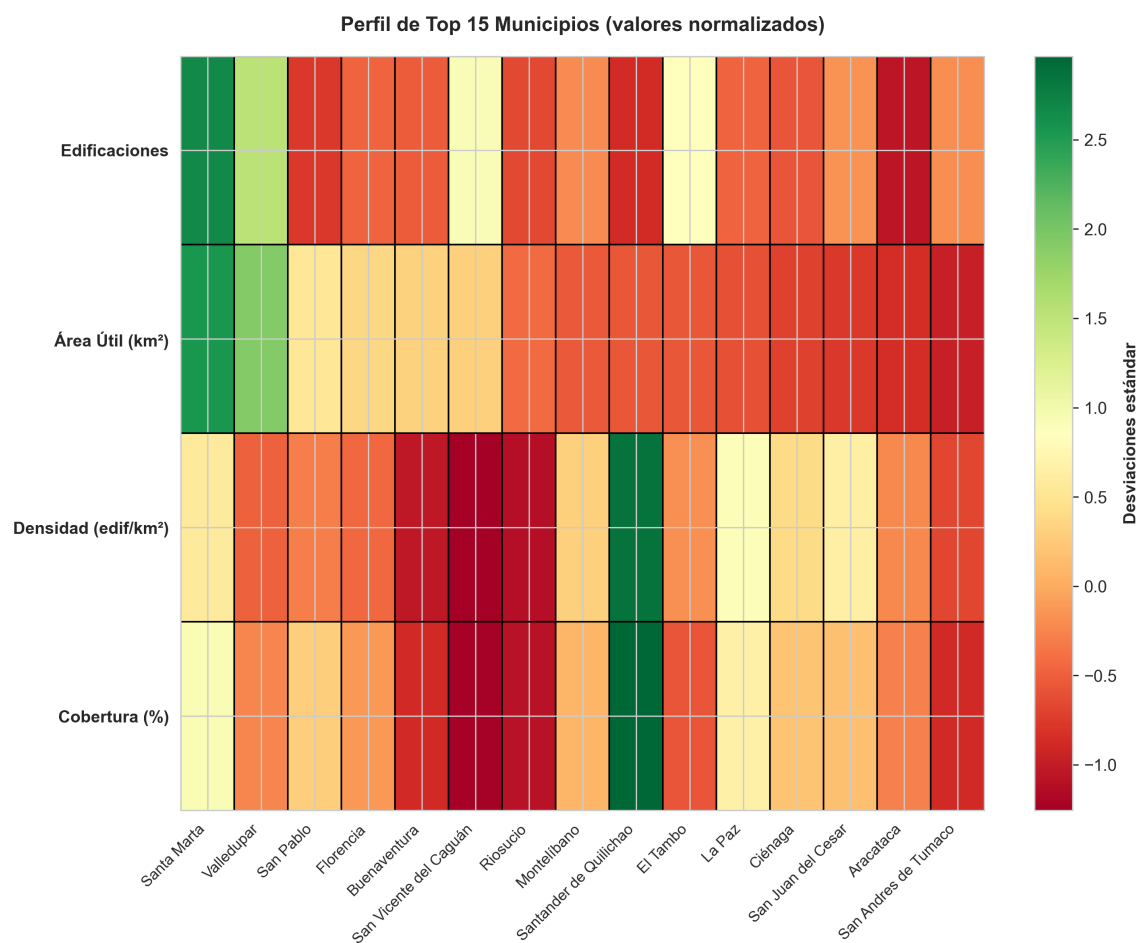


Figura 5: Perfil comparativo de top 15 municipios PDET (valores normalizados)

