

# Detección digital de Sistemas Convectivos de Mesoescala a partir de imágenes meteorológicas multiespectrales

Salomón Ramírez\*  
Ivan Lizarazo\*\*

*Recibido el 21 de noviembre de 2016; aceptado el 20 de abril de 2017*

## Abstract

The exact identification of Mesoscale Convective Systems (MCS) is not a simple task. In this work, Support Vector Machines (SVM), Decision Trees (DT) and Random Forests (RF) non-parametric machine learning algorithms (ML) were applied to detect MCS, from a series of GOES-13 weather satellite images acquired on April 03, 2013 every half hour from 11:45 to 22:15 hours, Coordinated Universal Time (UTC), covering the Colombian territory.

The results obtained by these methods were compared with a traditional method referred to as brightness temperature (BT). Accuracy assessment was conducted using STEP (*shape, theme, edge, position*), a method that evaluates geometric and thematic similarity between objects, using as reference a dataset of high accuracy data extracted from images of precipitation Tropical Rainfall Measuring Mission (TRMM).

The aim of this study was to determine whether using information from several spectral channels of weather images, rather than from a single infrared channel (IR) as traditional techniques do, allows accurate detection of MCS.

Experimental results show that the Decision Tree (DT) and Random Forest (RF) algorithms performed better than the IR-TB algorithm to detect MCS, while that the results of SVM algorithm, suggest that its use may not be favorable for practical applications.

\* Facultad de Ingeniería, Universidad Distrital Francisco José de Caldas, Bogotá, DC., Colombia, correo electrónico: seramirezf@correo.udistrital.edu.co

\*\* Facultad de Ciencias Agrarias, Universidad Nacional de Colombia, Bogotá, DC., Colombia, correo electrónico: ializarazos@unal.edu.co

Accuracy ranges at 95% confidence levels to DT, RF, TB and SVM algorithms, at evaluated time's instants [12:15, 15:15, 18:15, 21:15] by similarity metric (*[shape]*, *[theme]*, *[edge]*, *[position]*) in their respective order were: DT ([89%-98%], [75%-94%], [68%-96%], [80%-98%]), RF ([87%-95%], [73%-89%], [67%-87%], [78%-90%]), IR-TB ([83%-95%], [68%-87%], [60%-81%], [72%-91%]) y SVM ([86%-93%], [72%-83%], [65%-78%], [76%-81%]).

The decision criteria of the classification model yielded by DT could be replicated several times in different dates without performing visual interpretation in each image, being very useful for operational applications under the approach presented here.

Key words: *machine learning algorithms, weather images, decision trees, support vector machines, random forests, mesoscale convective systems.*

## Resumo

A identificação exata dos Sistemas Convectivos de Mesoescala (SCM) não é uma tarefa simples. No presente trabalho se aplicaram algoritmos não paramétricos de aprendizagem de máquina (ML) às Máquinas de Suporte Vetorial (SVM), Árvores de Decisão (DT) e Florestas Aleatórias (RF) para detectar SCM, a partir de uma série de subcenas de imagens meteorológicas tomadas pelo satélite GOES-13 em 3 de abril de 2013 a cada meia hora desde às 11:45 até às 22:15 horas, Tempo Universal Coordenado (UTC), que cobrem o território colombiano.

Os resultados obtidos por estes métodos foram comparados com o método tradicional denominado Temperatura de Brillo (TB). Para a avaliação dos resultados se empregou *STEP (shape, theme, edge, position)*, um método que avalia a semelhança geométrica e temática entre objetos, tomando como referência um conjunto de dados de alta precisão extraídos de imagens de precipitação da Missão para a Medição de Chuvas Tropicais (TRMM).

O objetivo deste trabalho foi determinar se, ao empregar informação de vários canais espectrais de imagens meteorológicas em lugar de usar um só canal infravermelho (IR), como fazem as técnicas tradicionais, é possível obter resultados mais exatos na detecção do SCM.

Os resultados experimentais mostraram que os algoritmos DT e RF proporcionam uma detecção mais exata que a obtida pelo algoritmo tradicional IR-TB para detectar SCM, a menos que os resultados do algoritmo SVM, sujam não sendo favorável seu uso para aplicações práticas.

As gamas de precisão obtidas com 95% de confiança para os algoritmos DT, RF, TB y SVM, nos instantes de tempo avaliados [12:15, 15:15, 18:15, 21:15] por métrica de semelhança (*[shape]*, *[theme]*, *[edge]*, *[position]*) foram em sua respectivo ordem: DT ([89%-98%], [75%-94%], [68%-96%], [80%-98%]), RF ([87%-95%], [73%-89%], [67%-87%], [78%-90%]), IR-TB ([83%-95%], [68%-87%], [60%-81%], [72%-91%]) y SVM ([86%-93%], [72%-83%], [65%-78%], [76%-81%]).

Os critérios de decisão do modelo de classificação obtido por DT poderiam ser replicados várias vezes em diferentes datas sem se realizar interpretação visual em cada imagem, sendo muito útil para aplicações operativas sob o enfoque que se apresenta aqui.

Palavras chave: *algoritmos de aprendizagem de máquina, imagens meteorológicas, árvores de decisão, máquinas de suporte vetorial, florestas aleatórias, sistemas convectivos de mesoescala.*

## Resumen

La identificación exacta de Sistemas Convectivos de Mesoescala (SCM) no es una tarea sencilla. En el presente trabajo se aplicaron los algoritmos de aprendizaje de máquina (ML) no paramétricos Máquinas de Soporte Vectorial (SVM), Árboles de Decisión (DT) y Bosques Aleatorios (RF) para detectar SCM, a partir de una serie de sub-escenas de imágenes meteorológicas tomadas por el satélite GOES-13 el 03 de abril de 2013 cada media hora desde las 11:45 hasta las 22:15 horas, Tiempo Universal Coordinado (UTC), que cubren el territorio colombiano.

Los resultados obtenidos por estos métodos fueron contrastados con el método tradicional denominado Temperatura de Brillo (TB). Para la evaluación de los resultados se empleó *STEP (shape, theme, edge, position)*, un método que evalúa la similitud geométrica y temática entre objetos, tomando como referencia un conjunto de datos de alta exactitud extraído de imágenes de precipitación de la Misión para la Medición de Lluvias Tropicales (TRMM).

El objetivo de este trabajo fue determinar si al emplear información de varios canales espectrales de imágenes meteorológicas en lugar de usar un solo canal infrarrojo (IR) como lo hacen las técnicas tradicionales, es posible obtener resultados más exactos en la detección de SCM.

Los resultados experimentales mostraron que los algoritmos DT y RF proporcionan una detección más exacta que la obtenida por el algoritmo tradicional IR-TB para detectar SCM, mientras que los resultados del algoritmo SVM, sugieren que no es favorable su uso para aplicaciones prácticas.

Los rangos de exactitud obtenidos con el 95% de confianza para los algoritmos DT, RF, TB y SVM, en los instantes de tiempo evaluados [12:15, 15:15, 18:15, 21:15] por métrica de similitud ([*shape*], [*theme*], [*edge*], [*position*]) fueron en su respectivo orden: DT ([89%-98%], [75%-94%], [68%-96%], [80%-98%]), RF ([87%-95%], [73%-89%], [67%-87%], [78%-90%]), IR-TB ([83%-95%], [68%-87%], [60%-81%], [72%-91%]) y SVM ([86%-93%], [72%-83%], [65%-78%], [76%-81%]).

Los criterios de decisión del modelo de clasificación obtenido por DT podrían ser replicados varias veces en diferentes fechas sin realizar interpretación visual en cada imagen, siendo muy útil para aplicaciones operativas bajo el enfoque que aquí se presenta.

Palabras clave: *algoritmos de aprendizaje de máquina, imágenes meteorológicas, árboles de decisión, máquinas de soporte vectorial, bosques aleatorios, sistemas convectivos de mesoescala.*

## Introducción

En latitudes tropicales, tanto ecuatoriales como medias, las nubes Cumulonimbus (Cb), son responsables de causar algunas condiciones climáticas extremas; estas nubes producen convección intensa que origina la formación de Sistemas Convectivos de Mesoescala (SCM) (Liu *et al.*, 2014). Un SCM es un sistema de nubes Cb de origen convectivo que produce un área de precipitación contigua de aproximadamente 100 km en por lo menos una dirección (Houze, 2004).

