
**EL CENTRO DE INFORMACIÓN GEOESPACIAL DE GEOCUBA, SOPORTE
PARA LA UTILIZACIÓN DE IMÁGENES DE SENTINEL 2 THE
GEOSPATIAL INFORMATION CENTER OF GEOCUBA, SUPPORT
FOR THE USE OF SENTINEL 2 IMAGES**

Dr C. José Luis Capote Fernández (1)

Dr C. Rafael Cruz Iglesias (2)

1. Unidad Científico Técnica GEOCUBA de Investigación y Consultoría.

capote@geomix.geocuba.cu

2. Unidad Científico Técnica GEOCUBA de Investigación y Consultoría.

rcruz@geomix.geocuba.cu

RESUMEN

El Centro de Información Geoespacial de GEOCUBA (CIG) tiene como principal misión la gestión de información geoespacial de la institución a lo interno y para la sociedad. Desde su concepción se establecieron las bases para una gestión basada en tecnologías de Big Data y servicios que garantizaran alta disponibilidad y desempeño. Una vez establecida la infraestructura del centro se comenzó un trabajo relacionado con la gestión completa de imágenes multiespectrales, incluyendo su almacenamiento, procesamiento y publicación en servicios geoespaciales basados en estándares de ISO y OGC. Se crearon los flujos de trabajo que descargan continuamente las imágenes desde su origen y se suben al almacén del centro, creando los correspondientes metadatos que garantizan el acceso a los mismos. Se cuenta con todas las imágenes Sentinel-2 sobre nuestro país con menos de un 60% de nubosidad desde enero del 2020, lo que permite hacer varios análisis espacio-temporales. Con esta infraestructura como soporte se han desarrollado un grupo de servicios, utilizando el estándar para Servicio de Procesamiento Web de OGC que permiten la interoperabilidad con otros proyectos. Con la utilización de estos servicios se generan productos que son utilizados por los clientes y entre ellos resaltan los mapas mensuales con mínimo de nubes que se vienen publicando desde enero del 2022.

Palabras Clave: Imágenes multiespectrales, Sentinel 2, Servicios de procesamiento Web

ABSTRACT

The main mission of the GEOCUBA Geospatial Information Center (CIG) is to manage the institution's geospatial information internally and for society. From its conception, the foundations were established for management based on Big Data technologies and services that guaranteed high availability and performance. Once the center's infrastructure was established, work began related to the complete management of multispectral images including their storage, processing and publication in geospatial services based on ISO and OGC standards. Workflows were created that continuously download images from their source and upload them to the center's warehouse, creating the corresponding metadata that guarantees access to them. All Sentinel-2 images are available over our country with less than 60% cloud cover since January 2020, which allows various spatio-temporal analyzes to be carried out. With this infrastructure as support, a group of services have been developed using the OGC Web Processing Service standard that allows interoperability with other projects. With their use, products are generated on the maps that are used by clients. Within these products, the monthly maps with minimum clouds that have been published since January 2022 stand out.

Key Words: Multispectral Imaging, Sentinel 2, Web Processing Services

Recibido: 13/09/2023

Aprobado: 13/11/2023

1. INTRODUCCIÓN

El grupo empresarial GEOCUBA posee experiencia en el trabajo con imágenes satelitales, convencionales y multiespectrales. Existen agencias de teledetección que trabajan frecuentemente con imágenes de satélites y las utilizan con diferentes propósitos.

En el pasado el acceso a dichas imágenes era un proceso engorroso, generalmente eran adquiridas y financiadas por determinados proyectos y más tarde replicadas entre los especialistas para su análisis. En los últimos años, con el desarrollo de Internet y el surgimiento de proyectos de observación a escala global, se ha facilitado el acceso a las imágenes y su utilización se ha hecho más frecuente entre los especialistas de la Geomática.

Desde enero del 2020, como parte del proyecto Centro de Información Geoespacial de GEOCUBA, se descargan y pre procesan todas las imágenes del satélite Sentinel 2 sobre nuestro país. Para facilitar el acceso interoperable a estas se han desarrollado servicios de procesamiento y se han generado productos listos para consumir por otros proyectos que demandan información Geoespacial.