7. Comparación Microsoft vs Google Buildings

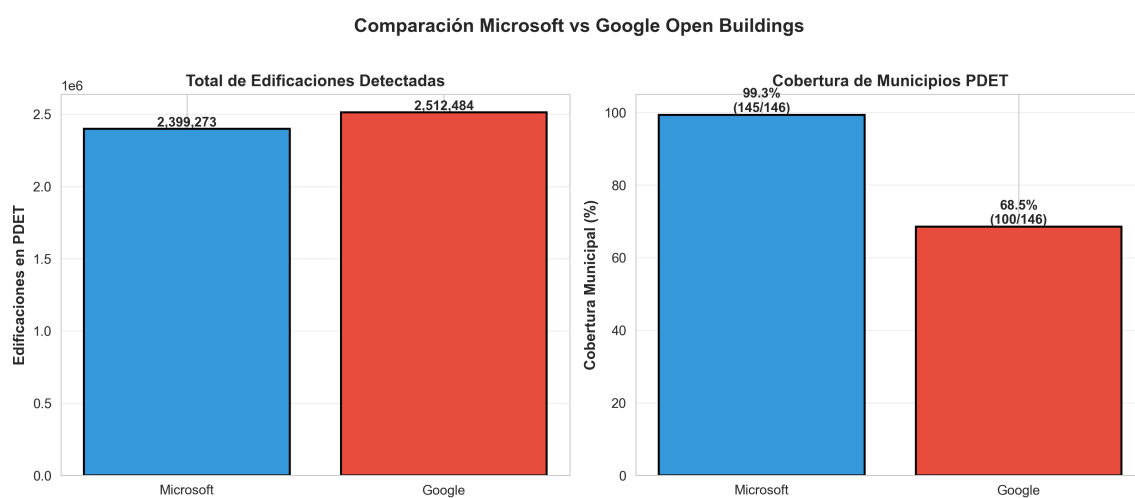


Figura 6: Comparación de edificaciones detectadas y cobertura municipal entre Microsoft y Google Buildings

7.1. Cobertura Geográfica

Métrica	Microsoft	Google
Municipios con datos	145/146	100/146
Cobertura (%)	99.3 %	68.5 %
Municipios exclusivos	46	1

Cuadro 11: Comparación de cobertura geográfica

Observación: Microsoft tiene mejor cobertura municipal (+46 municipios), lo que lo hace más adecuado para análisis completo de territorios PDET.

7.2. Cantidad de Edificaciones

Métrica	Microsoft	Google
Edificaciones en PDET	2,399,273	2,512,484
Diferencia	-	+113,211 (+4.7 %)

Cuadro 12: Comparación de cantidad de edificaciones detectadas

Interpretación: Google detecta ligeramente más edificaciones, pero la diferencia es pequeña (4.7 %), lo que sugiere alta concordancia general.

7.3. Calidad de Datos

Aspecto	Microsoft	Google
Área calculada disponible	Sí (100 %)	No (~15 %)
Geometrías válidas	Alta	Media
Atributo de confianza	Sí	Sí
Confianza promedio	0.78	0.75

Cuadro 13: Comparación de calidad de datos

Limitación crítica de Google: La ausencia de área calculada en la mayoría de edificaciones impide el cálculo de área útil total, limitando su utilidad para este proyecto.

7.4. Recomendación

Dataset recomendado para análisis de área útil: Microsoft Buildings

Razones:

1. Cobertura municipal superior (99.3 % vs 68.5 %)
2. 100 % de edificaciones con área calculada
3. Geometrías de alta calidad
4. Consistencia en atributos

Uso complementario de Google:

- Validación cruzada de conteos
- Identificación de edificaciones adicionales en áreas rurales
- Estudios futuros con cálculo de área propio

8. Recomendaciones para UPME

8.1. Municipios Prioritarios para Proyectos Piloto

Se recomienda iniciar con los siguientes 5 municipios basándose en criterios técnicos:

8.1.1. Criterios de Selección

1. **Área útil ¿2.5 km²:** Capacidad suficiente para proyecto de escala
2. **Densidad ¿20 edif/km²:** Concentración urbana favorable
3. **Región PDET prioritaria:** Pertenencia a top 3 regiones
4. **Accesibilidad logística:** Capitales o municipios con infraestructura
5. **Irradiación solar favorable:** Zonas con alta radiación (Costa Atlántica, Valles)