La presencia de SCM supone un grave riesgo para la aviación y puede afectar a los cultivos y las poblaciones urbanas debido a la rápida evolución de tiempo atmosférico que se produce en diferentes escalas espaciales y temporales (Liu *et al.*, 2014). En Colombia se localiza un gran número de SCM, los cuales de acuerdo a los datos de la Misión para la Medición de Lluvias Tropicales (TRMM) durante el período 1998-2002, contribuyeron aproximadamente al 70% de la precipitación anual del país (Sakamoto *et al.*, 2011). En este país particularmente, son recordados los desastres causados por las fuertes lluvias producidas por SCM, que azotaron a gran parte del territorio nacional entre el segundo semestre de 2010 y el primer semestre 2011, causando efectos devastadores causados sobre la población, así como sobre la economía local, regional y nacional (Colombia Humanitaria, 2010).

La observación de SCM puede ser una fuente importante de datos para su asimilación en la predicción del tiempo atmosférico y el seguimiento de las tendencias climáticas (Liu *et al.*, 2014). En tal sentido, las imágenes de los Satélites Geoestacionarios Operacionales Ambientales (GOES), han demostrado un alto potencial para la detección y seguimiento de SCM (Goyens *et al.*, 2012).

Una técnica bien conocida para llevar a cabo estas tareas, es el uso de umbrales de temperatura de brillo (TB) en el canal infrarrojo térmico (TIR) de 10.8 $\mu$ m, asumiendo que temperaturas bajas de brillo en este canal están asociadas con convección profunda que penetra en el interior de la troposfera superior (Goswami y Bhandari, 2012; Ravelo y Santa, 2000). Mientras esta técnica es simple de usar, varios estudios han mostrado que, en la atmósfera tropical, las estimaciones de precipitación basadas en este canal pueden no ser tan exactas como se espera, particularmente sobre latitudes ecuatoriales (Laurent *et al.*, 2002; Machado *et al.*, 1998). El método IR-TB es actualmente empleado por el Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM), servicio meteorológico colombiano, para realizar la detección y monitoreo de la evolución de SCM, usando como fuente de datos las imágenes meteorológicas GOES-13.

El propósito de este trabajo consiste en determinar si el empleo de información de varios canales espectrales de imágenes meteorológicas, en lugar del uso de un solo canal como lo hacen las técnicas tradicionales, permite obtener resultados confiables y consistentes en la detección de SCM.

### **Detección y seguimiento de SCM**

La detección y el seguimiento de SCM a partir de imágenes meteorológicas se realiza tradicionalmente de forma manual (Fiolleau y Roca, 2013), mediante interpretación visual (nefoanálisis) de imágenes del canal visible (VIS) y del canal infrarrojo (IR) (Desbois *et al.*, 1982), obtenidas de satélites de órbitas geoestacionarias (GEO, del inglés Geostationary Earth Orbit), cuyas tonalidades de grises permiten a los interpretes reconocer detalles inherentes a la estructura de los SCM. Sin embargo, este conocimiento no es fácil de adquirir y suele ser desarrollado después de muchos años de experiencia práctica (Peak y Tag PM, 1994), siendo un trabajo dispendioso y subjetivo (Machado *et al.*, 1998), cuyas limitaciones son más marcadas cuando el personal rota con frecuencia.

Para superar esas limitaciones, desde mediados de 1980 a la fecha, un número de autores han propuesto diferentes técnicas digitales para el análisis de SCM, las cuales en su gran mayoría se basan en la aplicación de algoritmos de detección a partir del infrarrojo térmico (TIR), mediante el uso de umbrales de temperatura de brillo (TB) asociados a bajas temperaturas (Tabla 1). El principal beneficio de usar umbrales de TB considerando tan solo los canales térmicos reside en su simplicidad (Morel y Senesi, 2002), y está basado en la premisa que el comienzo de las precipitaciones ocurre cuando la temperatura del tope de las nubes, es decir la más alta altitud de la parte visible de la nube, desciende de dicho umbral, y se considera que dichas nubes tienen gran desarrollo vertical y por lo tanto existe una alta probabilidad de que produzcan precipitaciones (Ravelo y Santa, 2000).

Sin embargo, las técnicas basadas en la TB del tope de la nube tienden a sobrestimar las precipitaciones en ambientes secos y a subestimarlas en ambientes más húmedos (Ba y Gruber, 2001), por ejemplo, no permiten distinguir nubes Cirrus (Ci) espesas de nubes convectivas profundas (Cb). Al respecto, algunos autores (Bader *et al.*, 1996) señalan que a partir de los otros canales es posible obtener información útil de las nubes; como por ejemplo su espesor y tamaño, derivados de los canales visibles (VIS); su temperatura y altura del tope, derivadas de los canales infrarrojos (IR) y su fase del ciclo de vida, derivada de canales de vapor de agua (WV) (Suvichakorn, 2007).

En la Tabla 1 se resumen los aspectos más relevantes de los principales estudios relacionados con el seguimiento y pronóstico de nubes convectivas, consignando características inherentes a las fuentes de datos empleadas, por ejemplo, la resolución

espacial ( $\Delta x$ ) y resolución temporal ( $\Delta t$ ), así como el método aplicado para la detección de SCM.

A partir de la información consignada en la Tabla 1, se evidencia que a excepción de Goswami y Bhandari, que emplean para la detección de SCM un método de clasificación no supervisado sobre el canal IR, la mayoría de autores emplean el método IR-TB cuyos umbrales varían en el intervalo [213 K, 258K].

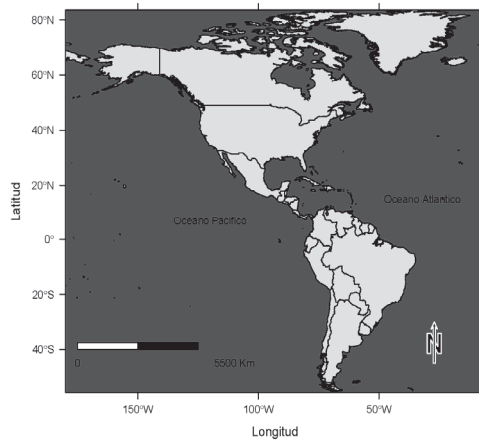
**Tabla 1**  
**Principales aspectos considerados durante detección de SCM**

<i>Autor</i>	<i>Satélite/Sensor</i>	$\Delta x$	$\Delta t$	<i>Canales</i>	<i>Método</i>
Woodley <i>et al.</i> (1980)	SMS-1	4 km	30 min	TIR VIS/TI	TB $\leq$ 253K
Klitch <i>et al.</i> (1985)	GOES-5	4 km	30 min	R	TB $\leq$ 243K
Williams y Houze (1987)	GMS-1	2.4 km	180 min	TIR	TB < 213K
Arnaud <i>et al.</i> (1992)	METEOSAT-4	5 km	30 min	TIR	TB $\leq$ 233K
Chen <i>et al.</i> (1996)	GMS-1	5 km	30 min	TIR	TB $\leq$ 235K
Chen y Houze (1997)	GMS-1	10 km	30 min	TIR	TB $\leq$ 235K
Hodges y Thorncroft (1997)	METEOSAT	150 km	180 min	TIR	TB $\leq$ 258K
Laurent <i>et al.</i> (1998)	METEOSAT-4	5 km	30 min	TIR	TB $\leq$ 233K
Machado <i>et al.</i> (1998)	GOES-7	30 km	180 min	TIR	TB $\leq$ 245K
Carvalho y Jones (2001)	GOES-8	4 km	30 min	TIR	TB $\leq$ 235K
Machado (2002)	GOES-8	4 km	30 min	TIR	TB $\leq$ 233K
Morel y Senesi (2002)	METEOSAT	6 km	30 min	TIR	TB $\leq$ 243K
Vila <i>et al.</i> (2008)	GOES-8	4 km	30 min	TIR	TB $\leq$ 235K
Goswami y Bhandari (2011)	KALPANA-1	8 km	30 min	TIR	K-Means
Sakamoto <i>et al.</i> (2011)	GOES-12	4 km	30 min	TIR	TB $\leq$ 235K
Sieglaff <i>et al.</i> (2011)	GOES-12	4 km	30 min	TIR	TB $\leq$ 235K
Goswami y Bhandari (2012)	KALPANA-1	8 km	30 min	TIR	K-Means
Goswami y Bhandari (2012b)	KALPANA-1	8 km	30 min	TIR	K-Means
Goyens <i>et al.</i> (2012)	METEOSAT-8	3 km	15 min	TIR	TB $\leq$ 235K
Shukla y Pal (2012)	KALPANA-1	8 km	30 min	TIR	TB $\leq$ 241K
Fiolleau y Roca (2013)	METEOSAT-7	3 km	30 min	TIR	TB $\leq$ 235K

