

Cubo de Datos Geoespaciales de El Salvador: arquitectura, procesos ETL y aplicaciones

Clímaco, J., ORCID: 0009-0006-2035-7965; Chávez, T., ORCID: 0000-0003-4525-5894;
Dirección General del Observatorio de Amenazas y Recursos Naturales del Ministerio de Medio Ambiente y Recursos Naturales, San Salvador, El Salvador.

`jclimaco1@ambiente.gob.sv`, `tchavez2@ambiente.gob.sv`

Resumen – El acceso y uso de imágenes satelitales constituyen una herramienta esencial para la gestión de recursos naturales, la planificación territorial y la evaluación ambiental. No obstante, la complejidad de los procesos de descarga, depuración, transformación y análisis de grandes volúmenes de datos representa una barrera significativa para investigadores, tomadores de decisiones y el público en general. Este artículo presenta el diseño, desarrollo y puesta en marcha del Cubo de Datos Geoespaciales de El Salvador, una plataforma informática que automatiza la adquisición, procesamiento y publicación de productos satelitales de nivel 2 provenientes de las misiones Landsat y Sentinel-2, así como productos de nivel 3 de MODIS. La arquitectura implementada integra bases de datos geoespaciales, procesos ETL automatizados y servicios interoperables de información geoespacial. Adicionalmente, se desarrolló un portal web que permite a los usuarios consultar, analizar y descargar productos como índices espectrales, estadísticas de temperatura superficial, coberturas de suelo y geomédianas anuales, sin necesidad de conocimientos avanzados en software SIG. Esta herramienta contribuye a la reducción de barreras técnicas, optimiza los tiempos de análisis y promueve la democratización del uso de datos satelitales para la gestión ambiental y territorial en El Salvador.

Palabras clave: cubo de datos geoespaciales de El Salvador, procesos ETL, imágenes satelitales ARD.

Abstract – Access to and use of satellite imagery are essential tools for natural resource management, territorial planning, and environmental assessment. However, the complexity of downloading, cleaning, transforming, and analysing large volumes of data remains a significant barrier for researchers, decision-makers, and the public. This article presents the design, development, and implementation of the Geospatial Data Cube of El Salvador, a computational platform that automates the downloading, processing, and dissemination of satellite products. The system integrates Level-2 data from Landsat and Sentinel-2 missions, as well as Level-3 MODIS products, including vegetation indices, land cover maps, and land surface temperature. The implemented architecture combines geospatial databases, automated ETL workflows, and interoperable geospatial services. In addition, a web portal was developed to enable users to query, analyse, and download products such as spectral indices, surface temperature statistics, land cover datasets, and annual geomédians without requiring advanced GIS expertise. The results demonstrate that this approach reduces technical barriers, optimizes analysis time, and promotes the democratization of satellite data use for environmental and territorial management in El Salvador.

Palabras clave: Geospatial data cube, El Salvador, ETL processes, ARD satellite imagery.

I. INTRODUCCIÓN

Las imágenes satelitales constituyen una fuente esencial de información para el monitoreo de los recursos naturales, la gestión ambiental y la planificación territorial. Desde hace décadas, programas como Landsat, SPOT, ASTER, MODIS y más recientemente Sentinel-2, han puesto a disposición series temporales extensas de observaciones que permiten generar indicadores clave como temperatura superficial, índices de vegetación, coberturas del suelo y dinámicas

territoriales a lo largo del tiempo (Roy et al., 2014; Drusch et al., 2012; Gómez et al., 2016).

No obstante, el aprovechamiento pleno de estas fuentes enfrenta limitaciones técnicas significativas. La descarga masiva de imágenes, su almacenamiento, la depuración de valores inválidos y cobertura nubosa, así como la generación de mosaicos e integración de diferentes resoluciones espaciales y espectrales, constituyen barreras que dificultan su uso extendido (Wulder & Coops, 2014; Gómez et al., 2016). A ello se suma la variabilidad en las condiciones de observación —atmósfera, nubosidad, geometría de toma— y las diferencias

instrumentales entre sensores, lo que incrementa la complejidad en su procesamiento y estandarización (Lewis et al., 2016).

Estos retos han impulsado el desarrollo de iniciativas internacionales que buscan procesar y reorganizar grandes volúmenes de datos satelitales en entornos de alto rendimiento computacional. Tal es el caso del Geoscience Data Cube de Australia, que permitió generar productos ambientales consistentes a partir de 27 años de observaciones Landsat, constituyéndose en referente global (Lewis et al., 2016). De manera similar, el Cubo de Datos Geoespaciales de México (CDGM), liderado por INEGI, ha generado productos históricos derivados de Landsat —como geomedianas e índices espectrales— que pueden ser integrados en análisis estadísticos para la caracterización de recursos naturales y variables agrícolas (Ornelas de Anda et al., 2019).

En El Salvador, estas limitaciones han restringido tradicionalmente el acceso a los datos satelitales a especialistas con experiencia en Sistemas de Información Geográfica (SIG), limitando su uso en instituciones públicas y privadas. Para superar estas barreras, se desarrolló el Cubo de Datos Geoespaciales de El Salvador, un sistema que automatiza la descarga, procesamiento y publicación de productos satelitales de Landsat, Sentinel-2 y MODIS, y que además pone a disposición un portal web de fácil acceso para usuarios finales.

