

Desarrollo de un Sistema de Monitoreo y Manejo Integral de Humedales a partir de Información Satelital

Natalia Soledad Morandeira², Francisco Grings³, Mariano Franco³, Adriana Rojas Barrios², Mariela Rajngewerc⁴, Matías Barber³, Mercedes Salvia³, Elio Campitelli^{5,6,7}, Esteban Roitberg⁸, Carlos David Clauser², Priscilla Minotti⁸, Pía Fernández⁹, Maira Gayol³ - DNI 33521319, Lautaro Cortés¹, Micaela Milano¹, Juan Pujó¹, Julia Martinuzzi¹, Ariel Anthieni¹, Walter Shilman¹, Maria Iannone¹

¹ Kan Territory & IT. Arenales 2370 Piso 2 Departamento D CP 1124 - CABA
administración@kan.com.ar

² IIIA-UNSAM-CONICET, Instituto de Investigación e Ingeniería Ambiental, Escuela de Hábitat y Sostenibilidad, Universidad Nacional de San Martín (UNSAM) - Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). iiia@conicet.gov.ar.

³ Instituto de Astronomía y Física del Espacio, Universidad de Buenos Aires (UBA) - CONICET. difusion@iafe.uba.ar

⁴ Sección Ciencias de la Computación, Facultad de Matemática, Astronomía, Física y Computación, Universidad Nacional de Córdoba - CONICET. <https://www.famaf.unc.edu.ar>

⁵ Departamento de Ciencias de la Atmósfera y los Océanos, FCEN-UBA. <http://www.at.fcen.uba.ar/>

⁶ Centro de Investigaciones del Mar y la Atmósfera (CIMA). UBA-CONICET. <https://www.cima.fcen.uba.ar/>

Resumen: El proyecto consiste en un sistema de monitoreo y gestión integral de humedales de Argentina a partir del desarrollo de modelos de áreas de estudio representativas de las unidades de paisaje de humedal piloto, en la Región de Humedales del corredor fluvial ChacoMesopotámico, a partir del uso de información satelital, productos de la misión SAOCOM, Sentinel 1 y 2 resultando como producto final la disponibilización de un Gestor de Contenido Geoespacial (GeoNode). El proyecto implica el desarrollo de modelos programados en Python, asimismo la automatización de los procesos ETL (sigla en inglés de extracción, transformación y carga) se llevaron adelante con la plataforma Airflow, implementando como solución específica para el manejo de información radar a la aplicación Snap. Como resultado marcado, el GeoNode, no solamente sirve en este proyecto como un webgis, sino como una herramienta integrada de manera transparente al Nodo IDE de CONAE, a través de los servicios interoperables OGC (siglas en inglés de Open Geographic Consortium), permitiendo que los datos, modelos y productos que se generen sean beneficiosos para las organizaciones y reutilizables en la integración de información para análisis de toma de decisiones a nivel del ecosistema de humedales y todos los actores involucrados.

Palabras Clave: conae, humedal, geonode, airflow, teledetección, imágenes satelitales, IDE, unsam.

1. INTRODUCCIÓN

El proyecto CONAE-PROSAT II Humedales “Desarrollo de un sistema de monitoreo y manejo integral de humedales a partir de información satelital” es desarrollado por consultores del Instituto de Investigación e Ingeniería Ambiental de la Universidad Nacional de San Martín y CONICET, del Instituto de Astronomía y Física del Espacio (UBA-CONICET) y del Canada Centre For Remote Sensing (Canadá), con participación de desarrolladores de la empresa argentina Kan - Territory I&T. El proyecto es ejecutado por la Fundación UNSAM Innovación y Tecnología.

El proyecto consiste en un sistema de monitoreo y gestión integral de humedales de Argentina a partir del desarrollo de modelos de áreas de estudio representativas de las unidades de paisaje de humedal piloto en la Región de Humedales del corredor fluvial Chaco-Mesopotámico ubicadas en las provincias de Entre Ríos y Santa Fe a partir del uso de información satelital, con foco en el aprovechamiento de productos de la misión SAOCOM, y resultando como producto final la disponibilización de un Gestor de Contenidos Geoespaciales.

