

**OBTENCIÓN DE CARACTERÍSTICAS BIOFÍSICAS PARA LA CLASIFICACIÓN  
DE COBERTURAS MEDIANTE EL ANÁLISIS TEMPORAL DE IMÁGENES SAR  
SOBRE LA CUENCA ALTA DEL RÍO CAUCA**



**Universidad  
del Cauca**

**CRISTIAN CAMILO MUÑOZ ORDÓÑEZ**

**UNIVERSIDAD DEL CAUCA  
FACULTAD DE CIENCIAS NATURALES, EXACTAS Y DE LA EDUCACIÓN  
DEPARTAMENTO DE FÍSICA  
INGENIERÍA FÍSICA  
POPAYÁN  
2018**

**OBTENCIÓN DE CARACTERÍSTICAS BIOFÍSICAS PARA LA CLASIFICACIÓN  
DE COBERTURAS MEDIANTE EL ANÁLISIS TEMPORAL DE IMÁGENES SAR  
SOBRE LA CUENCA ALTA DEL RÍO CAUCA**

**Trabajo de grado presentado como requisito parcial para optar al título de  
Ingeniero Físico**

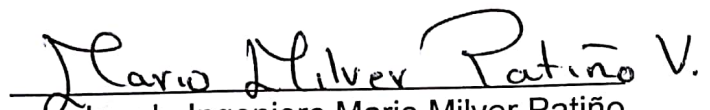
**Director**

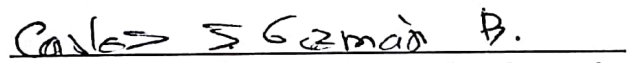
**Ing. Edgar Leonairo Pencue Fierro**


**UNIVERSIDAD DEL CAUCA  
FACULTAD DE CIENCIAS NATURALES, EXACTAS Y DE LA EDUCACIÓN  
DEPARTAMENTO DE FÍSICA  
INGENIERÍA FÍSICA  
POPAYÁN  
2018**

## NOTA DE ACEPTACIÓN

El Director y los Jurados han evaluado este documento, y escuchado la sustentación por el estudiante y lo encuentran satisfactorio, por lo cual autorizan al estudiante para que desarrolle las gestiones pertinentes para optar al título de Ingeniero Físico.

  
Jurado Ingeniero Mario Milver Patiño

  
Jurado Ingeniero Carlos Javier Guzmán

  
Director Ingeniero Leonairo Pencue Fierro

## CONTENIDO

	pág.
<b>1. INTRODUCCIÓN Y OBJETIVOS</b> .....	8
<b>2. MARCO TEÓRICO</b> .....	11
<b>2.1. Tipos de satélites</b> .....	11
2.1.1. Satélites de sensores Pasivos.....	11
2.1.2. Satélites de sensores Activos .....	12
<b>2.2. Principios básicos de Radar</b> .....	12
2.2.1. Ecuación de Radar .....	12
2.2.2. Efectos de la naturaleza coherente de la señal de Radar .....	13
2.2.3. Proceso de apertura sintética.....	14
<b>2.3. Satélite de radar Sentinel-1</b> .....	16
2.3.1. Distorsiones geométricas para el satélite Sentinel-1 .....	17
<b>2.4. Pre-procesamiento de imágenes Sentinel-1</b> .....	19
2.4.1. Corrección de orbita.....	19
2.4.2. Calibración .....	20
2.4.3. Deburst .....	20
2.4.4. Multilooking.....	21
2.4.5. Corrección de terreno por rango Doppler .....	21
<b>2.5. Obtención de características de datos Sentinel-1</b> .....	22
<b>2.6. Algoritmos de aprendizaje máquina: Bosque aleatorio</b> .....	35
<b>3. METODOLOGÍA</b> .....	37
3.1. Descarga de imágenes Sentinel-1 .....	37
3.2. Pre-procesamiento imágenes Sentinel-1.....	40
3.3. Extracción de características biofísicas.....	45
3.4. Clasificación de la CARC .....	47
3.5. Análisis temporal y espacial desde las características Sentinel-1 .....	56
<b>4. ANÁLISIS Y RESULTADOS</b> .....	57
<b>5. CONCLUSIONES Y TRABAJOS A FUTURO</b> .....	91
<b>6. BIBLIOGRAFÍA</b> .....	93

## LISTA DE FIGURAS

	pág.
<b>Figura 1.</b> Geometría de Radar mono-estática (Fuente: Tesis doctoral Adamo Ferro) .....	13
<b>Figura 2.</b> Efecto speckle, suma coherente de señales dispersadas por objetivos individuales (Fuente: Tesis doctoral Daniel C. Gallego Cabo) .....	14
<b>Figura 3.</b> La longitud de apertura sintética $L_s$ surge de iluminar durante un tiempo $t_i$ un objetivo puntual (Fuente: vialattea.net) .....	15
<b>Figura 4.</b> Distorsiones geométricas. a)foreshortening, b)desplazamiento por relieve, c)Sobreposición, d)sombras (Fuente: Tesis doctoral Adamo Ferro).....	18
<b>Figura 5.</b> Filtro de promedio basado en imágenes Sentinel-1 de la cuenca del Río Guachal .....	24
<b>Figura 6.</b> Filtro de mediana basado en imágenes Sentinel-1 de la cuenca del Río Molino.....	24
<b>Figura 7.</b> Filtro de desviación estándar basado en imágenes Sentinel-1 de la cuenca del Río Mondomo .....	25
<b>Figura 8.</b> Filtro Boxcar basado en imágenes Sentinel-1 en la cuenca del Río Palacé .....	26
<b>Figura 9.</b> Filtro Lee basado en imágenes Sentinel-1 en la cuenca del Río Palacé	27
<b>Figura 10.</b> Filtro Lee mejorado basado de imágenes Sentinel-1 en la cuenca del Río Palo .....	28
<b>Figura 11.</b> IDAN basado en imágenes Sentinel-1 en la cuenca del Río Quinamayó	29
<b>Figura 12.</b> Filtro Frost basado en imágenes Sentinel-1 en la cuenca del Río Molino.....	30
<b>Figura 13.</b> Filtro Gamma Map en la cuenca del Río Molino.....	31
<b>Figura 14.</b> Angular Second Moment (ASM) en la cuenca del Río Molino .....	33
<b>Figura 15.</b> Textura VA en la Cuenca del Río Guachal.....	34
<b>Figura 16.</b> Textura VL en la cuenca del Río Guachal .....	35
<b>Figura 17.</b> Mapa de modos, polarización y forma de observación de las imágenes obtenidas por el satélite Sentinel-1 (Fuente: Agencia Espacial Europea) .....	38
<b>Figura 18.</b> Herramienta de corrección de órbita para imágenes Sentinel-1 .....	42
<b>Figura 19.</b> Corrección de terreno sobre imágenes Sentinel-1 .....	43
<b>Figura 20.</b> Imagen Sentinel-1 con polarización VV de la Cuenca Alta del Río Cauca.....	44
<b>Figura 21.</b> Cuerpos de agua seleccionados en parte de la CARC .....	48
<b>Figura 22.</b> Zonas urbanas seleccionadas en parte de la CARC .....	49
<b>Figura 23.</b> Cultivos seleccionados en parte de la CARC .....	49
<b>Figura 24.</b> Pastizales seleccionados en parte de la CARC .....	50
<b>Figura 25.</b> Bosques seleccionados en parte de la CARC .....	50
<b>Figura 26.</b> Suelos desnudos seleccionados en parte de la CARC .....	51
<b>Figura 27.</b> Nubes seleccionadas en parte de la CARC .....	51
<b>Figura 28.</b> Sombras de nubes seleccionadas en parte de la CARC.....	52
<b>Figura 29.</b> Máscaras de coberturas extraídas sobre la sub-cuenca de Quinamayó	53

<b>Figura 30.</b> Clasificación de coberturas para seis sub-cuencas de la CARC.....	66
<b>Figura 31.</b> Clasificación de coberturas en la Cuenca del Río Guachal.....	66
<b>Figura 32.</b> Suelo desnudo, cultivos y bosques naturales.....	68
<b>Figura 33.</b> Zonas urbanas, suelo desnudo y cultivos en la zona de la Cuenca del río Guachal.....	69
<b>Figura 34.</b> Mapa de pendientes en la Cuenca del río Guachal.....	70
<b>Figura 35.</b> Clasificación de coberturas en la Cuenca del Río Molino.....	72
<b>Figura 36.</b> Clasificación de coberturas en la Cuenca del Río Mondomo .....	73
<b>Figura 37.</b> Clasificación de coberturas en la Cuenca del Río Palacé .....	74
<b>Figura 38.</b> Clasificación de coberturas en la Cuenca del Río Palo .....	75
<b>Figura 39.</b> Clasificación de coberturas en la Cuenca del Río Quinamayó .....	76
<b>Figura 40.</b> Valores de precipitación mensual y característica desviación estándar a lo largo de los meses indicados .....	79
<b>Figura 41.</b> Valores de precipitación mensual y característica IDAN a lo largo de los meses indicados .....	80
<b>Figura 42.</b> Valores de precipitación mensual y característica textura VL a lo largo de los meses indicados .....	81
<b>Figura 43.</b> Valores de precipitación mensual y característica Contraste a lo largo de los meses indicados .....	81
<b>Figura 44.</b> Valores de retrodispersión sobre la cobertura cultivos ubicada en la cuenca del río Guachal, para la característica Boxcar utilizando un perfil de línea.....	82
<b>Figura 45.</b> Valores de retrodispersión sobre la cobertura pastizales ubicada en la cuenca del río Guachal, para la característica Boxcar utilizando un perfil de línea.....	83
<b>Figura 46.</b> Valores de retrodispersión sobre la cobertura pastizales ubicada en la cuenca del río Guachal, para la característica textura VI utilizando un perfil de línea. ....	83
<b>Figura 47.</b> Valores de retrodispersión sobre la cobertura zonas urbanas ubicada en la cuenca del río Guachal, para la característica Entropía utilizando un perfil de línea. ....	84
<b>Figura 48.</b> Valores de intensidad promedio de la característica de radar Boxcar para cinco coberturas biofísicas .....	85
<b>Figura 49.</b> Valores de intensidad promedio de la característica de radar textura VI para cinco coberturas biofísicas .....	86
<b>Figura 50.</b> Valores de intensidad promedio de la característica de radar ASM para cinco coberturas biofísicas .....	86
<b>Figura 51.</b> Valores de intensidad promedio de la característica de radar Entropía para cinco coberturas biofísicas .....	87
<b>Figura 52.</b> Valores de intensidad promedio de la característica de radar MAX para cinco coberturas biofísicas .....	87
<b>Figura 53.</b> Imagen Falso color de la cuenca del río Palacé.....	89
<b>Figura 54.</b> Mascara de sombras de radar presentes sobre la cuenca del río Palacé .....	89

## LISTA DE TABLAS

	pág.
<b>Tabla 1.</b> Fecha exacta y representativa de las imágenes Sentinel-1 descargadas	39
<b>Tabla 2.</b> Código de colores para las diferentes máscaras de coberturas .....	54
<b>Tabla 3.</b> Recortes obtenidos de cinco coberturas para la cuenca del río Guachal.	56
<b>Tabla 4.</b> Caracterización de las coberturas clasificadas .....	57
<b>Tabla 5.</b> Datos de precisión y sensibilidad (recall) para cada cobertura usando características ópticas y de radar.....	59
<b>Tabla 6.</b> Matriz de confusión para todas las características inicialmente construidas.....	59
<b>Tabla 7.</b> Datos de precisión y sensibilidad (recall) usando selección de atributos.	60
<b>Tabla 8.</b> Matriz de confusión para el conjunto de datos con selección de atributos	61
<b>Tabla 9.</b> Caracterización de las coberturas clasificadas para la base de datos incrementada.....	62
<b>Tabla 10.</b> Datos de precisión y sensibilidad (recall) usando la base de datos incrementada.....	63
<b>Tabla 11.</b> Matriz de confusión para la base datos enriquecida e igualada en registros.....	63
<b>Tabla 12.</b> Datos estadísticos para las diferentes bases de datos (óptica, radar, radar-óptica con y sin selección de atributos) .....	64
<b>Tabla 13.</b> Área de ocupación de las coberturas sobre la cuenca del río Guachal	67
<b>Tabla 14.</b> Área de ocupación de las zonas urbanas seleccionadas sobre la cuenca del río Guachal.....	69
<b>Tabla 15.</b> Área de ocupación de las coberturas sobre la cuenca del río Molino ...	72
<b>Tabla 16.</b> Área de ocupación de las coberturas sobre la cuenca del río Mondomo	73
<b>Tabla 17.</b> Área de ocupación de las coberturas sobre la cuenca del río Palacé...	74
<b>Tabla 18.</b> Área de ocupación de las coberturas sobre la cuenca del río Palo .....	75
<b>Tabla 19.</b> Área de ocupación de las coberturas sobre la cuenca del río Quinamayó.....	76
<b>Tabla 20.</b> Área de ocupación total de las coberturas sobre las seis sub-cuencas del río Cauca.....	77
<b>Tabla 21.</b> Consolidado del valor promedio de retrodispersión para cada tipo de cobertura en las características de radar más representativas .....	87

## 1. INTRODUCCIÓN Y OBJETIVOS

El monitoreo de la evolución temporal y espacial de la tierra es de suma importancia para la administración de los recursos terrestres e hídricos [1], especialmente en países como Colombia, que presenta a lo largo de su territorio una notable biodiversidad agrícola a causa de sus diferentes pisos térmicos [2]. La percepción remota sobre la Cuenca Alta del Río Cauca cobra mayor importancia, ya que abarca casi todas las cadenas de producción priorizadas por el Ministerio de Agricultura de Colombia [3].

La Cuenca Alta del Río Cauca (CARC) abarca gran parte de la agricultura técnica y de alto valor en el país, se extiende desde el Macizo Colombiano hasta los límites de los departamentos de Caldas y Antioquia, entre las cimas de las cordilleras occidental y central. Su altitud varía desde los 5.400 msnm en las nieves perpetuas de la cordillera central y los 950 en el valle aluvial del Cauca. El área de estudio aloja 191 sub-cuencas que tributan al río Cauca, comprende 99 municipios y alberga una población de aproximadamente 7.056.000 habitantes. Hasta el momento es el corredor agrícola e industrial más importante de Colombia donde se desarrolla agricultura tradicional, cultivo de café, de caña de azúcar, frijol, plátano, cítricos, cacao, maíz, aguacate, entre otros [4]. Además, aproximadamente el 20% de la población nacional se beneficia de la oferta ambiental de los recursos de esta cuenca.

Actualmente la cuenca alta del río Cauca presenta un alto deterioro, debido principalmente al uso inadecuado del suelo, deforestación, contaminación hídrica de origen antrópico y desecamiento de ciénagas y humedales, para la expansión de la frontera agrícola especialmente en la zona plana y afectación constante y continua de sus coberturas vegetales, todos estos factores son determinantes para la sostenibilidad ambiental de la Cuenca [5].

Sus condiciones atmosféricas, su cercanía a la zona de convergencia intertropical (ZCIT) donde se une el aire cálido y húmedo de latitudes por encima y por debajo del Ecuador, es un rasgo atmosférico que favorece la distribución del calor y humedad, que es transportado verticalmente hacia arriba por la actividad convectiva



de las tormentas, haciendo que las regiones situadas sobre esta área perciban efectos de precipitación por más de 200 días al año [6]. Esta condición genera alto grado de nubosidad sobre la Cuenca Alta del Río Cauca (CARC) lo cual reduce drásticamente la resolución temporal de imágenes satelitales ópticas dificultando el análisis cuantitativo y cualitativo relacionado con el proceso de evolución de coberturas tales como: cultivos, cuerpos de agua, bosques naturales, páramos, presentes en la cuenca [1], revisando las bases de datos de acceso libre que dispone la **NASA** para la familia de satélites ópticos **Landsat y Modis**, desde Septiembre de 1984 y Febrero del 2000 hasta el presente año respectivamente, se observa que muy pocas escenas están libres de nubosidad y más del 70% de las imágenes presentan coberturas nubosas, lo que genera pérdida de información relevante de la cobertura del terreno dificultando un adecuado y preciso análisis temporal-espacial de la región, condición que demanda el uso de tecnologías alternativas para hacer el seguimiento y monitoreo de cambios de estas coberturas en los ecosistemas de la cuenca.

La nubosidad sobre la Cuenca Alta del Río Cauca reduce drásticamente la resolución temporal de las imágenes satelitales ópticas dificultando el análisis cuantitativo y cualitativo relacionado con el proceso de evolución de las coberturas tales como cultivos, cuerpos de agua, entre otros, presentes en la zona en cuestión, lo expuesto anteriormente genera la siguiente pregunta de investigación: ¿Cómo mejorar los procesos de clasificación sobre la Cuenca Alta del Río Cauca, usando imágenes satelitales que proporcionen datos relevantes para la extracción de características biofísicas, además que no se vean afectadas por coberturas nubosas?.

En este sentido las imágenes capturadas por un radar de apertura sintética trabajan en longitudes de onda superiores a las ópticas, evitando que la onda sea reflejada por las nubes y proporcionando características biofísicas adicionales y complementarias para los sistemas de clasificación. Logrando, de esta forma, mejorar los estudios realizados únicamente con imágenes ópticas, incrementando el con-

junto de herramientas para el desarrollo de estrategias de conservación y mitigación frente a posibles cambios que estén sucediendo sobre la CARC.

Para llevar a cabo el presente trabajo de grado se traza el siguiente objetivo general: Diseñar un sistema para la obtención de características biofísicas que permita la clasificación de coberturas mediante el análisis temporal de imágenes SAR sobre la Cuenca Alta del Río Cauca. Los respectivos objetivos específicos son los siguientes: en primer lugar, construir una base de datos conformada por imágenes SAR que contengan la información adecuada para la extracción de características en la región comprendida por la Cuenca Alta del Río Cauca; en segundo lugar, caracterizar la información obtenida para la extracción de muestras biofísicas que permitan construir un modelo de clasificación de coberturas; en tercer lugar, construir un sistema basado en aprendizaje máquina que permita obtener una clasificación de la región estudiada y finalmente sintonizar y validar la información obtenida con puntos foto-interpretados y estudios previos relacionados a procesos de clasificación sobre la zona.