Este artículo presenta la arquitectura, los procesos de ingeniería de software, el modelo de datos y las aplicaciones desarrolladas en el marco de dicho proyecto. El enfoque combina tres ejes principales: (i) el uso de imágenes satelitales y técnicas de sensoramiento remoto, (ii) el diseño del modelo de datos, procesos ETL y servicios GIS, y (iii) el desarrollo de aplicaciones web y servicios interoperables. El sistema integra múltiples fuentes satelitales —Landsat (4, 5, 7, 8 y 9), Sentinel-2 y MODIS—, con cobertura histórica del territorio salvadoreño desde 1986 con Landsat, desde 2000 con MODIS y, en el caso de Sentinel-2, desde 2016. Los datos procesados corresponden a colecciones Landsat Surface Reflectance, Sentinel-2 Nivel 2A y productos de MODIS Nivel 3, que incluyen índices de vegetación, mapas de cobertura y temperatura superficial.

Una vez descargados, estos insumos se someten a algoritmos de procesamiento que abarcan extracción de bandas, enmascaramiento de nubes, generación de índices y construcción de mosaicos a escala nacional. Posteriormente, los productos resultantes se ponen a disposición de los usuarios a través de un portal web y de servicios interoperables

soportados en GeoServer y una base de datos PostgreSQL/PostGIS. Así, el Cubo de Datos Geoespaciales de El Salvador constituye una respuesta nacional a los desafíos técnicos globales, y contribuye a la democratización del uso de datos satelitales en la gestión ambiental y territorial.

II. DATOS LISTOS PARA EL ANÁLISIS EN EL CUBO DE DATOS GEOESPACIALES DE EL SALVADOR

2.1 Cubo de datos

El término Big Data se refiere a la gestión, procesamiento y análisis de grandes volúmenes de datos (Tascón, 2013). Estos pueden clasificarse como estructurados, cuando siguen un modelo de organización y relación —por ejemplo, bases de datos, formularios web o encuestas—, o no estructurados, cuando carecen de dicho modelo, como ocurre con publicaciones en redes sociales o imágenes satelitales (Li, 2016). Si bien suele asociarse únicamente con la cantidad de datos, Big Data se caracteriza además por las denominadas 6V: volumen, variedad, velocidad, veracidad, visualización y valor (Gandomi & Haider, 2015).

En este sentido, Big Data implica la gestión eficiente de datos heterogéneos que son descargados, almacenados, procesados y analizados para generar productos con valor agregado. Estos productos pueden ser dispuestos en plataformas de visualización y consulta que apoyan la toma de decisiones en diversos sectores. El Cubo de Datos Geoespaciales de El Salvador (CDGES) constituye un ejemplo concreto de la aplicación de Big Data en la gestión ambiental y la reducción de riesgos de desastres, al poner a disposición de los usuarios observaciones satelitales listas para el análisis.

Un cubo de datos geoespaciales es una plataforma diseñada para organizar y procesar eficientemente grandes volúmenes de imágenes satelitales, de modo que estén disponibles a los usuarios finales en un formato Analysis Ready Data (ARD), ya corregidas radiométrica, atmosférica y geoméricamente (Juárez, 2021; Dwyer et al., 2018). Esto permite generar productos derivados de análisis históricos de la superficie terrestre, aplicables en estudios de suelo, vegetación, erosión costera, agricultura, deforestación, cuerpos de agua o asentamientos humanos.

De acuerdo con Lewis et al. (2017), un cubo de datos debe posibilitar flujos de trabajo automatizados y contar con características como: calibración de alta calidad en las imágenes, estándares de reflectancia,

georreferenciación precisa, controles de calidad a nivel de píxel, estructuración para análisis temporales y uso de formatos científicos que favorezcan el cálculo eficiente.

En el caso salvadoreño, el CDGES responde a la necesidad de:

- a. Gestionar y optimizar grandes volúmenes de datos.
- b. Poner a disposición, mediante una interfaz de usuario, productos accesibles que potencien el uso de observaciones satelitales y el desarrollo de investigaciones en recursos naturales y gestión de riesgos.

La Figura 1. presenta un esquema conceptual del cubo, en el cual se observa la relación entre fuentes de datos, procesos de transformación, generación de productos (índices, geomédianas, mapas de cobertura) y servicios geospaciales accesibles a usuarios institucionales y al público general. El diagrama muestra el flujo de productos según su nivel de procesamiento: desde indicadores básicos como la reflectancia superficial o los índices espectrales, que se transforman posteriormente en aplicaciones y servicios que permiten a los usuarios consultar estadísticas geospaciales sobre áreas geográficas y temáticas específicas de interés.