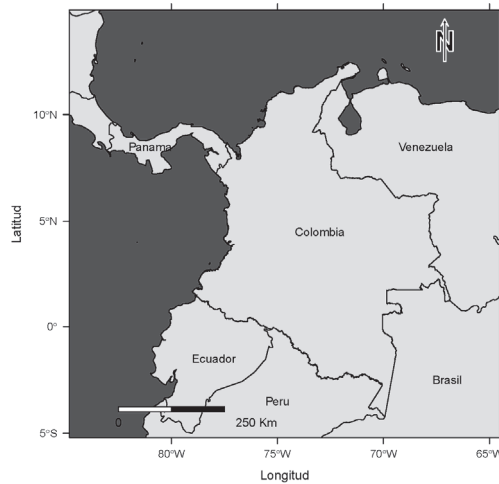
## Datos y métodos

### Área de estudio

El presente trabajo se llevó a cabo sobre la zona comprendida desde  $-84.81^\circ$  de longitud Oeste (W) y  $-5,21^\circ$  de latitud Sur (S), hasta  $-64.5^\circ$  de longitud W y  $14.92^\circ$  de latitud Norte (N), que comprende principalmente el territorio colombiano, el cual se encuentra ubicado en la esquina noroccidental de América del Sur (Figura 1a). Colombia ocupa una extensión de 2,070,408 km<sup>2</sup>, con una proporción aproximada de territorio marítimo del 45% y de territorio continental del 55% (Figura 1b).



(a) América



(b) Colombia

Figura 1. Localización área de estudio.

## Datos

Los datos usados para la detección de SCM comprenden una serie de 22 sub-escenas meteorológicas en modo-A (8 bits), cada una con 2240 x 2260 píxeles con resolución espacial de 1 km a 8 km, tomadas por el satélite GOES-13 el 03 de abril de 2013 desde las 11:45 hasta las 22:15 horas Tiempo Universal Coordinado (UTC).

El satélite GOES-13 transmite imágenes de la atmósfera cada 30 minutos, en 5 canales espectrales diferentes, uno visible (VIS), tres infrarrojos (IR) y uno de vapor de agua (WV), con resoluciones espaciales de 1 km en las longitudes de onda del visible, 4 km en las longitudes de onda del infrarrojo y 8 km en las longitudes de onda del vapor de agua (Tabla 2) (Xu, Wooster, Roberts y Freeborn, 2010).

**Tabla 2**  
**Características imágenes GOES-13**

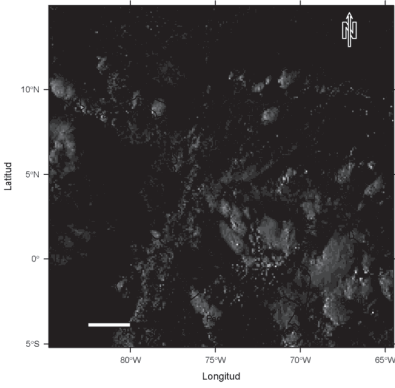
<i>Canal</i>	<i>Resolución espectral</i>	<i>Resolución espacial</i>	<i>Descripción</i>
1	0.52-0.72 $\mu\text{m}$	1 km	Visible (VIS)
2	3.78-4.03 $\mu\text{m}$	4 km	Infrarrojo medio (MIR)
3	6.47-7.02 $\mu\text{m}$	8 km	Vapor de agua (WV)
4	10.2-11.2 $\mu\text{m}$	4 km	Infrarrojo térmico (TIR1)
5	12.9-13.8 $\mu\text{m}$	4 km	Infrarrojo térmico (TIR2)

En las Figuras 2a-2e, se muestran los cinco canales que conforman las sub-escenas de GOES correspondientes al instante de tiempo (20130403-21:15 UTC) empleado como referencia durante el desarrollo del estudio.

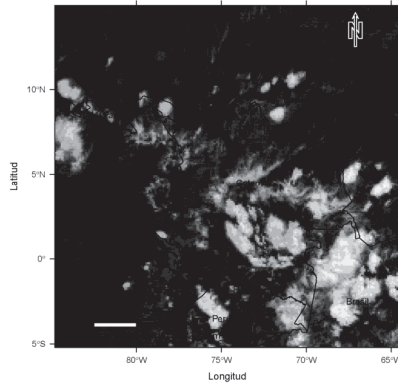
Algunas sub-escenas de imágenes correspondientes al producto 3B42 del satélite TRMM (Tropical Rainfall Measuring Mission) de la NASA (National Aeronautics and Space Administration), fueron empleadas para extraer objetos precipitantes de referencia, que permitieran evaluar la exactitud de los objetos precipitantes extraídos de las imágenes GOES-13.

En la Figura 2f, se muestra una sub-escena del producto TRMM-3B42, que combina datos de infrarrojos (escáner) y microondas (radar) (Doraiswamy, Natarajan y Nanjundiah, 2013), proporcionando la fracción de área de precipitación y estimaciones de la tasa media de precipitación expresada en mm/hr, en el rango de tiempo de  $\pm 90$  min cada 3 hr (Goyens *et al.*, 2012), para todas las longitudes de la tierra comprendidas entre los 50° de latitud N y los 50° de latitud S, en una grilla con resolución espacial de 25 km (Doraiswamy *et al.*, 2013). Este conjunto de imágenes TRMM, comprende una serie de 4 sub-escenas de precipitación, correspondientes al 03 de abril de 2013 desde las 12:00 hasta las 21:00 horas UTC.

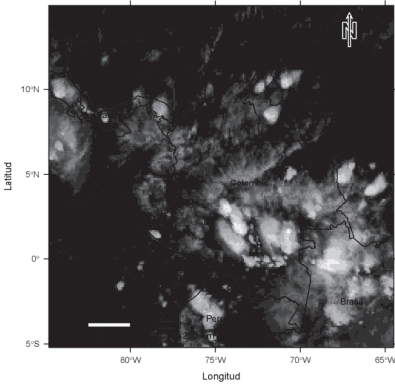




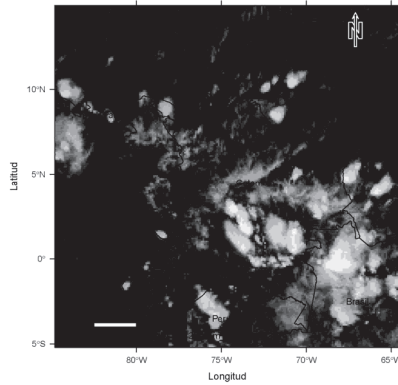
**a) Canal VIS**



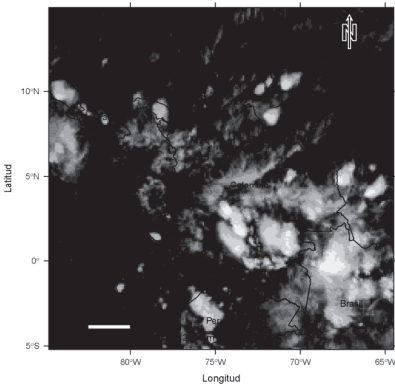
**(b) Canal MIR**



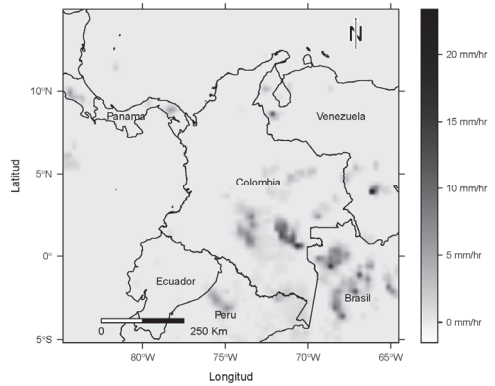
**(c) Canal WV**



**(d) Canal TIR1**



**(e) Canal TIR 2**



**(f) Precipitación TRMM (3B42)**

**Figura 2.** Imágenes canales GOES-13 y NASA-TRMM (3B42).

## Métodos

En este estudio se emplearon los algoritmos de Aprendizaje de Máquina (ML), Máquinas de Soporte Vectorial (SVM), Árboles de Decisión (DT) y Bosques Aleatorios (RF), para extraer nubes precipitantes a partir de una serie de imágenes meteorológicas. La Figura 3 muestra esquemáticamente la metodología seguida, que abarcó dos enfoques de análisis principales: (A) análisis basado en píxeles que incluye las fases 1 a 5 y (B) análisis orientado a objetos que incluye las fases 6 y 7. A continuación se describe cada una de las fases:

