

Almacenamiento y procesamiento automático de imágenes satelitales para proyectos de monitoreo de bosque nativo a escala regional y local

Viva Mayer, Francisco M.^{1,2}

vivamayer.francisco@uader.edu.ar

Walter F. Sione^{1,3}

wsione@gmail.com

Francisco D. Maldonado¹

francisco.dario.maldonado@gmail.com

¹ CEREGEO - Centro Regional de Geomática
Facultad de Ciencia y Tecnología. Universidad Autónoma de Entre Ríos.
Ruta 11, km 10,5, Oro Verde. Entre Ríos.

² CICyTTP - Centro de Investigaciones Científicas y Transferencia de Tecnología a la Producción
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
Matteri y España S/N° (CP 3105). Diamante. Entre Ríos.

³ PRODITEL- Departamento de Ciencias Básicas.
Universidad Nacional De Luján – UNLU
Ruta 5 (CP 6700), Lujan. Buenos Aires. Argentina.

Resumen

La necesidad de información sobre el estado de conservación y la dinámica de cambios en coberturas de bosque nativo de nuestro país implica el procesamiento de grandes volúmenes de datos, de diferentes fuentes y de manera oportuna, por lo que es indispensable tender a la automatización de procesos.

En este trabajo se presentan los objetivos, la metodología, se discuten las características y el impacto a escala local y regional que tienen las líneas de investigación sobre automatización de procesos de monitoreos de estado actual y cambios en la cobertura de suelo. La formación de recursos humanos en esta línea y el trabajo multidisciplinario se plantean como necesidades para abordar estas líneas a futuro.

Palabras claves: Imágenes satelitales, Teledetección, Modelo de datos.

1.- Contexto

Esta línea de investigación y desarrollo se enmarca en el Proyecto de Tesis Doctoral “Desarrollo y Validación de un Sistema Informatizado de Monitoreo Automático de los desmontes para un área del Espinal del Centro-Norte de Entre Ríos, en base a Grandes Volúmenes de Datos de Teleobservación” del autor, y en líneas de investigación del Centro Regional de Geomática de la Facultad de Ciencia y Tecnología de la Universidad Autónoma de Entre Ríos.

En este trabajo se avanza con el desarrollo de herramientas aplicables en el marco de la Ley de Presupuestos Mínimos de Protección Ambiental de los Bosques Nativos de la Ley Nacional de Bosques 26331 dependientes de los órganos de competencia de las provincias.

2.- Introducción

Los ecosistemas se encuentran en un estado de flujo permanente en una variedad de escalas espaciales y temporales en todo el mundo. Las causas de estos flujos pueden ser tanto naturales como antropogénicas, o pueden ser una combinación de ambas. En este sentido, los datos proporcionados por la tele-detección (*remote sensing*) desde plataformas satelitales activas (radar) y pasivas (ópticas) se han presentado durante mucho tiempo como un recurso crítico para monitorear, medir y explicar dichos fenómenos naturales y físicos (Baghdadi & Zribi, 2016; Campbell & Wynne, 2011; Chuvieco, 2016). De hecho, la tele-detección podría caracterizarse justamente como una de las primeras ciencias de "*big data*" (Goodchild et al., 2012).

Sin embargo, es en los últimos años se observa un aumento significativo en la provisión y uso de datos de imágenes satelitales del espectro óptico para el control y seguimiento de las dinámicas de los cambios de la superficie terrestre (natural y artificial) y de los cambios en el uso del suelo. Este aumento puede atribuirse a dos importantes hechos:

1. la necesidad de información precisa y oportuna de una sociedad que apunta hacia un desarrollo sustentable, que le permita mantener y/o mejorar el estilo de vida actual (Chuvieco, 2015; Hansen & Loveland, 2012);
2. los cambios en las políticas de acceso a la información de los organismos productores de datos de tele-detección desde plataformas satelitales (Soille et al., 2016; Woodcock et al., 2008), que ha resultado en un incremento de la disponibilidad (*data availability*) de la misma.

Estos avances han llevado a que la tele-detección produzca más información que la que resulta viable organizar y analizar mediante procedimientos tradicionales, que requieren de la intervención humana. Generando problemas de redundancia de información y/o desconocimiento de su existencia; o la adopción de técnicas y métodos de análisis que filtran, recortan o eliminan información a fin de volverlas humanamente manejables. Por ejemplo, para realizar un análisis activo de la cubierta vegetal de la provincia de Entre Ríos mediante imágenes de resolución media del espectro óptico de la plataforma Landsat8/OLI son necesarios ~ 6GB de datos cada 16 días. El procesado de esos datos $((7581 * 7301)_{pixel} * 4_{bandas} * 6_{tiles})$ incluyen billones de operaciones con números flotantes que requieren varios minutos o incluso hasta una hora para completar usando una sola estación de trabajo.