El primer objetivo fue elaborar un catálogo de coeficientes de retrodispersión (microondas activas, banda C y L) y de respuestas espectrales (visible e infrarrojo reflectivo) asociados a los tipos de humedal de las áreas de estudio. El segundo objetivo consistió en elaborar protocolos basados en datos satelitales SAR de la Misión SAOCOM para la clasificación automatizada de tipos de humedal. En tercer lugar, se generaron productos de actualización periódica que dan cuenta del estado de los humedales, asociados a su hidropériodo. Por último, se generó una plataforma WebSIG de gestión de contenidos geoespaciales, que permite ingerir, compartir, navegar, obtener vistas previas, visualizar, analizar y difundir datos de los productos de humedales elaborados, interoperable con la IDE de CONAE.

En el presente documento se detalla la selección de las áreas de estudio piloto, se resumen las principales fuentes de datos y escenas satelitales, se detalla el preprocesamiento de datos satelitales que se lleva a cabo y se resumen los requerimientos del sistema.

2. ÁREA DE TRABAJO

En la primera etapa del proyecto se definieron las tres áreas de trabajo piloto. En cada área de trabajo se delimitaron unidades de paisajes de humedal (Nivel III del Inventario Nacional de Humedales, INH) y se elaboró una tipología de humedales para la clasificación de unidades de humedal (Nivel IV del INH). La elección de las regiones de humedales y áreas de trabajo piloto en las que se desarrolla el PROSAT-Humedales se fundamenta en la diversidad de patrones de paisajes que presentan, involucrando tanto paisajes de mosaicos de humedales (es decir, áreas completamente dominadas por humedales de distinto tipo) como paisajes con humedales en una matriz terrestre (áreas en las que los humedales de distintos tipos son parches aislados en un entorno de áreas altas o ecosistemas terrestres) (Kandus et al., 2018). Así, los desarrollos generados en estos sitios piloto tienen potencial de aplicación en otras regiones de humedales, dado que estos tipos de patrones de paisaje constituyen unidades que se repiten a lo largo del país (independientemente del tipo de humedales presentes). Las áreas de trabajo son parte de la región Humedales del corredor fluvial Chaco-Mesopotámico. Las áreas 1 y 2 se encuentran en la Subregión Ríos, esteros, bañados y lagunas del río Paraná, mientras que el área 3 se encuentra en la Subregión Malezales, tembladerales y arroyos litoraleños.

Delta Superior (Área piloto 1).

Localizada en la provincia de Entre Ríos. Localidades cercanas: Diamante, Puerto Gaboto,

Las Cuevas. Coordenadas centrales: 32,4092° S 60,6053° O. El Delta Superior es parte del sistema "Humedales del río Paraná con grandes lagunas" descrito por Marchetti et al. (2013). La planicie aluvial presenta cuerpos lóticos y lénticos de variados tamaños con conectividades y dinámicas diferentes, en los que el régimen hidrológico está fuertemente condicionado por los pulsos de inundación del río Paraná. En cuanto a la vegetación, se presentan bosques nativos y bañados con herbáceas. En el área del Sitio Ramsar que abordamos en este PROSAT, se cuenta con un inventario de humedales (Kandus et al. 2019). Las unidades de paisaje típicas se caracterizan por incluir un mosaico de humedales, con predominancia de áreas naturales protegidas (Parque Nacional Pre-Delta) o de interés de conservación (Sitio Ramsar Delta del Paraná). Las actividades productivas predominantes son ganadería vacuna extensiva y pesca. No presenta áreas urbanas ni actividad agrícola intensiva a destacar.

Bajo de los Saladillos (Área piloto 2).

Ubicada en la provincia de Santa Fe. Localidades cercanas: San Javier, Helvecia, Cayastá. Coordenadas centrales: 30,8529° S 60,1479° O. El sistema "Humedales del Bajo de los Saladillos" descrito por Ramonell et al. (2013), constituye una franja inundable en la que los humedales corresponden a las planicies de inundación actuales de los arroyos Saladillo Dulce y Saladillo Amargo (únicos cauces activos del sistema), a las lagunas y esteros en antiguas fajas del Paraná y a las cañadas y esteros en paleocauces meandriformes y bajíos entre los cuerpos de dunas longitudinales. La fuente de agua es de origen pluvial, fluvial y/o subterránea de acuerdo al sector. El área presenta un mosaico de humedales naturales (pastizales, pajonales, sabanas y bosques) con presencia de ganadería vacuna extensiva y actividad pesquera recreativa, más una actividad agrícola intensa en la que predominan los cultivos de arroz.

Concepción del Uruguay (Área piloto 3).

