

# Memorias



Sociedad Latinoamericana en  
Percepción Remota y Sistemas  
de Información Espacial  
Capítulo Colombia

**30**  
AÑOS

Medellín, Colombia  
29 de Septiembre al 3 de Octubre de 2014



Evaluación de la calidad y selección de píxeles válidos de imágenes MODIS en los  
trópicos.

Quality assessment and selection of valid pixels of MODIS images in the tropics.

Jesús A. Anaya-Acevedo

Profesor Asociado, Universidad de Medellín, Cra 87 N° 30 – 65, Medellín, Colombia,

janaya@udem.edu.co

Germán M. Valencia-Hernández

Profesor, Universidad de San Buenaventura, Cra 56C N° 51 – 90, Medellín, Colombia,

especializacion.sig@usbmed.edu.co

## RESUMEN

Las imágenes obtenidas a partir de sensores ópticos instalados en satélites tienen un altísimo contenido de nubosidad en zonas tropicales y obstruyen la observación de la superficie terrestre. Por esta razón se hace necesario crear nuevas imágenes compuestas por píxeles sin nubes, donde cada pixel puede tener una fecha distinta para un período determinado. Sin embargo, no es fácil determinar cual debe ser la longitud de este periodo en la zona del Chocó (Colombia Sur América), esta región biogeográfica está sujeta al chorro del Chocó, a las lluvias de carácter orográfico por la cordillera de los Andes y a la formación de nubes de convección profunda. El objetivo de este estudio es

utilizar la información de calidad de las imágenes MODIS para seleccionar píxeles válidos en más de una década de datos MODIS, específicamente utilizando el producto MOD13Q1. Los resultados indican que utilizando una serie de tiempo de dos años se pueden reducir los datos inválidos desde un 90% hasta un 2%.

Palabras clave: Chocó-Colombia, estimación de la calidad, nubes, series de tiempo.

### ABSTRACT

Cloud content of optical satellite images for tropical regions is extremely high and the view of the earth surface is commonly obstructed. For this reason land surface studies require the selection of cloud free pixels from archive images to generate new images. Each pixel in a new image may have a different date for a given time period. Determining the length of this period is a challenge in the state of Chocó (Colombia South America), a biogeographic region subject to moisture flux from the Pacific Ocean known as the Chocó Jet, with moist air moving in orographic lift along the Andes Mountains and frequently formation of deep convective clouds. The aim of this study was to use MODIS quality data to select valid pixels from more than a decade of archive files, specifically using MOD13Q1 product. Results indicate that a time series of two years reduce invalid pixels from 90% to 2%.

Keywords: Chocó-Colombia, quality assessment, cloud cover, time series.

## INTRODUCCIÓN

Obtener reflectividad de la superficie terrestre en la parte óptica del espectro electromagnético es fundamental en el proceso de la teledetección, sin embargo la niebla, las nubes y sus sombras, limitan considerablemente la calidad de este proceso (Leinenkugel et al. 2013; Liu & Liu 2013). Esto puede corroborarse en las imágenes diarias obtenidas por los sensores *Moderate Resolution Imaging Spectroradiometer* (MODIS) a bordo de los satélites Terra y Aqua, donde la información de utilidad disminuye en la medida en que el contenido de nubes aumenta. Terra fue lanzado el 18 diciembre de 1999 y cruza el ecuador en forma descendente a las 10:30 am, mientras que Aqua, con las mismas características, fue lanzado el 4 de mayo de 2002 y cruza el ecuador a la 1:30 pm de forma ascendente (Savtchenko *et al.* 2004). Esta configuración tiene por objetivo capturar la rápida variación diurna (Salomonson et al. 2001) debido al calentamiento de la superficie terrestre por la radiación solar. En el Chocó, la formación de sistemas nubosos puede explicarse en parte como resultado del forzamiento termal causado por el sol, la humedad de superficie proveniente del océano Pacífico y el ascenso orográfico ocasionado por la cordillera de los Andes. En un análisis detallado de la distribución temporal horaria de la lluvia, Poveda y colaboradores (2005) encontraron que la precipitación en Colombia tiene un claro ciclo diurno, los valores mínimos de precipitación ocurren al final de la mañana, entre 9 am y 11 am, mientras que las lluvias en la tarde y noche son comunes. Por tanto es de esperarse que la ocurrencia de nubes también sea más baja en horas de la mañana. Otros fenómenos como el Niño y la oscilación Madden-Julian también afectan los patrones de precipitación tropical, el más reciente y de gran impacto fue el fenómeno de la niña en 2010 (Hoyos *et al.* 2013).