El orden de este trabajo está pactado de la siguiente manera: en la sección 2 se presenta el marco teórico desarrollado para una mejor interpretación de las temáticas involucradas en el actual trabajo de investigación. La Sección 3 enseña la metodología a seguir con la finalidad de extraer la información necesaria para la culminación de este proyecto de grado. La sección 4 muestra los resultados experimentales obtenidos sobre el conjunto de datos con los valores espectrales de imágenes Landsat 8 y Sentinel-1, exponiendo los resultados de las medidas de contraste obtenidas entre cada cobertura de la CARC a través de una reconstrucción y previa inspección visual de la clasificación obtenida, también se realiza un análisis temporal que correlaciona los valores de precipitación mensuales registrados para los años 2015 y 2016 con los valores de retrodispersión de las características de radar sobre un cultivo de caña y finalmente un análisis espacial es efectuado sobre algunas de las coberturas detectadas en esta investigación. Para culminar en la sección 6 se exponen algunas conclusiones y sugerencias para trabajos futuros.

## **2. MARCO TEÓRICO**

La obtención de características biofísicas para la clasificación de coberturas sobre la Cuenca Alta del Río Cauca se puede realizar mediante el análisis temporal de imágenes satelitales de radar obtenidas por medio del Sentinel 1, una aproximación de la base teórica que se tendrá en cuenta para realizar el proyecto se describe a continuación.

### **2.1. Tipos de satélites**

La percepción remota hace referencia, generalmente, al uso de tecnologías de sensores aéreos en la detección o clasificación de objetos en la Tierra (en la superficie, en la atmósfera y los océanos) a través del análisis de señales propagadas como la radiación electromagnética. La técnica de percepción remota permite adquirir información espacial, temporal y espectral de un objeto o escena sin entrar en contacto físico con ella. El sensado remoto juega un papel importante en la clasificación de coberturas terrestres permitiendo generar: estudios en las dinámicas de coberturas, análisis del uso de suelos, observación de los cambios en las coberturas vegetales y monitoreo y evaluación del recurso hídrico. La principal base de datos de acceso libre está constituida por imágenes de satelitales pasivos y activos tales como: Landsat, Modis, Sentinel 2 para la primera clase y Sentinel 1 para el segundo grupo, entre otros.

#### **2.1.1. Satélites de sensores Pasivos**

Aquellos que registran la radiancia reflejada o emitida por la superficie terrestre, llevan a cabo el "efecto pullover" donde el sol envía su radiación electromagnética a la Tierra y las interacciones de estas radiaciones con las partículas terrestres son recogidas por el satélite, uno de los más utilizados con este tipo de efecto es el Landsat-8 un satélite de observación terrestre estadounidense lanzado el 11 de febrero de 2013. Es el octavo y más reciente satélite del proyecto Landsat operado por la NASA y el Servicio Geológico de los Estados Unidos (USGS) desde 1972, con una resolución temporal de 16 días y la incorporación de dos sensores que permiten la captura de imágenes ópticas sobre la escena en

cuestión, el primero de ellos es el Operational Land Imager (OLI) que provee acceso a nueve bandas espectrales, cada una de ellas con resoluciones espaciales de 30 metros a excepción de una, la pancromática que posee una resolución de 15 metros, las cuales cubren el espectro desde los 0.433  $\mu\text{m}$  a los 1.390  $\mu\text{m}$ . El segundo sensor es el Thermal Infrared Sensor (TIRS) que registra imágenes de 10.30 $\mu\text{m}$  a 12.50 $\mu\text{m}$  con una resolución espacial de 100 metros para cada una de las dos bandas espectrales restantes de las 11 que entrega el satélite [7].

### **2.1.2. Satélites de sensores Activos**

El sistema consiste en un mecanismo generador de pulsos electromagnéticos que, por un lado controla la emisión de energía desde el transmisor y por otro accionan un mecanismo de registro, generan ellos mismos la radiación que miden tras ser reflejada. La antena emite un pulso y recibe la respuesta del terreno, un conmutador electrónico (duplicador) evita la interferencia entre ambos pulsos, el receptor es un amplificador de la señal captadas por la antena, así el pulso recibido es grabado para su procesamiento posterior en el segmento terrestre de la misión [8]. El sensor activo más utilizado es el radar y el satélite de radar utilizado para este proyecto fue el Sentinel-1.

## **2.2. Principios básicos de Radar**

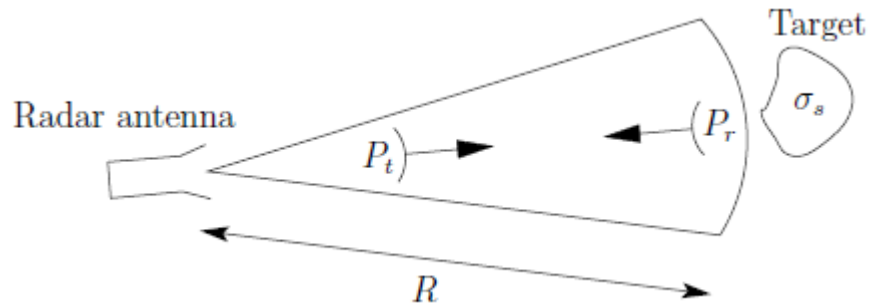
El Radar es un sistema activo que transmite pulsos electromagnéticos hacia un objetivo para posteriormente medir sus propiedades de dispersión, la frecuencia y el ancho de banda de la señal transmitida depende del tipo de radar utilizado para cada aplicación así mismo las características de dispersión del objetivo varían principalmente debido a su geometría, propiedades dieléctricas, la frecuencia de trabajo y la geometría de adquisición de radar [6]. En lo que transcurre de este capítulo se considerara geometría mono-estática, aquella en la que las antenas emisoras coinciden con las receptoras o su distancia aparente es muy mínima.

### **2.2.1. Ecuación de Radar**

La potencia que llega al dispositivo receptor viene dada por la siguiente ecuación, que llamamos ecuación de radar y se muestra aquí para el caso denominado

mono-estático el cual corresponde a utilizar el mismo sensor como trasmisor y receptor de la señal, un ejemplo de este tipo de geometría es mostrado en la Figura 1

**Figura 1.** Geometría de Radar mono-estática.



Fuente: Tesis doctoral Adamo Ferro.

$$P_r = \eta_{radar} * \sigma_s \quad (1)$$

$$\eta_{radar} = \frac{P_t \lambda_c^2 \eta_{ant}^2}{(4\pi)^3 R^4 \eta_{loss}^2} \quad (2)$$

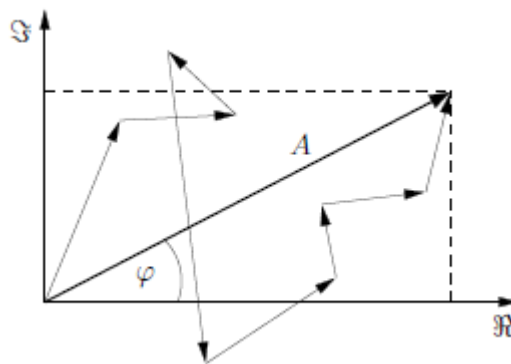
La potencia recibida ( $P_r$ ) puede ser expresada en términos del producto entre la ganancia del radar que incluye los principales parámetros del sistema como los son: la potencia transmitida ( $P_t$ ), la ganancia de la antena ( $\eta_{ant}$ ), la longitud de onda ( $\lambda_c$ ), la distancia entre el radar y el objetivo ( $R$ ), las pérdidas unidireccionales ( $\eta_{loss}$ ), y la sección transversal del radar ( $\sigma_s$ ) que puede ser expresado como el producto entre el coeficiente de retro-dispersión del objetivo ( $\sigma_0$ ) y su área efectiva.

### 2.2.2. Efectos de la naturaleza coherente de la señal de Radar

Un sistema de formación de imágenes que emplea radiación coherente en la región de radar es conocido como sistema activo debido a que genera de manera controlada la radiación necesaria, con una longitud de onda en la región de microondas, para iluminar la escena de interés haciendo factible el estudio de la superficie terrestre en cualquier situación atmosférica. La radiación utilizada por estos dispositivos es coherente ya que los paquetes de onda que forman el haz de

iluminación tienen la misma longitud de onda y la misma fase, uno de los principales efectos debido a este tipo de radiación es el denominado speckle; que es el resultado de interferencias constructivas y destructivas entre el número  $Q$  de objetivos (ver **Figura 2**), en una celda que con sus contribuciones individuales de retro-dispersión se suman de manera coherente para dar como resultado un único valor complejo [9].

**Figura 2.** Efecto speckle, suma coherente de señales dispersadas por objetivos individuales.



Fuente: Tesis doctoral Daniel C. Gallego Cabo.

El speckle es comúnmente modelado como un ruido multiplicativo aleatorio pero no es del todo así ya que este depende de la dispersión generada por los objetivos presentes en la escena, para reducir su efecto existen algunos métodos como: promediar muestras correlacionadas, en el caso de los radares de formación de imágenes conocido este método como multilooking, y la aplicación de filtros de speckle desarrollados para imágenes SAR procesadas [10].

### 2.2.3. Proceso de apertura sintética

La manera como el radar adquiere la imagen de la superficie terrestre se debe en gran medida a los procesos de apertura sintética o filtrado Doppler, es por esta razón que aquel radar que obtengan la información por este método es llamado radar de apertura sintética (SAR). Los procesos de apertura sintética solo son posibles cuando existe un movimiento relativo entre el radar y el objetivo, para el

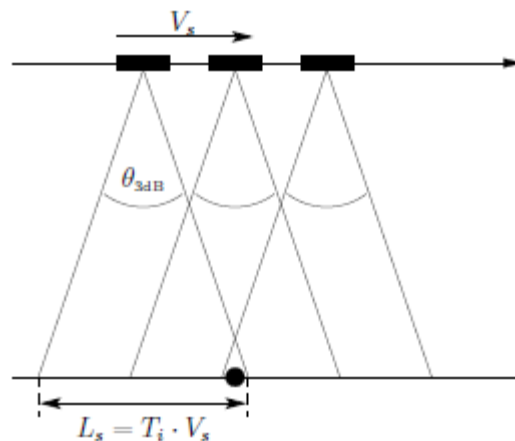
caso del satélite que se utilizó para este trabajo de grado consideramos que el objetivo estuvo estático mientras que la plataforma espacial no lo hizo, es por esta razón que a medida que el radar avanza sobre su trayectoria ilumina en un intervalo de tiempo el mismo objetivo, este intervalo es llamado tiempo de integración ( $t_i$ ), el espacio que recorre el radar durante  $t_i$  corresponde a una distancia  $L_s$  (véase Figura 3), los anteriores dos datos están dados teóricamente de la siguiente manera:

$$t_i = \frac{\theta_{3dB} * R}{V_s} \quad (3)$$

$$L_s = t_i * V_s \quad (4)$$

Donde  $\theta_{3dB}$  son los 3 dB de apertura de la antena de radar a lo largo del azimut y  $V_s$  es la velocidad del radar.

**Figura 3.** La longitud de apertura sintética  $L_s$  surge de iluminar durante un tiempo  $t_i$  un objetivo puntual.



Fuente: vialattea.net

Durante el tiempo de integración la plataforma espacial presenta un movimiento relativo con respecto a los objetivos de tal forma dos blancos puntuales separados ligeramente el uno del otro en la dirección acimutal mostrarán en cualquier instante velocidades relativas diferentes por esta razón el eco procedente de cada objetivo tendrá un desplazamiento en frecuencia Doppler distinto [11]. Posteriormente el radar coherente mide y registra el historial de fase de las señales recibidas y

mediante un algoritmo de enfoque identifica los objetivos dispuestos sobre la superficie. Como resultado del anterior proceso con la longitud real de la antena se puede generar una antena virtual en la cual la distancia acimutal sea mayor debido a la apertura sintética del radar.

### **2.3. Satélite de radar Sentinel-1**

El satélite es un cuerpo que gira en órbita alrededor de la tierra, situado a una cierta distancia desde la cual puede mantener siempre la misma posición con respecto al planeta. Existen dos tipos de satélites clasificados según su forma de adquirir datos: aquellos que utilizan la luz solar reflejada por la Tierra para la formación de imágenes de la superficie del planeta, conocidos como satélites ópticos y aquellos que emiten su propio haz conformado por ondas de microondas, del que posteriormente capturan la señal retro-dispersada proveniente de la Tierra, conocidos como satélites de radar.

Sentinel-1 es un satélite de radar de apertura sintética conformado por dos constelaciones la 1A lanzada el día 3 de Abril del 2014 y la 1B puesta en órbita el 25 de Abril de 2016 ambos con una expectativa promedio de vida de 7 años, elaborado por la Agencia Espacial Europea en conjunto con la Comisión Europea, la industria, los proveedores de servicios y los usuarios de datos. Diseñado y construido por un consorcio de alrededor de 60 compañías lideradas por Thales Alenia Space y Airbus Defence and Space.

El satélite Sentinel-1 trabaja con microondas en un ancho de banda comprendido entre 4 y 8 GHz (en longitud de onda igual a 3.75 – 7.5 cm) aproximadamente, correspondiente al intervalo específico de la banda C, posee también una resolución temporal de 12 días y una espacial de máximo 5x5 m y admite el funcionamiento del radar de apertura sintética con doble polarización (HH-HV ó VV-VH) [12]. Este satélite es perfecto para aplicaciones de monitoreo de la superficie terrestre ya que la imagen que se obtiene está libre de coberturas nubosas consecuencia de la longitud de onda que maneja, debido a este fenómeno físico la resolución temporal aumenta lo que hace más fácil observar las características biofísicas de las capas naturales de vegetación que recubren la superficie.



### 2.3.1. Distorsiones geométricas para el satélite Sentinel-1

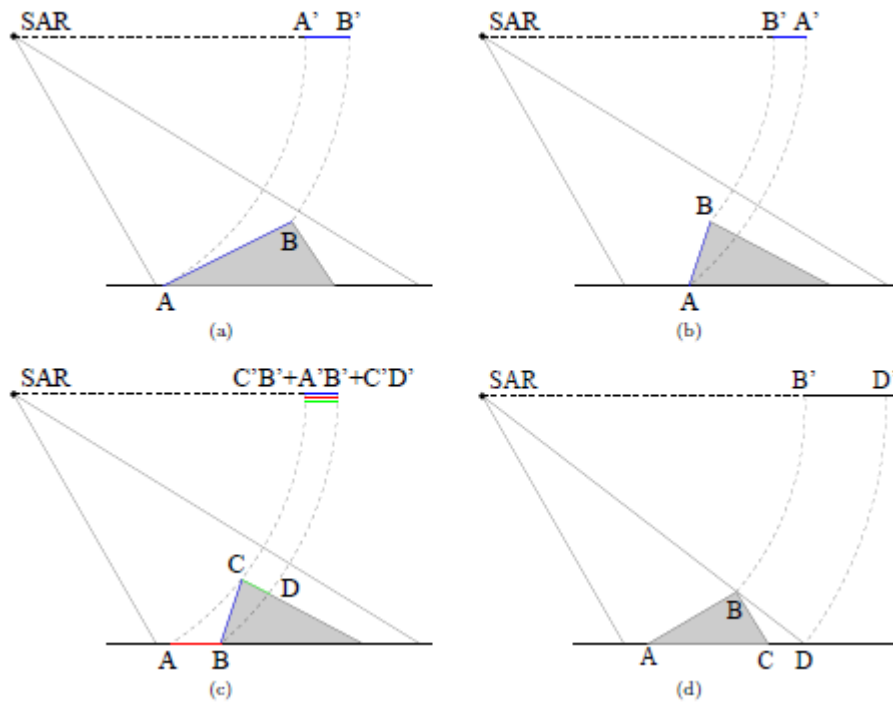
Los datos adquiridos por la plataforma satelital Sentinel-1 se pueden visualizar de dos maneras, la primera slant range, conocida por otorgar las mediciones de manera natural, es decir, como el radar percibe los objetivos (imagen de rango inclinado) la segunda ground range se obtiene a partir de la suposición de que la Tierra sea igual a una elipsoide, lo que es aproximadamente correcto donde no hay montañas, entonces las distancias se medirán entre la trayectoria del suelo justamente debajo del satélite hasta el objetivo.

Para este trabajo de grado se escogen las imágenes en geometría slant range dado que la Cuenca Alta del Río Cauca abarca sectores montañosos donde la geometría ground range no sería válida en un inicio, al trabajar con un terreno topográficamente variado la anterior elección al momento de visualizar la información recolectada traerá consigo los siguientes problemas de distorsión geométrica:

#### **Foreshortening**

El acortamiento es un efecto dominante en las imágenes SAR de áreas montañosas, especialmente en el caso de los sensores espaciales inclinados, la diferencia de rango oblicuo entre dos puntos ubicados en la pendiente de una montaña es menor de lo que sería en un área plana. Por ejemplo al observar la **Figura 4a**, la distancia  $\overline{AB}$  es mucho más larga que su proyección  $\overline{A'B'}$  en el espacio de imagen SAR de rango oblicuo. La compresión en el rango oblicuo da como resultado un área más brillante  $\overline{A'B'}$ , ya que contiene toda la energía dispersada por el área  $\overline{AB}$  más larga.

**Figura 4.** Distorsiones geométricas. a) foreshortening, b) desplazamiento por relieve, c) Sobreposición, d) sombras.



Fuente: Tesis doctoral Adamo Ferro.

### Desplazamiento por relieve

Si la inclinación de la superficie es mayor que el ángulo de incidencia, la parte superior de la estructura se desplazará en la imagen quedando más cerca al sensor, como se muestra en la **Figura 4b**, Aunque A está ubicado en primera posición frente al punto B, la proyección sobre la diagonal que es la línea sobre la que por geometría slant range mira el radar da como resultado un orden inverso, es decir, B' está más cerca del sensor que A'.