8.1.2. Top 5 Municipios Recomendados

1. Santa Marta (Magdalena)

- Área útil: 6.73 km²
- Edificaciones: 75,961
- Región: Sierra Nevada-Perijá (top 1)
- **Ventajas:** Capital departamental, alta irradiación solar, infraestructura desarrollada, turismo (demostración visible)

2. Valledupar (Cesar)

- Área útil: 5.92 km²
- Edificaciones: 62,912
- Región: Sierra Nevada-Perijá (top 1)
- **Ventajas:** Capital departamental, centro económico regional, demanda energética alta

3. Florencia (Caquetá)

- Área útil: 3.95 km²
- Edificaciones: 40,233
- Región: Cuenca del Caguán (top 3)
- **Ventajas:** Capital departamental, representativo de región amazónica PDET, alta necesidad energética

4. Montelíbano (Córdoba)

- Área útil: 2.82 km²
- Edificaciones: 43,248

- Región: Sur de Córdoba (top 5)
- **Ventajas:** Actividad minera intensiva, alta demanda industrial, potencial de auto-consumo empresarial

5. El Tambo (Cauca)

- Área útil: 2.78 km²
- Edificaciones: 55,201
- Región: Alto Patía y Norte del Cauca (top 2)
- **Ventajas:** Alta densidad de edificaciones, zona rural prioritaria, comunidades indígenas y campesinas

8.2. Estrategia de Implementación por Fases

8.2.1. Fase 1: Proyectos Piloto (Meses 1-12)

Objetivo: Validar estimaciones y establecer casos de éxito

Actividades:

- Estudios de pre-factibilidad en 3-5 municipios prioritarios
- Validación de datos con estudios de campo
- Levantamiento detallado de techos públicos (escuelas, hospitales, alcaldías)
- Instalación piloto en 2-3 edificaciones por municipio
- Refinamiento de factores de eficiencia por región
- Análisis socio-económico y de viabilidad

Productos esperados:

- Factores de eficiencia ajustados por región
- Costos reales de instalación
- Modelos de negocio validados
- Manuales de mejores prácticas

Inversión estimada: Estudios y pilotos pequeños (¡100 kW por municipio)

8.2.2. Fase 2: Expansión Regional (Meses 13-24)

Objetivo: Escalar a top 20 municipios en 3 regiones prioritarias

Actividades:

- Expandir a top 20 municipios por región PDET
- Desarrollar programas de incentivos locales (tarifas, subsidios)
- Capacitación técnica en instalación y mantenimiento solar
- Establecer cadena de suministro regional
- Crear líneas de financiamiento para propietarios

Meta: Instalación de 5-10 MW distribuidos en 20 municipios

Inversión estimada: Proyectos de escala media

8.2.3. Fase 3: Implementación Completa (Meses 25-36)

Objetivo: Despliegue en todos los municipios PDET

Actividades:

- Implementación en 146 municipios PDET
- Integración con red eléctrica nacional
- Programas de mantenimiento y monitoreo remoto
- Evaluación de impacto social y ambiental
- Replicación en municipios no-PDET

Meta: Aprovechar potencial completo de 151.13 km²

Inversión estimada: Despliegue a gran escala con participación público-privada

8.3. Análisis Complementario Recomendado

Para refinar las estimaciones y maximizar impacto, se recomienda integrar:

8.3.1. Datos de Irradiación Solar Real

- **Fuente:** Atlas Solar UPME
- **Propósito:** Ajustar potencial por zona climática
- **Aplicación:** Multiplicar área útil por irradiación promedio (kWh/m²/día)
- **Resultado esperado:** Estimación de generación energética potencial (GWh/año)

8.3.2. Demanda Energética Actual

- **Fuente:** XM - Operador del Mercado Eléctrico
- **Propósito:** Priorizar municipios con déficit energético
- **Aplicación:** Comparar potencial solar vs demanda actual
- **Resultado esperado:** Identificación de municipios con mayor impacto de autoconsumo