A pesar de la alta resolución temporal de MODIS, en teoría son dos observaciones diarias, en la práctica es inferior debido a la existencia de niebla, nubes y sombras generadas por nubes. Además también debe considerarse que la órbita de estos satélites es polar, por tanto el cubrimiento es mayor en los polos disminuyendo en la medida en que se aproxima al Ecuador. La información sobre el contenido de nubosidad está disponible tanto para el sensor *Advanced Very High Resolution Radiometer* (AHVRR) como para el sensor MODIS y se conoce como -máscara de nubes-, se realiza para cada píxel y es un indicador de la calidad del mismo. Se han seguido diversos métodos basados en rangos para generar esta máscara, cuyo objetivo es distinguir entre observaciones con algún contenido de nubosidad y observaciones con cielo completamente despejado (Sezgin & Sankur 2004; Yang et al. 2007). Otra forma de identificar la calidad por píxel es a través del método de LDOPE (*Land Data Operational Product Evaluation*) que permite, además de las nubes, evaluar otra serie de aspectos de calidad como sombras de nube, correcciones atmosféricas, contenido de aerosoles y geometría (ángulo de la observación respecto al nadir). Estos aspectos de estimación de calidad están incluidos en los productos MODIS y se conocen con el nombre de -conjunto de datos científicos de estimación de la calidad- QA-SDSs. Un total de 15 bits son dedicados a describir la calidad de cada píxel. Las Tablas 1 y 2 listan la posición del bit para cada aspecto de calidad y la combinación de bits para cada parámetro.

**Tabla 1.** Campos de bits con parámetros generales. Descripción de los datos de estimación de calidad de los índices de vegetación (IV) general. Disponible en <http://www.ctahr.hawaii.edu/grem/mod13ug/sect0005.html>

Bits	Nombre del parámetro	Valor	Descripción
0-1	Calidad general del IV	00	Producido con alta calidad
		01	Producido con alta calidad, revisar otros aspectos
		10	Pixel producido, pero probablemente con nubes
		11	Pixel no producido por razones distintas a las nubes
2-5	Utilidad del IV	0000	Perfecto
		0001	Alta calidad
		0010	Buena calidad
		0011	Calidad aceptable
		0100	Calidad aceptable mínima
		0101	Calidad intermedia
		0110	Inferior a calidad intermedia
		0111	Calidad promedio
		1000	Inferior a la calidad promedio
		1001	Calidad dudosa
		1010	Calidad superior a datos excluidos
		1011	Datos excluidos
		1100	Baja calidad
		1101	Sin corrección atmosférica
		1110	Calidad demasiado baja para ser útil
		1111	No es útil por otras razones, no se generó.

**Tabla 2.** Campos de bits con parámetros específicos. Descripción de los datos de estimación de calidad de los índices de vegetación (IV) por aerosoles, máscaras de agua, hielo, sombras y diferentes contenidos de nubes.

6-7	Cantidad de aerosoles	00	Clima
		01	Bajo
		10	Intermedio
		11	Alto
8	Detección de nubes adyacentes	0	No
		1	Si
9	Corrección atmosférica BRDF	0	No
		1	Si
10	Mezcla de nubes	0	No
		1	Si
11-13	Máscara tierra/agua	000	Océano superficial
		001	Solo tierra
		010	Costa de océanos y orilla de lagos
		011	Aguas continentales superficiales
		100	Aguas efimeras
		101	Aguas profundas continentales
		110	Océano próximo a la costa
111	Océano profundo		
14	Posibilidad de hielo o nieve	0	No
		1	Si
15	Posibilidad de sombras	0	No
		1	Si

Con el fin de disminuir el contenido de píxeles de baja calidad en la cadena de producción, MODIS utiliza los QA-SDSs para filtrar los mejores píxeles en un período de 16 días. Solo las mejores observaciones (para un mismo pixel) son seleccionadas en función del contenido de la nubosidad y de la geometría. Posteriormente, para observaciones en zonas con vegetación es de utilidad retener solo aquella observación con mayor valor de índice de vegetación. Dicha observación corresponderá entonces a

la observación seleccionada en un período de 16 días. Con la mejor observación en un período de 16 días para cada pixel se generan los productos MODIS denominados compuestos.