### Sobreposición

Su efecto es similar al anteriormente expuesto (desplazamiento por relieve), donde por la inclinación del objetivo se generan unos continuos solapamientos de las señales involucradas impidiendo separarlas al final del proceso, por ejemplo en la **Figura 4c**, la pendiente  $\overline{BC}$  es más pronunciada que el ángulo de incidencia del

haz del radar, de modo que  $\overline{AB}$ ,  $\overline{BC}$  y  $\overline{CD}$  se encuentran dentro de la misma distancia al sensor. Por lo tanto, sus superposiciones de retrodispersión en el área serán  $\overline{C'B'} + \overline{A'B'} + \overline{C'D'}$ .

### **Sombras**

Las sombras son áreas donde no se registra la retrodispersión en el sensor, ya que por la geometría del terreno no se pueden acceder a ellas con el rayo del radar. Las superficies que se alejan del sensor son más empinadas que la iluminación SAR, como se muestra en la **Figura 4d**, el haz de radar no puede iluminar el área entre  $\overline{BD}$ , ya que  $\overline{BC}$  se encuentra más inclinado que el haz del radar, lo que causa el área de sombra  $\overline{B'D'}$

## **2.4. Pre-procesamiento de imágenes Sentinel-1**

Como se ha mencionado anteriormente la CARC a lo largo de toda su superficie presenta continuamente montañas de distintos tamaños por esta razón se seleccionó la imagen satelital del tipo SLC que toma la información sin aproximar esto a un terreno plano pero que a su vez hace necesario aplicar una secuencia de pasos para lograr al final entregar una imagen orto-rectificada y calibrada.

### **2.4.1. Corrección de órbita**

Los vectores de estado de la órbita proporcionados en los metadatos de un producto SAR generalmente no son precisos y pueden refinarse con los archivos de órbita precisos que están disponibles días a semanas después de la generación del producto.

Para refinar los vectores de estado de la órbita, se realizan los siguientes pasos: se obtiene la hora de inicio del producto de origen, a continuación se encuentra el archivo del tipo de órbita especificada por el usuario por el usuario y la hora de inicio del producto, para cada vector de estado de órbita, se obtiene su tiempo Doppler cero, que se usara posteriormente para calcular el nuevo vector de estado de la órbita con la interpolación de Lagrange de 8vo orden usando los datos de las 9 posiciones de las órbitas más cercanas alrededor del tiempo Doppler cero.

El archivo de órbita proporciona información precisa de velocidad y posición del satélite. En base a esta información, se actualizan los vectores de estado de la órbita en los metadatos abstractos del producto.

#### **2.4.2. Calibración**

El objetivo de la calibración SAR es proporcionar imágenes en las que los valores de los píxeles se puedan relacionar directamente con la retrodispersión de la escena del radar. Aunque las imágenes SAR no calibradas son suficientes para el uso cualitativo, las imágenes SAR calibradas son esenciales para el uso cuantitativo de los datos radiométricos. Por lo tanto si se desea hacer un estudio con la segunda característica, es necesario aplicar la corrección radiométrica a la imagen SAR de modo que los valores del píxel realmente representen la retrodispersión del radar de la superficie reflectante. La corrección radiométrica también es necesaria para la comparación de imágenes SAR adquiridas con diferentes sensores o del mismo sensor pero en diferentes momentos.

Este operador promedia el brillo de cada línea con una constante, el valor de la constante suele ser el brillo medio de toda la imagen.

#### **2.4.3. Deburst**

Cada producto IW SLC consiste en una imagen de franjas de polarización. Los productos de IW tienen 3 hileras. Cada submatriz de imagen consta de una serie de ráfagas, donde cada ráfaga se procesó como una imagen SLC separada. Las imágenes complejas de ráfaga enfocadas individualmente se incluyen, en orden de acimut, en una sola imagen, con demarcación de relleno negro en el medio.

Las imágenes para todas las ráfagas en todas las franjas secundarias de un producto IW SLC se reexaminan en una cuadrícula de espaciado de píxeles común en el rango y acimut.

Para S-1 TOPSAR, el área de tierra con imágenes de ráfagas adyacentes solo se superpondrá marginalmente en acimut, lo suficiente como para proporcionar una cobertura contigua del terreno, en la dirección de acimut, las ráfagas se fusionan de acuerdo con su tiempo de Doppler.

La entrada al operador es el producto Sentinel-1 TOPSAR IW. El producto podría ser producto original Sentinel-1 o producto procesado. La salida de este operador es el producto unido en todas sus franjas.

#### **2.4.4. Multilooking**

Generalmente, una imagen original de SAR aparece con un ruido de moteado inherente, para reducir este comportamiento varias imágenes se combinan incoherentemente como si correspondieran a diferentes aspectos de la misma escena, este proceso se conoce como multilook. El resultado de este proceso brinda una imagen de salida que mejora su apariencia y por ende su interpretación se facilita. Además, el proceso de multilook se utiliza para producir una imagen con tamaño de pixel nominal.

Las imágenes multilook pueden ser producidas por el método de dominio de frecuencia usando el ancho de banda sub-espectral. Este operador implementa el método en el dominio del espacio promediando una sola imagen con una pequeña ventana deslizante, como resultado, se produce una imagen con apariencia moteada reducida y un espaciado de píxeles aproximadamente cuadrado en el suelo.

#### **2.4.5. Corrección de terreno por rango Doppler**

Debido a las variaciones topográficas de una escena y la inclinación del sensor del satélite, las distancias pueden distorsionarse en las imágenes SAR. Los datos de la imagen que no están directamente en la ubicación Nadir del sensor tendrán cierta distorsión. Las correcciones de terreno están destinadas a compensar estas distorsiones de modo que la representación geométrica de la imagen sea lo más parecida al mundo real.

El operador de corrección de terreno Doppler de rango implementa el método de ortorectificación de rango Doppler para geo-codificar imágenes SAR a partir de una única geometría de radar de trama 2D. Utiliza información de vectores de estado orbital disponible en los metadatos u órbitas precisas externas, las anotaciones cronológicas de radar, los parámetros de conversión de rango inclinado a tie-

rra junto con los datos DEM de referencia para obtener la información de geolocalización precisa.

## **2.5. Obtención de características de datos Sentinel-1**

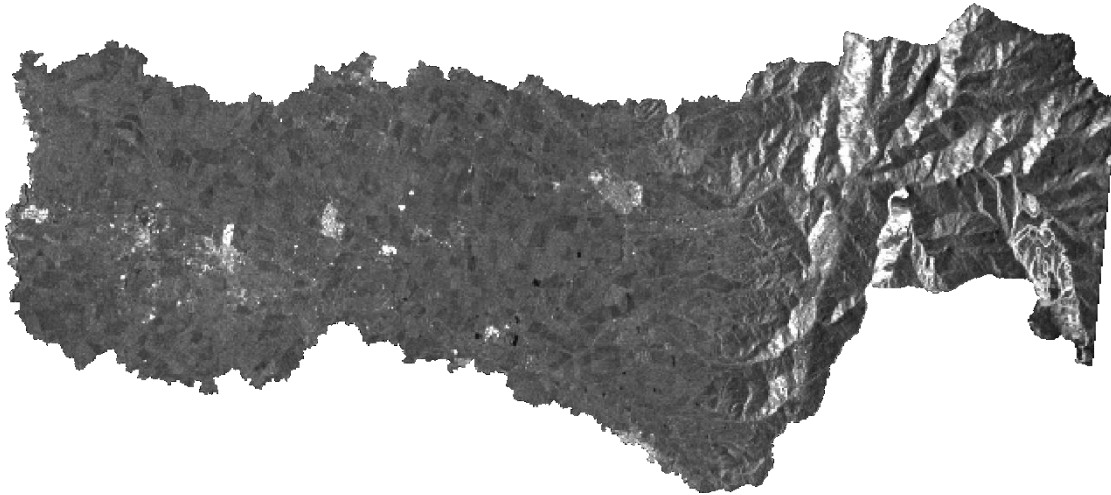
Para aprovechar la cantidad de información de los datos del satélite Sentinel-1 es necesario implementar métodos que permitan extraer patrones generados por las variaciones de intensidades de más de un píxel, dado que la textura por ejemplo, una característica de las imágenes de radar, no está definida para la unidad sino para un conjunto o una vecindad de píxeles, así analizando la distribución espacial de los valores de gris en la imagen, se extraen características locales de la textura.

En las imágenes SAR, la textura tiene dos componentes: la primera es la variabilidad espacial en las propiedades de reflexión difusa de la escena y la segunda texturas inherentes en la imagen como es el ruido sal y pimienta llamado también speckle que degrada la calidad de la imagen y dificulta la interpretación de sus características, estos puntos son causados por la interferencia constructiva y destructiva de las ondas de retorno desfasadas pero coherentes dispersadas por los objetivos dentro de cada región observada por el satélite [13]. La reducción del Speckle se puede llevar a cabo mediante filtrado espacial o procesamiento multi-look. Por esta razón se utilizan como características para una futura clasificación los resultados de aplicar distintos filtros espaciales de realce de texturas a las imágenes dispuestas para el trabajo de grado, en total se extrajeron 25 características Sentinel-1, incluida la imagen original, agrupadas la mayoría en tres grandes grupos: Filtros espaciales dentro de los cuales se encuentran: Mediana, Promedio y Desviación estándar, Filtros adaptativos como: Boxcar, Lee, LeeSigma, RefinedLee, Frost, IDAN, GammaMap y Filtros de textura extraídos de Grey Level Co-occurrence Matrix (GLCM) que abarca atributos de contraste como: Dissimilarity, Homogeneity, Contrast, características de orden como: Angular Second Moment (ASM), Maximum Probability (MAX), Entropy (ENT), Energy (EN) y características de estadísticas: GLCM Mean, GLCM Variance, GLCM Correlation, y finalizamos

con operaciones sobre la imagen que nos permiten calcular las últimas cuatro características  $U$ ,  $VA$ ,  $VI$ ,  $VL$ .

Los filtros para el realce espacial enfocan sus algoritmos en la conservación de la resolución espacial por célula de imagen, con frecuencia la característica más importante a realzar son las fronteras que definen los objetos presentes en la imagen, dentro de esta categoría se encuentran el filtro promedio, filtro mediana y desviación estándar, el primero de estos utiliza la intensidad en cada intervalo de muestreo en la imagen, reemplazándolos por el promedio de los valores de los píxeles localizados dentro de una ventana móvil que recorre toda la muestra, este filtro conserva bien las características radiométricas, pero induce un efecto borroso en las áreas con una cierta textura, el segundo de los procedimientos utilizados para filtrar y extraer información de la imagen fue el cálculo de la desviación estándar, que brindara información de que tan dispersos se encuentran los niveles de gris de la ventana móvil de su media, si el valor de la desviación es alta querrá decir que la ventana está situada sobre una textura heterogénea pero si por lo contrario el valor es bajo se puede inferir que el terreno presenta una homogeneidad. Por último se extrajo el filtro de mediana que asigna el valor de la mediana de la ventana a cada muestra, conserva de mejor manera la información de textura, modifica la información radiométrica de áreas homogéneas y no conserva la señal característica de los reflectores puntuales [14] , a continuación, se muestran ejemplos de estos filtros:

**Figura 5.** Filtro de promedio basado en imágenes Sentinel-1 de la cuenca del Río Guachal.



**Figura 6.** Filtro de mediana basado en imágenes Sentinel-1 de la cuenca del Río Molino.





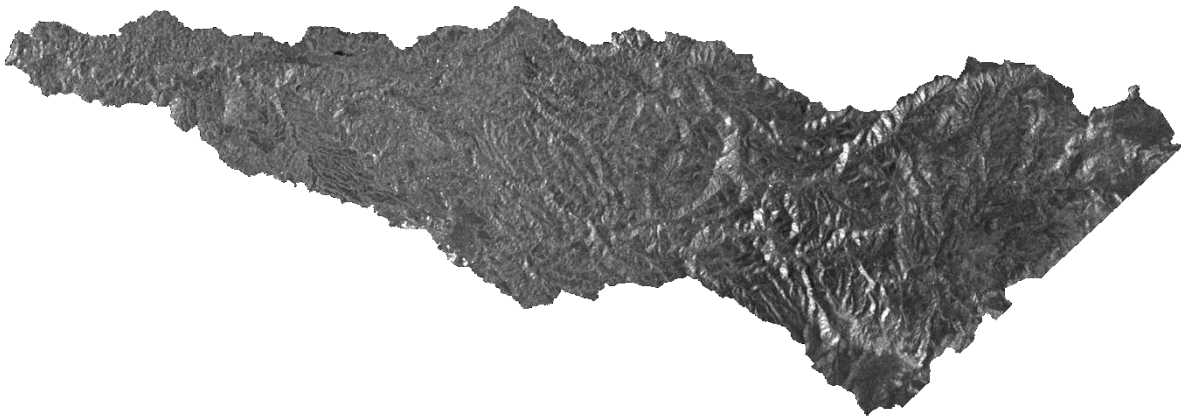
**Figura 7.** Filtro de desviación estándar basado en imágenes Sentinel-1 de la cuenca del Río Mondomo.



Existen situaciones donde no es conveniente aplicar filtros que afecten globalmente a la imagen, al haber particularidades topográficas sobre el terreno, por ejemplo, en imágenes donde se incluye una costa con gran cantidad de agua representada por valores bajos de niveles digitales y el terreno con valores altos, en estos casos es conveniente un filtro que se adapte a las distintas áreas de interés. Los filtros adaptativos son considerablemente más complejos ya que los coeficientes de ponderación se recalculan para cada uno de los píxeles en función del histograma de los niveles digitales que aparecen en la ventana. Se han utilizado con gran éxito filtros adaptativos para eliminar el speckle de las imágenes de radar y para detectar, con un solo filtro, diferentes elementos [15]. A diferencia de un filtro de realce espacial, el filtro adaptativo preserva los detalles de la imagen aun cuando se suprime en gran parte el ruido [16].

La primera característica adaptativa fue la obtenida del filtro Boxcar que realiza la estimación de los valores de la matriz de covarianza a partir del promediado de las matrices de  $N$  píxeles distintos generalmente cuadrada, centrada en el píxel de interés. Sin embargo esta técnica de filtrado tiene dos inconvenientes. En primer lugar, el resultado obtenido depende en gran medida del número de píxeles promediados, a mayor cantidad de píxeles mejor filtrado y reducción de la varianza del ruido speckle. En segundo lugar, produce una gran pérdida de resolución ya que en caso de situarse la ventana sobre una zona no homogénea o sobre un contorno entre dos regiones, los resultados que se obtendrán no serán válidos al mezclar píxeles de distribuciones estadísticas diferentes. Esto produce, por ejemplo, que los blancos puntuales brillantes se extiendan según la forma de la ventana o que los contornos entre los objetos de la escena se vuelvan borrosos. A continuación, se muestra un ejemplo de dicha característica:

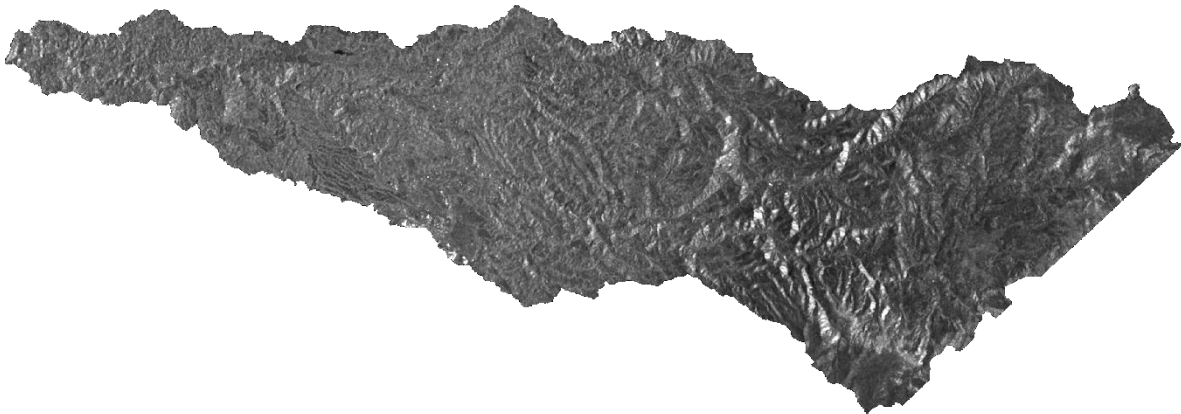
**Figura 8.** Filtro Boxcar basado en imágenes Sentinel-1 en la cuenca del Río Palacé.