Este contexto de alta demanda de información y alta disponibilidad de datos ha llevado al auge de sistemas de computación de alto rendimiento (\ll HPC \gg o *High Performance Computing*) y de plataformas de computación en La Nube. En Wang *et al.* (2016) encontramos un ejemplo donde se implementan múltiples paradigmas, técnicas y tecnologías de computación en la nube y HPC para un sistema de procesamiento de datos de tele-detección a escala global.

Otro ejemplo, lo encontramos en la plataforma de acceso gratuito Google Earth Engine (Gorelick et al., 2017). La cual ofrece a los investigadores y emprendedores una solución a los costos de infraestructura informática necesaria para el procesamiento y análisis de los datos de tele-detección; proporcionando acceso a un gran almacén de imágenes satelitales y la potencia

computacional necesaria para analizar esas imágenes.

Una característica de estas plataformas es que están orientadas a procesamientos a demanda (*on-demand*) de los usuarios. Por otro lado, no ofrecen estructuras sobre las cuales asentar las áreas de estudio u organizar los resultados de los procesos, quedando estas actividades en manos de los usuarios. Ambas características, son una ventaja para estudios de casos puntuales, pero se convierten en una desventaja al afrontar un estudio activo de resultados constantes para la toma de decisiones.

De los dos ejemplos presentados, solo Google Earth Engine se ofrece como una plataforma al público. Por otro lado, ambas plataformas no dan acceso al código fuente (*open source*) por consiguiente presenta dificultades para la implementación local en proyectos de pequeña y mediana escala.

En vista a las limitaciones actuales, en el marco de este proyecto se trabaja en la implementación de un sistema de gestión de información espacial, aplicable a escalas locales y regionales que organicen el almacenamiento, procesan de manera automática o semiautomática, y distribución de la información generada en base a datos de tele-observación y aplicables al monitoreo y gestión del bosque nativo.

Para el desarrollo de este proyecto, se optó por la utilización de soluciones software de código abierto (*open source*) o de libre distribución (*free software*). La utilización de herramientas open-source beneficia principalmente en la reducción de los costos de adquisición de licencias de software directas o indirectas previas y posteriores al proyecto; y también permite la libre distribución (adecuándose a las licencias existentes) de lo producido

Así, se decidió por el uso de PostgreSQL (PostgreSQL, 2018) como Sistema de Gestión de Base de Datos Relacional Orienta a Objetos (Object-Relational Database Management System, ORDBMS) y PostGIS (PostGIS, 2018) como extensión espacial para el mismo. PostgreSQL es un sistema de base de datos relacional de código abierto altamente escalable y compatible con SQL. Se ejecuta en todos los principales sistemas operativos, incluidos Linux, UNIX y Windows. Tiene una arquitectura comprobada que le ha valido una sólida reputación de fiabilidad, integridad de datos y corrección (Chen & Xie, 2008).

Por su parte, PostGIS, como una extensión espacial de PostgreSQL, cumple con los estándares OGC e ISO SQL/MM, que hacen posible el uso de PostgreSQL para el desarrollo moderno de aplicaciones espaciales. Así mismo, provee soporte para la gestión de información espacial en modelo ráster a partir de PostGIS 2.0+ (Dang, 2012).

La informatización de estas rutinas de procesamiento y transformación de datos se efectuara con Python (Python, 2020), haciendo uso de las librerías Geospatial Data Abstraction Library (GDAL, 2020), SciPy (SciPy, 2020) y de la rutina de extracción de nubes Fmask (Qiu et al., 2019; Zhu et al., 2015).

El resultado esperado de este trabajo de investigación y desarrollo es una estructura organizada permita establecer consultas elaboradas sobre múltiples capas de información; donde se reduzca la redundancia, se mejore la disponibilidad, la consistencia, y la integridad de los datos facilitando la interoperabilidad de los mismos con otros generadores de información y usuarios finales.

3.- Líneas de investigación y desarrollo

Las líneas son ejecutadas por un equipo multidisciplinario con capacidad para el procesamiento de imágenes, la adaptación y aplicación de técnicas de levantamiento de campo y comprender la dinámica de las coberturas vegetales mediante el uso de imágenes satelitales.

Una de las líneas en desarrollo es la estructura principal de la base de datos sobre la cual se construirá el sistema. La misma debe ser flexible y modular a fin de permitir el almacenamiento de una miríada de fuentes de información espacial, tanto de plataformas satelitales como de levantamiento de datos campo.

Otra línea se relaciona con el sistema que hace uso del modelo de base de datos y proporcione funcionalidades para la ingesta y consulta de los datos, así como una plataforma sobre la cual puedan construirse modelos automáticos o semiautomáticos de producción de información como por ejemplo para la detección de desmontes en tiempo casi-real y monitoreo del cambio de uso de suelo, entre otras.

4.- Resultados esperados

El principal resultado es una plataforma sobre la cual pueda implementarse una metodología automática operacional para la detección y monitoreo en tiempo casi-real de los desmontes en el bosque nativo del centro-norte de la provincia de Entre Ríos. Y a su vez permita producir periódicamente: alertas de desmontes, mapas de intensidad de los cambios del bosque natural, mapas de cambios de uso producidos en áreas de bosque natural y tablas cuantitativas de los desmontes en el centro-norte de Entre Ríos.