Una de las regiones con mayor nubosidad del mundo es el Chocó biogeográfico y por tanto los análisis de su superficie y las técnicas de monitoreo con datos ópticos se hace muy compleja. Los compuestos MODIS de 16 días son insuficientes para obtener observaciones válidas de la superficie y por lo tanto se propone un método que requiere períodos de observación más largos y técnicas estadísticas adicionales. El objetivo de este artículo es utilizar los estimados de calidad QA-SDSs con el fin de evaluar más de una década de observaciones de MODIS y generar imágenes ópticas libres de nubes. Esto permitirá tener una idea clara de las potencialidades y limitaciones de los sensores ópticos en el seguimiento de regiones tropicales bajo condiciones de alta nubosidad.

## METODOLOGÍA

- Área de Estudio

La región biogeográfica del Chocó fue seleccionada para este estudio por sus condiciones extremas de precipitación y nubosidad. El clima del Chocó es bastante complejo debido a las diferencias altitudinales de las cordilleras costeras de los Andes y de Centro América, al clima que se ve condicionado en gran parte por la corriente de chorro superficial del oeste (*low level jet stream*) sobre el océano Pacífico a 5°N y a las bajas presiones de la zona de convergencia inter-tropical (ITCZ) (Poveda et al. 2006). La mayor parte del área se encuentra en el departamento del Chocó, pero condiciones climáticas y topográficas similares se extienden al sur, hacia los departamentos de Valle

del Cauca, Cauca y Nariño; y hacia el norte al departamento de Antioquia hasta los límites de la cuenca del Río León donde el cambio en relieve y la proximidad con el mar Caribe generan condiciones distintas. El área delimitada para los análisis es de 141654 km<sup>2</sup>. Figura 1.

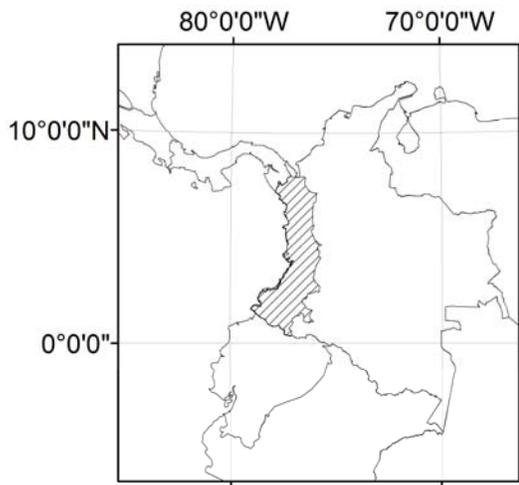


Figura 1. La zona sombreada corresponde a los límites del área de estudio

- Datos utilizados

Los productos MODIS seleccionados para generar imágenes libres de nubes son los índices de vegetación MOD13Q1, estos productos tienen la mejor observación en un periodo de 16 días siguiendo el método de ángulo de visión restringido (asociado a la geometría) y máximo valor del compuesto- (CV-MVC) asociado al índice de vegetación (Huete *et al.* 2002). La resolución espacial de MOD13Q1 es de 250 m para las bandas ( $\lambda$ ) Rojo 0.620-0.670  $\mu\text{m}$  e Infrarrojo cercano (IRC) 0.841-0.876  $\mu\text{m}$ ; aunque también incluye las bandas azul 0.459-0.479  $\mu\text{m}$  e infrarrojo medio (IRM) 2.105-2.155  $\mu\text{m}$ , originalmente de 500 m pero remuestreadas a 250 m. En los análisis se utilizaron 276 imágenes de la ventana H10V08 entre el año 2000 y 2011, ambos incluidos.