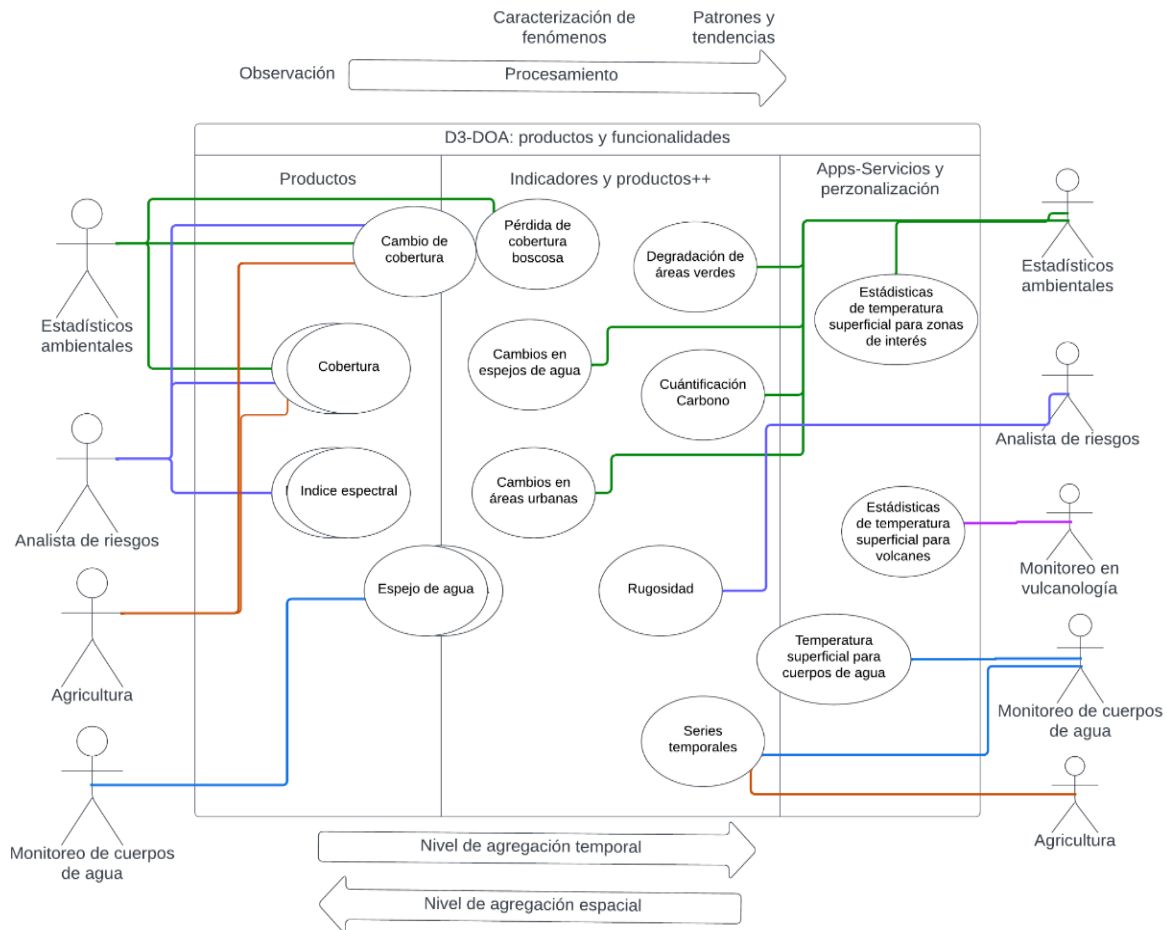


Fig 1. Esquema conceptual y casos de uso
Fuente: Tomado de (Climaco y Chávez, 2023)

2.2 Datos listos para el análisis

Los datos listos para el análisis (ARD) son productos satelitales previamente corregidos que permiten su uso directo en estudios ambientales sin requerir preprocesamiento complejo. En Landsat y Sentinel-2, los ARD de Nivel 2 proporcionan reflectancia de superficie con correcciones atmosféricas, radiométricas y geométricas, además de máscaras de

calidad (Louis et al., 2016; Dwyer et al., 2018). Por su parte, MODIS ofrece productos de Niveles 3 y 4, como índices de vegetación, temperatura superficial y mapas de cobertura del suelo (Justice et al., 2002). Estos datos estandarizados facilitan la construcción de series temporales consistentes y la generación de indicadores clave para la gestión ambiental y territorial.

2.3 Datos de MODIS

MODIS (Moderate Resolution Imaging Spectroradiometer) es un instrumento a bordo de los satélites Terra y Aqua, que orbitan a unos 700 km de altitud, pasando en la mañana y la tarde respectivamente. Captura información en 36 bandas del espectro electromagnético cada dos días como máximo, con resoluciones de 250, 500 y 1000 metros, lo que permite estudiar procesos de la superficie terrestre, la atmósfera baja y los océanos (Justice et al., 2002; Masuoka et al., 2011).

Los datos MODIS se organizan en distintos niveles de procesamiento, desde datos crudos y de geolocalización (Niveles 0 y 1), hasta productos corregidos y calibrados (Niveles 1A y 1B), variables geofísicas (Nivel 2) y productos de valor agregado derivados de estas (Niveles 3 y 4) (Mas, 2011).

El Cubo de Datos de El Salvador integra principalmente productos de los niveles 3 y 4, tales como:

- Índices de Vegetación (MOD13)
- Temperatura y emisividad de la superficie (MOD11)
- Cobertura de suelo (MOD12)

Estos productos son esenciales para caracterizar dinámicas ambientales, modelar procesos y apoyar la toma de decisiones en la gestión de recursos naturales.

2.4 Datos de Landsat

El programa Landsat, desarrollado conjuntamente por la NASA y el USGS, se inició en 1965 bajo el nombre ERTS y lanzó su primer satélite en 1972. Desde entonces ha operado ocho misiones (el Landsat 6 falló en su lanzamiento), de las cuales actualmente están en operación Landsat 8 y 9, mientras que Landsat 7 continúa con cobertura reducida tras ser relegado en 2022 a una órbita menor (USGS, 2021; 2023).

Cada misión ha incorporado sensores con distintas capacidades: desde el MSS en Landsat 1-5, el TM en Landsat 5, el ETM+ en Landsat 7 (con banda pancromática de 15 m de resolución), hasta los sensores OLI y TIRS en Landsat 8 y 9 (INEGI, 2020).

El Cubo de Datos de El Salvador utiliza imágenes Landsat desde 1984, basadas en la Colección 2 de USGS, que mejora la precisión geométrica y provee productos ARD (Figura 2). Estos incluyen:

- Reflectancia superficial (SR): 6 bandas en Landsat 4-7 y 7 bandas en Landsat 8-9.
- Temperatura superficial (ST): 1 banda térmica

en Landsat 4-7 y la banda 10 del TIRS en Landsat 8-9.

- Bandas auxiliares (8): transmitancia, emisividad, radiancias térmicas, incertidumbre, entre otras.
- Máscaras de calidad (3): control de uso de píxel, aerosoles y saturación radiométrica.
- Metadatos (txt y xml).