Se espera que la misma sirva como trampolín para futuros desarrollos y avances en el libre acceso a la información.

5.- Formación de Recursos Humanos

El Magister Lic. Francisco M. Viva Mayer lleva adelante su tesis de doctorado en Ciencias Aplicadas, con el título “*Desarrollo y Validación de un Sistema Informatizado de Monitoreo Automático de los desmontes para un área del Espinal del Centro-Norte de Entre Ríos, en base a Grandes Volúmenes de Datos de Teleobservación*”. En el marco de los proyectos de equipo de trabajo se insertan también Licenciados en Biología, Licenciados en Sistemas, y alumnos de carrera de ambas carreras quienes se inician en actividades de investigación en estas líneas.

6.- Bibliografía

- Baghdadi, N., & Zribi, M. (2016). *Optical Remote Sensing of Land Surface. Techniques and Methods* (1st ed.). ISTE Press - Elsevier.
- Campbell, J. B., & Wynne, R. H. (2011). *Introduction to Remote Sensing* (5th ed.). Guilford Press.
- Chen, R., & Xie, J. (2008). Open Source Databases and Their Spatial Extensions. In G. B. Hall & M. G. Leahy (Eds.), *Open Source Approaches in Spatial Data Handling* (pp. 105–129). Springer. https://doi.org/10.1007/978-3-540-74831-1_6
- Chuvienco, E. (2015). *Teledetección Ambiental: la observación de la Tierra desde el espacio* (2nd ed.). Ariel Ciencia.
- Chuvienco, E. (2016). *Fundamentals of satellite remote sensing: an environmental approach* (Second edition). Taylor & Francis.
- Coppin, P. R., Jonckheere, I., Nackaerts, K., Muys, B., & Lambin, E. (2004). Digital change detection methods in ecosystem monitoring: a review. *International Journal of Remote Sensing*, 25(9), 1565–1596.

- <https://doi.org/10.1080/0143116031000101675>
- Dang, T. N. U. (2012). *PostGIS Raster: Extending PostgreSQL for The Support of Continuous Fields* [Master Thesis]. Université Libre de Bruxelles.
- GDAL. (2020). *Geospatial Data Abstraction Library* (Version 3.0.4) [Computer software]. Open Source Geospatial Foundation. <https://gdal.org>
- Goodchild, M. F., Guo, H., Annoni, A., Bian, L., Bie, K. de, Campbell, F., Craglia, M., Ehlers, M., Genderen, J. van, Jackson, D., Lewis, A. J., Pesaresi, M., Remeteý-Fülöpp, G., Simpson, R., Skidmore, A., Wang, C., & Woodgate, P. (2012). Next-generation Digital Earth. *Proceedings of the National Academy of Sciences*, 109(28), 11088–11094. <https://doi.org/10.1073/pnas.1202383109>
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>
- Hansen, M. C., & Loveland, T. R. (2012). A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment*, 122, 66–74. <https://doi.org/10.1016/j.rse.2011.08.024>
- PostGIS. (2018). *PostGIS: Homepage*. <https://postgis.net/>
- PostgreSQL. (2018). *PostgreSQL: Homepage*. <https://www.postgresql.org/about/>
- Python. (2020). *Python* (Version 3.8.2) [C, C++]. The Python Software Foundation. <https://www.python.org/>
- Qiu, S., Zhu, Z., & He, B. (2019). Fmask 4.0: Improved cloud and cloud shadow detection in Landsats 4–8 and Sentinel-2 imagery. *Remote Sensing of Environment*, 231, 111205. <https://doi.org/10.1016/j.rse.2019.05.024>
- SciPy. (2020). *SciPy Ecosystem* [Python]. <https://www.scipy.org/>
- Soille, P., Burger, A., Rodriguez, D., Syrris, V., & Vasilev, V. (2016). Towards a JRC Earth Observation Data and Processing Platform. *Proceedings of the 2016 Conference on Big Data from Space (BiDS'16)*, 4. <https://doi.org/10.2788/854791>
- Wang, L., Ma, Y., Yan, J., Chang, V., & Zomaya, A. Y. (2016). pipsCloud: High performance cloud computing for remote sensing big data management and processing. *Future Generation Computer Systems*. <https://doi.org/10.1016/j.future.2016.06.009>
- Woodcock, C. E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., Gao, F., Goward, S. N., Helder, D., Helmer, E., Nemani, R., Oreopoulos, L., Schott, J., Thenkabail, P. S., Vermote, E. F., Vogelmann, J., Wulder, M. A., & Wynne, R. (2008). Free Access to Landsat Imagery. *Science*, 320(5879), 1011–1011. <https://doi.org/10.1126/science.320.5879.1011a>
- Zhu, Z., Wang, S., & Woodcock, C. E. (2015). Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sensing of Environment*, 159, 269–277. <https://doi.org/10.1016/j.rse.2014.12.014>