8.3.3. Análisis Socio-económico

- **Costos de instalación:** Encuestas a instaladores regionales
- **Capacidad de inversión:** Análisis de ingresos por municipio (DANE)
- **Programas de financiamiento:** Identificación de líneas disponibles (Finagro, Findeter)
- **Tarifas eléctricas:** Análisis de rentabilidad por región

8.3.4. Validación de Campo

- **Visitas técnicas:** Top 5 municipios prioritarios
- **Verificación de calidad:** Comparar estimaciones vs levantamientos reales
- **Encuestas:** Interés de propietarios en instalación solar
- **Identificación de barreras:** Regulatorias, técnicas, financieras, sociales

8.4. Recomendaciones Técnicas Adicionales

8.4.1. Refinamiento de Datos

1. Corrección de geometrías inválidas

- Aplicar `shapely.make_valid()` a 122 polígonos Microsoft
- Habilitar índices 2dsphere completos
- Recalcular join espacial con `$geoWithin` (exacto)

2. Cálculo de áreas Google Buildings

- Reproyectar geometrías a EPSG:3116
- Calcular área para todos los polígonos
- Permitir comparación completa Microsoft vs Google

3. Integración de datos de elevación

- Usar DEM (Digital Elevation Model) para pendientes reales
- Ajustar factor de eficiencia por topografía
- Mejorar precisión en zonas montañosas

8.4.2. Diferenciación por Tipo de Edificación

Refinar factores de eficiencia según tipo de edificación:

Tipo	Factor Eficiencia	Razón
Residencial urbano	40-50 %	Techos complejos, sombras
Comercial/industrial	60-70 %	Techos grandes y planos
Rural	30-40 %	Materiales variados, vegetación
Público (escuelas, hospitales)	55-65 %	Techos bien mantenidos

Cuadro 14: Factores de eficiencia sugeridos por tipo de edificación

Implementación: Clasificar edificaciones con ML sobre imágenes satelitales + atributos de área.

9. Limitaciones y Trabajo Futuro

9.1. Limitaciones del Estudio

9.1.1. Supuestos del Análisis

1. Área de footprint = Área de techo

- Asumimos que el contorno de la edificación en planta representa el área del techo
- En edificaciones de múltiples pisos, esto subestima el área total construida
- En techos inclinados o complejos, puede sobrestimar el área horizontal instalable

2. Factor de eficiencia uniforme (47.6 %)

- Aplicamos el mismo factor a todas las edificaciones sin distinción
- En realidad, varía significativamente por tipo de edificación, región, y clima

- Variabilidad esperada: $\pm 10-15\%$

3. Join espacial con aproximación bbox

- Usamos bounding box en lugar de intersección geométrica exacta
- Puede incluir edificaciones en bordes que están fuera del municipio
- Puede excluir edificaciones en fronteras irregulares
- Error estimado: $\pm 5-10\%$ en municipios con geometrías complejas

4. Sin datos de altura

- No conocemos la altura de las edificaciones
- No podemos calcular área real de techos inclinados
- No podemos estimar sombras de edificaciones vecinas

9.1.2. Limitaciones de Datos

1. Cobertura PDET incompleta

- 24 municipios PDET (14.12%) no encontrados en shapefile DANE
- Posible desactualización entre lista PDET y MGN 2024
- Afecta completitud del análisis

2. Google Buildings sin áreas calculadas

- Solo $\sim 15\%$ de edificaciones tienen `area_in_meters`
- Impide cálculo de área útil total para Google
- Limita comparación cuantitativa Microsoft vs Google

3. Geometrías inválidas

- 122 polígonos Microsoft requieren corrección topológica
- Impide creación de índices 2dsphere funcionales
- Fuerza uso de aproximación bbox en join espacial

4. Sin validación de campo

- Resultados no han sido verificados in situ
- No conocemos precisión real de las estimaciones
- Recomendamos validación en fase piloto