Continuando con la extracción de características se obtuvieron las provenientes del filtro Lee, en el cual se definen ventanas direccionales para localizar el área más homogénea dentro de un vecindario determinado para cada píxel. Para seleccionar la sub-ventana más homogénea se utiliza la información de la diagonal de la matriz de covarianza. Los píxeles de la sub-ventana seleccionada se em-

plean para calcular la matriz de covarianza filtrada, a partir del estimador de mínimo error cuadrático medio localmente lineal de la matriz de covarianza. Para reducir la sensibilidad del coeficiente de filtrado adaptativo se emplean máscaras direccionales, las sub-ventanas mencionadas anteriormente, para determinar la parte más homogénea dentro de la ventana deslizante, dentro de la cual se estima el valor del píxel actual. Esto permite reducir la pérdida de resolución que obtiene el filtrado Boxcar, de forma que cerca de los contornos se selecciona una sub-ventana con píxeles homogéneos. Sin embargo, sigue teniendo el inconveniente de la dependencia con el tamaño de la ventana. Además, las sub-ventanas son fijas, lo que limita en gran medida su capacidad de filtrado o de preservación de contornos en función de la forma de las regiones presentes en la imagen [17], la Figura 9 muestra un ejemplo del filtro Lee:

**Figura 9.** Filtro Lee basado en imágenes Sentinel-1 en la cuenca del Río Palacé.



Los filtros Lee mejorado y Lee sigma son una adaptación del filtro Lee y de manera similar usan estadísticas locales (coeficientes de variación) dentro de ventanas de filtros individuales [18]. Cada píxel se coloca en una de tres clases: la homogénea donde el valor del píxel se reemplaza por el promedio de la ventana del filtro, la heterogénea en la que el valor de píxel se reemplaza por un promedio ponderado y la clase objetivo de punto donde el valor de píxel no cambia. Los an-

teriores filtros son utilizados para reducir el moteado en las imágenes de radar y, al mismo tiempo, conservar la información de textura [19]:

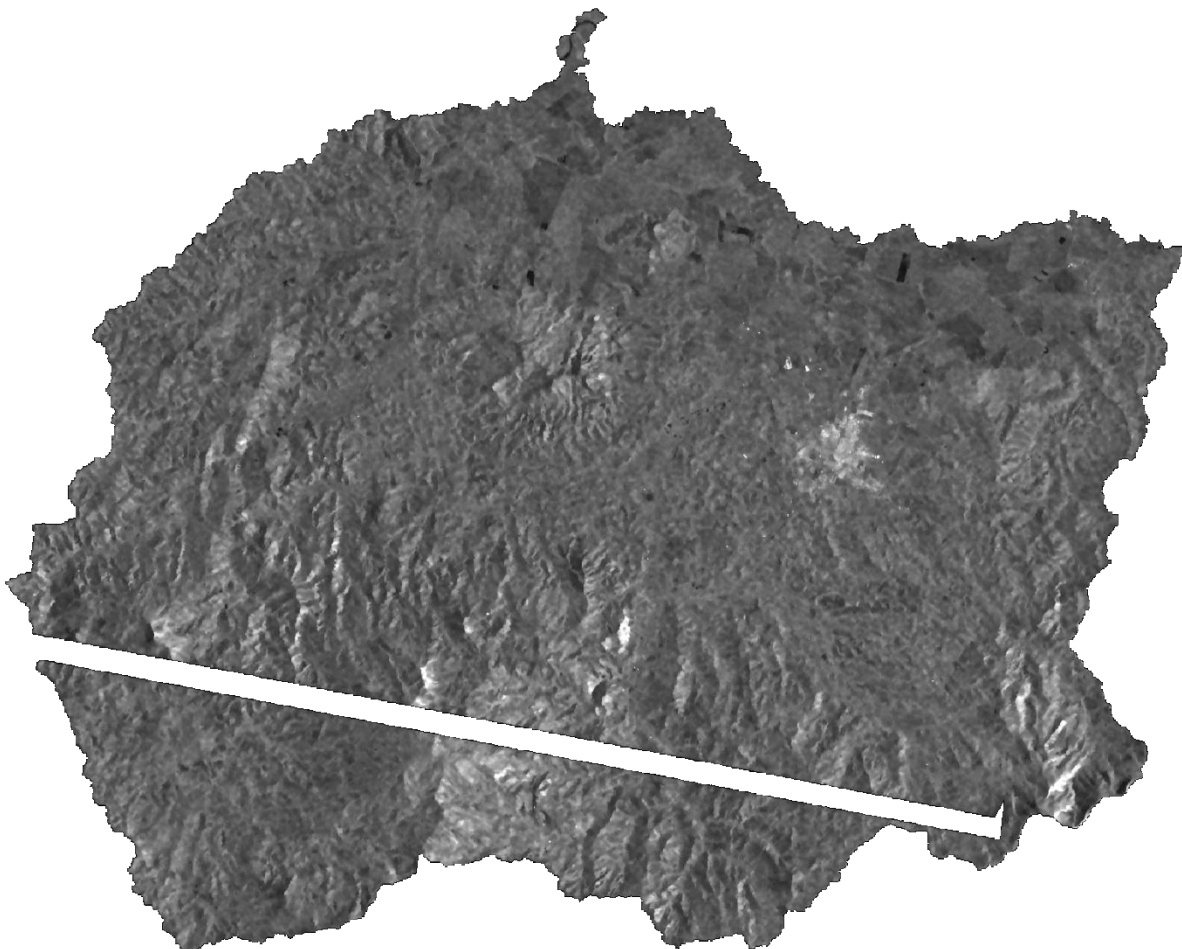
**Figura 10.** Filtro Lee mejorado basado de imágenes Sentinel-1 en la cuenca del Río Palo.



El filtro IDAN (Intensity-Driven Adaptive-Neighborhoog), perteneciente a los filtros adaptativos, usa una técnica para la estimación de las matrices de coherencia adaptativas en el espacio, alrededor de cada píxel se define un vecindario adapta-

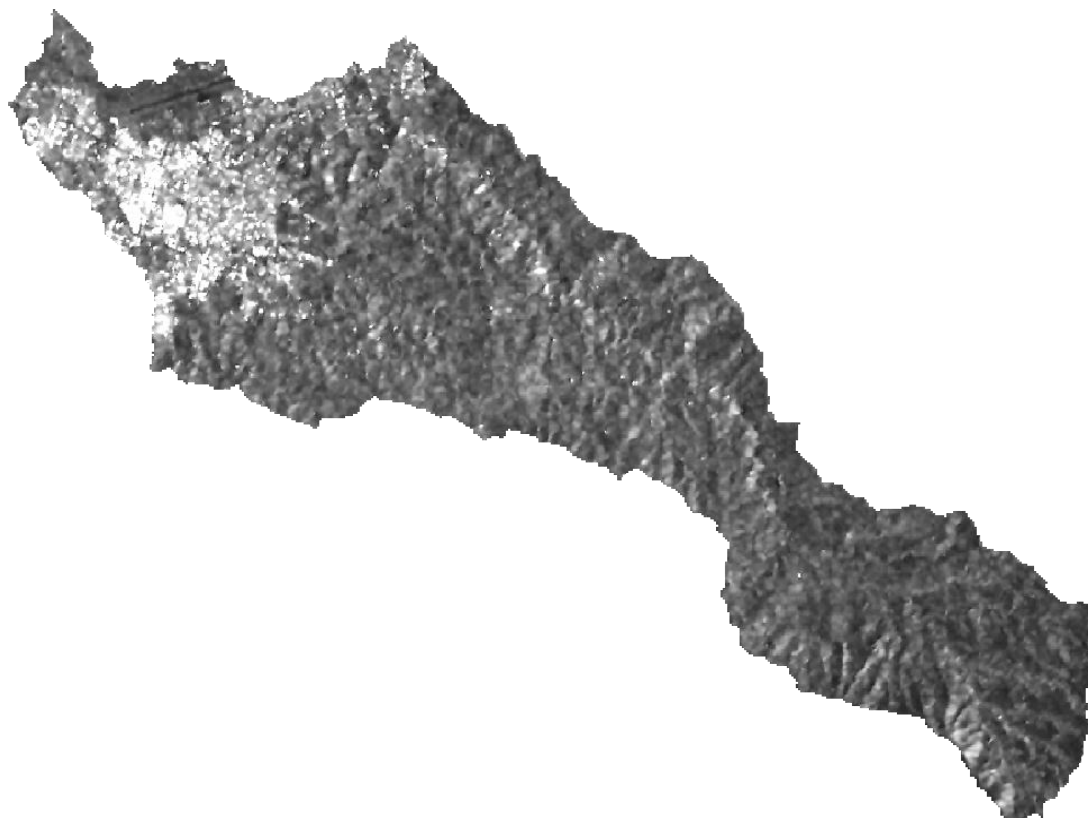
tivo (AN), utilizando un proceso de crecimiento de regiones o Region Growing (RG), obteniendo un vecindario homogéneo mucho más flexible y preciso comparado con el uso de las sub-ventanas del filtro Lee. Luego, para la estimación de la matriz de coherencia pueden utilizarse sólo los píxeles presentes en el vecindario adaptativo, que para cada píxel, llamado semilla, mediante el crecimiento de regiones construye un vecindario de forma y dimensiones variables, que contiene únicamente píxeles pertenecientes a la misma clase estadística que la semilla. Para la primera estimación de la semilla se calcula la mediana del conjunto de píxeles inmediatamente adyacentes al inicial [20], esto se hace para evitar la pérdida de resolución que conlleva el uso de la media. La principal ventaja de esta técnica es que permite acumular un gran número de muestras homogéneas en el vecindario adaptativo.

**Figura 11.** IDAN basado en imágenes Sentinel-1 en la cuenca del Río Quinamáyó.



Usualmente en las imágenes SAR se habla de áreas homogéneas y heterogéneas. Las áreas homogéneas son aquellas donde la reflectividad del radar es considerada constante (lagos, zonas de agricultura uniforme, etc.), en las áreas heterogéneas la reflectividad varía pixel a pixel (costas, texturas, etc.). Frost propone un filtro adaptativo lineal bajo un modelo de ruido multiplicativo con estadísticas estacionarias, suponiendo homogeneidad en las imágenes SAR y luego transfiriendo el filtro a uno adaptativo basado en el error medio cuadrático mínimo, el cual se debe adaptar a las estadísticas locales de la imagen, estas proporcionan el peso de los parámetros que provocan la respuesta del filtro. En este filtro existe una clara situación de compromiso entre eliminación del ruido y difuminado de la imagen. A diferencia de otros filtros, la transferencia de este es menos prominente en los bordes, preservando así la forma de estos [21] , la siguiente figura muestra el comportamiento del filtro Frost.

**Figura 12.** Filtro Frost basado en imágenes Sentinel-1 en la cuenca del Río Molino.



Los filtros Frost y Lee están basados en modelos que no usan las propiedades estadísticas de la escena bajo procesamiento sino la suposición de una homogeneidad topográfica. El filtro Gamma Map está basado en la suposición de que la intensidad sin moteado de la escena tiene una distribución Gamma, intenta estimar el nivel digital original del píxel, que se asume debe estar entre la media local y el nivel digital del píxel degradado. La lógica del MAP maximiza la función de densidad de probabilidad a posteriori respecto a la imagen original. Este filtro minimiza la pérdida de información de textura aún mejor que los filtros Frost y Lee, en casos de escenas con distribución Gamma tales como áreas arboladas, áreas agrícolas y océanos. Esto permite una reducción más efectiva del moteado y mejor preservación de las variaciones espaciales de la escena (textura, bordes, objetos puntuales). El filtro conserva el valor observado del píxel en las escenas que no tienen distribución Gamma.

**Figura 13.** Filtro Gamma Map en la cuenca del Río Molino.



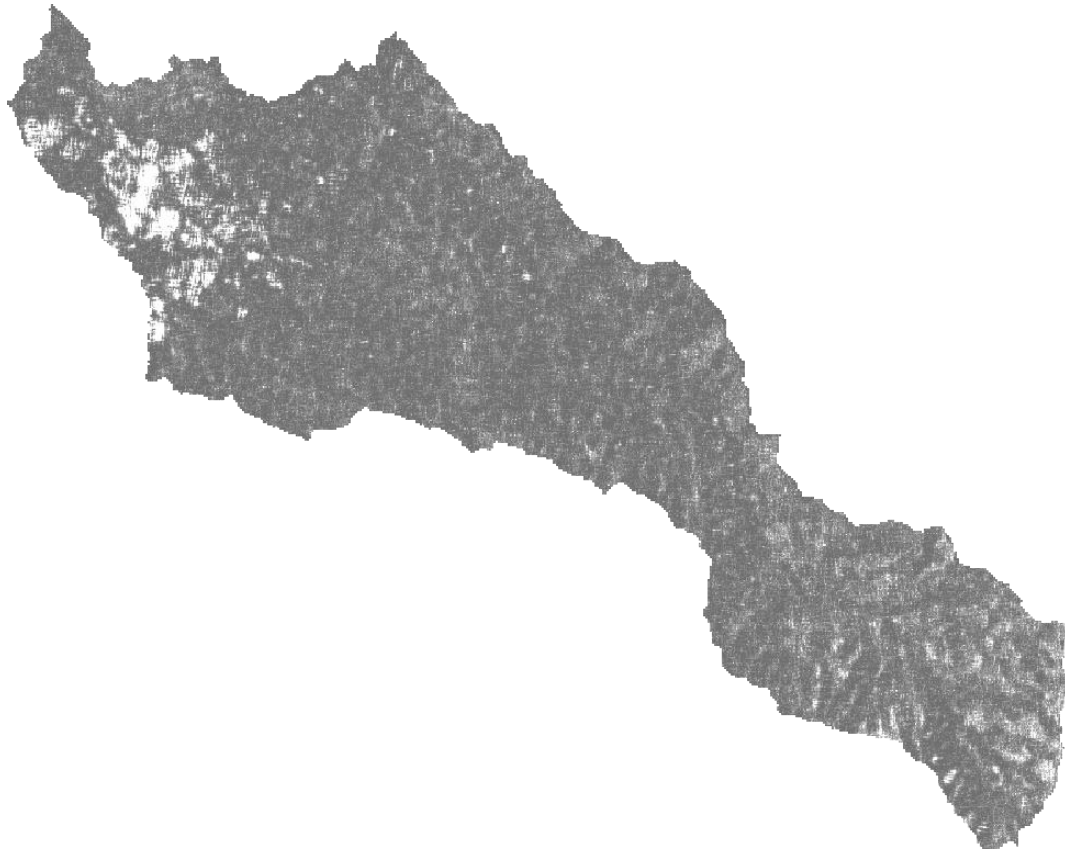
Para estimar las características de textura restantes, se usa de nuevo como en casos anteriormente enunciados una ventana móvil para el cálculo de las métricas de segundo orden (también denominadas coocurrencia), las cuales analizan la repetición de la distribución de los niveles de gris que se presentan en la ventana, en función del pixel de interés y el pixel adyacente, lo anterior para deducir valores de textura, mediante la relación angular y la distancia entre dos píxeles vecinos, así al final se mostrara el número de apariciones de la relación entre un píxel y su vecino especificado. Haralick, Shanmugan y Dinstein se refieren a esto como una "matriz de dependencia espacial de tonos grises". Su implementación considera cuatro direcciones ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  y  $135^\circ$ ) entre celdas vecinas que están separadas por cierta distancia [22]. Algunas medidas texturales posibles de obtener mediante el software Sentinel Application Platform (SNAP) son las siguientes: homogeneidad la cual presenta valores altos cuando la matriz de coocurrencia se concentra a lo largo de la diagonal, esto ocurre cuando la imagen es localmente homogénea sobre la ventana seleccionada, también se obtiene información sobre un patrón opuesto al anteriormente descrito, llamado contraste, una medida de la variación local que tiene un valor alto cuando la región dentro de la escala de la ventana tiene un alto contraste, de manera análoga se realiza el análisis para disimilaridad otra característica de textura, donde los pixeles del kernel se evalúan bajo los mismos parámetros de contraste [23].

Otro parámetro extraído de la matriz de coocurrencia es el GLCM (Grey Level Co-occurrence Matriz) Media que no es simplemente la media aritmética de los valores dentro del kernel, donde el valor del pixel es ponderado por la frecuencia de coocurrencia en combinación de un determinado valor del pixel vecino. GLCM Variance se suma a las medidas de textura extraídas para este trabajo, se basa en la media y la dispersión alrededor de la media de los valores de las celdas de la matriz de coocurrencia utilizando específicamente la combinación entre los píxeles de referencia y vecino. En cuanto a la entropía, un valor más sumado a la lista de características de textura, se puede decir que es alta cuando los elementos de la matriz de coocurrencia tienen valores iguales y es baja cuando los elementos son cercanos a 0 o 1, situación que se presenta cuando la imagen es uniforme



dentro de la ventana seleccionada, en suma este descriptor nos mide la aleatoriedad de la imagen, alcanzando su máximo cuando todos los elementos de la matriz de coocurrencia son iguales, a mayor entropía mayor complejidad. Continuando con un valor más encontramos la correlación, que nos mide la dependencia lineal de los tonos de gris en la imagen, escrito de otra forma la semejanza de la imagen consigo misma, algunas de sus propiedades son: un objeto tiene más alta correlación dentro de él que entre objetos adyacentes y píxeles cercanos están más correlacionados entre sí que los píxeles más distantes. El Angular Second Moment (ASM) es una medida de textura de homogeneidad local que arroja valores altos cuando la matriz de coocurrencia tiene pocas entradas de gran magnitud y valores bajos cuando todas las entradas son similares, a la raíz cuadrada del ASM se le denomina Energía o Uniformidad, y también es usada como medida textural [24].