2. MATERIALES Y MÉTODOS

2.1 Centro de Información Geoespacial de GEOCUBA

El proyecto comenzó su ejecución en el año 2018 con el objetivo de garantizar la existencia de una infraestructura de hardware y software que soporte el despliegue de los datos, servicios y aplicaciones, patrimonio espacial digital del grupo empresarial GEOCUBA cuya visión es:

- La gestión de la información geoespacial de la Institución, a partir de las necesidades demandadas por el Estado, la sociedad y otras empresas externas o internas de la organización.
- La existencia de una plataforma de estándares, normas de calidad, tecnologías y reglas de operación que integran la plataforma y satisfacen los requerimientos para recopilar, catalogar, distribuir y acceder a la información geoespacial.
- La garantía de una alta disponibilidad, redundancia, replicación, conectividad y calidad en el servicio, para el manejo de grandes volúmenes de información geoespacial con facilidad de manejo de datos en tiempo real.
- El establecimiento de los modelos de negocios internos y externos sobre el uso de los datos, servicios y aplicaciones geoespaciales del Centro que garantizan su sostenibilidad. Varios clientes hacen uso de los servicios del centro y los procesos diseñados garantizan la producción requerida; creación, actualización, publicación de Datos, Servicios y Aplicaciones.

Cada uno de estos aspectos tiene esta estrategia de trabajo con retos particulares. En trabajos anteriores (1,2) se expuso la evaluación de varias tecnologías relacionadas con la gestión de datos ráster en entornos de Big Data. En este trabajo se presentan algunos productos informáticos obtenidos en el marco del proyecto Centro de Información Geoespacial de GEOCUBA basados en la gestión de las imágenes multiespectrales del satélite Sentinel 2.

2.2 MrGeo como soporte a imágenes

MrGeo (MapReduce Geoespacial) es un proyecto que contiene un conjunto de herramientas diseñado para realizar un procesamiento geoespacial ráster a escala regional o global (3). Al distribuir el cómputo entre varias computadoras puede reducir considerablemente el procesamiento en comparación a los sistemas tradicionales. MrGeo proporciona dos formas de visualizar estos productos a través de los estándares Consorcio Geoespacial Abierto (OGC), utilizando el software GeoServer por medio de una extensión o por una implementación propia de los servicios WMS, WMTS y WCS.

Dentro de las funcionalidades que presenta está el almacenamiento y procesamiento escalable de datos ráster. Propone almacenar los datos en un formato listo para el cálculo, eliminando varios pasos de preprocesamiento del flujo de trabajo de producción. Además, proporciona un conjunto de métodos de análisis robustos para Apache Spark que incluyen operaciones matemáticas algebraicas, operaciones focales (por ejemplo, pendiente y aspecto). MrGeo presenta una interfaz de álgebra de mapas que permite el desarrollo de algoritmos personalizados en una simple API de interpretación de guiones de ejecución.

Una de las ventajas de MrGeo es que garantiza un flujo de trabajo desde el almacenamiento del ráster hasta su publicación como servicio OGC. MrGeo se ha utilizado para almacenar, indexar, crear mosaicos y bases de datos de imágenes a escala de múltiples terabytes en pirámide. Una vez almacenados los datos, están disponibles mediante un servicio de mosaicos de mapas (TMS), servicio de mapa en Web (WMS) y servicio de coberturas en Web (WCS).

El proyecto MrGeo incluye un subproyecto que permite su utilización desde el lenguaje Python llamado pymrgeo. Este proyecto hace simple la ejecución de las funcionalidades de álgebras de mapas dentro de un guión de ejecución.

En el caso de MrGeo se realizaron modificaciones a la versión del código fuente para resolver algunos problemas detectados en la ejecución de las pruebas con los juegos de datos y para la creación de un producto más afín, tanto con las tecnologías y las capacidades de hardware disponibles en nuestra empresa, como a los objetivos propios del proyecto.

2.3 Imágenes del Satélite Sentinel 2

La Agencia Espacial Europea (ESA), en el marco del programa Copérnico para la Vigilancia Mundial del Medio Ambiente y la Seguridad, ha creado un portal de acceso a los productos de las diferentes misiones que desarrolla entre las que se encuentra el satélite Sentinel 2. El satélite Sentinel 2 lleva una cámara multispectral de alta resolución, con 13 bandas que aportan una nueva perspectiva de la superficie terrestre y la vegetación. La combinación de la alta resolución y las nuevas capacidades espectrales, así como un campo de visión que abarca 290 kilómetros de anchura y sobrevuelos frecuentes (de 4 a 5 días), proporciona vistas de la Tierra sin precedentes. La misión se basa en una constelación de dos satélites idénticos en la misma órbita, separados por 180 grados, para lograr una cobertura y una descarga de datos óptima. Cada cinco días los satélites cubren todas las superficies terrestres, grandes islas y aguas costeras. La misión proporciona información útil para las prácticas agrícolas y forestales, y para gestionar la seguridad alimentaria. Las imágenes pueden ser descargadas desde el propio portal

de Copérnico utilizando un navegador o utilizando interfaces para aplicaciones disponibles para este fin (SentinelSat). El formato original es un archivo comprimido con amplia información sobre los datos (metadatos) y las imágenes en formato JPEG2000 para cada una de las bandas en resoluciones de 10m, 20m y 60m (4,5).