- Selección de observaciones válidas utilizando QA-SDSs a partir de compuestos

#### MOD13Q1

Con el fin de implementar la calidad del píxel se utilizó el constructor de series de tiempo TiSeG (Colditz *et al.* 2008). Esta herramienta programada en IDL se fundamenta en que la construcción de series de tiempo solo se pueden realizar con observaciones válidas, y por tanto, solo después un análisis riguroso de la calidad de cada una de las observaciones. TiSeG a su vez utiliza la estructura de QA-SDSs y la combinación de bits para cada parámetro de calidad según las Tablas 1 y 2. La estructura de TiSeG tiene toda la flexibilidad que QA-SDSs ofrece para marcar las observaciones como válidas o inválidas. Este punto es de especial cuidado pues la selección exclusiva de observaciones de ‘alta calidad’ hace que la disponibilidad de observaciones para el análisis sea mínima, y a su vez, se eliminan observaciones que aún con aspectos de calidad bajos pueden aportar información útil. Por esto es necesario ensayar múltiples alternativas de configuración para cada parámetro hasta obtener un equilibrio entre la calidad de los datos resultantes y el número de observaciones válidas. La configuración seleccionada para seleccionar observaciones válidas se lista en la Tabla 3.

**Tabla 3.** Configuración de TiSeG para selección de observaciones válidas.

Bits	Nombre del parámetro	Valor	Descripción
0-1	Calidad general del IV	00	Producido con alta calidad
		01	Producido con alta calidad, revisar otros aspectos
2-5	Utilidad del IV	0000	Perfecto
		0001	Alta calidad
		0010	Buena calidad
		0011	Calidad aceptable
		0100	Calidad aceptable mínima
		0101	Calidad intermedia
10	Mezcla de nubes	0	No

Las observaciones válidas se conservaron con su respectivo valor de reflectividad mientras que las observaciones inválidas fueron reclasificadas con el valor -1000 para luego excluirlos de cualquier operación estadística. Luego se agruparon 46 observaciones con el fin de definir un periodo de dos años como unidad para obtención de compuestos libres de nubes. Para cada período de dos años se consideraron diferentes estadísticos para hacer un nuevo filtro de la mejor observación, esto es, para cada pixel  $X_i$ ,  $Y_i$  y para cada banda  $\lambda$ . Las operaciones utilizadas en la serie de tiempo de 276 fechas fueron la media, la mediana, el mínimo y el máximo.

## RESULTADOS

Se generaron 6 series de tiempo, cada una de 46 fechas (2 años) para un total de 276 fechas (12 años) siguiendo la selección de píxeles expuesta en la Tabla 3. Observaciones entre Perfectas y Calidad intermedia fueron aceptadas como válidas para la construcción de las series de tiempo. Debido a la alta humedad y nubosidad de la zona de estudio se hace necesario disminuir la rigurosidad en la selección para obtener observaciones de la superficie terrestre con la menor contaminación posible. En la medida en que se aumenta el número de compuestos analizados de la serie de tiempo se aumenta la probabilidad de encontrar un pixel libre de nubes.

En este sentido se creó una serie de tiempo a partir del último compuesto de 2011 con día juliano 353. Para este compuesto el área cubierta con píxeles inválidos fue de 127905 km<sup>2</sup> (90% del área de estudio), adicionando los 5 compuestos de fechas inmediatamente anteriores, hasta el día juliano 273, la nueva disponibilidad de píxeles libres de nubes disminuye el área cubierta por píxeles inválidos a 79685 km<sup>2</sup>. Este

proceso se repitió hasta encontrar que la inclusión de nuevos compuestos en la serie de tiempo no reportaba mayores beneficios en la disminución de píxeles inválidos. Con una serie de tiempo de dos años se redujo el área con píxeles inválidos desde el 90% hasta el 2% Figura 2.