**Figura 14.** Angular Second Moment (ASM) en la cuenca del Río Molino.

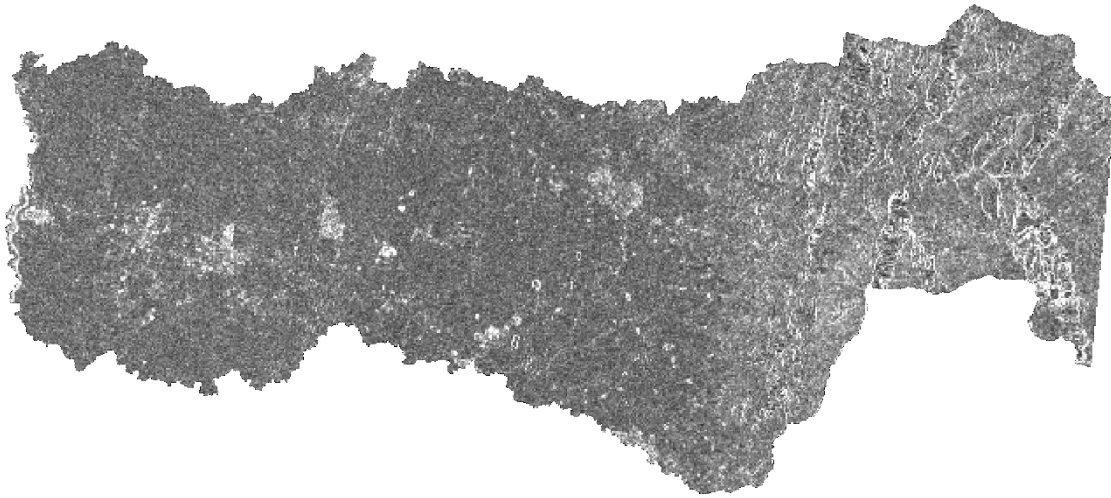


La textura es una de las características importantes que se utilizan para identificar objetos o regiones de interés en una imagen. A diferencia de las características espectrales, que describen la variación tonal promedio en las diversas bandas de una imagen, las características de textura contienen información sobre la distribución espacial de las variaciones tonales dentro de una banda. Además de las características mencionadas anteriormente existen cuatro medidas más de textura obtenidas a partir de ventanas móviles que aplican operaciones de realce sobre los objetivos, estas son: la varianza de la media del logaritmo para intensidad (VL) definida como la diferencia del valor medio del logaritmo de intensidad al cuadrado y el cuadrado del logaritmo de intensidad media, la cual es aplicable cuando la intensidad del píxel tiene una distribución log-normal. Las siguientes dos características de textura son más representativas cuando presentan una distribución gaussiana, estas son: Varianza de intensidad (VI), una relación entre la media de la intensidad al cuadrado y la intensidad media al cuadrado, y la Varianza de amplitud (VA) una relación entre la intensidad media y la amplitud media al cuadrado. Como ultima característica de textura se obtuvo la medida del logaritmo normalizado (U) una medida de registro de textura normalizada. Las últimas cuatro características explican la formación de imágenes de radar y las propiedades estadísticas del speckle; por lo tanto, es esencial que no se haya realizado un filtrado previo en el canal de entrada para calcularlas. A continuación, se presentan algunos ejemplos de dichas texturas:

**Figura 15.** Textura VA en la Cuenca del Río Guachal.



**Figura 16.** Textura VL en la cuenca del Río Guachal.



### **2.6. Algoritmos de aprendizaje máquina: Bosque aleatorio**

Una vez se extraen todas las características de las imágenes de radar se procede a darles uso aplicando algoritmos de aprendizaje máquina que son procesos de inducción al conocimiento por medio de los cuales se intentan crear mecanismos capaces de generalizar comportamientos y reconocer patrones a partir del suministro de información en forma de muestras. Estos algoritmos de aprendizaje se basan en una serie de datos sobre los cuales aprender para luego aplicar la experiencia adquirida en otros conjuntos. Es necesario evaluar su rendimiento sobre un conjunto distinto a aquel con el que el sistema fue entrenado para obtener una estimación válida de su capacidad de generalización ante nuevos ejemplos. Tradicionalmente estos métodos se han dividido en dos grandes grupos: aprendizaje supervisado, en el cual se genera una función que establece una correspondencia entre la entrada y salida deseada, donde la base del sistema de aprendizaje está formado por muestras etiquetadas a priori, es decir, muestras de las que ya sabemos su clasificación correcta, y los no supervisados, donde el proceso de modelado se lleva a cabo sobre un conjunto de muestras formadas únicamente por entradas al sistema, sin conocer su clasificación correcta, por lo que se busca que el sistema sea capaz de reconocer patrones para poder etiquetar las nuevas entradas [25]. Un tercer grupo intermedio lo conformaran los métodos semi-

supervisados que son una combinación de los dos algoritmos anteriores [26], algunos de los clasificadores más utilizados son las redes neuronales, soporte de máquinas vectoriales y los bosques aleatorios, para este trabajo de grado se utilizó el último de los mencionados anteriormente.

El clasificador semi-supervisado de bosque aleatorio es una técnica de aprendizaje máquina que nace para mejorar los resultados obtenidos por los árboles de decisión. Se parte como en todos los métodos de aprendizaje supervisado, de un conjunto de características o atributos, las variables explicativas, y, por otro lado, asociadas a ellas, las clases o respuestas, lo que se pretende identificar, para que de manera interna el clasificador realice un proceso de división que le permita aprender e identificar las clases correctamente (parte no supervisada del clasificador) [27]. El bosque aleatorio es un método que combina una cantidad grande de árboles de decisión independientes, utilizados sobre un conjunto de datos aleatorios con igual distribución, donde cada árbol de decisión contribuye con un voto para la asignación de la clase más frecuente.

Un árbol de decisión se construye mediante sucesivas particiones binarias del conjunto de datos de entrenamiento en subconjuntos (nodos) de homogeneidad creciente. La homogeneidad se mide con el índice de Gini (G) [28] que se reduce al mínimo cuando todas las observaciones pertenecen a la misma clase. En cada paso, el nodo con mayor valor de G se divide en función de que los valores en su condicional para una determinada variable superen o no un determinado umbral. Dicho umbral y variable se seleccionan para minimizar el valor de G en los nodos resultantes. El proceso continúa hasta que ninguna subdivisión puede reducir más el índice de Gini [29].

La fase de aprendizaje radica en crear muchos árboles de decisión independientes, construyéndolos a partir de datos de entrada ligeramente distintos. Se altera, por tanto, el conjunto inicial de partida, seleccionando aleatoriamente un porcentaje de la muestra total de los datos. Esta modificación, algo antiintuitiva, añade aleatoriedad y disminuye la correlación entre árboles, esta falta de correlación es una propiedad deseable que da sentido a la última fase del proceso, la entrega

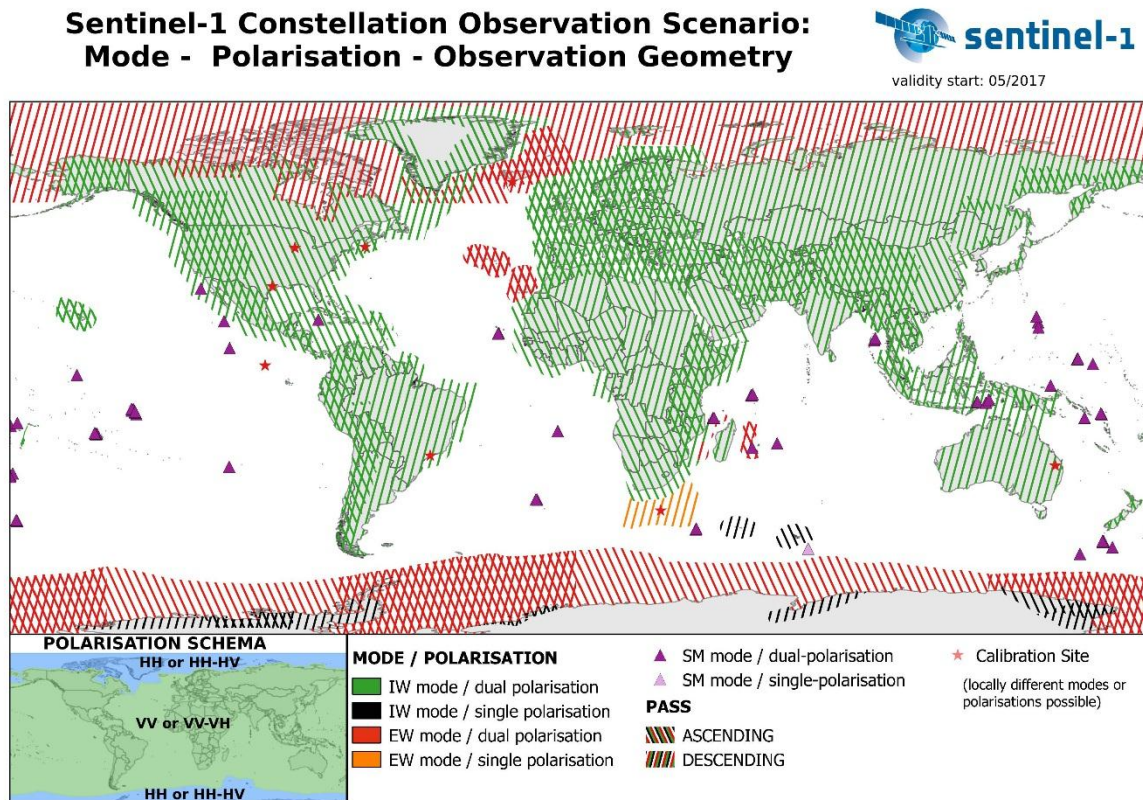
del resultado, donde cada uno de los árboles de manera independiente se evalúa y la predicción del bosque aleatorio será la media de las respuestas de los árboles creados [30]. El porcentaje de datos que no se utilizan para la clasificación se usarán posteriormente para validar los resultados, así mismo, la proporción de árboles que concuerdan con una misma respuesta se interpreta como la probabilidad de la misma [27].

### **3. METODOLOGÍA**

#### **3.1. Descarga de imágenes Sentinel-1**

La conformación de la base de datos de imágenes de radar Sentinel-1 se hace a través de la siguiente página en internet: <https://scihub.copernicus.eu/dhus/#/home> donde seleccionamos con anterioridad el tipo de satélite del cual deseamos obtener información, para este proyecto fue el Sentinel-1, posteriormente los parámetros sensor mode y polarisation deben ser: Interferometric Wide Swath y polarización simple VV respectivamente puesto que por condiciones geográficas para Colombia el satélite Sentinel-1 solo captura imágenes con esas características como se puede verificar en la Figura 17.

**Figura 17.** Mapa de modos, polarización y forma de observación de las imágenes obtenidas por el satélite Sentinel-1.



Fuente: Agencia Espacial Europea.

Para culminar con el último parámetro que es tipo de producto elegimos Single Look Complex (SLC), se decide trabajar con este tipo de imagen, puesto que los productos GRD Sentinel-1 se obtienen a partir del sistema “slant range” después de haberlos multilooked y proyectado en el rango terrestre utilizando un modelo donde se supone que la Tierra es igual a una elipsoide (Earth ellipsoid model), esto es aproximadamente verdadero donde no hay montañas, para el caso de la CARC esto no se cumple, por tal razón la imagen SLC es seleccionada, un proceso de corrección de terreno es realizado sobre la imagen SAR para corregir los datos teniendo en cuenta la topografía del suelo.

Cada producto con las especificaciones descritas anteriormente pesa alrededor de 4GB por lo que descargar y procesar una imagen de tan gran tamaño consume

demasiado espacio y recursos computacionales, es por esta razón que se conformó una base de datos de los años 2015 y 2016 únicamente, para no exceder el espacio en memoria del disco con el que se contaba para realizar este trabajo de investigación, en la **Tabla 1** se presentan las fechas exactas de las imágenes del satélite Sentinel-1 descargadas.

**Tabla 1.** Fecha exacta y representativa de las imágenes Sentinel-1 descargadas

Fecha exacta (dd/mm/aaaa)	Fecha representativa (dd/mm/aaaa)
18/04/2015	01/04/2015
07/05/2015	01/05/2015
05/06/2015	01/06/2015
29/06/2015	01/07/2015
11/08/2015	01/08/2015
09/09/2015	01/09/2015
28/09/2015	01/10/2015
22/10/2015	01/11/2015
09/12/2015	01/12/2015
07/01/2016	01/01/2016
19/02/2016	01/02/2016
14/03/2016	01/03/2016
07/04/2016	01/04/2016
01/05/2016	01/05/2016
30/05/2016	01/06/2016
17/07/2016	01/07/2016
05/08/2016	01/08/2016
29/08//2016	01/09/2016
16/10/2016	01/10/2016
09/11/2016	01/11/2016
09/12/2016	01/12/2016

Finalmente se descargaron dos imágenes más provenientes del satélite óptico Landsat 8, orto-rectificadas y corregidas radiométricamente mediante las herramientas dispuestas por el software Q-GIS, correspondientes a los path/row 009/057 y 009/058 del 29 de Diciembre del 2015. Las escenas del Landsat 8 y Sentinel-1 abarcan la mayor parte de la CARC.

Al realizar satisfactoriamente lo descrito anteriormente, se da por culminado el primer objetivo específico de este trabajo de grado, que hace referencia a la construcción de una base de datos conformada por imágenes SAR que contengan la

información adecuada para la extracción de características en la región comprendida por la Cuenca Alta del Río Cauca.

### 3.2. Pre-procesamiento imágenes Sentinel-1

Sentinel-1 brinda una imagen en banda C a 5.405 GHz con un ángulo de incidencia entre 20° y 45° con una órbita circular, a una altura de 693 km, la plataforma 1A tiene una resolución temporal de 12 días respecto al ecuador y tiene cuatro modos de adquisición de imágenes: Stripmap (SM), Interferometric Wide Swath (IW), Extra Wide (EW) y Wave, con diferentes resoluciones, amplitudes, ángulos de incidencia y polarizaciones.

Este estudio se enfocó en el uso de datos del modo IW en polarización simple VV (transmisión vertical, recepción vertical) para la zona de la CARC. Todos los datos están disponibles en ESA Data Hub, para este trabajo de grado se conformó una base de datos de todas las imágenes Sentinel-1 que estaban contenidas en la Cuenca Alta del Río Cauca, el formato descargado para las imágenes fue Standard Archive Format for Europe (SAFE), que contiene información general del producto en XML para toda la región. La metadata contiene la información de las principales características: fecha de adquisición, propiedades de la imagen, polarización, fusión de bandas, calibración y ubicación geográfica. Los productos SAFE se obtuvieron como Single Look Complex (SLC).

Los algoritmos de pre-procesamiento aplicados en la imagen SLC fueron: corrección de órbita, calibración, deburst, multilooking y finalmente corrección del terreno, los pre-procesos a las imágenes Sentinel-1 se realizaron con la ayuda del programa que dispone de manera gratuita la Agencia Espacial Europea para el tratamiento de imágenes satelitales llamado SNAP por sus siglas en inglés Sentinel Application Platform, en la **Figura 18** se presenta parte de la interfaz gráfica necesaria para realizar el pre-procesamiento de las imágenes Sentinel-1, dentro de las ventajas más notorias para este software se encuentran la rápida visualización y navegación sobre imágenes de gran tamaño (giga-píxeles), la administración avanzada de capas que permite agregar y manipular nuevas

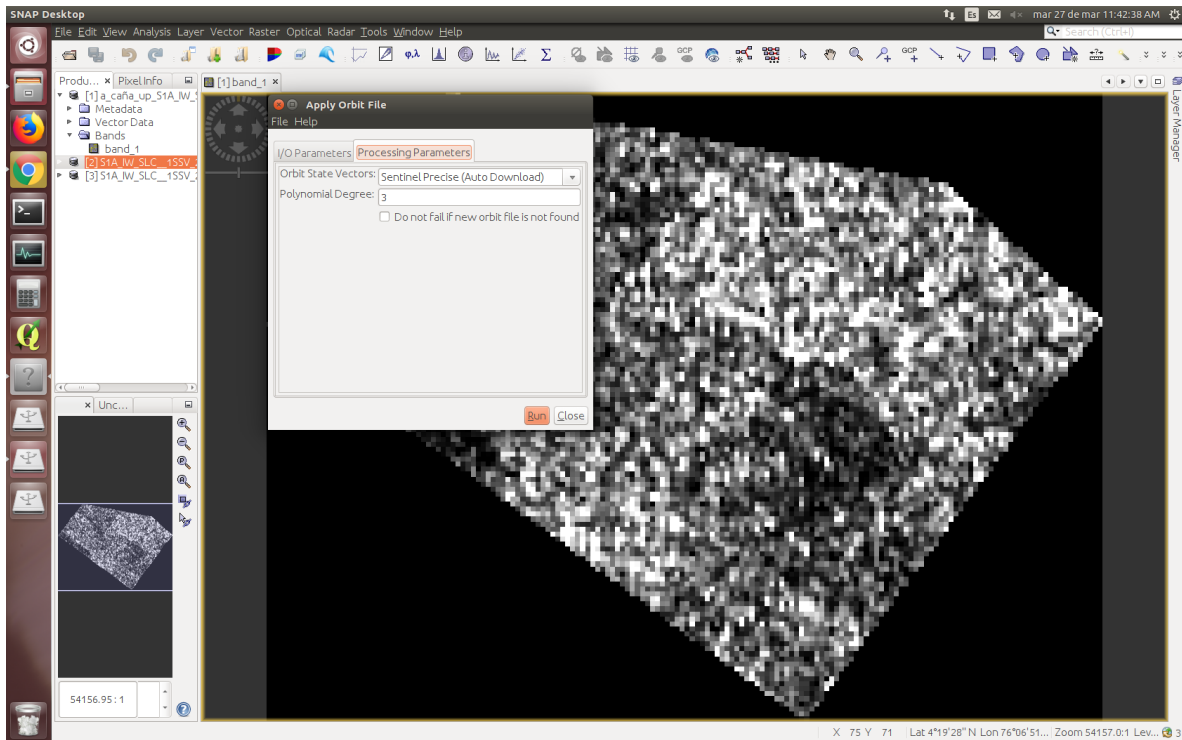


superposiciones, como imágenes de otras bandas, incorporación de aritmética de banda flexible usando expresiones matemáticas arbitrarias, precisa reproyección y orto-rectificación usando puntos de control sobre tierra y específicamente trae consigo una caja de herramientas diseñada para realizar los diferentes procesos mencionados al inicio, necesarios para proyectar la imagen SLC proveniente del satélite de radar Sentinel-1 y convertirla en una GRD.

Para ingresar la imagen de entrada al primer pre-proceso no es necesario descomprimir el archivo descargado desde la página de la ESA, pues el software SNAP permite hacer su lectura de manera indirecta desde el archivo comprimido, solo se necesita arrastrarlo a la barra de productos para poder manipularlo.