Localizado en la provincia de Entre Ríos. Localidad cercana: Concepción del Uruguay. Coordenadas centrales: 32,4639° S 58,2868° O. El sistema presenta tanto unidades de paisaje en las que predominan humedales como unidades de paisaje en la que los humedales se encuentran distribuidos en una matriz terrestre (Rojas y Kandus 2021). El régimen hidrológico está condicionado por las precipitaciones y por la dinámica del río Uruguay y sus tributarios. Presenta áreas urbanas, humedales naturales y actividad agrícola en las tierras altas.

3 FUENTES DE DATOS Y ESCENAS SATELITALES

A continuación se listan las principales fuentes de datos geográficos y satelitales a ser utilizados en el marco del proyecto. En el catálogo SAR y espectral se ingestan datos satelitales ópticos de Sentinel-2 y datos SAR de Sentinel-1 y SAOCOM StripMap QP (polarización completa). Las escenas ópticas se utilizan corregidas a reflectancia de superficie. Las escenas SAR se procesan calibran y procesan a intensidad (σ^0). En las clasificaciones, se utilizan sólo escenas SAOCOM StripMap QP, de las cuales se extraen un conjunto de 20 indicadores polarimétricos. Finalmente, para la construcción de los indicadores de estado del hidoperíodo, se utilizan escenas SAR Sentinel-1. Las/los usuarios/os del proyecto cargarán las delimitaciones de las áreas piloto y de las unidades de paisaje de humedales, en formato vectorial. El pre-procesamiento SAR se realizó utilizando el software libre SNAP v.8 (European Space Agency), en lotes con grafos .xml y automatizado con Python; mientras que los indicadores polarimétricos se generan con scripts de Python elaborados por el equipo.

4 REQUERIMIENTOS DE LA SOLUCIÓN

La estructura tecnológica del proyecto está basada en los siguientes componentes:

Plataforma WebGIS y catálogo de metadatos

La plataforma webGIS está integrada por GeoNode, un sistema de gestión de contenido geoespacial que facilita la publicación y acceso a datos georreferenciados y metadatos de varias fuentes, mejorando el intercambio de información entre organizaciones y diversos actores, usando las capacidades de la Internet. Asimismo, posibilita configurar fácilmente un portal geoespacial permitiendo a los usuarios ingerir, compartir, navegar, obtener una vista previa, visualizar y difundir datos. GeoNode está integrado por dos herramientas principales, GeoServer y Postgres, que brindan un servicio robusto de datos al usuario. Además, está integrado a un catálogo de metadatos (administrador de recursos referenciados espacialmente cuyas funciones más potentes son las de edición y búsqueda) contemplando los estándares de calidad nacionales e internacionales. Toda la interfaz está destinada a perfiles de usuarios tanto técnicos, como a aquellos que no son especializados en geomática. Entre sus principales características se destacan las siguientes:

- Cargar, administrar y compartir datos en línea.
- Crear mapas interactivos.
- Visualizaciones sobre el mapa (Widgets/Geo Stories).
- Permisos sobre capas y mapas de lectura, escritura y edición.
- Descarga de datos.
- Carga de datos como shapefiles, archivos CSV y archivos GeoTiff, así como documentos auxiliares como PDF, imágenes, videos (con la posibilidad de habilitar visualizaciones basadas en el tiempo).
- Integración con aplicaciones de terceros.
- Plataforma escalable.
- Gestión de metadatos.
- Front-end amigable.
- Interoperabilidad.
- API para personalización y conexión.
- Geoservicios OGC para la provisión de interacción con la IDE CONAE a través de los diferentes servicios, considerando el desarrollo de la funcionalidad de CSW bajo norma 19115-3 / 19139.

Orquestador de modelo (Airflow)

Es el entorno para la gestión de la información entre distintas plataformas que brinda la posibilidad de automatizar y monitorear los diferentes procesos para la disponibilización de la información.