En la primera fase, identificada con el número 1 en la Figura 3, se realizó el pre-procesamiento de la imagen, que incluyó: (i) lectura de cada canal que conforma cada imagen de la serie temporal; (ii) asignación del sistema de referencia espacial MAGNA-SIRGAS (EPSG:4686) a cada canal de la serie de imágenes; (iii) remuestreo del tamaño de los píxeles de los canales MIR, WV, TIR1 y TIR2 de cada imagen, al tamaño del píxel del canal VIS (1 km) de cada imagen, aplicando el método bilineal; (iv) recorte de la extensión geográfica cubierta por los canales MIR, WV, TIR1 y TIR2 de cada imagen ( $-88.23^\circ$  W,  $-9.14^\circ$  S,  $-51.58^\circ$  E,  $18.34^\circ$  N), a la extensión geográfica cubierta por el canal VIS de la serie de imágenes ( $-84.81^\circ$  W,  $-64.5^\circ$  E,  $-5.21^\circ$  S,  $14.92^\circ$  N); y (v) agrupación de todos los canales en una sola imagen para cada instante de la serie.

Durante la segunda fase, identificada con el número 2 en la Figura 3, se realizó la delimitación de los objetos de entrenamiento, adoptando como clases objetivo  $\omega_j$  con  $j \in \{1, \dots, R\}$  (Lizarazo, 2008) y  $R = 3$ , correspondientes a la leyenda de clasificación propuesta por Tsonis (1984) y Tsonis e Isaac (1985), siendo establecidas tres clases principales: Cielo despejado (Cd), Nubes que no precipitan (Nn) y Nubes que precipitan (Np).

A partir de esta leyenda, se realizó la interpretación visual de los canales VIS y TIR1, tomando como imagen de referencia la correspondiente al instante 2013-04-03-21:15 UTC. Se obtuvieron 18 objetos de entrenamiento, mostrados en la Figura 4, divididos como sigue: siete pertenecientes a la clase Cd (verde), seis a la clase Nn (amarillo) y cinco a la clase Np (rojo), los cuales comprenden un total de 270,563 píxeles.

La tercera fase, identificada con el número 3 en la Figura 3, consistió en realizar un diseño de muestreo y su implementación para establecer un modelo de entrenamiento del proceso de clasificación, a partir de muestras representativas de las clases seleccionadas (Lizarazo, 2008), que comprendió: (i) definición del tamaño de la muestra, siendo aplicado el criterio de Tso y Mather (2009), que sugiere que el número mínimo de datos de píxeles por clase objetivo  $\omega_j$  debe ser 30 veces el número de atributos (canales)  $x_i$  con  $i \in \{1, \dots, n\}$  (Lizarazo, 2008) y  $n = 5$ , utilizados para identificar las clases, obteniendo así un tamaño de 450 muestras de entrenamiento;

(ii) distribución espacial de la muestra de entrenamiento, realizada de manera aleatoria sobre los objetos de entrenamiento; y (iii) extracción de la respuesta multispectral  $x_i$  de la imagen de referencia y la etiqueta de la clase objetivo  $x_j$  asociada a los objetos de entrenamiento en cada localización  $j$  de la muestra.

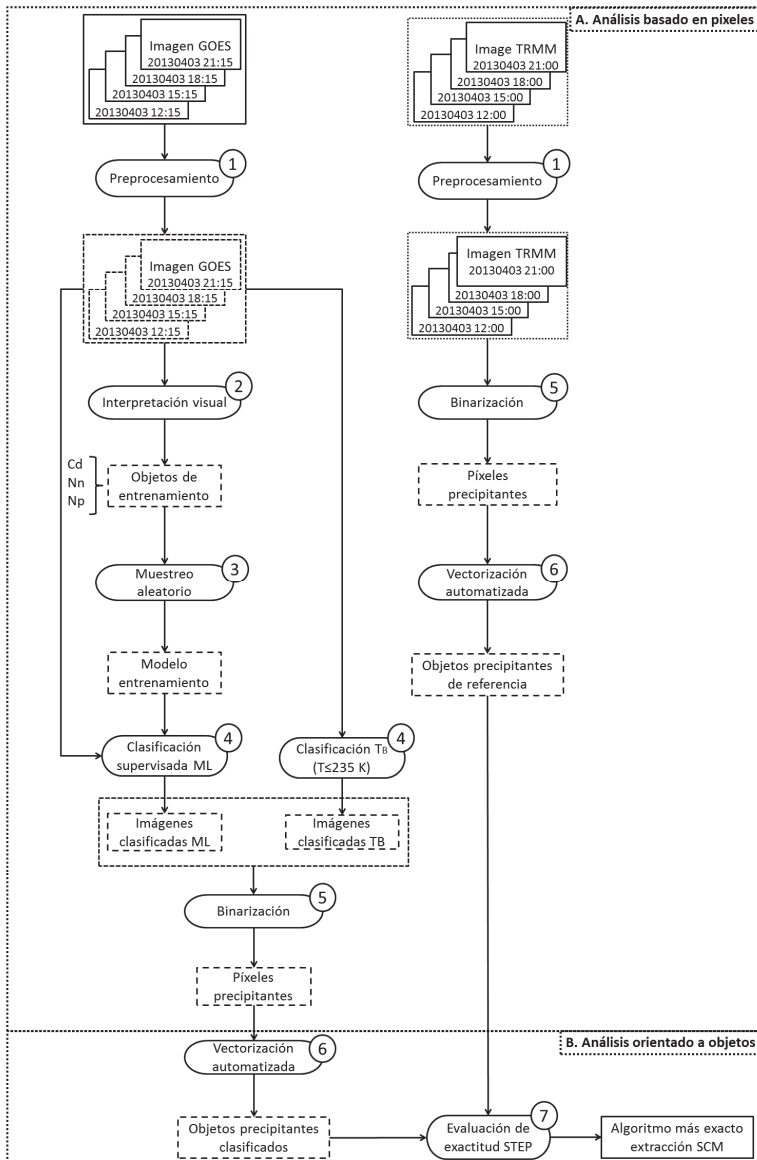
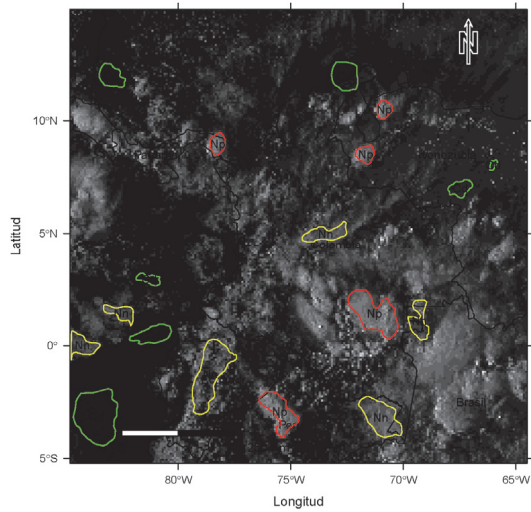
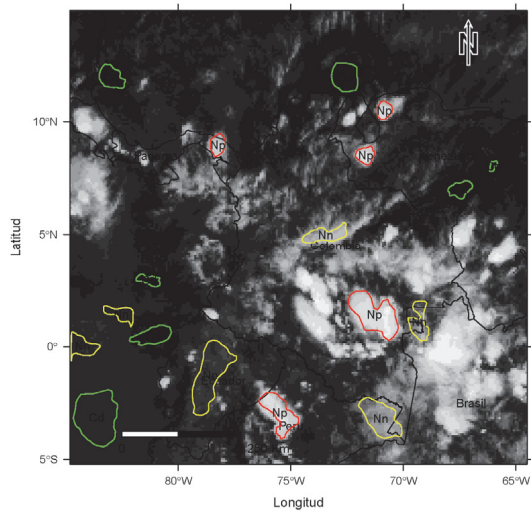


Figura 3. Metodología.