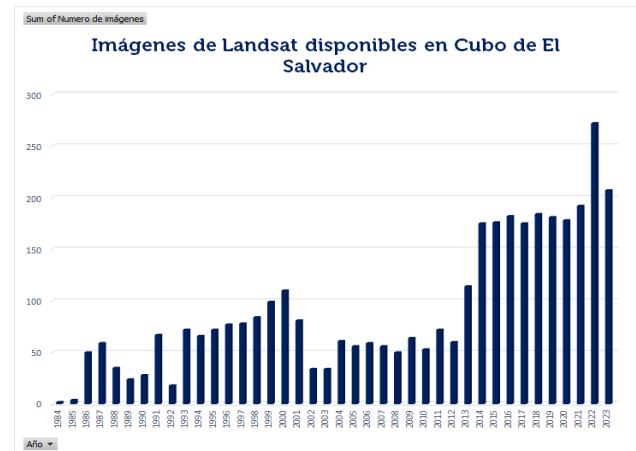


Fig. 2. Imágenes Landsat por año

Fuente: Tomado de (Clímaco y Chávez, 2023)

Los productos de Landsat 7 se generan con LEDAPS v3.4.0, y los de Landsat 8-9 con LaSRC v1.5.0 (USGS, 2021; 2023).

2.5 Datos de Sentinel 2

El Sentinel-2 es una constelación de satélites de observación óptica desarrollada en el marco del programa Copernicus, gestionado por la Agencia Espacial Europea (ESA). Actualmente está integrada por tres satélites: Sentinel-2A (lanzado en 2015), Sentinel-2B (2017) y Sentinel-2C (2024), este último concebido para sustituir progresivamente a Sentinel-2A y garantizar la continuidad del servicio (ESA, 2024).

Cada satélite orbita a una altitud aproximada de 786 km en órbita heliosíncrona y transporta el sensor Multispectral Instrument (MSI), diseñado para la adquisición de imágenes en 13 bandas espectrales que abarcan el rango visible, infrarrojo cercano e infrarrojo de onda corta. Estas bandas ofrecen resoluciones espaciales de 10 m, 20 m y 60 m, con un ancho de franja de 290 km, lo que proporciona cobertura global frecuente y de alta calidad (Drusch et al., 2012).

Los productos de Nivel 2A, generados a través del procesador Sen2Cor, corresponden a reflectancia de superficie corregida atmosféricamente, con georreferenciación precisa y capas auxiliares que

incluyen máscaras de nubes, sombras y aerosoles (Louis et al., 2016). En el Cubo de Datos Geoespaciales de El Salvador (CDGES), estos productos se integran para facilitar análisis multitemporales y comparables, sin requerir preprocesamiento adicional por parte de los usuarios.

A partir de estas imágenes, el CDGES genera productos derivados como índices espectrales (NDVI, NDWI, NBR, entre otros), geomedianas anuales, y mapas de cobertura del suelo y dinámica de cuerpos de agua, insumos fundamentales para la planificación territorial, la gestión ambiental y la reducción de riesgos de desastres en El Salvador.

III ARQUITECTURA DEL SISTEMA Y MODELO DE DATOS.

3.1 Diseño de la base de datos y almacenamiento

El sistema se apoya en una base de datos PostgreSQL/PostGIS, diseñada para registrar los metadatos de cada escena satelital descargada y permitir la gestión integrada de información espacial y alfanumérica. La estructura de la base de datos contempla las siguientes tablas principales:

- **Tiles satelitales:** definen la cobertura del territorio salvadoreño. Se incluyen cuatro tiles para Landsat (019050, 019051, 018051 y 018050), diez para Sentinel-2 (15PZQ, 16PBB, 15PZR, 16PDV, 16PBA, 16PBV, 16PCV, 16PCA, 16PDA y 15PZS), y un único tile en el caso de MODIS.
- **Productos satelitales por fuente:** MOD11A1, MOD13Q1 y MOD12Q1 en MODIS; distintas generaciones de Landsat (4–5, 7, 8–9); y bandas multiespectrales de Sentinel-2.
- **Parámetros de procesamiento:** especificaciones técnicas como bandas a extraer, valores válidos y configuración de máscaras de nubes.
- **Metadatos temporales:** incluyen fechas de adquisición, resolución espacial y fuente original.
- **Registro de descargas y procesamiento:** permite la trazabilidad de cada imagen y de los productos generados a partir de ella.

La estrategia de almacenamiento físico de archivos combina la organización automática de directorios con las convenciones de las fuentes oficiales, indexando y registrando este almacenamiento en la base PostgreSQL. Al extraerse

los archivos comprimidos descargados de cada escena, estos se almacenan en carpetas identificadas con el nombre de la escena definido por la fuente satelital. Posteriormente, cuando se generan mosaicos nacionales, los resultados se guardan bajo una estructura jerárquica que facilita la consulta histórica: una carpeta por año, dentro de ella carpetas por cada mes y dentro de cada mes subcarpetas para cada día. En otros casos, la fecha queda codificada en el nombre del archivo resultante. Esta organización asegura un control sistemático, facilita búsquedas temporales y optimiza la automatización de procesos.

La estructura está optimizada para soportar altos volúmenes de datos y consultas eficientes. Una política de manejo de almacenamiento elimina archivos originales descargados una vez procesados, conservando solo productos derivados y datos esenciales.