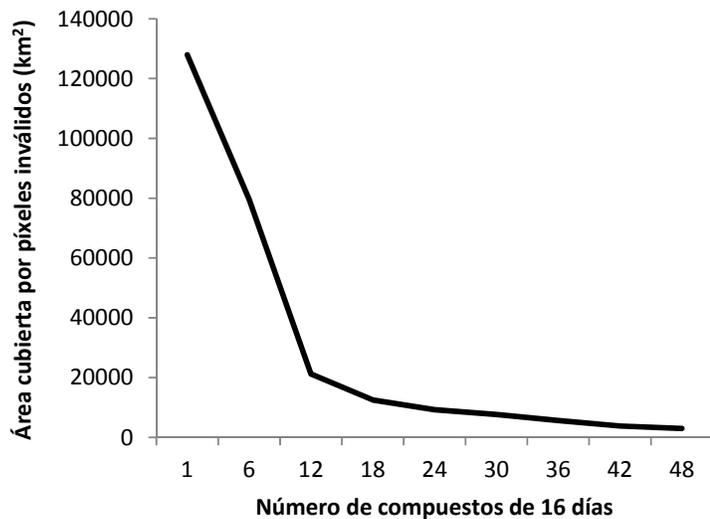


Figura 2. Disminución de observaciones inválidas en la medida en que se aumenta el número de compuestos en la serie de tiempo. Este análisis inicia en el compuesto 2011353 (18 de diciembre 2011) con 127905 km<sup>2</sup> de área con observaciones inválidas hasta 2009337 (2 de diciembre de 2009) con 3020 km<sup>2</sup> de área con observaciones inválidas.

Utilizando toda la serie de tiempo disponible se pueden lograr las mejores observaciones en términos de calidad, pero queda restringido cualquier análisis de cambios interanuales. La Figura 3 muestra el contenido de nubosidad en uno de los compuestos de menor contenido de nubosidad encontrada en los últimos 10 años según Anaya y Valencia (2013).