(a) Canal VIS



(b) Canal TIR1

**Figura 4.** Objetos de entrenamiento sobre imagen GOES 2013-04-03-21:15 UTC.

En la cuarta fase, identificada con el número 4 en la Figura 3, se realizó la clasificación supervisada de SCM basada en píxeles para la serie de imágenes, aplicando el modelo de entrenamiento establecido previamente, con el objetivo dividir su espacio de atributos en compartimentos que englobaran cada una de las clases objetivo

(Lizarazo, 2008) mediante la implementación de los algoritmos de aprendizaje máquina (ML) no paramétricos Máquinas de Soporte Vectorial (SVM), Árboles de Decisión (DT) y Bosques Aleatorios (RF) (Boulesteix *et al.*, 2012; Hastie *et al.*, 2009; Tan *et al.*, 2005; Tso y Mather, 2009; Vapnik, 2000), los cuales han sido ampliamente aplicados para resolver problemas científicos y de ingeniería durante casi dos décadas, sin embargo la aplicación de estos métodos en el área de geociencias y teledetección es bastante reciente y limitada (Lary *et al.*, 2016), sin evidencia aparente de haber sido utilizados en aplicaciones meteorológicas, ganando popularidad en los últimos años en la clasificación de datos de sensores remotos para uso del suelo y cobertura de la tierra, debido a que sus resultados experimentales sugieren que computacionalmente son más eficientes y exactos que otros métodos (Lizarazo y Elsner, 2009; Lizarazo, 2008, 2010).

El método tradicional de detección de nubes precipitantes conocido como Temperatura de Brillo (TB) también fue usado como referencia para comparación de resultados. Este método asume que la precipitación está altamente correlacionada con la fracción de área cubierta por píxeles más fríos que un umbral de temperatura dado. En el presente estudio (i) los niveles digitales (ND) Modo-A del canal TIR1 de GOES (10.8 $\mu$ m) fueron convertidos en temperatura de brillo aplicando las ecuaciones (1) y (2) (Tsonis, 1984); y (ii) luego para delimitar áreas precipitantes por TB se empleó un umbral de temperatura 235K (-38,15° C), por parecer apropiado para la detección de nubes precipitantes en diferentes regiones tropicales de América del Sur (Vila *et al.*, 2008), separando así píxeles más fríos que dicho umbral de otros píxeles.

$$TB = 330K - \frac{ND_{TIR1}}{2}; ND_{TIR1} \leq 176 \quad (1)$$

$$TB = 418K - ND_{TIR1}; ND_{TIR1} > 176 \quad (2)$$

Donde:

TB es la temperatura de brillo calculada a partir de los niveles digitales (ND) del canal TIR1.

Durante la quinta fase, identificada con el número 5 en la Figura 3, se realizó la binarización de las imágenes clasificadas, para separar píxeles que representaban precipitaciones de otros píxeles. Asignando el valor 1 a los píxeles correspondientes a la clase *nubes precipitantes* y el valor 0 en caso contrario, para las clasificaciones obtenidas mediante los algoritmos ML, mientras que a las imágenes de TB se les asignó el valor de 1 a los píxeles más fríos que 235K y el valor de 0 en caso contrario. Las cantidades de precipitación de TRMM también fueron binarizadas mediante la conversión de intensidades de lluvia no nulas en 1 y 0 de otro caso, obteniendo así un conjunto exacto de áreas precipitantes de referencia para llevar a cabo la evaluación de los métodos de clasificación.

La sexta fase, identificada con el número 6 en la Figura 3, consistió en extraer áreas precipitantes a partir de la vectorización automatizada de las imágenes binarizadas, agregando así píxeles contiguos de precipitación para crear polígonos (Lizarazo y Barros, 2010), que representaran objetos precipitantes.

En la séptima fase, se llevó a cabo la evaluación de la exactitud temática y geométrica de los objetos precipitantes obtenidos por cada método, aplicando la matriz de similitud STEP propuesta recientemente (Lizarazo, 2014), que provee cuatro (4) métricas de similitud: (i) *shape* (forma); (ii) *theme* (tema); (iii) *edge* (frontera); y (iv) *position* (posición).

A partir de las métricas expuestas anteriormente se construyó una matriz de similitud STEP, la cual fue empleada para realizar el cálculo de la exactitud global (PCC), de manera similar a como se calcula dicha exactitud en una matriz de error tradicional (Congalton y Green, 2008), pero incorporando en cada celda de la matriz de error el área de los objetos en lugar de su conteo, que corresponde a un área ponderada para compensar la desigualdad entre el área los SCM y el área de otros objetos (Lizarazo, 2014), siendo esa última mucho mayor que la primera, por lo que la probabilidad de clasificar correctamente estos últimos sea mucho mayor.

También se calculó el intervalo de confianza (IC) para PCC, con un nivel de confiabilidad del 95%, empleando la Ecuación 3 (Tan *et al.*, 2005):

$$IC = \frac{2 \times N \times PCC + Z^2 \mp Z \sqrt{Z^2 + 4 \times N \times PCC - 4 \times N \times PCC^2}}{2(N + Z^2)} \quad (3)$$

donde *PCC* es la exactitud, *N* es el tamaño de la muestra de validación y *Z* es el valor de probabilidad a dos colas de la distribución normal estándar para el nivel de confiabilidad evaluado.

La evaluación fue realizada para los instantes 12:15, 15:15, 18:15 y 21:15 horas UTC, de acuerdo a la disponibilidad de los objetos precipitantes de referencia extraídos de las imágenes de precipitación de TRMM. Los datos brindados por TRMM son de alta exactitud temática, lo cual en términos prácticos resulta ser favorable para realizar la validación de resultados. Sin embargo, dicho conjunto de datos es de menor resolución espacial (625 km<sup>2</sup>) que el conjunto de datos de GOES-13 (1 km<sup>2</sup>), empleado para extraer los objetos precipitantes por diferentes métodos e instantes de tiempo. Por lo tanto, para minimizar el efecto de desajuste entre las dos escalas espaciales, la evaluación se enfocó en los objetos precipitantes que exhibían una organización de mesoescala, es decir,  $\geq 1000$  km<sup>2</sup> (Houze, 2004).

## Resultados

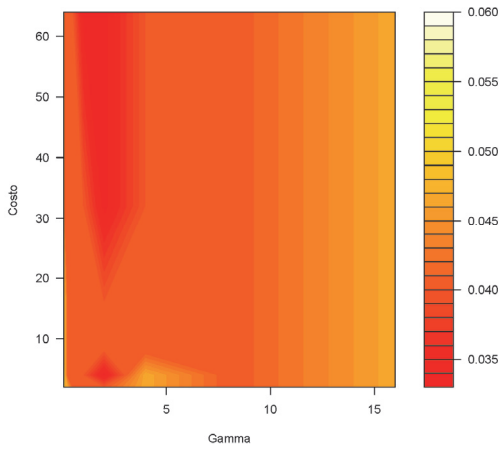
La implementación de “búsqueda mediante retícula” para el algoritmo SVM permitió obtener como parámetros óptimos del modelo de clasificación los valores de  $C = 4$

y  $\gamma = 2$ , para los cuales se minimiza el error como se muestra en la Figura 5a, donde las tonalidades rojas indican los parámetros para los cuales el error del modelo de clasificación disminuye, mientras que las tonalidades amarillas indican parámetros para los cuales el error del modelo de clasificación aumenta.

