

Generación de un banco de áreas de reflectividad pseudoinvariante para la Península Ibérica mediante imágenes MODIS

Lluís Pesquer⁽¹⁾, Oscar González⁽²⁾, Cristina Domingo⁽²⁾, Jordi Cristóbal⁽³⁾ y Xavier Pons⁽²⁾

⁽¹⁾ Grumets Research Group, CREAM. Campus de la Universitat Autònoma de Barcelona, Bellaterra, España.

Email: l.pesquer@creaf.uab.cat

⁽²⁾ Grumets Research Group, Dep. Geografía. Universitat Autònoma de Barcelona. Bellaterra, España.

Emails: oscar.gonzalez.guerrero@uab.cat, cristina.domingo@uab.cat, xavier.pons@uab.cat

⁽³⁾ Grumets Research Group, Geophysical Institute and Institute of Northern Engineering, University of Alaska, United States of America. Email: j.cristobal@alaska.edu

Resumen: La reflectividad derivada de imágenes satelitales sigue generalmente el ciclo fenológico de las cubiertas presentes en el territorio. A pesar de ello, es posible encontrar zonas donde la reflectividad es prácticamente invariante. Estas áreas definidas como pseudoinvariantes (API) permiten comparar y calibrar imágenes provenientes de distintos sensores y procesar series temporales con una elevada coherencia. Se presenta un nuevo método automático (especialmente útil en entornos Big Data) para seleccionar API a partir del producto diario MOD09GA derivado de imágenes Terra-MODIS, utilizando una serie temporal de 14 años y las bandas del espectro solar (visible, infrarrojo cercano y de onda corta) con una resolución espacial de 500 m. Dicha metodología consta de dos etapas de filtrado, una primera que evalúa la calidad de las imágenes de la serie mediante técnicas geoestadísticas, seleccionando las mejores, y una segunda que define umbrales específicos para cada banda espectral, en función de la dispersión que presentan los datos en la selección previa de imágenes. La aplicación de este método sobre ámbitos de características topográficas y estructura de paisaje diferenciados en la Península Ibérica ha permitido la obtención de más de 12 000 API en una superficie asimilable a 9 escenas Landsat (WRS-2). Los resultados muestran que la metodología aplicada contempla la adecuada distribución tanto interanual como intraanual de las imágenes, dando lugar a API que abarcan una amplia variedad de cubiertas con reflectividades diversas, ubicadas principalmente en zonas boscosas o seminaturales (77%), zonas agrícolas (21 %), así como en otros tipos de cubiertas no vegetales.

Palabras clave: MODIS, reflectividad, área pseudoinvariante, series temporales, Big Data.

Generation of pseudoinvariant reflectance areas bank through MODIS data for the Iberian Peninsula.

Abstract: Reflectance captured by satellite imagery usually follows land cover phenology dynamics. However, it is possible to find areas where reflectance remains almost invariant. These areas, called pseudoinvariant areas (PIA), allow comparing and calibrating images taken from different sensors as well as processing highly coherent time series from remote sensing data. This work presents a new automatic method (especially useful in Big Data frameworks) for selecting PIA from a 14 year time series from the daily product of Terra-MODIS MODIS09GA (from visible to short wave infrared) at 500 m spatial resolution. The implemented methodology consists of two image filtering steps: the first evaluates the image quality of the series using geostatistical techniques, where the best are selected; a second defines specific thresholds for each spectral band based on the standard deviation of the filtered imagery. Applying this methodology on several topographic and landscape features for the Iberian Peninsula obtained more than 12,000 PIA in an area around 9 Landsat (WRS-2) scenes. Results yielded high consistency in the two step filtering process providing a proper inter- and intra-annual PIA distribution and leading to a wide range of PIA reflectance over forests and natural areas (77%), agricultural areas (21%) and other types of non vegetation covers.

Keywords: MODIS; reflectance, pseudoinvariant area, time series, Big Data

1. INTRODUCCIÓN

La Teledetección ha entrado también en la era del Big Data y progresivamente va aumentando el volumen de información satelital disponible. A las largas series históricas ya existentes, se suman los datos de los nuevos satélites, pero de forma significativa contribuye a este Big Data, la nueva política de distribución de datos de las agencias espaciales y otras instituciones (en este sentido, debemos aplaudir de forma especial al USGS).

Entre la amplia gama de productos de teledetección, las imágenes de reflectividad de la superficie terrestre son tal vez de las más utilizadas debido a su gran versatilidad

para distintas aplicaciones (Lillesand et al. 2004): generación de cartografía temática, análisis de incendios, seguimiento de sequías, caracterización del estado de la vegetación, dinámica de usos del suelo, etc.