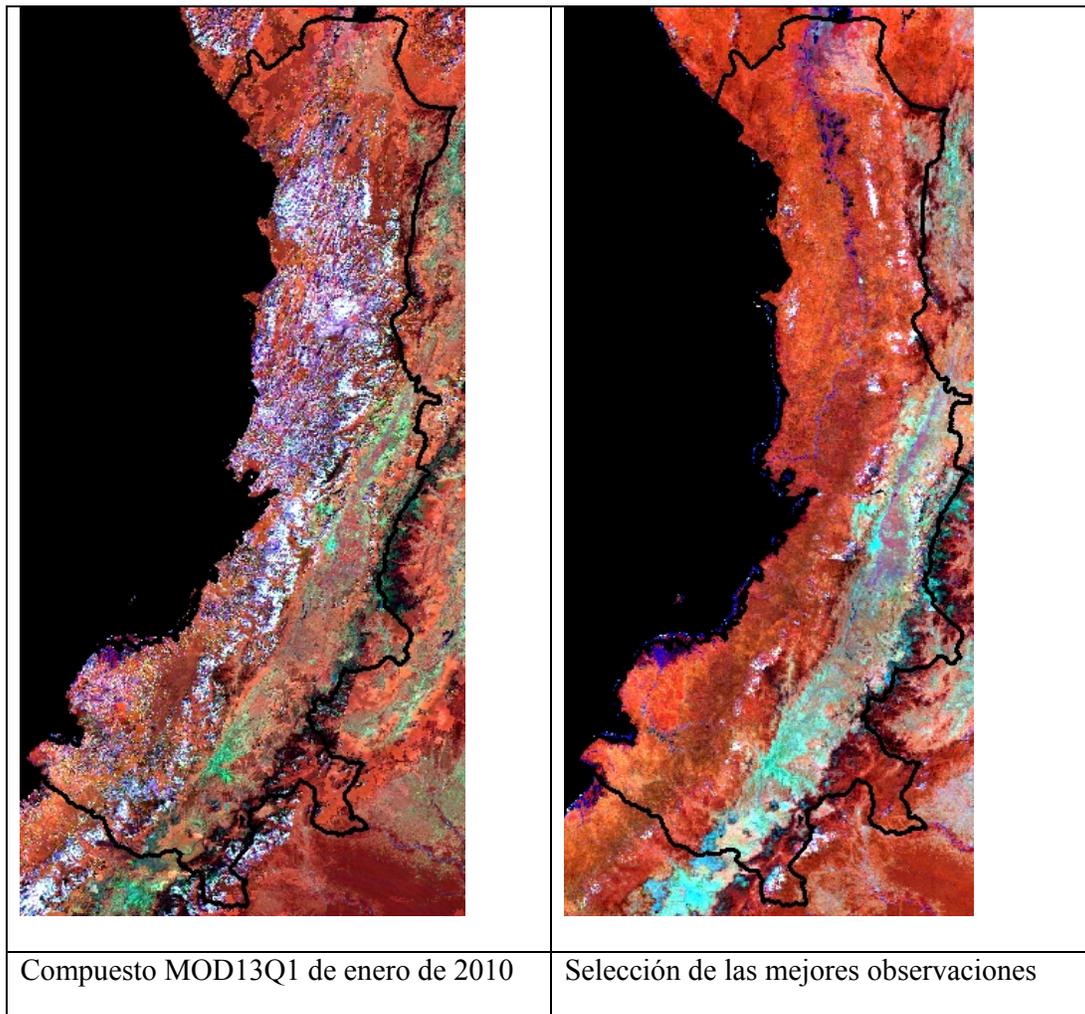


Figura 3. Imagen correspondiente al compuesto MOD13Q1 de enero de 2010, época seca y de menor contenido de nubosidad del año (izquierda). Nótese el alto contenido de nubosidad al compararla con la selección de las mejores observaciones (derecha).

Después de conformar las series de tiempo de dos años se pueden presentar tres casos para cada pixel: una sola observación válida, ninguna observación válida o varias observaciones válidas. En el primer caso se selecciona la única observación válida para el compuesto de salida, en el segundo caso es necesario aumentar la serie de tiempo hasta encontrar observaciones válidas para ese pixel y en el tercer caso se debe seguir

un procedimiento estadístico para seleccionar la mejor observación con el fin de generar los compuestos de dos años (Figura 4). La selección de observaciones de la serie de tiempo con el estadístico de “mínimo” evitó la inclusión de observaciones con alto contenido de nubosidad remanente de los filtros utilizados, dado que las nubes se caracterizan por su alto albedo. También favoreció la inclusión de los cuerpos de agua con bajo contenido de sedimentos, las superficies cubiertas por agua se consideran sumideros de energía y por tanto es de esperar que el uso del estadístico “mínimos” favorezca la identificación de cuerpos de agua, incluyendo ríos, ciénagas, canales y cualquier extensión inundada en el período de observación. También se encontró la inclusión de observaciones asociadas a la sombra ocasionada por las nubes o a la sombra ocasionada por efectos de la topografía. Por otro lado, al utilizar el estadístico de máximo se excluyó todo el efecto de las sombras caudas por las nubes y el relieve, pero se favoreció la selección de píxeles remanentes con alto contenido de nubosidad. Por tanto se hace evidente que existe un alto remanente de información no deseada, que puede identificarse con máximos y con mínimos, lo cual además es un indicador que cualquier intento de utilizar el cálculo con valores promedio por píxel implicaría la inclusión de datos inválidos en el cálculo de la media.