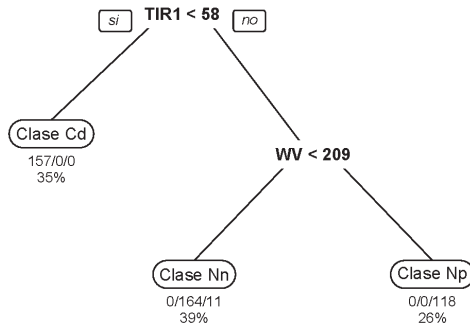
Mediante la implementación del algoritmo DT se obtuvo el árbol de reglas presentado en Figura 5b, donde se indica la proporción de la muestra de entrenamiento que fue asignada a cada clase objetivo. Este árbol exhibe una estructura muy simple que consta tan solo de dos particiones, seleccionando los canales TIR1 (10.8 $\mu$ m) y WV (6.5 $\mu$ m) como las que aportan mayor información en el modelo de clasificación. El canal IR permite discriminar entre cielo despejado (Cd) y nubes (N), hecho que está asociado a que las nubes son los objetos más fríos en este canal (Hillger y Ellrod, 2003), y la discriminación entre las nubes que no precipitan (Nn) y las nubes que precipitan (Np), está asociada a las altas concentraciones de gotículas presentes en el canal de WV (Ferreira, 2006). Las reglas del modelo de clasificación DT son las siguientes: (i) si los ND del canal TIR1 son  $< 58$  entonces es la clase es Cd; (ii) si los ND del canal TIR1 son  $\geq 58$  y los ND del canal WV son  $< 209$  entonces la clase es Nn; y (iii) si los ND del canal TIR1 son  $\geq 58$  y los ND del canal WV son  $\geq 209$  entonces la clase es Np.

En la Figura 5c, se muestra la contribución de cada canal en el modelo de clasificación en función de la medida de importancia Disminución Media de Gini, obtenida a partir de la implementación del algoritmo RF, evidenciando similitud entre los canales más importantes elegidos por este modelo de clasificación y los canales elegidos por el modelo de clasificación DT. Siendo el canal más importante el TIR1 (10.8 $\mu$ m) con una contribución al modelo de clasificación del 98.7%, seguido por el TIR2 (13.3 $\mu$ m) con una contribución del 85.1%, sin embargo, al ser canales similares con una diferencia 2.5 $\mu$ m en su longitud de onda media, existe una diferencia de su contribución de apenas 13.6%, lo cual, sugiere que quizás su aporte sea redundante y por lo tanto no fue elegido por DT. Mientras que el canal WV (6.5 $\mu$ m), elegido por DT, contribuye en un 66%. En contraste los canales VIS (0.63 $\mu$ m) y MIR (3.9 $\mu$ m) son los que menos contribuyen con el 39.1% y 8.9% respectivamente, la escasa contribución del canal MIR quizás se deba a su baja emisividad, lo que a temperaturas de brillo muy frías genera ruido.

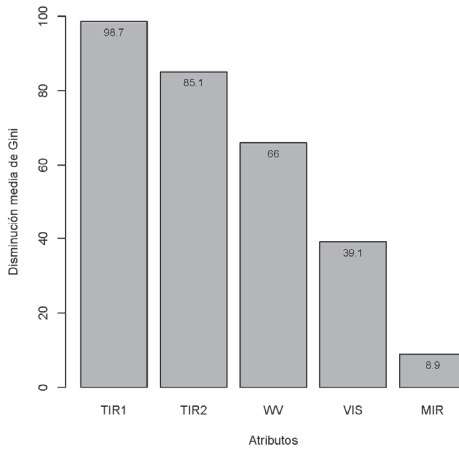
Al observar las salidas gráficas de los SCM detectados por cada uno de los métodos aplicados, visualmente se evidencia que existe coincidencia en la localización de las áreas de precipitación. Sin embargo, para los cuatro instantes de validación presentados en la Figura 5 se puede observar de manera visual que el método SVM permite obtener áreas de precipitación cuyo tamaño es más cercano a los SCM de referencia que las áreas obtenidas mediante los métodos DT y RF. Por otro lado, aunque algunas de las formas del método RF son similares a las de los SCM de referencia, en particular en la región central del país, en general la similitud de forma no es sobresaliente para ninguno de los métodos.



a) “búsqueda mediante retícula” algoritmo SVM



b) Árbol de decisión obtenido para el algoritmo DT



c) Importancia atributos para el algoritmo RF

Figura 5. Salidas de los modelos de clasificación para los algoritmos ML implementados.



Los resultados de evaluación de exactitud mediante el método STEP, consignados en la Tabla 3, muestran que los índices que exhiben mayor exactitud en su respectivo orden son: *Shape*, *Theme*, *Position* y *Edge*; siendo *Edge* la métrica más problemática en la detección de SCM. Mientras que la exactitud de los métodos, en orden descendente, fue DT, RF, TB y SVM. Para la métrica *Shape*, el método DT presenta variaciones en el intervalo de validación comprendidas entre el 89% y el 98%, el método RF entre el 87% y el 95%, el método TB entre el 83% y 95% y el método SVM entre el 86% y 93%. Para la métrica *Theme*, el método DT presenta variaciones en el intervalo de validación comprendidas entre el 75% y el 94%, el método RF entre el 73% y el 89%, el método TB entre el 68% y 87% y el método SVM entre el 72% y 83%. Para la métrica *Edge*, el método DT presenta variaciones en el intervalo de validación comprendidas entre el 68% y el 96 %, el método RF entre el 67% y el 87%, el método TB entre el 60% y 81% y el método SVM entre el 65% y 78%. Para la métrica *Position*, el método DT presenta variaciones en el intervalo de validación comprendidas entre el 80% y el 98%, el método RF entre el 78% y el 90%, el método TB entre el 72% y 91% y el método SVM entre el 76% y 81%.

**Tabla 3**  
PCC global por métrica de similitud STEP

<i>Shape</i>	12:15	15:15	18:15	21:15	<i>Theme</i>	12:15	15:15	18:15	21:15
SVM	93%	93%	90%	86%	SVM	83%	83%	77%	72%
DT	97%	98%	94%	89%	DT	94%	94%	85%	75%
RF	95%	95%	92%	87%	RF	89%	89%	82%	73%
IR-TB	95%	94%	91%	83%	TB	87%	86%	78%	68%
<i>Edge</i>	12:15	15:15	18:15	21:15	<i>Position</i>	12:15	15:15	18:15	21:15
SVM	78%	74%	70%	65%	SVM	87%	88%	81%	76%
DT	96%	91%	78%	68%	DT	98%	97%	89%	80%
RF	87%	87%	76%	67%	RF	90%	93%	86%	78%
IR-TB	81%	77%	70%	60%	IR-TB	91%	91%	82%	72%

En la Tabla 4 se presentan los intervalos de confianza calculados al 95% de confiabilidad para cada métrica de similitud y para la exactitud general (PCC). Sin embargo, como puede observarse en esa tabla, no es posible establecer de manera categórica un mejor desempeño de alguno de los métodos aplicados puesto que los intervalos de exactitud presentan solapamiento entre sí. No obstante, se observa un comportamiento ligeramente superior de los métodos DT e IR-TB respecto a los métodos SVM y RF.