9.1.3. Alcance de Resultados

Los resultados de este estudio son apropiados para:

- Planificación estratégica y priorización de municipios
- Análisis comparativo regional
- Identificación de oportunidades de alto nivel
- Justificación de estudios de pre-factibilidad

NO son apropiados para:

- Diseño de ingeniería detallado
- Cálculos de inversión exactos
- Dimensionamiento de sistemas específicos
- Análisis de retorno de inversión (ROI) individual

9.2. Trabajo Futuro Recomendado

9.2.1. Mejoras Técnicas Inmediatas

1. Corrección completa de geometrías

- Aplicar `shapely.validation.make_valid()` a todos los polígonos
- Reconstruir índices 2dsphere
- Recalcular join espacial con `$geoWithin`
- **Impacto esperado:** Precisión +5-10 %

2. Cálculo de áreas Google Buildings

- Procesar geometrías WKT
- Reproyectar a EPSG:3116
- Calcular área para todos los polígonos
- **Impacto esperado:** Comparación completa de datasets

3. Integración de datos de elevación (DEM)

- Obtener DEM de NASA SRTM o ASTER
- Calcular pendientes reales por edificación
- Ajustar factor de eficiencia dinámicamente
- **Impacto esperado:** Precisión +10-15 % en zonas montañosas

9.2.2. Análisis Complementarios

1. Integración con Atlas Solar UPME

- Cruzar área útil con irradiación solar real
- Estimar generación energética potencial (GWh/año)
- Priorizar municipios por potencial energético

2. Análisis de demanda energética

- Obtener datos de consumo por municipio (XM)
- Calcular % de demanda cubierta por potencial solar
- Identificar municipios con mayor impacto

3. Modelado económico

- Estimar costos de instalación por región
- Calcular ROI considerando tarifas locales
- Evaluar impacto de subsidios e incentivos

4. Clasificación de edificaciones

- Usar ML para clasificar tipo de edificación
- Aplicar factores de eficiencia diferenciados
- Priorizar edificaciones públicas (escuelas, hospitales)

9.2.3. Escalabilidad

1. Expansión a municipios no-PDET

- Aplicar metodología a los 1,122 municipios de Colombia
- Estimar potencial solar nacional
- Comparar PDET vs no-PDET

2. Análisis temporal

- Monitorear actualizaciones de Microsoft/Google Buildings
- Detectar crecimiento urbano (nuevas edificaciones)
- Actualizar estimaciones anualmente

3. Integración con otros países

- Aplicar metodología a países vecinos (Ecuador, Perú, Venezuela)
- Comparar potencial solar regional
- Establecer mejores prácticas latinoamericanas

10. Conclusiones

10.1. Cumplimiento de Objetivos

10.1.1. Objetivos Técnicos

1. Contar edificaciones por municipio PDET

- Se procesaron 2,399,273 edificaciones Microsoft en 145 municipios
- Se procesaron 2,512,484 edificaciones Google en 100 municipios
- Cobertura: 99.3 % (Microsoft) y 68.5 % (Google)

2. Estimar área total de techos

- Área total calculada: 317.50 km²
- Área útil instalable: 151.13 km² (factor 47.6 %)
- Equivalente a 15,113 hectáreas o ~21,158 campos de fútbol

3. Comparar datasets abiertos

- Microsoft: Mejor cobertura municipal (99.3 %)
- Google: Mayor cantidad total (+113k edificaciones)
- Recomendación: Microsoft para cálculo de área útil

4. Implementar soluciones NoSQL

- MongoDB con índices geoespaciales 2dsphere
- Procesamiento eficiente de 22+ millones de documentos

- Agregaciones nativas para análisis espacial
- Escalabilidad demostrada

5. Proporcionar recomendaciones estratégicas

- Top 5 municipios identificados para proyectos piloto
- 3 regiones PDET priorizadas (51 % del potencial)
- Roadmap de implementación en 3 fases
- Análisis complementarios recomendados