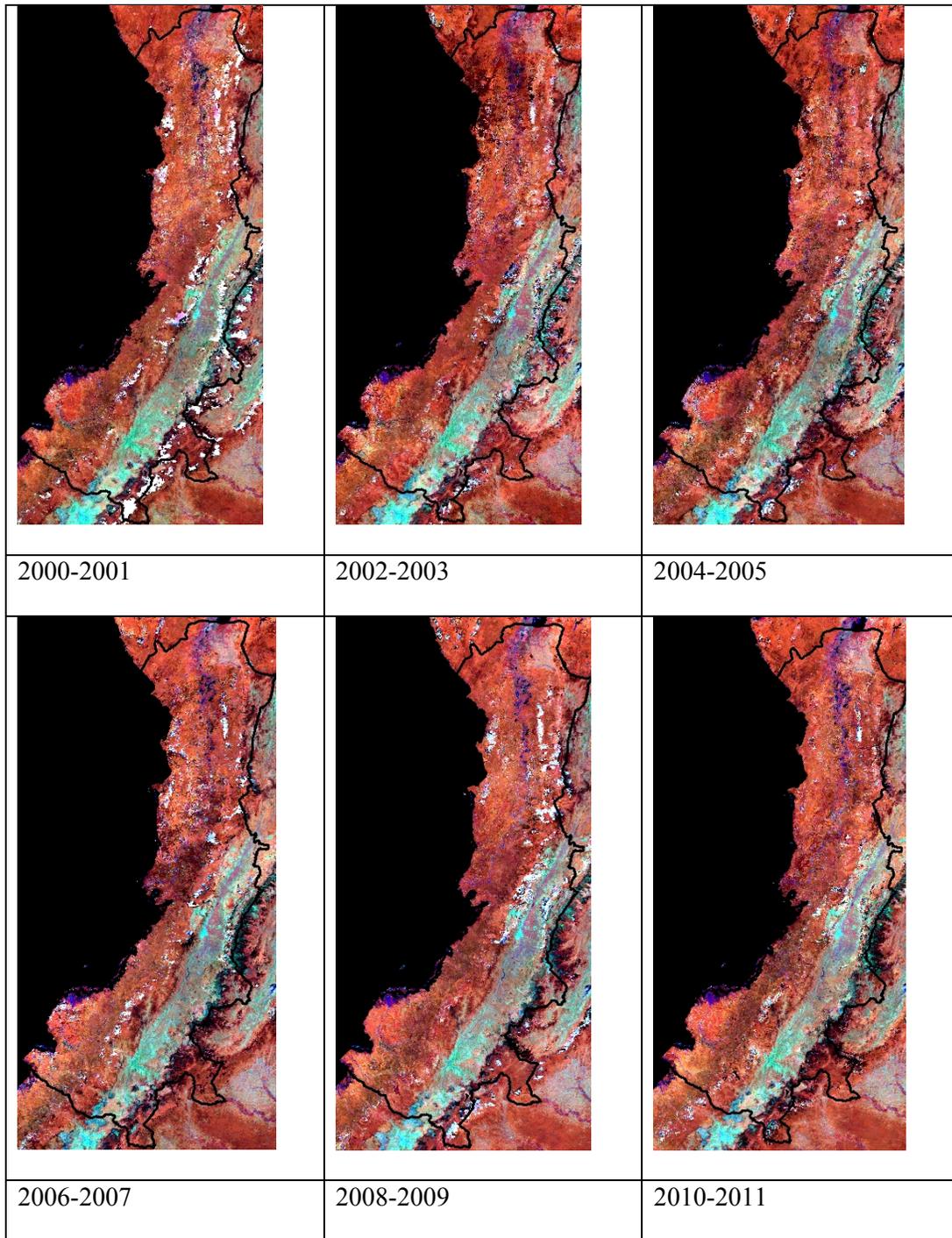


Figura 4. Cada imagen representa la mediana para 46 compuestos de 16 días, es decir, un compuesto dos años en falso color RGB-(IRC, IRM, Rojo)

## DISCUSIÓN DE RESULTADOS

A pesar de que la construcción de los compuestos MOD13A1 tiene un filtro de nubes y luego se aplica un nuevo filtro para la generación de series de tiempo (TiSeG) aún se conservan píxeles con alto contenido de nubosidad. La configuración de TiSeG y QA SDSs utilizada en este trabajo es aún muy permisiva en la inclusión de nubes para la generación de las series de tiempo. Si bien, desde el método se argumenta la necesidad de ser flexible con la selección de píxeles válidos, los datos resultantes en esta serie de tiempo no deben ser utilizados para el cálculos estadísticos como la media o la desviación estándar, pues se incluyen píxeles altamente contaminados por valores extremos de reflectividad, bien sea muy baja de sombras o muy alta de nubes. Haciendo un análisis de la serie de tiempo se encontró que la disminución de observaciones inválidas es drástica con una serie de tiempo de 23 fechas y llega a valores asintóticos con 46 fechas.

Tanto por el análisis temporal como por el análisis estadístico se optó por generar el compuesto con series de tiempo de dos años utilizando la mediana como método estadístico de selección. Este método reporta dos beneficios, el primero es que evita las observaciones extremas en la serie de datos de dos años, y el segundo es que el valor de reflectividad obtenido corresponde a un valor real, medido por el sensor y no a un valor resultante de un promedio. De esta forma se generaron 6 compuestos de dos años cada uno utilizando los datos MODIS disponibles entre 2000 y 2011 (Figura 4).