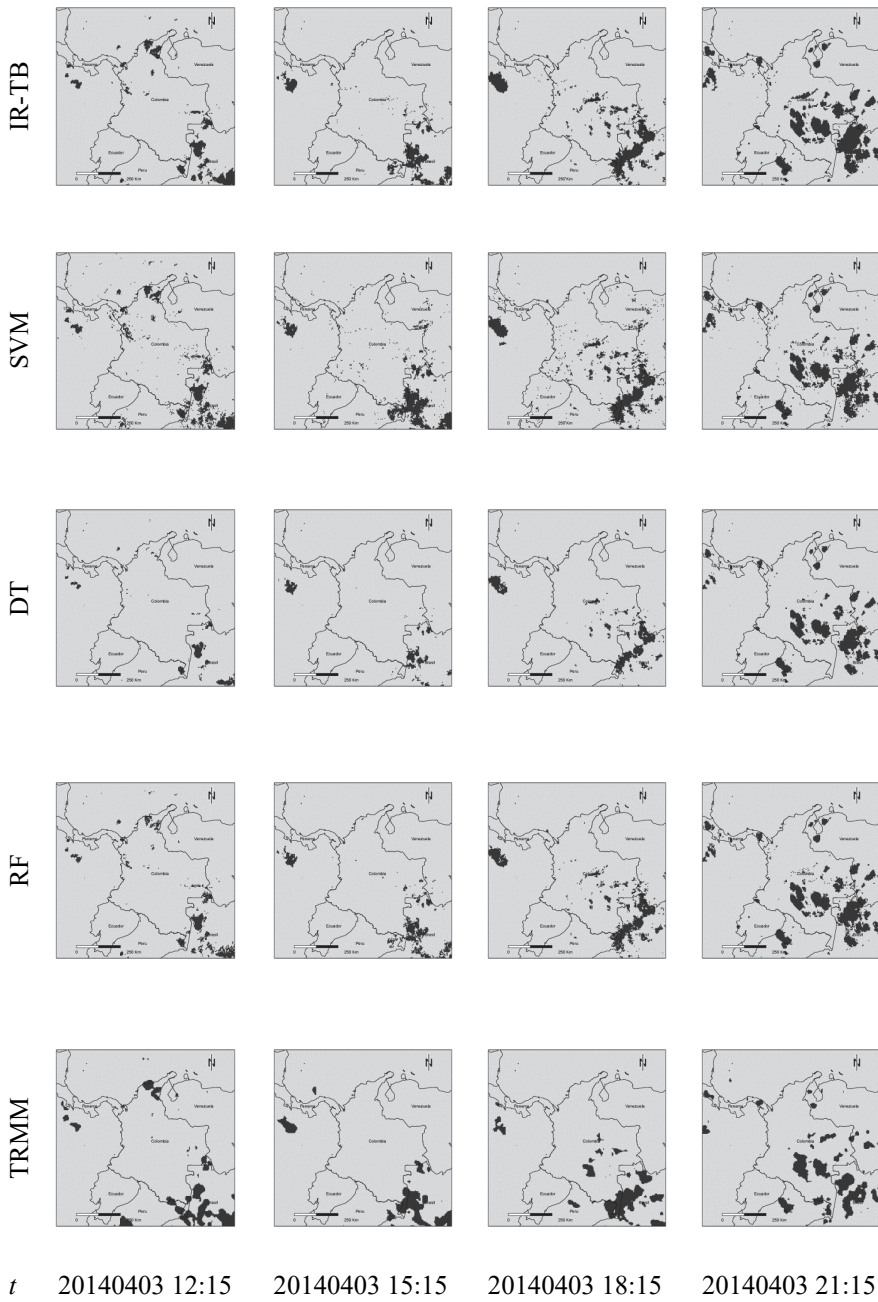


Figura 5. SCM detectados sobre una serie de imágenes meteorológicas.

**Tabla 4**  
**PCC global por métrica de similitud STEP**

<i>Shape</i>	<i>12:15</i>	<i>15:15</i>	<i>18:15</i>	<i>21:15</i>	<i>Theme</i>	<i>12:15</i>	<i>15:15</i>	<i>18:15</i>
SVM	93%	93%	90%	86%	SVM	83%	83%	77%
DT	97%	98%	94%	89%	DT	94%	94%	85%
RF	95%	95%	92%	87%	RF	89%	89%	82%
IR-TB	95%	94%	91%	83%	TB	87%	86%	78%
<i>Edge</i>	<i>12:15</i>	<i>15:15</i>	<i>18:15</i>	<i>21:15</i>	<i>Position</i>	<i>12:15</i>	<i>15:15</i>	<i>18:15</i>
SVM	78%	74%	70%	65%	SVM	87%	88%	81%
DT	96%	91%	78%	68%	DT	98%	97%	89%
RF	87%	87%	76%	67%	RF	90%	93%	86%
IR-TB	81%	77%	70%	60%	IR-TB	91%	91%	82%

En términos de eficiencia computacional, el método propuesto no es muy exigente. Usando un computador Intel® Core™ i5 de 64 bits a 1,60 GHz, con 6 Gb de RAM y sistema operativo Windows 8, el tiempo de cómputo para los algoritmos de ML fue de 235 segundos para SVM, 81 segundos para DT y 248 segundos para RF. Sin embargo, estos tiempos no son competitivos con el algoritmo tradicional IR-TB que tardó tan sólo 24 segundos para realizar la tarea de detección.

## Conclusiones

Se cumplió el objetivo principal de este trabajo, estableciendo que: (i) el uso de cinco canales espectrales meteorológicos, en lugar de un único canal de IR, mejora la detección de SCM; y (ii) el uso de los métodos no paramétricos, es decir, métodos no basados en supuestos estadísticos sobre la distribución de datos, no garantizan siempre ser más eficientes que el algoritmo tradicional de IR-TB. Sin embargo, el algoritmo DT permite obtener mejores exactitudes que el algoritmo IR-TB para detectar SCM.

En términos generales los cuatro métodos aplicados en este trabajo exhiben una efectividad favorable, para la extracción de SCM a partir de imágenes meteorológicas con respecto a la similitud y exactitud de los cuatro aspectos inherentes a un objeto: *shape, theme, edge, position*. Sin embargo, se evidencia la presencia de errores tanto por omisión como por comisión en cada uno de los métodos. Dichos errores sugieren que existe algún grado de dificultad por parte de los métodos empleados, en extraer objetos precipitantes preservando sus características temáticas y geométricas. Al respecto se considera, que los resultados obtenidos mediante los métodos no paramétricos dependen en gran medida de la optimización o ajuste de los modelos empleados por cada uno de ellos, lo cual permite obtener resultados más confiables.

De acuerdo con los resultados de este estudio, los algoritmos de aprendizaje máquina muestran ligeras diferencias en la clasificación exactitud de SCM en comparación con el método de IR-BT. Si bien se obtuvieron mejores exactitudes utilizando el algoritmo DT, no se encontraron diferencias estadísticamente significativas en las métricas de exactitud entre DT y el método tradicional IR-TB. Esto sugiere la necesidad de realizar pruebas adicionales para DT y IR-TB en otras fechas y horas con el fin de confirmar si tales diferencias se mantienen o no significativamente.

Los métodos no convencionales de aprendizaje máquina DT y RF, exhiben resultados favorables en la extracción de SCM, frente a los obtenidos para el método convencional IR-TB. Por otra parte, los resultados obtenidos por el método no convencional SVM, no sugieren favorabilidad para su uso en aplicaciones prácticas.

Desde el punto de vista práctico, la contribución científica de este trabajo tiene que ver con el diseño e implementación de un nuevo método para extraer SCM a partir de imágenes meteorológicas, lo cual es el primer paso hacia la cuantificación y

la predicción de precipitación integrando técnicas de teledetección y aprendizaje de máquina.

Los trabajos futuros de los autores de este documento estarán orientados a investigar si el modelo de clasificación obtenido por el algoritmo DT a partir de GOES-13 puede aplicarse en diferentes fechas y permite detectar sistemas convectivos de mesoescala sin realizar interpretación visual en cada nueva imagen. Esta eventual transferibilidad de los criterios de decisión de DT sería muy útil para aplicaciones operativas bajo el enfoque presentado en este artículo.

## Bibliografía

- Arnaud, Y.; Desbois, M. y Maizi, J. (1992). "Automatic Tracking and Characterization of African Convective Systems on Meteosat Pictures", *Journal of Applied Meteorology*, 31(5):443-453.
- Ba, M.B. y Gruber, A. (2001). "GOES Multispectral Rainfall Algorithm (GMSRA)", *Journal of Applied Meteorology*, 40(8):1500-1514.
- Bader, M.J.; Forbes, G.S.; Grant, J.R.; Lilley, R.B.E. y Waters, A.J. (1996). "A review of: Images in Weather Forecasting—a practical guide for interpreting satellite and radar imagery", *International Journal of Remote Sensing*, 17(16):3292.
- Boulesteix, A.L.; Janitza, S.; Kruppa, J. y König, I.R. (2012). "Overview of Random Forest Methodology and Practical Guidance with Emphasis on Computational Biology and Bioinformatics", *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2:493-507.
- Carvalho, L. y Jones, C. (2001). "A satellite method to identify structural properties of mesoscale convective systems based on the maximum spatial correlation tracking technique (MASCOTTE)", *Journal of Applied Meteorology*, 40:1683-1701.
- Chen, S. y Houze, R. (1997). "Diurnal variation and life-cycle of deep convective systems over the tropical pacific warm pool", *Quarterly Journal of the Royal Meteorological Society*, 123(538):357-388.
- Chen, S., Houze, R. y Mapes, B. (1996). "Multiscale Variability of Deep Convection In Relation to Large-Scale Circulation in TOGA COARE", *Journal of the Atmospheric Sciences*, 53(10):1380-1409.
- Colombia Humanitaria (2010). "Reporte diario ola invernal 2010-2011", recuperado 25 de septiembre de 2013, a partir de <<http://www.colombiahumanitaria.gov.co/Cifras/Paginas/Cifras2011.aspx>>.
- Congalton, R.G. y Green, K. (2008). *Assessing the accuracy of remotely sensed data: principles and practices* (Second), New York, CRC Press, USA.
- Desbois, M.; Seze, G. y Szejwach, G. (1982). "Automatic Classification of Clouds on METEOSAT Imagery: Application to High-Level Clouds", *Journal of Applied Meteorology*, 21(3):401-412.
- Doraiswamy, H.; Natarajan, V. y Nanjundiah, R.S. (2013). "An exploration framework to identify and track movement of cloud systems", *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2896-2905.