3.2 Procesos ETL

Los procesos ETL se implementaron en Python, haciendo uso en mayor parte de librerías de GDAL, para automatizar los procesos de Extracción, Transformación y Carga (ETL):

- **Extracción:** las descargas se realizan a través de las APIs oficiales (NASA EarthData, Copernicus Dataspace, USGS EROS) para descargar automáticamente los productos definidos en la base de datos. La descarga y obtención de los archivos de imágenes para cada fuente de observación espacial se gestionan de manera diferenciada, ya que cada proveedor aplica distintos niveles de preprocesamiento y ofrece variaciones en resolución y frecuencia temporal. Por esta razón, la lógica algorítmica de descarga se ha implementado de forma independiente para cada fuente. Sin embargo, todas las implementaciones comparten un mismo principio: se apoyan en la parametrización definida en la base de datos, la cual especifica los tiles que deben descargarse y la última fecha registrada, garantizando así consistencia y trazabilidad en el proceso.
- **Transformación:** integra un conjunto de operaciones informáticas que convierten los datos brutos en productos listos para análisis. Entre las principales actividades se encuentran:
 - **Extracción de bandas relevantes:** a partir de los archivos descargados de cada fuente satelital (MODIS, Landsat,

- Sentinel 2), se seleccionan únicamente las bandas que aportan a los análisis de interés.
- Limpieza de píxeles inválidos mediante máscaras de calidad: se aplican máscaras provistas por cada misión satelital que identifican nubes, sombras, saturaciones radiométricas o valores fuera de rango.
 - Construcción de stacks multibanda: las bandas relevantes son apiladas en archivos multibanda, consolidando la información espectral de cada fecha en un único archivo.
 - Cálculo de índices espectrales (NDVI, EVI, NBR): a partir de las bandas del rojo, infrarrojo cercano y otras específicas según cada índice, se generan productos derivados que permiten caracterizar la vegetación, el estado del suelo y procesos como incendios o degradación.
 - Generación de geomedianas anuales con Landsat: para cada año se construye una composición representativa del territorio basada en la geomedia, un estadístico robusto que integra múltiples observaciones temporales y mitiga la influencia de nubes, sombras y valores atípicos. (Mueller et al., 2016)
- **Carga:** Los productos generados se publican automáticamente en GeoServer, mediante servicios estándar como WMS, WFS y WCS, garantizando la interoperabilidad. La automatización se realiza mediante un wrapper en Python que gestiona la publicación y optimiza almacenamiento eliminando duplicados innecesarios. Dado que GeoServer crea una copia del archivo al momento de publicarlo, la lógica de carga incluye un procedimiento adicional: se elimina la copia innecesaria y se modifica el archivo XML de configuración de la capa para que apunte directamente a la ruta original del producto. Con esta estrategia se optimiza el uso de almacenamiento, especialmente crítico en el manejo de archivos de gran tamaño como las geomedianas y los índices espectrales.

Los procesos de ETL se ejecutan de forma automatizada mediante crontab en un servidor Linux, lo que garantiza la actualización periódica de los productos sin necesidad de intervención manual.

3.3 Tecnologías y arquitectura de implementación

La plataforma del cubo de datos geoespaciales se construyó con tecnologías que garantizan eficiencia, interoperabilidad y escalabilidad.

- Backend y procesamiento de datos: implementado en Python 3.8, utilizando librerías especializadas (GDAL/Rasterio para rásteres, Fiona/Shapely para vectores) y herramientas estándar para manejo de JSON, APIs y compresión de ZIP/RAR.
- Base de datos: PostgreSQL 10 con PostGIS 3.2 para almacenamiento y análisis espacial.
- Arquitectura de infraestructura: desplegada en la nube sobre Ubuntu 20.04 LTS, con tres instancias de máquinas virtuales:
 - Servidor web y procesos ETL para automatizar la obtención, transformación y carga de datos.
 - Base de datos, centralizando la gestión de la información espacial.
 - GeoServer, publicando productos geoespaciales mediante WMS, WFS y WCS.

Se configuró un almacenamiento compartido montado en las tres instancias, optimizando el uso de recursos y asegurando consistencia de los datos.

- Frontend: desarrollado en Laravel 6, con soporte de JavaScript, jQuery, Leaflet) para mapas interactivos) y Bootstrap 5 (para estilos responsivos).

La Figura 3 presenta el esquema de componentes del sistema y sus protocolos de comunicación. La arquitectura modular garantiza un funcionamiento eficiente, facilita su escalabilidad y asegura la interoperabilidad con otras aplicaciones y servicios.

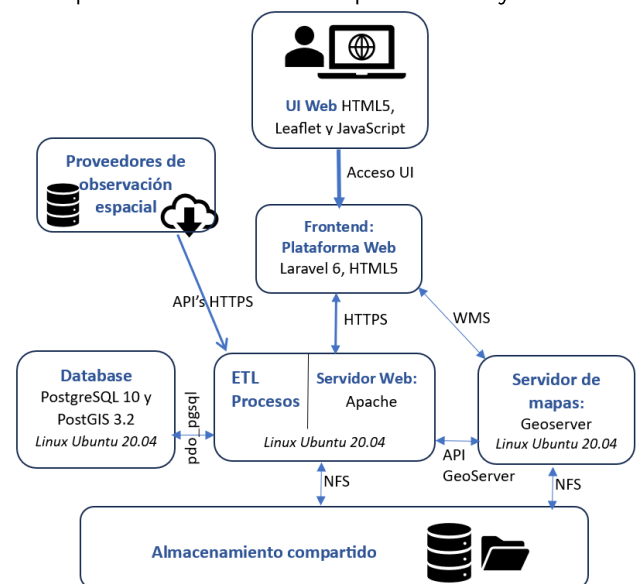


Figura 3. Componentes de software.
Fuente: elaboración propia

IV. APLICACIONES Y SERVICIOS PARA USUARIOS

El portal web del Cubo de Datos Geoespaciales de El Salvador (Figura 4) fue desarrollado para acercar los productos satelitales a usuarios no especializados. Entre sus funcionalidades se destacan:

- Consultar la temperatura superficial en áreas predefinidas o en Áreas de Interés (AOIs) definidas por el usuario.
- Visualizar mapas de cobertura de suelo derivados MODIS.
- Calcular y comparar en el tiempo índices de vegetación (NDVI, EVI, NBR) de Landsat y Sentinel-2.
- Generar series temporales de índices en puntos definidos por el usuario.
- Descargar imágenes para AOI, permitiendo análisis posteriores en software especializado.