Conceptualmente la reflectividad, definida como el cociente entre el flujo energético reflejado por una superficie y el flujo energético incidente en ella (Pons y Arcalís 2012) es característica para el tipo de cubierta de la superficie captada, y es una magnitud espectral, pero idealmente debe ser independiente del sensor usado y de las condiciones de iluminación (Roy et al., 2002). En la práctica, los métodos de corrección atmosférica y topográfica (Vanonckelen et al. 2013) consiguen generar productos que se acercan a estas características dando

como resultado un enorme volumen de información cuyo mayor potencial reside en el hecho de poder utilizarse sinérgicamente.

Para armonizar correctamente y de forma automática largas series de imágenes de teledetección, o para intercalibrar productos entre distintos sensores, es necesario disponer de un conjunto de valores de referencia. Una posible fuente de valores de referencia son las áreas pseudo-invariantes (API), cuyos valores para una zona delimitada son casi constantes durante el tiempo (Hadjimitsis *et al.*, 2009). Naturalmente, deberemos definir adecuadamente el concepto de “pseudo” variabilidad (“casi” sin variabilidad), el cual dependerá de nuestros objetivos. Un margen muy estricto de “pseudo” variabilidad generará zonas muy estables a lo largo del tiempo pero poco abundantes, y tal vez no suficientemente significativas; mientras que un margen más amplio nos permitirá obtener mayor número de áreas con una distribución espacial y estacional más completa a costa de una menor estabilidad. A la vez, es indudable que algunas cubiertas son por naturaleza altamente inestables y no son adecuadas como API.

El principal objetivo del presente trabajo es generar un extenso banco de zonas pseudoinvariantes para el rango solar a partir de datos del sensor MODIS (en concreto entre 459 nm y 2155 nm) y que cubra toda la Península Ibérica. Este objetivo se está desarrollando principalmente en el contexto del proyecto DinaCliVe (*vid.* Agradecimientos) y en el presente trabajo se explica la metodología y los resultados hasta el momento obtenidos, ya claramente significativos.

2. MATERIAL Y MÉTODOS

2.1. Material

El material base es una serie de 14 años (2000-2014) de las bandas 1, 2, 3, 4, 6 y 7 del espectro solar (visible, infrarrojo cercano y de onda corta) del producto diario MOD09GA *Surface Reflectance Daily L2G Terra* del sensor MODIS (MODerate resolution Imaging Spectroradiometer) de 500 m de resolución espacial. Dicho producto se ha descargado del *NASA Earth Observing System Data and Information System* a través del *REVERB Data and Service Access Client*.

El ámbito geográfico de estudio, la Península Ibérica, corresponde a cuatro escenas del sistema de distribución de MODIS (h17/18v04/05) detallado en la Figura 1. En la misma figura se aprecia el distinto grado de finalización del proceso de generación de API (en diferentes colores).

2.2. Metodología

El principal requerimiento que necesita la presente metodología es la automatización, ya que deben procesarse 14 años de productos diarios con 6 bandas de reflectividades y unas 15 bandas complementarias (máscaras y productos de calidad asociados) para las 4 escenas de MODIS, idealmente fragmentadas en ámbitos menores, como por ejemplo el ámbito del sistema WRS-2 de Landsat. Debido al gran volumen de datos a

procesar, del orden de 1 839 600 imágenes, podemos considerarlo un proyecto Big Data.

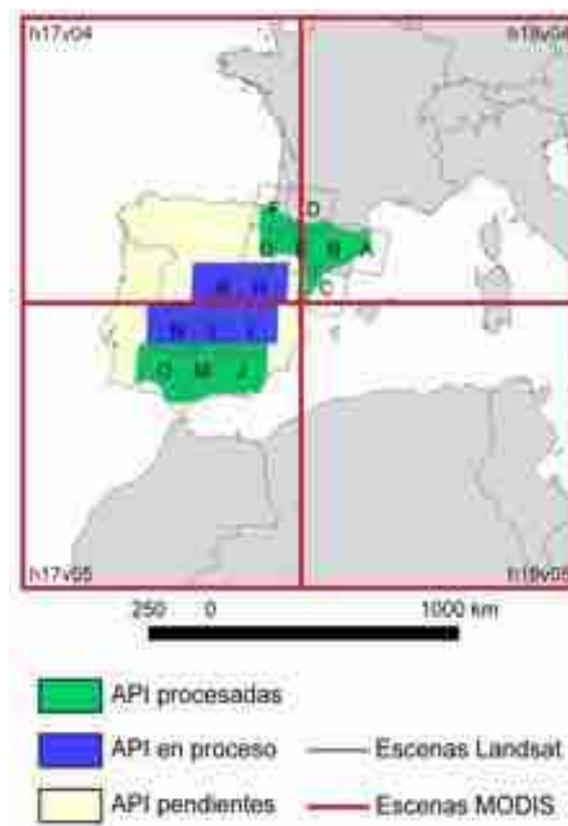


Figura 1. Área de estudio y estado del proceso de generación de API de las zonas de estudio.