## CONCLUSIONES

En este artículo se presenta posiblemente la mejor observación óptica (Rojo, IRC, IRM) jamás obtenida de la región biogeográfica del Chocó. Esta parte del espectro

electromagnético es fundamental en la percepción remota dado que está directamente relacionado con la capacidad fotosintética de las plantas, la estructura de las hojas y el contenido de humedad de la superficie (Lillesand & Kiefer 2000). Los análisis climáticos de ciclo diurno sobre distribución de la precipitación y de la nubosidad son importantes para definir la hora en que se pueden obtener las mejores observaciones sobre el Chocó biogeográfico. En este sentido, los valores de reflectividad captados tanto desde Terra y como desde Aqua están disponibles para los usuarios, sin embargo, los análisis realizados por Poveda *et al.* (2005) indican que las imágenes con menor contenido de nubosidad serán aquellas captadas en la mañana, es decir, por el sensor instalado en Terra, con lo cual podría argumentarse que en la práctica la resolución temporal de Terra es mayor a la de Aqua para esta zona de estudio. Se observó que la construcción de series de tiempo de dos años es suficiente para disminuir casi en su totalidad el contenido de nubes. Para los años 2010-2011 se pasó de un contenido de nubes del 90% a un contenido de nubes del 2% del área de estudio. Si bien estos datos pueden variar si seleccionan otros años, también se ha demostrado que la distribución temporal y espacial año a año de las nubes en Colombia es muy consistente (Anaya & Valencia 2013). Con estos resultados se puede también concluir que el seguimiento anual de las coberturas vegetales en este tipo de ambientes y utilizando métodos ópticos desde satélites es difícil de lograr. En su lugar se recomiendan compuestos de dos años. Para los sitios remanentes con nubosidad persistente o perpetua, se sugiere el uso de datos radar o la toma de imágenes a baja altura.

## REFERENCIAS BIBLIOGRÁFICAS

- Anaya, J., Valencia, G.M. (2013). Fenología de ambientes tropicales en el marco de la teledetección. *Geofocus*, 13(2), 195-211.
- Colditz, R.R., Conrad, C., Wehermann, T., Schmidt, M., Dech, S. (2008). TiSeG: A flexible software tool for time-series generation of MODIS data utilizing the quality assessment science data set. *IEEE*, 46(10), 3296-3308.
- Hoyos, N., Escobar, J., Restrepo, J.C., Arango, A.M., Ortiz, J.C. (2013). Impact of the 2010–2011 La Niña phenomenon in Colombia, South America: The human toll of an extreme weather event. *Applied Geography*, 39(0), 16-25.
- Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1-2), 195-213.
- Leinenkugel, P., Kuenzer, C., Dech, S. (2013). Comparison and enhancement of MODIS cloud mask products for Southeast Asia. *International Journal of Remote Sensing*, 34(8), 2730-2748.
- Lillesand, T.M., Kiefer, R.W. (2000). Remote sensing and image interpretation, Fourth ed. Wiley.
- Liu, R., Liu, Y. (2013). Generation of new cloud masks from MODIS land surface reflectance products. *Remote Sensing of Environment*, 133(0), 21-37.
- Poveda, G., Mesa, O., Salazar, L.F., Arias, P.A., Moreno, H.A., Vieira, S.C., Agudelo, P.A., Toro, V., G., Alvarez, F.J. (2005). The diurnal cycle of precipitation in the tropical Andes of Colombia. *Monthly weather review*, 133, 228-240.
- Poveda, G., Waylen, P.R., Pulwarty, R.S. (2006). Annual and inter-annual variability of the present climate in northern South America and southern Mesoamerica. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 234, 3-27.
- Salomonson, V., Guenther, B., Masucka, E. (2001). A summary of the status of the EOS Terra Mission Moderate Resolution Imaging Spectroradiometer (MODIS) and attendant data product development after one year of on-orbit performance, Proceedings of the International Geoscience and Remote Sensing Symposium/IGARSS'2001, Sydney, Australia.

Savtchenko, A., Ouzounov, D., Ahmad, S., Acker, J., Leptoukh, G., Koziana, J., Nickless, D. (2004). Terra and Aqua MODIS products available from NASA GES DAAC. *Advances in Space Research*, 34(4), 710-714.

Sezgin, M., Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *Journal of electronic imaging*, 13, 146-168.

Yang, Y., Di Girolamo, L., Mazzoni, D. (2007). Selection of the automated thresholding algorithm for the Multi-angle Imaging SpectroRadiometer Radiometric Camera-by-Camera Cloud Mask over land. *Remote Sensing of Environment*, 107(1-2), 159-171.