El Centro descarga solamente imágenes que presentan corrección atmosférica, nivel L1A y L1B. El territorio de la isla de Cuba está cubierto por 29 imágenes. Aunque la frecuencia de cada una es de 4 a 5 días, en las zonas de solapamiento la frecuencia puede ser mayor permitiendo un mejor monitoreo (Figura 1).



Figura 1: Cartograma de imágenes del satélite Sentinel 2 para Cuba.

Aunque en el cartograma la imagen tiene una apariencia cuadrada, en la realidad existen varios problemas que pueden afectar la utilidad de la imagen. Dos de los factores que más afectan el flujo que se realiza en el CIG son la nubosidad y la falta de cubrimiento. La información de las nubes se obtiene de la capa MASK (Figura 2). Se han desarrollado algoritmos que minimizan los efectos de ambos sobre los resultados.

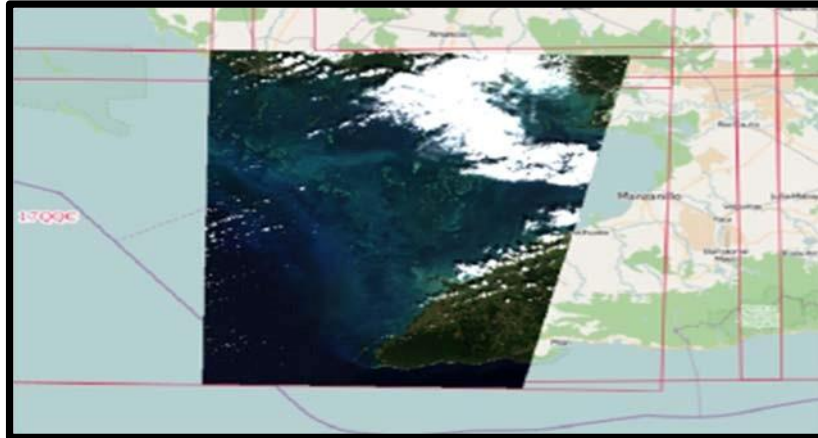


Figura 2: Nubosidad y falta de cubrimiento de una imagen de Sentinel 2.

Actualmente para el CIG se descargan solo las imágenes que tienen menos de un 60% de nubosidad, y como parte del metadato original proporcionado por la misión se obtiene el área útil de la imagen, que se incorpora al metadato propio que se genera.

3. RESULTADOS Y DISCUSIÓN

3.1 Generación de imágenes con mínimo de nubes

En potentes plataformas de procesamiento como Google Earth Engine se ejecutan procesos de generación de imágenes sin nubes, utilizando una gran cantidad de imágenes que combinan máscaras de nubes y sombras, obtenidas de la información del producto L1C del Sentinel 2. Como parte de este trabajo se ha diseñado un procedimiento para generar imágenes, con un mínimo de nubes y utilizando -como base- un cartograma 1:25 000 (Figura 3).



Figura 3: Cartograma 1:25 000 como índice de procesamiento para buscar imágenes con mínimo de nubes.

Los pasos del Procedimiento son los siguientes:

1. Para cada polígono del cartograma 1: 25 000 se buscan las imágenes con un porcentaje de nubes aceptable (por defecto 5%), ordenadas comenzando por las más recientes. Este proceso se realiza a partir de calcular la estadística zonal de la capa de nubes para el polígono. Si el polígono no es interior al borde útil (footprint) de la imagen, se calcula la unión geométrica y si con esta unión se satisface el requerimiento de nubosidad se almacena el polígono con los datos estadísticos. Con la parte restante del proceso anterior se realiza la búsqueda, partiendo desde las más recientes. Un polígono de entrada puede devolver varios polígonos de salida, referenciando diferentes imágenes. Al cartograma original se le añaden como atributos el porcentaje de nubes, la imagen identificada, la fecha de la misma y otros datos estadísticos.
2. Una vez detectada la imagen con un mínimo de nubes se calcula el proceso objetivo. Por defecto este proceso es el mapa de color natural, pero es posible generar imágenes más complejas como índices o clasificaciones. Un proceso WPS desarrollado para este fin es invocado para realizar la operación de cálculo utilizando álgebra de mapas con MrGeo, generar la imagen, ingestarla en el clúster de datos y agregar al archivo índice inicial un campo con la URL de la imagen correspondiente e información referente al proceso objetivo.
3. Teniendo para cada polígono una imagen que representa el proceso objetivo queda el proceso de componer el gran mosaico, generar una imagen única, ingestarla en el clúster para que quede disponible para otros procesos y generar los metadatos para que pueda ser encontrada.