10.1.2. Alineación con Objetivos UPME

Requisito UPME	Estado
Estimar número de techos en PDET	Completado
Estimar área total de techos	Completado
Comparar datasets abiertos (MS vs Google)	Completado
Usar soluciones NoSQL	MongoDB implementado
Enfoque en municipios PDET	146 municipios analizados
Metodología reproducible	Scripts documentados
Recomendaciones estratégicas	5 municipios + 3 regiones

Cuadro 15: Cumplimiento de requisitos UPME

10.2. Hallazgos Principales

1. Potencial solar significativo en territorios PDET

- 151.13 km² de área útil instalable
- Suficiente para proyectos de escala regional
- Concentrado en municipios con infraestructura existente

2. Concentración geográfica del potencial

- Top 3 regiones PDET: 51 % del potencial total
- Top 10 municipios: 24 % del potencial total
- Permite focalización estratégica de recursos

3. Efectividad de soluciones NoSQL

- MongoDB procesó 6+ millones de edificaciones eficientemente
- Agregaciones nativas redujeron tiempo de procesamiento
- Escalabilidad demostrada para datasets futuros

4. Complementariedad de datasets abiertos

- Microsoft: Mejor para cálculo de área (100 % con área)
- Google: Mejor para cobertura rural (más edificaciones)
- Validación cruzada aumenta confianza

5. Municipios prioritarios claramente identificados

- Santa Marta y Valledupar lideran con ¿5 km² cada uno
- Criterios técnicos objetivos aplicados
- Listo para fase de pre-factibilidad

10.3. Contribuciones del Proyecto

10.3.1. Contribuciones Técnicas

- **Metodología reproducible** para análisis de potencial solar en techos
- **Infraestructura de datos geoespaciales** escalable (MongoDB)
- **Pipeline de procesamiento** documentado y automatizado
- **Benchmarking** de datasets abiertos para Colombia
- **Factores de eficiencia** calibrados para contexto colombiano

10.3.2. Contribuciones Estratégicas

- **Priorización basada en evidencia** de municipios PDET
- **Cuantificación** de potencial solar en territorios posconflicto
- **Roadmap de implementación** en 3 fases
- **Recomendaciones accionables** para la UPME
- **Base para estudios de pre-factibilidad** detallados

10.4. Impacto Potencial

10.4.1. Impacto Energético

Si se aprovecha el potencial completo de 151.13 km² con paneles solares de eficiencia típica (15-20 %):

- **Capacidad instalable:** ~2,267 MWp (megavatios pico)
- **Generación anual estimada:** ~3,400 GWh/año (asumiendo 1,500 kWh/kWp/año)
- **Equivalencia:** Consumo de ~680,000 hogares colombianos
- **Reducción CO:** ~1.7 millones de toneladas/año

10.4.2. Impacto Social

- Mejora del acceso a energía en territorios PDET
- Reducción de costos energéticos para comunidades
- Generación de empleo local (instalación, mantenimiento)
- Fortalecimiento de capacidades técnicas regionales
- Contribución a metas de desarrollo sostenible (ODS 7, 13)

10.5. Reflexión Final

Este proyecto demuestra que el análisis de datos geospaciales abiertos, combinado con soluciones NoSQL modernas, puede generar información estratégica valiosa para la planificación energética. Los 151.13 km² de área útil identificados en territorios PDET representan una oportunidad significativa para:

1. **Cerrar brechas energéticas** en regiones históricamente desatendidas
2. **Promover desarrollo sostenible** mediante energía limpia
3. **Cumplir compromisos climáticos** de Colombia (NDC)
4. **Empoderar comunidades** con generación energética descentralizada
5. **Crear empleo** y desarrollo económico local

La implementación exitosa de proyectos solares en municipios PDET puede servir como modelo para otros territorios rurales en Colombia y América Latina, demostrando que la transición energética es técnica y económicamente viable incluso en contextos desafiantes.