- Ferreira, A.G. (2006). *Meteorología práctica*, Oficina de Textos, São Paulo.
- Fiolleau, T. y Roca, R. (2013). "An Algorithm for the Detection and Tracking of Tropical Mesoscale Convective Systems Using Infrared Images From Geostationary Satellite", *IEEE Transactions on Geoscience and Remote Sensing*, 51(7):4302-4315.
- Goswami, B. y Bhandari, G. (2011). "Automatically adjusting cloud movement prediction model from satellite infrared images", en *India Conference (INDICON), 2011 Annual IEEE* (pp. 1-4), <<http://doi.org/10.1109/INDCON.2011.6139604>>.
- (2012a). "Convective Cloud Detection and Tracking from Series of Infrared Images", *Journal of the Indian Society of Remote Sensing*, 41(2), 291-299.
- (2012b). "Temperature induced mean based cloud motion prediction model from satellite infrared images", *India Conference (INDICON), Annual IEEE* (pp. 719-723), Springer, Verlag.
- Goyens, C.; Lauwaet, D.; Schröder, M.; Demuzere, M. y Van Lipzig, N.P.M. (2012). "Tracking mesoscale convective systems in the Sahel: relation between cloud parameters and precipitation", *International Journal of Climatology*, 32(12), 1921-1934.
- Hastie, T.; Tibshirani, R. y Friedman, J. (2009). *The elements of statistical learning: Data Mining, Inference and Prediction* (Second). Stanford, California, Springer.
- Hillger, D. y Ellrod, G. (2003). "Detection of important atmospheric and surface features by employing principal component image transformation of GOES imagery", *Journal of Applied Meteorology*, pp. 611-629.
- Hodges, K.I. y Thorncroft, C.D. (1997). "Distribution and Statistics of African Mesoscale Convective Weather Systems Based on the ISCCP Meteosat Imagery", *Monthly Weather Review*, 125(11):2821-2837.
- Houze, R. (2004). "Mesoscale convective systems", *Reviews of Geophysics*, 42:1-43.
- Klitch, M.; Weaver, J.; Kelly, F. y Vonder Haar, T. (1985). "Convective Cloud Climatologies Constructed from Satellite Imagery", *Monthly Weather Review*, 113(3):326-337.
- Lary, D.J.; Alavi, A.H.; Gandomi, A.H. y Walker, A.L. (2016). "Machine learning in geosciences and remote sensing", *Geoscience Frontiers*, 7(1):3-10, <<http://doi.org/10.1016/j.gsf.2015.07.003>>.
- Laurent, H.; D'Amato, N.; y Lebel, T. (1998). "How important is the contribution of the mesoscale convective complexes to the Sahelian rainfall?", *Physics and Chemistry of the Earth*, 23(5-6):629-633.
- Laurent, H.; Machado, L.A.T., Morales, C.A. y Durieux, L. (2002). "Characteristics of the Amazonian mesoscale convective systems observed from satellite and radar during the WETAMC/LBA experiment", *Journal of Geophysical Research*, 2(1):1-43.
- Liu, Y.; Xi, D.; Li, Z. y Shi, C. (2014). "Automatic Tracking and Characterization of Cumulonimbus Clouds from FY-2C Geostationary Meteorological Satellite Images", *Advances in Meteorology*, 2014, <<http://doi.org/10.1155/2014/478419>>.

- Lizarazo, I. (2008). "Clasificación de la cobertura y del uso del suelo urbano usando imágenes de satélite y algoritmos supervisados de inteligencia", *UD y la GEOMÁTICA*, 2:4-18.
- (2010). "Fuzzy image regions for quantitative land cover analysis", *GEOBIA 2010, Geographic Object-Based Image Analysis*, p. 15, Ghent, Belgium.
- (2014). "Accuracy assessment of object-based image classification: another STEP", *International Journal of Remote Sensing*, 35(16):6135-6156.
- Lizarazo, I. y Barros, J. (2010). "Fuzzy image segmentation for urban land-cover classification", *Photogrammetric Engineering and Remote Sensing*, 76(2):151-162.
- Lizarazo, I. y Elsner, P. (2009). "Fuzzy segmentation for object based image classification", *International Journal of Remote Sensing*, 30(6):1643-1649.
- Machado, L. (2002). "Diurnal march of the convection observed during TRMM-WETAMC/LBA", *Journal of Geophysical Research*, 107(D20), 8064.
- Machado, L.A.T.; Rossow, W.B.; Guedes, R.L. y Walker, A.W. (1998). "Life Cycle Variations of Mesoscale Convective Systems over the Americas", *Monthly Weather Review*, 126(6):1630-1654.
- Morel, C. y Senesi, S. (2002). "A climatology of mesoscale convective systems over Europe using satellite infrared imagery. I", *Methodology. Quarterly Journal of the Royal Meteorological Society*, 128(584):1953-1971.
- Peak, J. y Tag P.M. (1994). "Segmentation of satellite imagery using hierarchical thresholding and neural networks", *Journal of Applied Meteorology*, 33:605-616.
- Ravelo, A.C. y Santa, J.A. (2000). "Estimación de las precipitaciones utilizando información satelital y terrestre en la provincia de Córdoba (Argentina)", *Agriscientia*, XVII:21-27.
- Sakamoto, M.; Ambrizzi, T. y Poveda, G. (2011). "Moisture sources and life cycle of convective systems over Western Colombia", *Advances in Meteorology*, 2011:1-11.
- Shukla, B.P. y Pal, P.K. (2012). "A Source Apportionment Approach to Study the Evolution of Convective Cells: An Application to the Nowcasting of Convective Weather Systems", *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, 5(1):242-247, <<http://doi.org/10.1109/JSTARS.2011.2170661>>.
- Sieglaflaff, J.M.J.; Cronce, L.L.M.; Feltz, W.F.; Bedka, K.M.; Pavolonis, M.J. y Heidinger, A.K. (2011). "Nowcasting Convective Storm Initiation Using Satellite-Based Box-Averaged Cloud-Top Cooling and Cloud-Type Trends", *Journal of Applied Meteorology and Climatology*, 50(1):110-126.
- Suvichakorn, A. (2007). *The use of the texture and motion of clouds from geostationary satellite images in rain rate estimation and prediction*. University of Southampton.
- Tan, P.; Steinbach, M. y Kumar, V. (2005). "Classification: Basic Concepts, Decision Trees, and Model Evaluation", *Introduction to Data Mining* (1st ed., pp. 145-205), Addison-Wesley.

- Tso, B. y Mather, P. (2009). *Classification methods for remotely sensed data* (2nd ed.). New York, US: CRC Press.
- Tsonis, A.A. (1984). "On the separability of various classes from the GOES visible and infrared data", *Journal of Climate and Applied Meteorology*, 23(10):1393-1410.
- Tsonis, A.A. y Isaac, G.A. (1985). "On a New Approach for Instantaneous Rain Area Delineation in the Midlatitudes Using GOES Data", *Journal of Climate and Applied Meteorology*, 24(11):1208-1218.
- Vapnik, V. (2000). *The Nature of Statistical Learning Theory* (Second), Springer, Red Bank, New Jersey, USA.
- Vila, D.; Machado, L.A.T.; Laurent, H. y Velasco, I. (2008). "Forecast and Tracking the Evolution of Cloud Clusters (ForTraCC) Using Satellite Infrared Imagery: Methodology and Validation", *Weather and Forecasting*, 23(2):233-245.
- Williams, M. y Houze, R. (1987). "Satellite-observed characteristics of winter monsoon cloud clusters", *Monthly weather review*, 115:505-519.
- Woodley, W.; Griffith, C.; Griffin, J. y Stromatt, S. (1980). "The Inference of GATE Convective Rainfall from SMS-1 Imagery", *Journal of Applied Meteorology*, 19(4):388-408.
- Xu, W.; Wooster, M.J.; Roberts, G. y Freeborn, P. (2010). "New GOES imager algorithms for cloud and active fire detection and fire radiative power assessment across North, South and Central America", *Remote Sensing of Environment*, 114(9):1876-1895.