La corrección de órbita es el primer proceso que se lleva a cabo en la cadena de pre-procesamiento y adaptación de la imagen SAR, pues se necesita antes que nada realizar una reproyección de la imagen sobre el terreno de donde se descargó, para estar totalmente seguros que las dimensiones entregadas por el satélite concuerdan con la realidad topográfica. Mediante la herramienta Apply Orbit File, que se despliega después de ingresar en la pestaña de Radar dispuesta en la plataforma SNAP, se selecciona en Vectores de estado de la órbita la opción Sentinel Precise (Auto Download) que obtiene las órbitas necesarias para el ajuste de semanas después de la fecha de adquisición de la imagen. La siguiente figura permite observar la ubicación de la herramienta:

**Figura 18.** Herramienta de corrección de órbita para imágenes Sentinel-1.

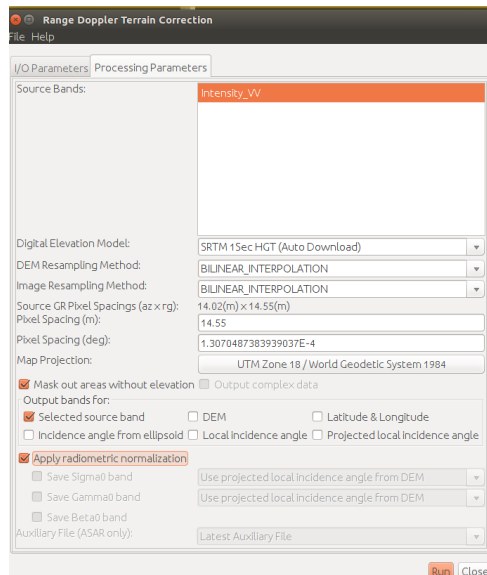


Usando la imagen de salida como banda de entrada para el siguiente procedimiento, se continua con la calibración de la imagen mediante la opción Calibrate disponible dentro del menú desplegado al ingresar a la pestaña Radiometric encontrada dentro de Radar, una vez se despliegan las opciones a seleccionar, se escoge en los parámetros de procesos aquella que guarda la imagen de salida como datos complejos con el fin de poder ejecutar el siguiente proceso de Deburst donde se utilizará la información de amplitud y ángulo como entrada polarimétrica para que de esta manera la imagen que viene particionada en sub-huellas quede unida en una sola, accedemos a esta herramienta a través de la pestaña S-1 TOPS Deburst de la cual también se puede seleccionar la extensión del archivo de salida de la imagen resultante y el directorio donde se desea agregarla, de los formatos sobresalientes para guardar los resultados se encuentran: JPG, PNG, GeoTIFF, BEAM-DIMAP, entre otros, para la sección de pre-procesamiento se decide usar la última opción de extensión (BEAM-DIMAP).

El penúltimo paso de esta etapa de pre-procesos se lleva a cabo mediante el multilooking que además de minimizar el ruido sal y pimienta inherente a las imágenes SAR también brinda como resultado una imagen con pixeles cuadrados, de esta manera se procede a seleccionar una vez se despliega el menú de la herramienta la opción GR Square Pixel, que permite de forma automática calcular insertando el número de vistas de Rango el número de vistas del Azimuth y por ende el valor del Pixel Cuadrado promedio GR.

Para culminar con los pasos de pre-procesamiento y dejar la imagen lista para la extracción de características, es necesario ejecutar el último proceso llamado corrección del terreno usando el rango Doppler (Range Doppler Terrain Correction), dentro de los parámetros modificables se selecciona para Digital Elevation Model la opción de SRTM 1SecHGT (Auto Download) que descarga de manera automática un DEM con resolución de 30m situado sobre Colombia mediante las coordenadas UTM Zone 18/ World Geodetic System 1984 escogidas en la pestaña Proyección de Mapa e interpoladas con un modelo Bilineal, finalmente se selecciona en aplicar normalización radiométrica, la cual la imagen orto-rectificada. El resto de parámetros seleccionados se ubican en la **Figura 19**.

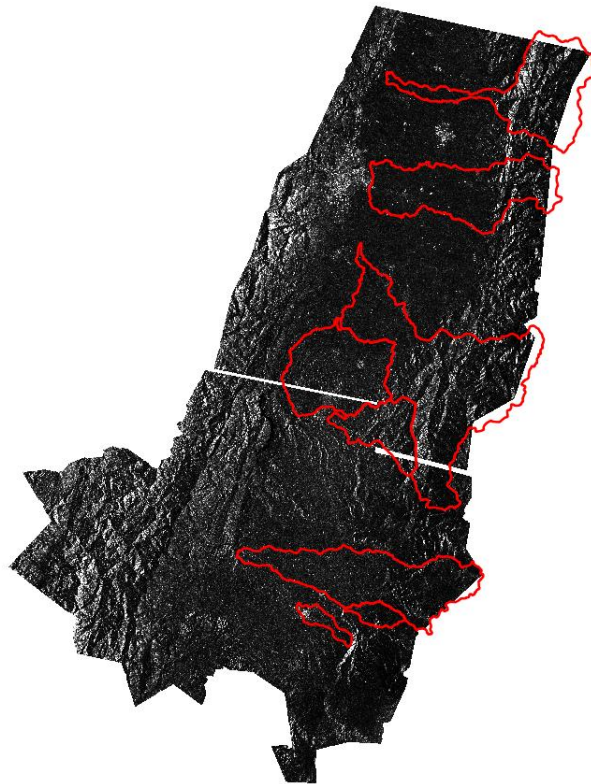
**Figura 19.** Corrección de terreno sobre imágenes Sentinel-1.



Al finalizar todos los procesos descritos con anterioridad se tienen las imágenes Sentinel-1 listas para procesar de forma individual o en conjunto con otros tipos de imágenes satelitales debido a que tienen las mismas geometrías de las imágenes ópticas.

A continuación, se presenta un ejemplo de cómo queda una imagen Sentinel-1 completamente procesada sobre gran parte de la Cuenca Alta del Río Cauca:

**Figura 20.** Imagen Sentinel-1 con polarización VV de la Cuenca Alta del Río Cauca.



Las imágenes Sentinel-1 muestran mediante el patrón de speckle características del terreno, por ejemplo, un patrón muy brillante muestra donde se encuentran las ciudades y patrones muy oscuros muestran por lo general cuerpos de agua. Con el procesamiento final de las imágenes Sentinel-1 se abre la posibilidad de extraer nuevas características, en el proyecto de grado se logran construir 25 características Sentinel-1.

Antes de iniciar con la extracción de características es necesario verificar el resultado del pre-procesamiento, más exactamente lo que compete al tamaño del pixel y los valores no definidos (nan) dentro de la zona de trabajo. Suele suceder que al terminar los procedimientos anteriormente descritos, no todas las imágenes quedan con el mismo tamaño de pixel, variando unas de otras en magnitudes del 1%, valor relativamente pequeño pero que significa la cuota de diferenciación entre las distintas coberturas, también es un punto importante tener en cuenta que a la hora de reducir la zona por medio de un recorte, los valores de los niveles de grises deben existir para cada uno de los pixeles presentes en la zona de interés, siendo necesario la implementación de tres algoritmos que entreguen: una escena totalmente recorrida con la certeza que ningún valor de pixel dentro de la misma va tener un valor no definido, el segundo algoritmo debe verificar y si es necesario cambiar los valores del pixel de las imágenes Sentinel-1 que no concuerden con la resolución espacial trabajada, que para este caso fue de 14.5228 en X y 14.529 en Y, finalmente se deben recortar a través de un polígono de la CARC o de la zona de interés, la sección de la cuenca que se desea trabajar. Todo lo anterior con el ánimo de minimizar el tiempo y los recursos computacionales necesarios para la extracción de características biofísicas sobre las zonas de interés.

### **3.3. Extracción de características biofísicas**

La plataforma de aplicaciones Sentinel (SNAP) reúne todas las herramientas necesarias para realizar un análisis sobre las imágenes de esta familia de satélites entre las funciones básicas que ofrece destacan las siguientes: abrir y explorar las componentes del producto, como bandas, máscaras y cuadrículas de puntos de unión. Las herramientas de navegación y la funcionalidad de información de píxeles también representan algunas de las capacidades básicas. De manera específica para el satélite Sentinel-1 posee un conjunto de herramientas de radar que facilitan la extracción de las diferentes características biofísicas de la imagen, entendiendo estas como aquellas que permiten describir el estado y los ritmos de cambio de una cobertura vegetal o cuerpo de agua. Algunos parámetros biofísicos son: la respuesta espectral, radiación fotosintéticamente absorbida por la cobertura, índice de vegetación de radar y humedad del suelo.

Por medio de los diferentes módulos que ofrece SNAP para el tratamiento de las imágenes de Radar se extraen las características que provienen de la reducción del ruido speckle a través de la opción Single Product Speckle Filter, escogiendo los filtros y sintonizando sus parámetros, trabajando con ventanas de 5x5, tamaño justo para que los resultados de los filtros no afecten la información dispuesta en la imágenes y a su vez cumpla con la propiedad de atenuar el moteado, este parámetro es el único que se debe alterar en los siguientes filtros: Boxcar, Median, GammaMap, Lee, Frost con un factor de amortiguamiento de 2, LeeSigma con un sigma de 0.9 y un número de vistas igual a 1, para los siguientes dos filtros no es posible asignar un número de filas y columnas para la ventana, se selecciona para el IDAN un número de vista igual a 1 y un vecindario adaptativo de 50, para finalizar solo es necesario situarse sobre Refined Lee ya que la configuración de sus parámetros se hace de manera interna. De igual forma con las herramientas dispuestas por SNAP se obtiene un conjunto de características extraídas a partir de los niveles de gris de la matriz de coocurrencia a través del análisis de textura dispuesto dentro de la opción análisis de imagen dentro del menú Raster.

Finalmente mediante algoritmos implementados en base a la teoría de la extracción de información en imágenes de Radar, se obtuvieron siete características más, fundamentadas en el realce de bordes, tonalidades, difuminación del moteado e identificación de texturas, estas son desviación estándar, promedio, mediana, textura U, textura VA, textura VI y textura VL. Como información adicional se generaron mapas de sombras para ubicar zonas con pérdida de información en las imágenes Sentinel-1, sobre estas zonas no se extraen puntos de muestras de coberturas.

Hasta el momento con los resultados alcanzados de los procesos descritos anteriormente se puede dar por culminado el segundo objetivo específico del presente trabajo de grado que consiste en caracterizar la información obtenida para la extracción de muestras biofísicas que permitan construir un modelo de clasificación de coberturas.

Para la base de datos óptica se trabajó con un total de 2 escenas Landsat 8, ortorectificadas, correspondientes a los path/row 009/057 y 009/058 del 29 de Diciembre del 2015, las cuales se usaron para obtener los índices de vegetación e información de reflectancia de la superficie y así caracterizar el área de estudio. Todos los datos del Landsat-8 se obtuvieron del Earth Explorer de United States Geological Survey (USGS).

Los índices de vegetación son menos sensibles al ruido de la imagen, a la geometría de visualización y a la atenuación atmosférica, haciéndolos particularmente ventajosos sobre los productos que usan únicamente la información reflectancia. De las imágenes Landsat-8 se extraen los siguientes índices de vegetación: Normalized difference water index (NDVI), Soil adjusted vegetation index (SAVI), Enhanced vegetation index (EVI), Normalized difference water index (NDWI), Simple ratio (SR), Transformed vegetation index (TVI), Atmospheric resistance vegetation index (ARVI), Structure insensitive pigment index (SIPI), Carotenoid reflectance index (CRI) además de las bandas 2,3,4,5,6,7 del satélite los cuales han demostrado utilidad en la cartografía de la biomasa (coberturas vegetales) y condiciones superficiales, compensando el brillo del suelo y los artefactos de fondo [31].

#### **3.4. Clasificación de la CARC**

Una vez se obtienen todas las características descritas anteriormente, se procede hacer una selección de coberturas por medio de la creación de polígonos usando algunas herramientas del software libre Q-GIS, que permite cargar imágenes georeferenciadas de la zona de donde se quieren extraer las muestras, lo correcto de este procedimiento sería llevarlo a cabo a lo largo de toda la CARC pero debido a lo extenso de este territorio se realiza sobre dos sub-cuencas representativas, Quinamayo y Palace, escogidas principalmente por su variedad en coberturas y centralidad en cada uno de los dos Path usados en este trabajo de grado (path/row 009/057 y 009/058).

La extracción de muestras consiste en obtener la mayor cantidad de puntos de cada una de las 7 coberturas a reconocer en este trabajo, las cuales son: cuerpos de agua, zonas urbanas, pastizales, cultivos, suelo desnudo, nubes y sombras de nube, siendo estas dos últimas necesarias en una clasificación óptica, así pues con la ayuda del satélite Landsat 8 se genera una imagen falso color de la zona con el siguiente orden de bandas: infrarrojo cercano. SWIR y rojo, y con las imágenes entregadas a través de Q-GIS con el plugin de vista de satélite se identifican las siguientes coberturas:

**Figura 21.** Cuerpos de agua seleccionados en parte de la CARC.

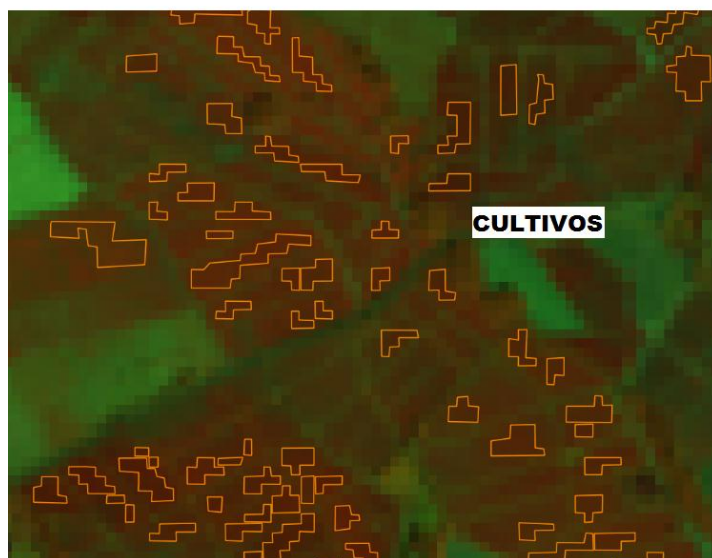




**Figura 22.** Zonas urbanas seleccionadas en parte de la CARC.



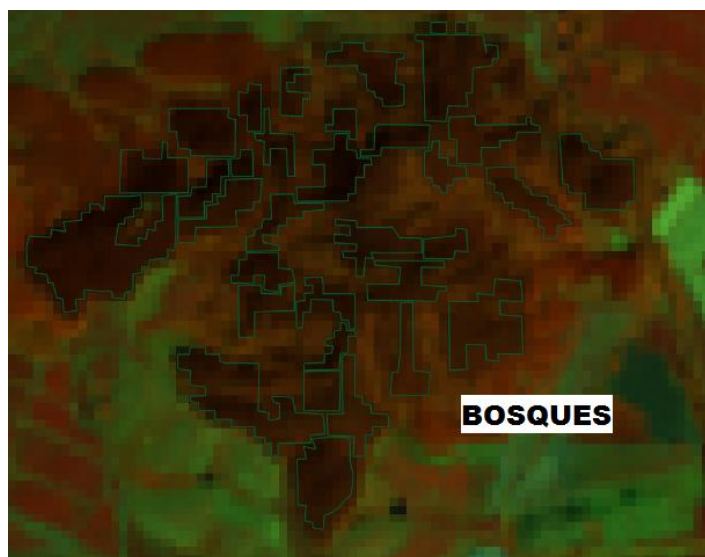
**Figura 23.** Cultivos seleccionados en parte de la CARC.



**Figura 24.** Pastizales seleccionados en parte de la CARC.



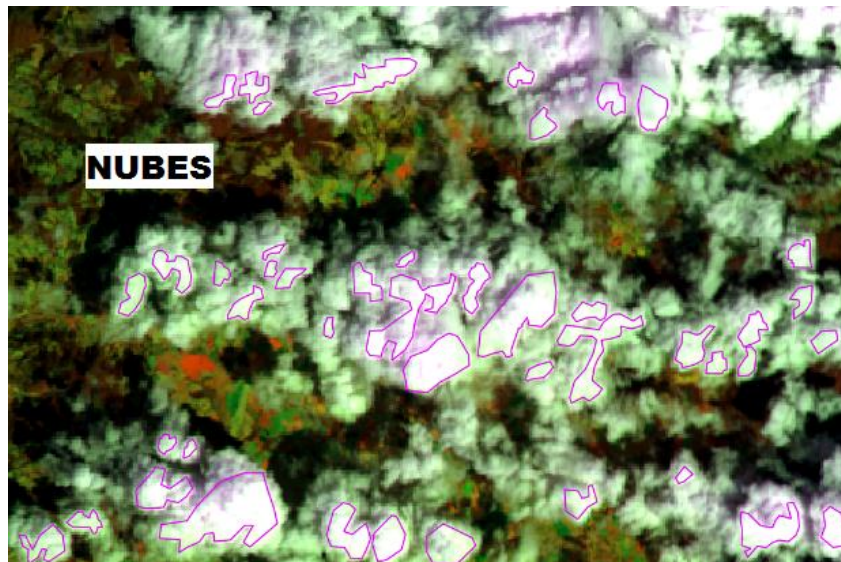
**Figura 25.** Bosques seleccionados en parte de la CARC.