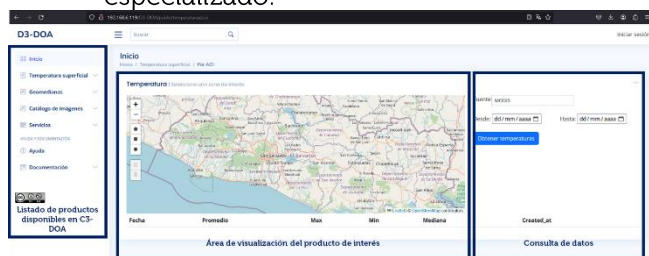


Fig. 4. Interfaz de usuario del cubo de datos disponible en <https://d3.snet.gob.sv/>.

Fuente: Tomado de (Climaco y Chávez, 2023)

El valor agregado principal radica en que los usuarios acceden a datos satelitales procesados y listos para el análisis, sin necesidad de conocimientos avanzados en SIG.

El Cubo de Datos Geoespaciales se extienden a diferentes ámbitos:

- Gestión ambiental: monitoreo de deforestación, cambios de uso de suelo, áreas naturales protegidas.
- Agricultura: seguimiento de vigor vegetativo, detección de estrés hídrico, análisis de sequías.
- Planificación territorial: evaluación de dinámicas urbanas y expansión de la frontera agrícola.
- Clima y energía: estudios de variabilidad térmica y análisis de islas de calor urbanas.

Cada producto tiene un propósito específico: por ejemplo, las geomedianas facilitan la comparación multitemporal al ofrecer una representación estable de la superficie terrestre, mientras que los índices de vegetación y los mapas de coberturas, permiten monitorear dinámicas ecológicas y procesos de degradación ambiental.

V RESULTADOS Y DISCUSIÓN

El desarrollo del Cubo de Datos Geoespaciales de El Salvador ha enfrentado retos significativos en:

- Almacenamiento: optimización del espacio mediante eliminación de datos redundantes y priorización de productos derivados.
- Rendimiento: procesamiento eficiente de grandes volúmenes de datos satelitales, especialmente Sentinel 2 y Landsat.
- Calidad de datos: implementación de algoritmos robustos de limpieza y enmascaramiento.
- Escalabilidad: diseño modular que permite integrar nuevos productos o fuentes en el futuro.

Entre las principales innovaciones destacan la automatización completa del flujo de trabajo, la integración entre backend (ETL) y frontend (portal web), y la democratización del acceso a datos satelitales en El Salvador.

VI CONCLUSIONES Y TRABAJO FUTURO

El Cubo de Datos Geoespaciales de El Salvador constituye una plataforma integral que automatiza la adquisición, procesamiento y publicación de imágenes satelitales, facilitando el acceso a información crítica para la gestión ambiental y territorial. Al reducir las barreras técnicas y ofrecer productos listos para su uso, el sistema impulsa la toma de decisiones informada y la investigación científica.

Entre las proyecciones futuras se incluyen: - Incorporación de nuevas fuentes satelitales, como Sentinel-3 y productos climáticos. - Integración de algoritmos de inteligencia artificial para clasificación automática. - Optimización de infraestructura para manejo de big data geoespacial. - Ampliación de casos de uso en ámbitos como gestión de riesgos y adaptación al cambio climático.

El proyecto constituye un paso clave hacia la democratización del uso de imágenes satelitales en El Salvador, con potencial de ser replicado en otros contextos regionales.