La presente metodología está basada en el trabajo de Pesquer *et al.* (2013) con dos nuevas contribuciones ensayadas en la parte preparatoria de la metodología de Pons *et al.* (2014) y ha sido implementada básicamente en el software MiraMon (Pons 2004). Esencialmente existe un primer proceso de selección de las imágenes de mayor calidad de la serie, teniendo en cuenta que se han fragmentado las escenas originales en ámbitos aproximadamente coincidentes al sistema WRS-2 de Landsat. Esta selección tiene por objetivo reducir las posibles fuentes de error, de incoherencia y de variabilidad adicional para la detección, con la mayor precisión posible, de zonas de reflectividad estable. En esta selección juegan un papel importante las máscaras y las bandas de calidad del propio producto MOD09GA, entre ellas el ángulo cenital para filtrar distorsiones en zonas alejadas del nadir; pero nuestra propuesta añade un filtro propio basado en el análisis comparativo del patrón geoespacial de la región de estudio mediante técnicas geoestadísticas (análisis del variograma empírico y ajuste de un variograma teórico que determina sus parámetros estructurales).

Una vez seleccionadas las mejores imágenes se aplican unos umbrales (bajos) de variabilidad máxima respecto a toda la serie temporal de imágenes seleccionadas. Estos umbrales son particulares para cada una de las bandas espectrales (ver Tabla 1). Esta no-uniformidad está

basada en los resultados comparativos de Feng *et al.* (2013) y en los análisis de calidad de Pons *et al.* (2014). Adicionalmente se descartan aquellos píxeles con poca variabilidad pero que provienen de un pobre número de muestras, exigiendo un número mínimo de muestras de alrededor de 1/3 de las imágenes de la selección total.

Tabla 1. Umbral de máxima desviación estándar para cada banda para considerar un píxel de MODIS como API en reflectividades (en %).

| Banda | 1 | 2 | 3 | 4 | 6 | 7 |
|-------|------|------|------|------|------|------|
| Desv. | 1.93 | 2.70 | 2.41 | 1.99 | 3.09 | 2.12 |

3. RESULTADOS Y DISCUSIÓN

3.1. Resultados

Los resultados de este trabajo consisten en diversas capas vectoriales de polígonos correspondientes a las áreas pseudoinvariantes. Estos polígonos contienen, en su base de datos asociada y para cada banda espectral, la reflectividad media y valores de desviación a lo largo de toda la serie temporal.

La Tabla 3 resume estos resultados mostrando el número de polígonos obtenidos para cada ámbito de procesamiento, así como descriptores estadísticos básicos de las reflectividades para cada una de estas zonas. Adicionalmente se muestran los parámetros estructurales de los variogramas modelo ajustados para cada zona. Los distintos ámbitos corresponden aproximadamente a escenas Landsat (ver Tabla 2).

Tabla 2. ID correspondientes a las escenas Landsat según el sistema WRS-2.

| ID | Path-row | ID | Path-row | ID | Path-row |
|----|----------|----|----------|----|----------|
| A | 197-31 | F | 200-30 | K | 201-32 |
| B | 198-31 | G | 200-31 | L | 201-33 |
| C | 198-32 | H | 200-32 | M | 201-34 |
| D | 199-30 | I | 200-33 | N | 202-33 |
| E | 199-31 | J | 200-34 | O | 202-34 |

3.2. Discusión

Los resultados muestran un número muy variable de polígonos, según las zonas, no relacionado con el área útil (excluyendo el mar) del ámbito de estudio. Las zonas de menor número son las D-E-F-G, y la M, con menos de 300 API. Esencialmente ello es debido a una de las dos siguientes causas: cubierta de nubes extensa y recurrente, o paisaje dominante estacionalmente variable como por ejemplo cultivos de temporada o vegetación caducifolia. En cualquier caso, para determinados usos de las API, éste es un número más que suficiente (por ejemplo como valor de referencia y soporte a la corrección radiométrica de otros sensores: Pons *et al.* 2014) y, si no lo fuera, podrían definirse unos umbrales (Tabla 1) algo superiores, logrando así un mayor número de API, menos estables, pero suficientes para su uso.