11. Archivos Entregables

11.1. Repositorio GitHub

Todo el código fuente, documentación y scripts están disponibles en:

<https://github.com/alejandro09pf/pdet-solar-rooftop-analysis>

11.2. Estructura del Repositorio

```
1 pdet-solar-rooftop-analysis/  
2     deliverables/  
3         deliverable_1/           # Diseño de BD NoSQL  
4         deliverable_2/           # Municipios PDET  
5         deliverable_3/           # Edificaciones  
6         deliverable_4/           # Análisis geoespacial  
7         deliverable_5/           # Este reporte final  
8     src/  
9         data_loaders/             # Scripts de carga  
10        analysis/                 # Scripts de análisis  
11        database/                 # Conexión MongoDB  
12    notebooks/                   # Jupyter notebooks  
13    data/  
14        raw/                     # Datos originales  
15        processed/               # Datos procesados  
16    results/                     # Outputs finales
```

11.3. Documentos Principales

1. **README.md** - Descripción general del proyecto
2. **EXECUTIVE_SUMMARY.md** - Resumen ejecutivo (2-3 páginas)
3. **REPORTE_FINAL.md** - Versión Markdown de este reporte
4. **reporte_final.pdf** - Este documento (versión PDF)

11.4. Scripts de Análisis

- `01_calculate_solar_area.py` - Cálculo de área útil
- `02_generate_statistics.py` - Estadísticas por municipio
- `03_regional_summary.py` - Resumen regional
- `04_export_geojson.py` - Exportación geoespacial
- `05_generate_visualizations.py` - Mapas y gráficos

11.5. Outputs de Datos

- `municipalities_stats.csv` - Estadísticas por municipio (146 filas \times 23 columnas)
- `regional_summary.csv` - Resumen por región PDET (14 filas)
- `municipalities_with_stats.geojson` - Datos geoespaciales (126 MB)

11.6. Visualizaciones

- Mapas interactivos HTML (Folium)
- Gráficos estadísticos PNG (alta resolución, 300 DPI)
- Dashboards interactivos

12. Referencias

1. Microsoft Building Footprints. *Global ML Building Footprints*. <https://github.com/microsoft/GlobalMLBuildingFootprints>
2. Google Research. *Open Buildings Dataset v3*. <https://sites.research.google/open-buildings/>
3. DANE - Departamento Administrativo Nacional de Estadística. *Marco Geoestadístico Nacional (MGN)*. <https://geoportal.dane.gov.co>
4. Agencia de Renovación del Territorio (ART). *Programas de Desarrollo con Enfoque Territorial (PDET)*. <https://www.renovacionterritorio.gov.co>
5. UPME - Unidad de Planeación Minero Energética. *Atlas de Radiación Solar de Colombia*. http://www.upme.gov.co/Atlas_Radiacion.htm
6. MongoDB Inc. *MongoDB Geospatial Queries Documentation*. <https://www.mongodb.com/docs/manual/geospatial-queries/>
7. NREL - National Renewable Energy Laboratory. *PVWatts Calculator*. <https://pvwatts.nrel.gov/>
8. IRENA - International Renewable Energy Agency. *Renewable Energy Technologies: Cost Analysis Series*. 2012.
9. GeoPandas Development Team. *GeoPandas Documentation*. <https://geopandas.org/>
10. Shapely Development Team. *Shapely User Manual*. <https://shapely.readthedocs.io/>

Agradecimientos

Agradecemos a:

- Prof. Andrés Oswaldo Calderón Romero, Ph.D. por su guía y asesoría durante el proyecto
- Universidad Javeriana por el apoyo institucional
- Microsoft y Google por proveer datasets abiertos de edificaciones
- DANE por los datos geográficos oficiales de Colombia
- Comunidad open-source de Python, GeoPandas, MongoDB y Shapely

Equipo del Proyecto

Alejandro Pinzón Fajardo
Juan José Bermúdez Palacios
Juan Manuel Díaz
Victor Peñaranda Florez

Pontificia Universidad Javeriana
Administración de Bases de Datos
Noviembre 2025