REFERENCIAS

- Appel, M.; Pebesma, E. On-Demand Processing of Data Cubes from Satellite Image Collections with the gdalcubes Library. *Data* 2019, 4(3), 92; <https://doi.org/10.3390/data4030092>.
- Augustin, H.; Sudmanns, M.; Tiede, D.; Lang, S.; Baraldi, A. Semantic Earth Observation Data Cubes. *Data* 2019, 4(3), 102; <https://doi.org/10.3390/data4030102>
- Chatenoux, Bruno & Röösl, Claudia & Wingate, Vladimir & Poussin, Charlotte & Rodila, Denisa & Peduzzi, Pascal & Steinmeier, Charlotte & Ginzler, Christian & Psomas, Achilleas & Schaeppman, Michael & Giuliani, Gregory. (2021). The Swiss data cube, analysis ready data archive using earth observations of Switzerland. *Scientific Data*. 8. 295. 10.1038/s41597-021-01076-6.
- Cheng, M-C., Chang, L-Y., Kawakita, S., Chen, B., Liu, C., Lin, H-C., Lin, L-C (2020). 'Development of Open Data Cube to Facilitate Disaster Risk Reduction'. IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium, 26 Sept-2 Oct 2020. DOI: 10.1109/IGARSS39084.2020.9324214.
- Christopherson, J.B., Ramaseri Chandra, S.N., and Quanbeck, J.Q. (2019). 2019 Joint Agency Commercial Imagery Evaluation—Land remote sensing satellite compendium: U.S. Geological Survey Circular 1455, 191 p., <https://doi.org/10.3133/cir1455>.
- Climaco, J., Chávez, T. (2023). Cubo de Datos de El Salvador. Diseño e implementación Año 1. Ministerio de Medio Ambiente de El Salvador.
- Copernicus Programme. (2023). Sentinel Data Access. <https://dataspace.copernicus.eu>
- NASA EarthData. (2023). <https://earthdata.nasa.gov>
- Coronado, A. [Facultad de Geografía UAEMex](27 de abril de 2021). Plática "El Cubo de Datos Geoespaciales de México una perspectiva de Big Data y Machine Learning"
- Daleroberts. (s. f.). GitHub - Daleroberts/HDMedians: High-dimensional medians (Medoid, Geometric Median, etc.). Fast implementations in Python. GitHub. <https://github.com/daleroberts/hdmedia>
- Dhu, T., Giuliani, G., Juárez, J., Kavvada, A., Killough, B., Merodio, P., Minchin, S., Ramage, S (2019). 'National Open Data Cubes and Their Contribution to Country-Level Development Policies and Practices'. *Data*, 4(4):144, <https://doi.org/10.3390/data4040144>.
- Didan, K. (2015). MODIS Vegetation Index Products (MOD13): Algorithm Theoretical Basis Document. NASA.
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., et al. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, 120, 25–36.
- Dwyer, J. L., et al. (2018). *Analysis Ready Data: Enabling analysis of the Landsat archive*. *Remote Sensing*, 10(9), 1363.
- ESA (2021). *Sentinel-2 User Handbook* European Space Agency.
- ESA, 2024. Sentinel-2C se une a la familia Copernicus en órbita. Recuperado de: [https://www-es-int.translate.google.com/Applications/Observing_the_Earth/Copernicus/Sentinel-2/Sentinel-2C_joins_the_Copernicus_family_in_orbit?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=tc](https://www.esa-int.translate.google.com/Applications/Observing_the_Earth/Copernicus/Sentinel-2/Sentinel-2C_joins_the_Copernicus_family_in_orbit?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=tc)
- Gebbert, S.; Leppelt, T.; Pebesma, E. A Topology Based Spatio-Temporal Map Algebra for Big Data Analysis. *Data* 2019, 4(2), 86; <https://doi.org/10.3390/data4020086>.
- Giuliani, G.; Masó, J.; Mazzetti, P.; Nativi, S.; Zabala, A. Paving the Way to Increased Interoperability of Earth Observations Data Cubes. *Data* 2019, 4(3), 113; <https://doi.org/10.3390/data4030113>.
- Giuliani, G., Câmara, G., Killough, B., & Minchin, S. (2019). Earth Observation Open Science: Enhancing Reproducible Science Using Data Cubes. *Data*, 4(4), 147. <https://doi.org/10.3390/data4040147>.
- Gorelick, N., et al. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18–27.
- Gómez, C., White, J. C., & Wulder, M. A. (2016). *Optical remotely sensed time series data for land cover classification: A review*. *ISPRS Journal of Photogrammetry and Remote Sensing*, 116, 55–72.
- INEGI (2023, 12 de julio). Índice de Vegetación de diferencia normalizada (NDVI) [Comunicado de Prensa]. <https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2023/IVDN/IVDN84-22.pdf>
- INEGI (2022). Índice de Vegetación Normalizada NDVI. Landsat, 1984-2021. Recuperado de: <https://www.inegi.org.mx/app/biblioteca/ficha.html?upc=889463908272>
- INEGI (2021). Producción del índice de clasificaciones de Agua Superficial desde el Espacio. ICASE. Landsat. Documento metodológico. México.
- INEGI (2020). Producción y publicación de la Geomediana Nacional a partir de imágenes del Cubo de Datos Geoespaciales de México : documento metodológico. México.
- INEGI (2020). Cubo de datos de México (CDGM). Aguascalientes.
- Juárez Carrillo, O.J., Merodio Gómez, P., del Socorro Ponce Medina, M., Ornelas de Anda, J.L., Coronado Iruegas, A.A (2020). 'Cubo de datos geoespaciales para el uso de las imágenes satelitales en la generación de información geográfica y estadística'. *Realidad, Datos y*