Airflow posibilita crear, programar y monitorear flujos de trabajo de manera programática, además, planificarlos y monitorearlos de forma centralizada. Entre las prestaciones de Airflow se destacan:

- Gestión de fallos en data sources o sinks: Gestión integrada de errores en el acceso a datos pudiendo definir diferente comportamiento para cada caso.
- Reprocesamiento de jobs históricos: ordenar reprocesar jobs del pasado directamente desde el GUI de forma muy sencilla.
- Parámetros entre jobs: pasar parámetros entre diferentes jobs a través de un almacenamiento intermedio, esto nos permite dotar a los workflows de un estado.
- Reintentos automáticos: gestión automática de reintentos en base a la configuración que hagamos para cada DAG (en inglés, Directed Acyclic Graphs).
- Soporte CI (en español, Integración Continua) y CD (en español, Distribución Continua) en workflows: integrar fácilmente Airflow con nuestros sistemas de integración o despliegue continuo, ya sea con herramientas como Jenkins o gestión de DAGs en Git.
- Muchas integraciones: Airflow tiene integraciones con multitud de plataformas y software de terceros a través de los Operators.
- Data sensors: son un tipo particular de operadores que nos permiten esperar hasta que se cumpla cierta condición, por ejemplo, que se escriba un fichero en HDFS o en S3.
- Capacidades para testing de jobs: permite testear y validar nuestros tests antes de que sean desplegados.
- Logs y metadatos desde el GUI web: un punto centralizado desde donde acceder a todos los logs y metadatos asociados a cada tarea desde el interfaz web.
- Triggers que disparen tareas: tenemos diferentes opciones para disparar la ejecución de nuestros workflows de forma automática en base a una planificación temporal, o también podemos dispararlos de forma manual.
- Monitoreo en tiempo real y alertas: podemos monitorear el estado de ejecución de nuestros workflows en tiempo real desde el GUI.

Procesamiento de imágenes

De acuerdo a la escena satelital a procesar se generaron distintos scripts que realizan el pre-procesamiento de la imagen (de ser necesario), junto con la extracción de subproductos y muestras para la clasificación de los humedales. Todos los scripts fueron elaborados por los equipos del Instituto de Investigación e Ingeniería Ambiental de la Universidad Nacional de San Martín y CONICET, del Instituto de Astronomía y Física del Espacio (UBA-CONICET) y luego acondicionados para su funcionamiento dentro de Airflow por el equipo de KAN. Cada uno de estos scripts fue incorporado a un flujo de trabajo distinto con el formato de un DAG de Airflow.

El procesamiento de imágenes se realizó principalmente en python, utilizando en algunos pasos scripts de bash y la API para python de Google Earth Engine. Las librerías ráster utilizadas fueron en su mayoría GDAL y Rasterio en menor medida. Una vez obtenidos los índices necesarios, se elabora un tablero de control mediante ShinyApps con R.

Para imágenes SAOCOM, el flujo de automatización de las escenas SAOCOM comienza con un script que conecta y descarga del servicio FTP las imágenes requeridas. Posteriormente, un segundo paso ejecuta el pre-procesamiento de cada escena descargada utilizando grafos .xml con las herramientas disponibles en el software SNAP. Una vez que se obtiene cada escena pre-procesada se continúa con una traslación y corrección de las escenas para su completo alineamiento, de manera tal de poder obtener muestras e indicadores uniformes a través de las distintas imágenes. El siguiente script dentro del DAG, realiza la extracción de muestras de cada región de interés, definidas en archivos geográficos de tipo Shapefile, para calcular los índices polarimétricos. El último paso del flujo corresponde a la importación de los sets de datos resultantes y las escenas procesadas a Geonode junto con sus metadatos. Para esto se renderizan las escenas con una composición de falso color compuesto.

En el caso de Sentinel-1, el flujo empieza de manera similar a SAOCOM con un script que conecta y descarga del servicio del Alaska Satellite Facility (ASF) las imágenes requeridas. El segundo paso ejecuta el pre-procesamiento con las herramientas de SNAP. La extracción de muestras de cada región de interés se realiza con el mismo script, utilizando parámetros particulares para la cantidad de bandas de este tipo de escena. Finalmente se realiza la importación de los set de datos resultantes y las escenas procesadas a Geonode junto con sus metadatos.

Para el caso de las Sentinel-2, las imágenes son obtenidas y procesadas mediante scripts de python que hacen uso de la API de Google Earth Engine. En primer lugar se obtiene la colección de imágenes de acuerdo a los parámetros de fecha, tipo de producto y el filtro espacial dado por el área piloto buscada. Luego se construyen los diversos índices requeridos (NDVI, NWI, EVI, entre otros) para las regiones de interés sobre el conjunto de imágenes y se almacena todo en un mismo *dataframe* que es exportado como CSV. El último paso comprende un DAG que importa el resultado en Geonode.