Tabla 3. Resumen estadístico de los polígonos de API procesados hasta el momento (ámbitos en verde en la Figura 1).

| ID (escena) | A | B-C | D-E-F-G | J | M | O | |
|------------------------------|--------------|---------|---------|---------|---------|---------|-------|
| Extensión (km ²) | 19683.3 | 42315.5 | 29786.3 | 49413.1 | 31794.8 | 47832.8 | |
| Núm API | 444 | 2101 | 270 | 8736 | 174 | 460 | |
| Reflectividades MODIS en % | B1 Mín | 2.3 | 2.5 | 1.1 | 3.8 | 4.8 | 3.7 |
| | B1 Máx | 19.8 | 18.3 | 16.4 | 18.3 | 11.4 | 11.8 |
| | B1 Media | 5.2 | 5.7 | 5.2 | 9.6 | 7.9 | 5.6 |
| | B2 Mín | 3.2 | 2.4 | 4.2 | 14.1 | 17.2 | 6.9 |
| | B2 Máx | 29.3 | 30.5 | 27.0 | 28.3 | 26.1 | 24.3 |
| | B2 Media | 20.3 | 20.6 | 19.4 | 19.4 | 20.8 | 18.7 |
| | B3 Mín | 1.5 | 1.3 | 0.7 | 1.7 | 2.5 | 1.9 |
| | B3 Máx | 10.6 | 10.4 | 8.4 | 9.7 | 6.0 | 6.6 |
| | B3 Media | 2.7 | 3.0 | 2.7 | 4.9 | 4.1 | 2.9 |
| | B4 Mín | 3.3 | 3.3 | 1.6 | 4.4 | 5.1 | 3.9 |
| | B4 Máx | 17.2 | 16.0 | 13.9 | 15.4 | 9.8 | 10.8 |
| | B4 Media | 5.5 | 6.0 | 5.5 | 8.3 | 7.2 | 5.6 |
| | B6 Mín | 3.8 | 2.2 | 3.3 | 10.5 | 13.4 | 5.0 |
| | B6 Máx | 30.9 | 27.2 | 27.1 | 28.2 | 24.7 | 22.4 |
| | B6 Media | 15.8 | 16.8 | 15.0 | 18.9 | 17.8 | 14.8 |
| | B7 Mín | 2.1 | 1.3 | 1.8 | 4.9 | 6.7 | 2.9 |
| | B7 Máx | 27.0 | 24.8 | 23.0 | 22.3 | 14.0 | 17.1 |
| | B7 Media | 8.5 | 9.4 | 8.2 | 12.4 | 10.7 | 8.0 |
| | Nugget Media | 0.59 | 0.84 | 7.96 | 1.1 | 1.16 | 0.74 |
| | Nugget Desv | 0.27 | 0.27 | 4.07 | 2.8 | 0.29 | 0.17 |
| | Range Media | 19671 | 23094 | 78804 | 63476 | 28867 | 22147 |
| Range Desv | 4642 | 4482 | 132811 | 8722 | 4389 | 4725 | |
| Sill Media | 3.87 | 4.45 | 8.81 | 7.75 | 3.05 | 2.08 | |
| Sill Desv | 1.49 | 1.42 | 12.33 | 1.72 | 0.79 | 0.67 | |

Es interesante observar que se han conseguido detectar tanto API con reflectividades bajas como elevadas. Para aplicaciones como la normalización de series de imágenes o el apoyo a la corrección radiométrica sería un inconveniente que los polígonos se caracterizaran por tener, p. ej., valores sólo bajos, ya que no podría asegurarse que el modelo ajustado se comportara correctamente en un amplio rango de reflectividades. La distribución del conjunto de API hasta el momento obtenidas respecto a su tipo de cubierta según CORINE 2006 se muestra en la Tabla 4. También es de interés comprobar cómo el tipo de cubierta para los polígonos es diverso: aproximadamente el 77 % de API en bosques y zonas seminaturales (principalmente sobre especies perennifolias, arbóreas y arbustivas) y, aunque en menor número, también en zonas agrícolas (21 %) o sobre superficies artificiales, masas de agua, etc. En el caso concreto de las API en zonas agrícolas según Corine, se observa que mayoritariamente corresponden a zonas de cultivos leñosos (típicamente con una elevada proporción de suelo desnudo) y a zonas cultivadas con una elevada proporción de vegetación natural.