- Espacio, *Revista Internacional de Estadística y Geografía*, 11:3, pp. 124-139.
- Kopp, S.; Becker, P.; Doshi, A.; Wright, D.; Zhang, K.; Xu, H. Achieving the Full Vision of Earth Observation Data Cubes. *Data* 2019, 4(3), 94; <https://doi.org/10.3390/data4030094>.
- Landsat Collection 2 U.S. analysis ready data. (2023). Fact sheet /. <https://doi.org/10.3133/fs20233015>
- Landsat NASA. (2023, 13 septiembre). History | Landsat Science. Landsat Science | A joint NASA/USGS Earth observation program. <https://landsat.gsfc.nasa.gov/satellites/history/>
- Landsat NASA. (2023a, septiembre 11). Timeline | Landsat Science. Landsat Science | A joint NASA/USGS Earth observation program. <https://landsat.gsfc.nasa.gov/satellites/timeline/>
- Leith, A (2019). What is the Open Data Cube? Medium blog post. Accessible via: <https://medium.com/opendatacube/what-is-open-data-cube-805af60820d7> (accessed on 21/05/2021).
- Leith, A. (2018, 11 agosto). What is the Open Data Cube? - OpenDataCube - Medium. Medium. <https://medium.com/opendatacube/what-is-open-data-cube-805af60820d7>
- Lewis, A., Oliver, S., Lymburner, L., Evans, B., Wyborn, L., et al. (2017). 'The Australian Geoscience Data Cube-Foundations and lessons learned'. *Remote Sensing of Environment*, 2020, 276-292, <http://dx.doi.org/10.1016/j.rse.2017.03.015>.
- Louis, J., et al. (2016). Sentinel-2 Sen2Cor: L2A Processor for Users. *Proceedings of the Living Planet Symposium*.
- Lucas, R.; Mueller, N.; Siggins, A.; Owers, C.; Clewley, D.; Bunting, P.; Kooymans, C.; Tissott, B.; Lewis, B.; Lymburner, L.; Metternicht, G. Land Cover Mapping using Digital Earth Australia. *Data* 2019, 4(4), 143; <https://doi.org/10.3390/data4040143>.
- Mcdougall, K., & Koswate, S. (2018). The future of authoritative geospatial data in the big data world—Trends, opportunities and challenges. *Proceedings of the FIG Commission: Spatial Information in an Era of Data Science: Challenges and Practical Solutions*, Naples, Italy, 3-6.
- Mas, Jean. (2011). Aplicaciones del sensor MODIS para el monitoreo del territorio.
- Maso, J.; Zabala, A.; Serral, I.; Pons, X. A Portal Offering Standard Visualization and Analysis on top of an Open Data Cube for Sub-National Regions: The Catalan Data Cube Example. *Data* 2019, 4(3), 96; <https://doi.org/10.3390/data4030096>.
- Mueller, N., et al. (2016). Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia. *Remote Sensing of Environment*, 174, 341–352.
- Ornelas de Anda, J. L., Silván Cárdenas, J. L., Molina-Villegas, A., Hernández, G., López-Ramírez, P., Tapia-McClung, R., González Zuccolotto, K., & Chirinos Colunga, M. (2019). *Open Data Cube for Natural Resources Mapping in Mexico*. En: *Proceedings of the 1st International Conference on Geospatial Information Sciences (iGISc 2019)*, Kalpa Publications in Computing, 13, 70-78.
- Opendatacube. (s. f.). GitHub - OpenDatacube/Datacube-Stats: Data Cube Temporal Statistic Tools. GitHub. <https://github.com/opendatacube/datacube-stats/>
- Open Geospatial Consortium (2020). *Geospatial Coverages Data Cube Community Practice*. Version 1.0, OGC Document 18-095r7. Available at: <https://www.ogc.org/docs/cp>.
- Plag, H.; Jules-Plag, S. A Transformative Concept: From Data Being Passive Objects to Data Being Active Subjects. *Data* 2019, 4(4), 135; <https://doi.org/10.3390/data4040135>.
- Poussin, C.; Guigoz, Y.; Palazzi, E.; Terzago, S.; Chatenoux, B.; Giuliani, G. Snow Cover Evolution in the Gran Paradiso National Park, Italian Alps, Using the Earth Observation Data Cube. *Data* 2019, 4(4), 138; <https://doi.org/10.3390/data4040138>.
- Roy, D. P., Wulder, M. A., Loveland, T. R., Woodcock, C. E., Allen, R. G., Anderson, M. C., et al. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, 145, 154–172.
- Siqueira, A., Lewis, A., Thankappan, M., Szantoi, Z., Gory, P. et al. (2019). CEOS Analysis Ready Data for Land-An Overview on the Current and Future Work. *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pp.5536-5537.DOI: 10.1109/IGARSS.2019.8899846.
- Schubert, C.; Seyerl, G.; Sack, K. Dynamic Data Citation Service—Subset Tool for Operational Data Management. *Data* 2019, 4(3), 115; <https://doi.org/10.3390/data4030115>.
- Tascón, M.B. (2013). *Introducción: Big Data. Pasado, presente y futuro*.
- Ticehurst, C.; Zhou, Z.; Lehmann, E.; Yuan, F.; Thankappan, M.; Rosenqvist, A.; Lewis, B.; Paget, M. Building a SAR-Enabled Data Cube Capability in Australia Using SAR Analysis Ready Data. *Data* 2019, 4(3), 100; <https://doi.org/10.3390/data4030100>.
- Truckenbrodt, J.; Freemantle, T.; Williams, C.; Jones, T.; Small, D.; Dubois, C.; Thiel, C.; Rossi, C.; Syriou, A.; Giuliani, G. Towards Sentinel-1 SAR Analysis-Ready Data: A Best Practices Assessment on Preparing Backscatter Data for the Cube. *Data* 2019, 4(3), 93; <https://doi.org/10.3390/data4030093>.

- USGS (2018). MOD13Q1. recuperado de: <https://lpdaac.usgs.gov/products/mod13q1v006/>
- USGS Earth Resources Observation and Science (EROS) Center. (2023). <https://eros.usgs.gov>
- Wulder, M. A., & Coops, N. C. (2014). Making full use of the Earth Observation record: recent advances in land remote sensing and challenges ahead. *Remote Sensing of Environment*, 118, 127–137.
- Roy, D. P., et al. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, 145, 154–172.
- Zuo, Yijun. (2013). Multidimensional medians and uniqueness. *Computational Statistics & Data Analysis*. 66. 82–88. 10.1016/j.csda.2013.03.020.
- Small, C. G. (1990). A Survey of Multidimensional Medians. *International Statistical Review / Revue Internationale de Statistique*, 58(3), 263–277. <https://doi.org/10.2307/1403809>
- INEGI, 2020. Producción y publicación de la Geomediana Nacional a partir de imágenes del Cubo de Datos Geoespaciales de México. Documento Metodológico. México
- USGS (2021) a. Landsat 4-7. Collection 2 (C2). Level 2 Science Product (L2SP) Guide. Sioux Falls, South Dakota.
- USGS (2021b). Landsat 8-9. Collection 2 (C2). Level 2 Science Product (L2SP) Guide. Sioux Falls, South Dakota.
- Wulder, M. A., & Coops, N. C. (2014). Satellites: Make Earth observations open access. *Nature*, 513(7516), 30–31.