En cuanto al análisis de inundabilidad se utilizan imágenes Sentinel-1 procesadas mediante scripts de python que hacen uso de la API de Google Earth Engine de acuerdo a los parámetros de fecha, tipo de producto y el filtro espacial dado por el área piloto buscada. Previo a realizar el análisis, se separa cada imagen en tiles, a los cuales se les calcula la probabilidad de inundación para el periodo temporal estudiado. Luego, se realiza un re-tiling (mosaico) para recomponer la imagen con los indicadores ya obtenidos y su resultado es luego importado en Geonode.

Finalmente, para el tablero de control, el flujo para construirlo comienza importando los resultados de los anteriores DAGs (SAOCOM, Sentinel-1 y Sentinel-2), a partir de los cuales mediante un script de R se construye la ShinyApp que contiene al tablero, el cual es embebido posteriormente en Geonode.

5 OTROS REQUERIMIENTOS

Licencia SAOCOM

Para la planificación y descarga de adquisiciones de escenas SAOCOM se solicitó una licencia en el marco de este AO, a nombre de Natalia Morandeira. Las/los Expertos Claves y colaboradores técnicos se suscribieron a la licencia como Usuarios finales.

Requerimientos para interoperabilidad e integración a CONAE

Se considerarán los requerimientos mínimos del documento “*Referencia de Productos para integración a CONAE*”, en cuanto a: **a)** las características de los productos a disponibilizar en una plataforma WebGIS; **b)** el metadato de los productos integrados en base a los estándares de IDERA (ISO 19115-3) con capacidad de ser consumidos como servicios por un catálogo bajo el estándar CWS; **c)** los requerimientos de automatización del procesamiento (desarrollos de código abierto, procesadores con contenedores docker y docker-compose, distribución en repositorio git requisitos de automatización); **c)** la documentación; **d)** los geoservicios WMS, WFS, WCS, CSW.

Requerimientos de seguridad informática

Se considerarán los requerimientos de la Subgerencia de Informática, Servicios de Datos y Telecomunicaciones, Gerencia de Gestión Tecnológica de CONAE, especificados en el documento “*Consideraciones de seguridad informática para el desarrollo de software y su puesta en producción*”.

6. EQUIPOS DE TRABAJO

Para la realización del trabajo se cuenta con un equipo de profesionales interdisciplinario formados en disciplinas como la Biología, Ingeniería en Agrimensura, Geografía, Ecología y Sistemas con amplios conocimientos en teledetección, sistemas de información geográfica y scripting y notable experiencia tanto en el estudio de humedales como en soluciones técnicas y arquitectura de datos.

7. CONCLUSIONES

La generación de un sistema que consiste en el monitoreo y gestión integral de humedales de Argentina a partir del desarrollo de modelos de áreas de estudio representativas de la Región de Humedales del corredor fluvial ChacoMesopotámico ubicadas en las provincias de Entre Ríos y Santa Fe, haciendo uso de información satelital (productos de la misión SAOCOM, Sentinel 1 y 2) llevado a cabo con tecnologías y herramientas desarrolladas de punta a punta en código abierto es una novedad sin antecedentes.

La integración de un gestor de contenidos geospaciales (GeoNode) con el desarrollo de modelos programados en python y la automatización de los procesos ETL (Airflow) como base del presente sistema establece un caso fácilmente replicable a lo largo y ancho del territorio nacional para realizar análisis, monitoreo y gestión integral de otras unidades de paisaje de distinta naturaleza.

Este sistema permite que los datos, modelos y productos que se generen sean beneficiosos para las organizaciones y reutilizables en la integración de información para el análisis y la toma de decisiones a nivel del ecosistema de humedales y de todos los actores involucrados.

Por último, cabe destacar, una vez más, que el sistema está realizado en código abierto, con su respectiva documentación, logrando una integración robusta de las soluciones y los requerimientos propuestos por los diferentes actores y permitiendo su réplica y programación para otros proyectos de interés.

AGRADECIMIENTOS

Se agradece a la Universidad Nacional de San Martín (UNSAM), al Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) y a la Universidad de Buenos Aires (UBA) por el interés, el fomento y financiamiento de actividades de investigación en torno a los humedales.

REFERENCIAS

Todas las referencias se listarán en orden alfabético. Las referencias se escribirán de acuerdo a las Normas APA 7ma. Ed.

Airflow

Apache License Version 2.0, (January)2004, <http://www.apache.org/licenses/>

Geonode

GNU GENERAL PUBLIC LICENSE, Version 3, 29 June 2007 Copyright (C) 2007 Free Software Foundation, Inc. [<http://fsf.org/>](http://fsf.org/)