Tabla 4. Distribución de las API en función del correspondiente tipo de cubierta según CORINE 2006.

| % | Tipo de cubierta (CORINE niv. 3) |
|--------------|--|
| 27.5 | Zonas de vegetación dispersa |
| 17.9 | Bosques de coníferas |
| 14.3 | Transición bosque-matorral |
| 10.2 | Vegetación esclerófila |
| 7.0 | Zonas de cultivos con patrones complejos |
| 4.7 | Zonas principalmente agrícolas con porción de vegetación natural significativa |
| 4.2 | Bosques de hoja ancha |
| 3.6 | Frutales |
| 2.0 | Zonas cultivables no irrigadas |
| 1.5 | Superficies artificiales |
| 1.5 | Olivos |
| 1.4 | Bosques mixtos |
| 0.9 | Masas de agua |
| 0.8 | Zonas permanentemente irrigadas |
| 0.5 | Viñedos |
| 0.4 | Zonas agro-forestales |
| 1.5 | Otros |
| 100.0 | Total |

4. CONCLUSIONES

Los resultados mostrados en la obtención de API en distintas zonas de la Península Ibérica indican que la metodología presentada es útil para paisajes y orografías muy distintas, y proporciona un amplio rango geográfico y de valores de reflectividad, también con una variabilidad adecuada. La metodología necesita del procesamiento y análisis geoestadístico de un enorme conjunto de datos de teledetección, de manera que el uso de procesos y toma de decisiones automatizados, y la inclusión de técnicas computacionales avanzadas es indispensable para que la presente metodología sea viable.

Actualmente se está ya trabajando sobre la zona central de la Península cubriendo un área de 5 escenas Landsat (Figura 1.). En breve, se completará el análisis de imágenes para determinar las API sobre toda la Península Ibérica.

5. AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por el Gobierno de España (Ministerio de Economía y Competitividad), FEDER a través de la ayudas CGL2012-33927 correspondiente al proyecto DinaCliVe y por el Ministerio de Ciencia e Innovación mediante una ayuda FPU (AP2008-02016) a Cristina Domingo, así como con recursos del Grupo de Investigación Consolidado "Grumets" de la Generalitat de Catalunya (SGR 2014, núm. 1491). Xavier Pons es beneficiario de un ICREA Academia Excellence in Research grant (2011-2015).

6. BIBLIOGRAFÍA

- Feng, M., Sexton, J.O., Huang C., Masek, J.G., Vermote, E.F., Gao, F., Narasimhan, R., Channan, S., Wolfe, R.E., Townshend, J.R., 2013. *Global surface reflectance products-from Landsat-Assessment using coincident MODIS observations*. Remote Sensing of Environment 134, 276-293.
- Hadjimitsis, D.G., Clayton, C.R.I., Retalis, A., 2009. *The use of selected pseudo-invariant targets for the application of atmospheric correction in multi-temporal studies using satellite remotely sensed imagery*. International Journal of Applied Earth Observation and Geoinformation 11, 192-200.
- Lillesand, T.M, Kiefer, R.W. Chipman, J.W., 2004. *Remote Sensing and Image Interpretation* (5th edition), John Wiley and Sons, New York.
- Pesquer, L., Domingo, C., Pons, X., 2013. *A Geostatistical Approach for Selecting the Highest Quality MODIS Daily Images*. Springer Lecture Notes in Computer Science Series, 7887 LNCS, 608-615.
- Pons, X., 2004. "MiraMon. Sistema de Información Geográfica y software de Teledetección" CREA. Bellaterra. ISBN: 84-931323-4-9.
- Pons, X., Arcalís, A., 2012 *Diccionari terminològic de teledetecció*. Barcelona, Institut Cartogràfic de Catalunya y Enciclopèdia Catalana. 597 p. ISBN: 978-84-412-2249-6.
- Pons, X., Pesquer, L., Cristóbal, J., González-Guerrero, O., 2014. *Automatic and improved radiometric correction of Landsat imagery using reference values from MODIS surface reflectance images*. International Journal of Applied Earth Observation and Geoinformation 33: 243-254.
- Roy, D., Borak, J., Devadiga, S., Wolfe, R., Zheng, M., Desclotres, J., 2002. *The MODIS land product quality assessment approach*. Remote Sensing of Environment 83, 62-76.
- Vanonckelen, S., Lhermitte, S., Van Rompaey, A., 2013. *The effect of atmospheric and topographic correction methods on land cover classification accuracy*. International Journal of Applied Earth Observation and Geoinformation 24, 9-21.