**Figura 26.** Suelos desnudos seleccionados en parte de la CARC.



**Figura 27.** Nubes seleccionadas en parte de la CARC.



**Figura 28.** Sombras de nubes seleccionadas en parte de la CARC.

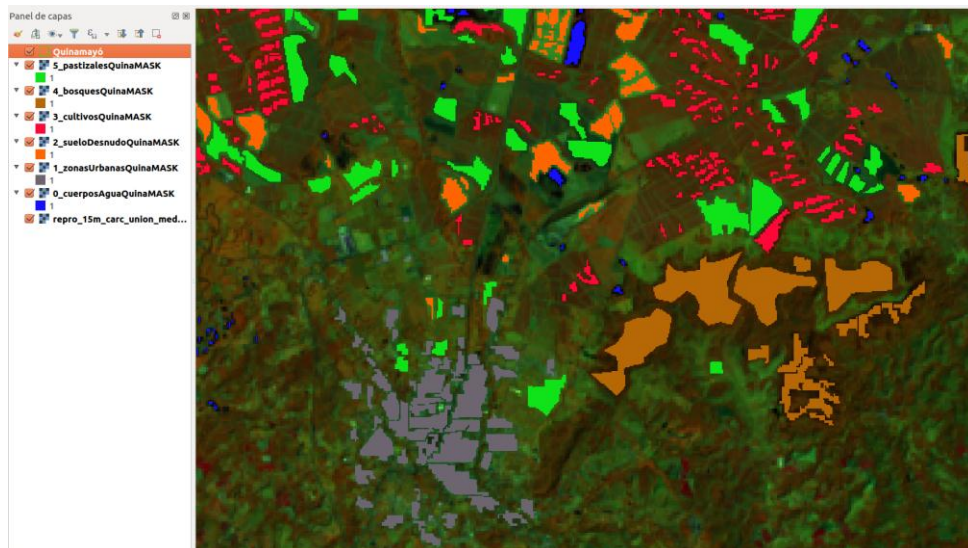


Este trabajo se desarrolló a partir de dos tipos de análisis para cada una de las imágenes presentadas anteriormente, donde se diferencian las zonas por colores. En primer lugar, se realizó un análisis cualitativo, abarcando de manera homogénea para cada cobertura la mayor cantidad de píxeles que cumplieran con la característica de ser parte de la misma, según la variedad espectral; por ejemplo, saber de antemano que se sitúa sobre una zona de pastizales que a su vez contenga diferentes tonalidades que representen esa cobertura, de tal forma que exista una variedad de coloración que se verá reflejada en los valores del píxel dentro de los polígonos. En segundo lugar, se llevó a cabo un análisis cuantitativo, debido que en algunas ocasiones es posible observar tonos similares para diferentes coberturas; por lo tanto, con el apoyo de los índices de vegetación ópticos SAVI, SIPI y NDVI, se logró estimar un intervalo de confianza entre los cuales se pudo diferenciar entre cada una de ellas y además tener certeza sobre qué zona se está trabajando; por ejemplo, en el caso de los cultivos el valor del NDVI se encuentra entre 0.65 – 1.00, para pastizales entre 0.30 – 0.50 y suelo desnudo entre 0.07 – 0.20 [32] [33]. El resultado de la combinación de estos dos tipos de análisis: cualitativo y cuantitativo, sobre la CARC otorgará un mayor grado de confiabilidad a la hora de realizar la extracción de píxeles sobre cada

cobertura, así la obtención de un modelo de clasificación se llevará a cabo sobre unas bases firmes de zonas con anterioridad bien identificadas.









Después de obtener los polígonos de cada cobertura, se extrajeron los valores de los píxeles usando la herramienta Clipper del programa Q-GIS, encontrada en el menú desplegado por Raster en opción de Extracción, como archivo de entrada se recibió la imagen en escala de grises que representa la zona trabajada; al ingresar el nombre de la capa de salida se continuó seleccionando en capa de máscara el shape de cada cobertura, manteniendo la resolución del raster de entrada sin crear una banda alfa de salida; cabe resaltar, que por defecto se trabajó el valor sin datos como cero y se realizó el anterior procedimiento para cada una de las áreas detectadas. El resultado de la extracción de coberturas se observa en la Figura 29, diferenciando con colores todas las máscaras extraídas para una parte de la sub-cuenca de Quinamayó.

**Figura 29.** Máscaras de coberturas extraídas sobre la sub-cuenca de Quinamayó.



En la **Tabla 2** se observan las máscaras de coberturas con sus respectivos colores.

**Tabla 2.** Código de colores para las diferentes máscaras de coberturas

Mascaras de coberturas	Identificador de la cobertura	COLOR
Cuerpos de agua	0	
Zonas urbanas	1	
Suelos desnudos	2	
Cultivos	3	
Bosques	4	
Pastizales	5	
Nubes	6	
Sombra de nubes	7	

Una vez obtenidas las máscaras, se procedió a la extracción de los valores del pixel y asignación de la etiqueta numérica para cada cobertura, sobre las 25 imágenes satelitales de radar que representan las características Sentinel-1 y las 15 imágenes complementarias del satélite óptico Landsat-8 extraídas para la zona de estudio. Lo anterior se llevó a cabo mediante un algoritmo que alberga en un archivo de configuración tres parámetros: el primero de ellos PathFileImage, recibe la ubicación de un documento de texto que contiene todas las direcciones de las 40 imágenes satelitales ópticas y de radar, ordenadas de la forma en la que se desea obtener el archivo final de pixeles; el segundo parámetro PathDirMascara, recibe la dirección de la carpeta donde se almacenan cada una de las máscaras con el nombre de la etiqueta numérica, unida mediante un guion bajo al nombre de la cobertura, de la siguiente forma: 1\_ZonasUrbanas; en el último parámetro FileName, se debe colocar el nombre del archivo de salida que almacenará todos los pixeles de interés.

Todos los procedimientos anteriormente descritos hacen parte de la metodología para una clasificación supervisada, con los resultados previamente obtenidos se llevarán a cabo los últimos pasos para conseguir la clasificación semi-supervisada de la CARC empezando por la creación del archivo de extensión .arff, que posteriormente recibirá la herramienta WEKA; el documento está compuesto por un encabezado con el nombre de todos los atributos, seguido por las etiquetas numéricas de las clases detectadas, finalizando con la base de datos conformada

por los valores de los píxeles de cada cobertura en cada una de las características trabajadas.

Weka (Waikato Environment for Knowledge Analysis, en español entorno para análisis del conocimiento de la Universidad de Waikato) es un software libre, desarrollado para tareas de minería de datos a través de la colección de algoritmos de aprendizaje automático. También contiene herramientas para el pre-procesamiento de datos, regresión, clustering y clasificación [34]; esta última utilizada para obtener el modelo que representó de manera acertada la distribución de la base de datos descrita anteriormente.

Utilizando la interfaz de usuario de Experimentador dispuesta por Weka en su menú de inicio, cargamos en la pestaña de Procesos la base de datos integrada en el archivo con extensión .arff seleccionando la opción Abrir documento, una vez visualizados los datos en pantalla, se usó la instancia Clasificar, escogiendo como método de aprendizaje maquina bosque aleatorio, el cual, el cual arroja óptimos resultados a diferencia de otros sistemas de clasificación [30] [35]; incluso, cuando existen más variables que observaciones o en su mayor parte son ruido. Es interesante el hecho de que no sobreajusta el modelo obteniendo gran capacidad de generalización. Además, aporta un ranking de importancia a las variables que determina cuál es la que más pesa en todo el proceso de decisión [36]. Por otra parte, tiene como ventajas: funcionar eficientemente en grandes bases de datos y permitir calcular proximidades entre pares de casos que sirven para identificar valores atípicos [37]. Las anteriores razones sumadas con la capacidad de Weka para entrenar el clasificador con el 70% de la base de datos ingresada y posteriormente verificarlo con el 30% restante mediante la opción de Porcentaje de división, hace que el modelo que caracteriza la dispersión de datos de entrada sea más fiable; al culminar el anterior proceso el software trabajado entrega como resultado del entrenamiento un archivo de salida con extensión .model, el cual se usó posteriormente para realizar la clasificación de las sub-cuencas seleccionadas para representar la CARC.






Con lo descrito anteriormente se pueden dar por culminado los dos últimos objetivos específicos del presente trabajo de investigación los cuales radican en construir un sistema basado en aprendizaje máquina que permita obtener un modelo de clasificación sobre la CARC para posteriormente sintonizar y validar los resultados del proceso sobre la región estudiada.

### 3.5. Análisis temporal y espacial desde las características Sentinel-1

Haciendo uso del portal en internet del Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM), se solicitan los datos de precipitación de la estación meteorológica Hacienda Lucerna ubicada a 12,8 km de los cultivos de caña estudiados en esta etapa de la metodología, para mediante la traza de graficas de precipitación y atributos Sentinel-1 a lo largo de los meses trabajados en este proyecto se realizara su respectivo análisis temporal.

Por último se efectúa un análisis espacial sobre las siguientes áreas: cultivos, suelos desnudos, pastizales, zonas urbanas y bosques, con la finalidad de obtener un valor representativo para cada cobertura en cada característica de radar implementada. Se procedió a ubicar sobre cada zona de interés un perfil de línea, un trazo que sirve como guía para posteriormente mediante un programa extraer los valores de los pixeles en la área dibujada, es muy importante recalcar que el perímetro sobre el cual se trazaron los perfiles de línea intentan ser los más homogéneos posibles, en la **Tabla 3** se pueden observar los recortes sacados para la cuenca del río Guachal, cabe aclarar que sobre las cinco sub-cuencas restantes también se disponen las anteriores muestras.

**Tabla 3.** Recortes obtenidos de cinco coberturas para la cuenca del río Guachal

Nombre de la Cuenca	Cobertura	Recorte	Longitud real del recorte (m)
Guachal	Cultivos		465
	Pastizales		465
	Suelos desnudos		465
	Zonas urbanas		465
	Bosques		465








Una vez se adquieren las muestras de cada cobertura se extraen los valores de los píxeles que están sobre la línea que atraviesa por la mitad el ancho del rectángulo usado para representar cada tipo de vegetación, posteriormente se realizan las gráficas de cómo cambian los valores de retrodispersión para cada característica de radar sobre la longitud del perfil de línea y finalmente se calcula un valor promedio sobre los píxeles extraídos de cada cobertura.




#### 4. ANÁLISIS Y RESULTADOS

Tres procesos de clasificación fueron realizados con las características ópticas y de radar extraídas a saber: clasificación de características ópticas, clasificación de características de radar y clasificación con características ópticas y de radar, el algoritmo de clasificación utilizado fue bosque aleatorio, este método fue escogido por su capacidad robusta a la hora de manejar conjuntos de datos grandes y diversos de manera eficiente, como los trabajados en este proyecto, además de ser una técnica altamente efectiva e implementada para el mapeo de cultivos, humedales, proliferación de algas, expansión urbana, biomasa, plantaciones, entre otras. basado en estudios previos los árboles de decisión también presentan mejores resultados frente a otros métodos de aprendizaje máquina [38]. Posteriormente, se realiza un proceso de selección de atributos para el análisis con full características, se lleva a cabo con la finalidad de disminuir el conjunto de características, el coste computacional e incrementar el porcentaje de acierto en la clasificación [39].

La cantidad de muestras extraídas para cada clase, la etiqueta y color asignado para cada una de las coberturas se presentan en la **Tabla 4**.

**Tabla 4.** Caracterización de las coberturas clasificadas

Cobertura	Cantidad de muestras	Etiqueta asignada por clase	Color asignado por clase
Agua	2041	0	
Zonas urbanas	2704	1	
Suelo desnudo	2146	2	
Cultivos	3582	3	
Bosque natural	2649	4	

Pastizales	3143	5	
Nubes	--	6	
Sombras de nubes	--	7	

Del análisis de las imágenes ópticas son extraídas 15 características y con el algoritmo de clasificación bosque aleatorio se procede a clasificar las coberturas presentes sobre la CARC con una resolución espacial de 15m, de esta manera para un total de 16625 puntos seleccionados para la evaluación, los cuales se distribuyeron un 70% de entrenamiento un 30% para prueba, los resultados fueron los siguientes: Instancias correctamente clasificadas igual a 77.22%, estadístico Kappa igual a 0.72, error medio absoluto de 0.10 y sensibilidad (recall) de 0.77.

Del análisis de la imagen de radar se obtienen 25 características con una resolución espacial de 15m, los resultados fueron los siguientes: Instancias correctamente clasificadas igual a 60.41%, estadístico Kappa igual a 0.51, error medio absoluto de 0.16 y sensibilidad (recall) 0.60.

Hasta el momento los mejores resultados se presentan para la clasificación con la base de datos óptica superando en cada indicador de precisión a la base de datos de radar, sin embargo se plantea la necesidad de fusionar ambas bases de datos en esta investigación. El objetivo es conseguir mejorar la precisión y poder de clasificación general del algoritmo. Un dataset con un total de 40 atributos es creado, este tiene todas las características extraídas de las imágenes Sentinel-1 y Landsat 8, el dataset contiene 16625 muestras, un 70% de ellas se utiliza para entrenamiento y un 30% para prueba. Los resultados obtenidos de la clasificación fueron los siguientes: Instancias correctamente clasificadas igual a 88.18%, estadístico Kappa igual a 0.85, error medio absoluto de 0.07 y sensibilidad (recall) de 0.88 respectivamente. La precisión y sensibilidad (recall) de cada clase clasificada (cobertura) obtenida del entrenamiento y prueba del modelo se presenta en la **Tabla 5**.

**Tabla 5.** Datos de precisión y sensibilidad (recall) para cada cobertura usando características ópticas y de radar

Cobertura	Etiqueta	Precisión	Recall
Cuerpos de Agua	0	0.912	0.902
Zonas Urbanas	1	0.926	0.948
Suelo Desnudo	2	0.779	0.716
Cultivos	3	0.884	0.904
Bosques	4	0.847	0.920
Pastizales	5	0.914	0.862

La **Tabla 6** Presenta la matriz de confusión del proceso de clasificación basado en el conjunto de características ópticas y de radar.

**Tabla 6.** Matriz de confusión para todas las características inicialmente construidas

Clases reales	Clasificadas como						Total	Precisión (%)
	0	1	2	3	4	5		
0	645	9	14	0	23	24	715	91.2
1	6	787	16	0	4	17	830	92.6
2	22	21	451	104	22	10	630	77.9
3	0	0	89	1002	17	0	1108	88.4
4	3	0	6	28	709	25	771	84.7
5	31	33	3	0	62	804	993	91.4
Total	707	850	579	1134	837	880	4987	88.2
Recall(%)	90.2	94.8	71.6	90.4	92.0	86.2	88.2	
Kappa	0.857							

El proceso de clasificación logra detectar con mayor precisión la cobertura correspondiente a zonas urbanas (clase 1), siendo el valor más alto alcanzado en todo el proceso y para todas las clases (0.926), la sensibilidad (recall) también es el más alto 0.948, lo que hace que las instancias incorrectamente clasificadas no se

distribuyan entre todas las coberturas, este caso es similar a lo que ocurre con las clases correspondientes a pastizales (clase 5) y cuerpos de agua (clase 0) que ocupan el segundo y tercer lugar respectivamente con los valores más altos de precisión. La cobertura de suelo desnudo (clase 2) tiene una precisión y sensibilidad (recall) buena con valores de 0.779 y 0.716 respectivamente, pero puede ser mejorada enriqueciendo el dataset para esta clase.

En los resultados conseguidos con la unión de los dos dataset óptico y radar se supera en cada instancia a la clasificación obtenida individualmente para cada uno de ellos. Con la finalidad de incrementar aún más la precisión y sensibilidad (recall) de los resultados y reducir el costo computacional que implica entrenar un dataset de 40 características y 16625 registros, se procede a realizar una selección de atributos, escogiendo únicamente aquellas características necesarias para obtener una mejor clasificación, como consecuencia de este proceso las características que se deben tener en cuenta son aquellas que superaran el 70% de importancia, por parte del satélite Sentinel-1 fueron: Idan, textura VL, GLCMCorrelation y GLCMMean y por parte del satélite Landsat-8 tenemos: banda 2,3,4,6 y 7, índice de vegetación CRI y NDWI. Usando el conjunto de datos entregado por la selección de atributos se obtuvieron los siguientes resultados: instancias correctamente clasificadas de 88.91%, estadístico kappa igual a 0.86, error absoluto medio igual a 0.06, con una precisión y sensibilidad (recall) general de 0.88 respectivamente.

La precisión y sensibilidad (recall) de cada cobertura obtenida mediante el clasificador se presenta en la **Tabla 7**.

**Tabla 7.** Datos de precisión y sensibilidad (recall) usando selección de atributos

Cobertura	Etiqueta	Precisión	Recall
Cuerpos de Agua	0	0.911	0.917
Zonas Urbanas	1	0.928	0.946
Suelo Desnudo	2	0.786	0.737
Cultivos	3	0.896	0.904
Bosques	4	0.861	0.925

Pastizales	5	0.920	0.872
------------	---	-------	-------

En la **Tabla 8**, se observa la matriz de confusión con selección de atributos, los estadísticos de precisión para cada clase también son mostrados:

**Tabla 8.** Matriz de confusión para el conjunto de datos con selección de atributos









Clases reales	Clasificadas como						Total	Precisión (%)
	0	1	2	3	4	5		
0	656	6	13	0	21	19	715	91.1
1	11	785	15	0	2	17	830	92.8
2	20	23	464	93	18	12	630	78.6
3	0	0	89	1002	17	0	1108	89.6
4	5	2	5	23	713	23	771	86.1
5	28	30	4	0	57	814	933	92.0
Total	720	846	590	1118	828	885	4987	88.9
Recall(%)	91.7	94.6	73.7	90.4	92.5	87.2	88.9	
Kappa	0.866							

Se observa que la selección de atributos mejora las estadísticas generales de la clasificación incrementando la precisión general en un 0.72%, junto con el estadístico Kappa de 0.866 y la disminución del error medio absoluto a 0.06, la selección de características aumenta la precisión del algoritmo y disminuye el costo computacional en el momento de trabajar con 12 características. Se evidencia en la Tabla 8 que los porcentajes de precisión y sensibilidad (recall) se incrementan para cada cobertura al aplicar la selección de características, las zonas urbanas (clase1) es la clase con mayor precisión en la detección. En general, con la selección de características los porcentajes de precisión y sensibilidad (recall) aumentaron para todas las clases.

Con los resultados obtenidos anteriormente, se optó por incrementar la base de datos de 16625 a 147953 registros, que incluyeron la mezcla de las imágenes

ópticas y de radar y por ende las 7 coberturas trabajadas con ellas, con la premisa de que el aumento traería consigo mejores efectos en el modelo de clasificación. La selección de puntos se hizo procurando mantener el equilibrio entre cada clase, por consiguiente se alcanzaron 18494 puntos por cobertura trabajada, esto con la finalidad de que en el momento de encontrar el nuevo modelo el clasificador no tienda a sobre-estimar la cobertura con una cantidad de puntos muy superior a las demás, los registros alcanzados para cada clase se observan en la Tabla 9.