3.2 Servicios de procesamiento

Para implementar los nuevos procesos se utilizó el proyecto PyWPS(6), una implementación del estándar para Servicio de Procesamiento Web (WPS) de OGC(7) en lenguaje Python. En función de la obtención de los productos con un mínimo de nubes con las imágenes de Sentinel 2 se desarrollaron los siguientes procesos:

1. CIGSentinelZonalStats: Estadísticas de una capa de una imagen Sentinel 2 para una zona. Tiene como entradas a la capa para obtener las estadísticas relativas al polígono, el intervalo de fechas para buscar las imágenes, así como el mínimo porcentaje de nubes para parar el proceso, los polígonos de entrada en formato geojson(8) y el sistema de referencia espacial

de los mismos. Como salida se obtiene un archivo geojson, con uno o más polígonos por cada polígono de entrada, el nombre de la imagen identificada, la fecha de la misma y las estadísticas de la capa para el polígono de entrada.

2. CIGGenericExpression: Calcular índice a partir de la expresión genérica sobre imagen multiespectral. Ejemplos:

Í 08 03 / 08 03 Í 08/041 / 08/041

Los operandos son los nombres de las bandas B02, B03, B04, los operadores posibles son: +, *, /, sqrt, etc. Como parámetro de entrada recibe un geojson con polígonos, una expresión y una imagen o las fechas para encontrarla. Devuelve el mismo geojson con la URL a una imagen con la expresión calculada y la fecha de la imagen con la imagen utilizada. Con la combinación de estos dos procesos se ejecuta el procedimiento descrito anteriormente.

3.3 Imágenes con mínimo de nubes

A partir del procedimiento descrito, desde enero de 2022 se vienen generando imágenes con mínimo de nubes de todo el territorio nacional, así como un mapa con las mejores imágenes de los últimos tres meses. Esta forma de presentar los resultados con un mapa mensual tiene gran utilidad para los usuarios, ya que tienen una temporalidad fija permitiendo realizar análisis más complejos. Además, el archivo índice utilizado como base, permite saber la fecha exacta de la imagen utilizada. Estos productos se utilizan desde diferentes proyectos de IDE desarrollados por GEOCUBA: MINAG, AZCUBA, Tarea VIDA (Figura 4).



Figura 4: Imagen con mínimo de nubes desde el visor de la Infraestructura de Datos Espaciales de la Tarea VIDA (<https://idevida.geocuba.cu/visor>).

4. CONCLUSIONES

En el Centro de Información Geoespacial de GEOCUBA se ha desarrollado una infraestructura que permite el procesamiento y análisis de imágenes del satélite Sentinel 2 de todo el país. Se han desarrollado servicios de procesamiento interoperables para interactuar con los datos y extraer información relevante a diferentes objetivos de investigación, tales como mapas de color natural e índices de vegetación.

Como parte de este procesamiento continuo de imágenes de Sentinel 2 se obtienen capas, con un mínimo de nubes y con una frecuencia mensual, que hoy forman parte de varios productos de GEOCUBA.

5. REFERENCIAS

1. Fernández JLC, Iglesias RC. Tecnologías de Big Data geoespacial en el Centro de Información Geoespacial de Geocuba. *Rev Cuba Transform Digit.* 2020;1(2):74–85.
2. Cruz Iglesias R, Capote Fernández JL, González Suarez G, Batule Domínguez M. Arquitectura escalable y de alto desempeño para la provisión de servicios geoespaciales. In: *Convención y Feria Informática 2022, CONGRESO INTERNACIONAL GEOMÁTICA 2022.* Cuba: Ministerio de Comunicaciones; 2022.
3. National Geospatial-Intelligence Agency. GitHub. 2017 [cited 2023 Apr 21]. MrGeo on GitHub. Available from: <https://github.com/ngageoint/mrgeo/wiki/Home>
4. ESA. El programa Copérnico [Internet]. 2017 [cited 2023 Apr 21]. Available from: https://www.esa.int/Space_in_Member_States/Spain/El_programa_Copernico
5. ESA - Introducing Sentinel-2 [Internet]. [cited 2023 Jul 5]. Available from: https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2/Introducing_Sentinel-2
6. Čepický J, de Sousa LM. New implementation of OGC Web Processing Service in Python programming language. PyWPS-4 and issues we are facing with processing of large raster data using OGC WPS. *Int Arch Photogramm Remote Sens Spat Inf Sci.* 2016;41:927–30.
7. De la Beaujardiere J. OGC Implementation Specification 05-007r7: OpenGIS Web Processing Service [Internet]. Open Geospatial Consortium; 2003 [cited 2009 Sep 22].

Available from: <http://www.opengeospatial.org/standards/wms>

8. Butler H, Daly M, Doyle A, Gillies S, Hagen S, Schaub T. The gejson format. 2016.
9. Beaujardiere, J. de la. 2003. "OGC Implementation Specification 05-007r7: OpenGIS