**Tabla 9.** Caracterización de las coberturas clasificadas para la base de datos incrementada

Cobertura	Cantidad de muestras	Etiqueta asignada por clase	Color asignado por clase
Agua	18494	0	
Zonas urbanas	18494	1	
Suelo desnudo	18494	2	
Cultivos	18494	3	
Bosque natural	18494	4	
Pastizales	18494	5	
Nubes	18494	6	
Sombras de nubes	18494	7	

La base de datos incrementada se distribuyó de la siguiente forma: 70% para entrenamiento y 30% de prueba, los resultados fueron los siguientes: Instancias correctamente clasificadas igual a 99.49%, estadístico Kappa igual a 0.994, error medio absoluto de 0.005 y sensibilidad (recall) de 0.99, el modelo de clasificación se llevó a cabo mediante el algoritmo bosque aleatorio, lo descrito anteriormente se observa en la Tabla 11. Con los resultados obtenidos se logra mejorar la precisión y poder de clasificación general del algoritmo comparado con los modelos previos, en la **Tabla 10** se presentan algunas estadísticas obtenidas de cada cobertura.

**Tabla 10.** Datos de precisión y sensibilidad (recall) usando la base de datos incrementada

Cobertura	Etiqueta	Precisión	Recall
Cuerpos de Agua	0	0.996	0,993
Zonas Urbanas	1	0.992	0,990
Suelo Desnudo	2	0.997	0,993
Cultivos	3	0.995	0,997
Bosques	4	0.997	0,996
Pastizales	5	0.985	0,993
Nubes	6	0.999	1.000
Sombras de nubes	7	0.998	0.998

En la Tabla 11, se observa la matriz de confusión para el modelo obtenido de la base de datos enriquecida e igualada en registros, los estadísticos de precisión para cada clase también son mostrados:

**Tabla 11.** Matriz de confusión para la base datos enriquecida e igualada en registros

Clases reales	Clasificadas como								Total	Precisión (%)
	0	1	2	3	4	5	6	7		
0	5549	15	0	8	3	9	0	5	5589	99.6
1	7	5577	5	1	4	35	4	0	5633	99.2
2	1	8	5436	1	0	26	0	0	5472	99.7
3	3	0	0	5553	3	8	0	0	5567	99.5
4	2	0	0	11	5469	4	0	6	5492	99.7
5	0	22	9	5	5	5500	0	0	5541	98.5
6	0	0	0	0	0	0	5506	0	5506	99.9
7	9	0	0	0	3	1	0	5573	5586	99.8
Total	5571	5622	5450	5579	5487	5583	5510	5584	44386	99.5
Recall(%)	99.3	99.0	99.3	99.7	99.6	99.3	100.0	99.8	99.5	
Kappa	0.9943									

El modelo de clasificación con la base de datos incrementada e igualada en registros tiene una muy notoria mejoría en los procesos de detección de coberturas, pues no solo se logra una precisión en cada una de ellas mayor al 98% sino también una sensibilidad (recall) por encima del 99% lo que asegura que en las instancias donde el algoritmo se equivoque no lo hará eligiendo entre muchas clases sino entre unas pocas, mientras más cercano sea este valor a 1.00 disminuirá el número de instancias con las que la clasificación se confundirá.

El modelo de clasificación logra detectar con mayor precisión la cobertura correspondiente a nubes (clase 6), siendo el valor más alto alcanzado en todo el proceso y para todas las clases (0.999), la sensibilidad (recall) también es la más alto 1.00, este caso viene acompañado de los datos otorgados por la clase 7 sombras de nubes que con un 0.998 de precisión ocupa el segundo lugar, seguido de cerca por las coberturas bosques (clase 4) y suelo desnudo (clase 2). Cabe resaltar que el resto de coberturas tienen precisiones específicas altas superiores al 98%. En el caso del suelo desnudo (clase 2) se observa que adicionando más de sus muestras y trabajando con todas las características ópticas y de radar, su precisión aumenta considerablemente, pasando de 78.3% a 99.2%, con respecto al modelo clasificado con selección de atributos. La precisión y sensibilidad (recall) de todas las características mejora enriqueciendo el conjunto de datos para cada clase.

En la Tabla 12, se presenta en resumen todos los resultados de las estadísticas generales de cada proceso: individual y uniendo las dos bases de datos: óptica y radar.

**Tabla 12.** Datos estadísticos para las diferentes bases de datos (óptica, radar, radar-óptica con y sin selección de atributos)

Data	Instancias Correctamente clasificadas (%)	Estadístico Kappa	Error medio absoluto	Precisión	Recall
Óptica	77.22	0.72	0.100	0.77	0.77

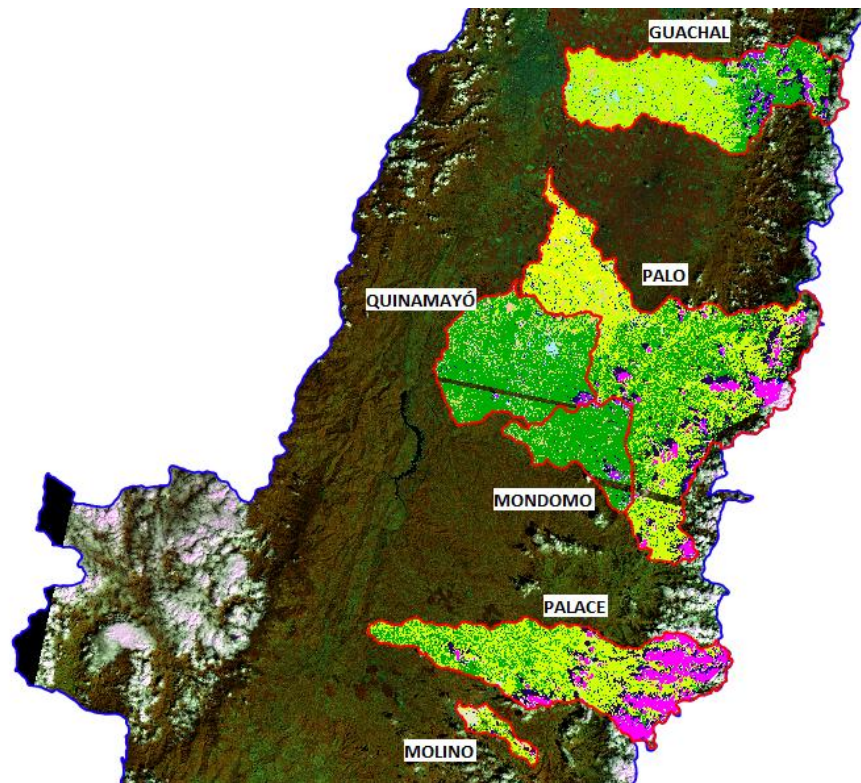


Radar	60.41	0.51	0.160	0.59	0.60
Óptica y radar sin selección de atributos	88.18	0.85	0.070	0.88	0.88
Óptica y radar con selección de atributos	88.91	0.86	0.060	0.88	0.88
Óptica y radar con incremento e igualación de registros	99.49	0.99	0.005	0.99	0.99

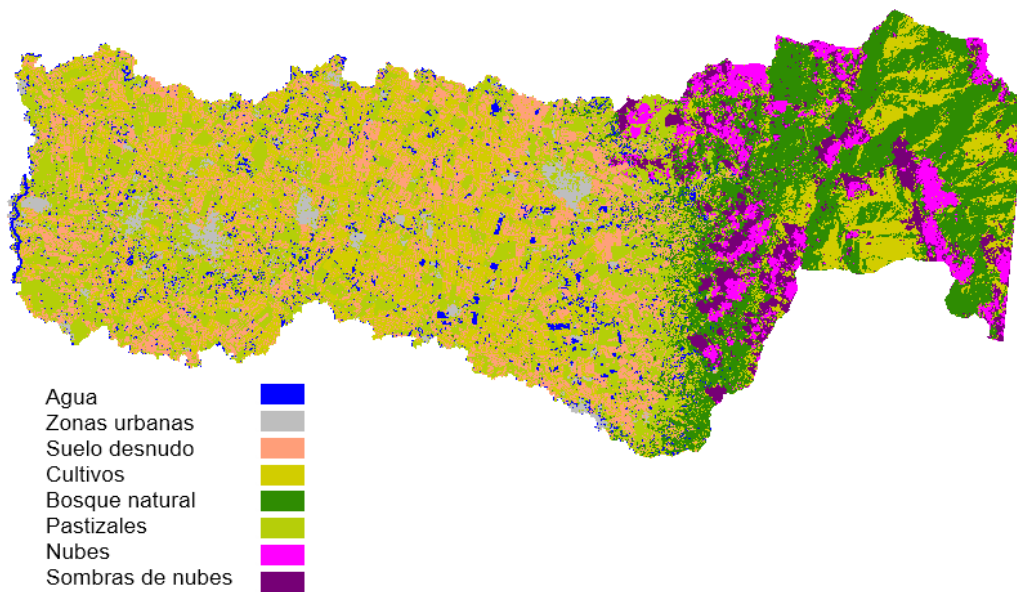
---

La aplicación de la mejor metodología (modelo óptico y radar con incremento e igualación de registros) se realiza en la cuenca del río Guachal como piloto para realizar la reconstrucción de la clasificación a 15m (ver **Figura 30**), con el objetivo de verificar la ubicación y área de las coberturas el estudio se enfoca en gran medida en los cultivos, puesto que estos presentan un alto nivel de variabilidad espacial, y están asociados con la principal actividad social y económica de la región como es la agricultura. También se adiciona la clasificación de 5 subcuencas: Quinamayó, Palo, Mondomo, Palacé y Molino (ver **Figura 31**) con la finalidad de estudiar a fondo la variabilidad de coberturas y lograr que estas sean una muestra representativa de la extensión territorial de la CARC.

**Figura 30.** Clasificación de coberturas para seis sub-cuencas de la CARC.



**Figura 31.** Clasificación de coberturas en la Cuenca del Río Guachal.



La clasificación que se observa en la **Figura 30** tiene una resolución espacial de 15m por las características del satélite Sentinel-1, con lo cual se logra detectar de mejor forma los límites de separación entre los distintos cultivos, además se observa visualmente que las clases pastizales, suelo desnudo y cultivos superan en tamaño a los bosques, esta información se puede corroborar calculando el área y porcentaje de ocupación sobre la clasificación final, los datos se muestran a continuación:

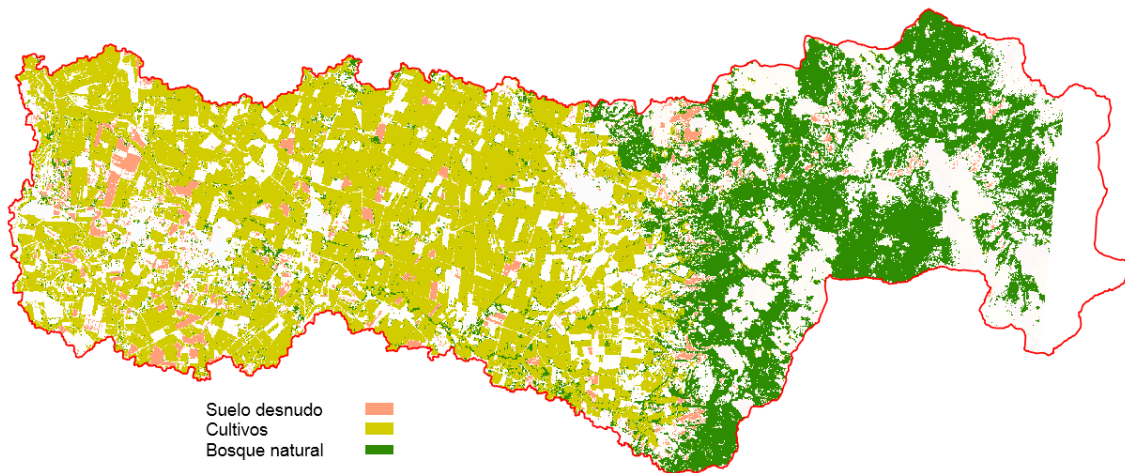
**Tabla 13.** Área de ocupación de las coberturas sobre la cuenca del río Guachal

Cobertura	Área (ha)	Porcentaje de ocupación efectivo (%)
Cuerpos de agua	1091,00	1,85
Zonas urbanas	2772,50	4,71
Suelo desnudo	2334,67	3,97
Cultivos	22112,60	37,62
Bosque natural	13612,60	23,16
Pastizales	12256,50	20,85
Nubes	2773,82	4,71
Sombra de nubes	1815,39	3,08
Total	58769,08	100,00
Área real	58398,90	

Como se verifica visualmente y cuantitativamente los Bosques solo ocupan el 23.16% del total de la cuenca del río Guachal, un resultado que inquieta puesto que una cobertura tan importante para los diferentes procesos ambientales como la conservación de los recursos hídricos, la disponibilidad y conservación de hábitats para plantas y animales, el control de los niveles de erosión, humedad y calidad del aire, y la nutrición del caudal viviente de la naturaleza sirviendo como fuente continua de proteínas, son algunas de las dinámicas dominantes sobre los bosques naturales, que deben ser tenidas en cuenta al momento de implementar otra clase de actividades que alteren el uso del suelo sobre esta cobertura. Como

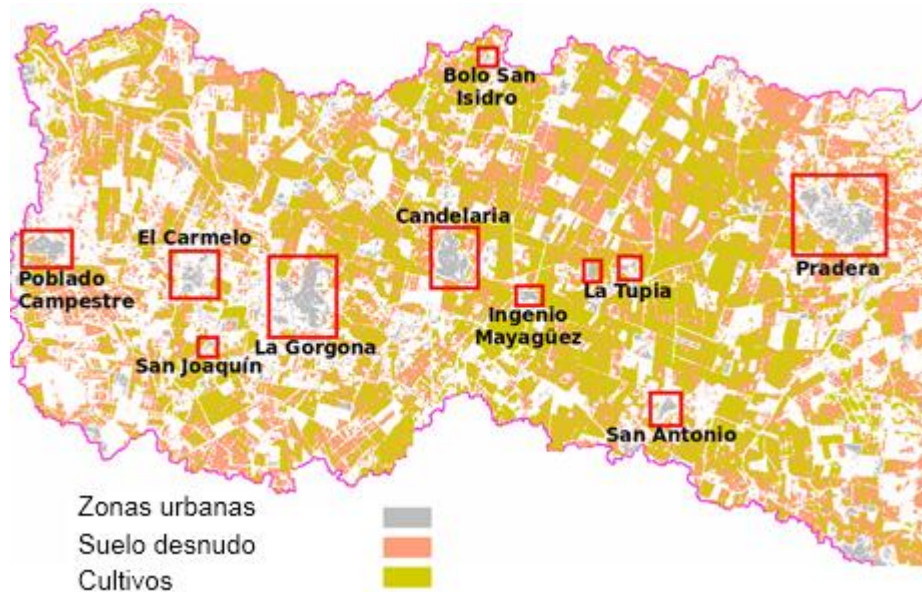
se puede observar en la **Figura 32** la expansión agrícola comienza amenazar el área boscosa sobre la cuenca del río Guachal, los suelos desnudos (color rosado) originados por la dinámica de los cultivos tienden a ocupar más área, haciendo evidente la diferencia en uso del suelo entre la zona plana y la zona de pendientes (ver Figura 34). Lo anterior también se puede verificar con las estadísticas obtenidas en la Tabla 13 donde los suelos desnudos, pastizales, cultivos y zonas urbanas (actividades humanas) tienen 25863,67 ha más que el área de los bosques naturales.

**Figura 32.** Suelo desnudo, cultivos y bosques naturales.



Basados en los resultados de la clasificación, los suelos desnudos y cultivos tienen una presencia importante sobre la cuenca del río Guachal, lo cual potencia los procesos erosivos debido a que la primera cobertura es una respuesta de la dinámica de cambios de la segunda, en la Figura 33 se resaltan los suelos desnudos, cultivos y zonas urbanas, junto con esta imagen se presenta el área de cada zona urbana en la Tabla 13.

**Figura 33.** Zonas urbanas, suelo desnudo y cultivos en la zona de la Cuenca del río Guachal.

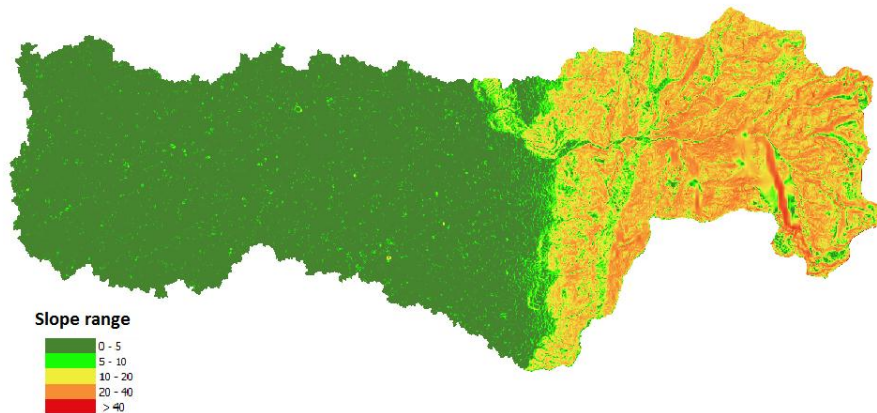


**Tabla 14** Área de ocupación de las zonas urbanas seleccionadas sobre la cuenca del río Guachal

Zonas Urbanas seleccionadas	Área (ha)
Pradera	154.395
Ingenio Mayagüez	14.535
La Tupia	23.152
Candelaria	87.885
La Gorgona	156.285
El Carmelo	57.465
San Joaquín	7.042
Poblado Campestre	56.79
Bolo San Isidro	3.712

En la Figura 34 un mapa de pendientes se anexa, donde se observa que las zonas urbanas y cultivos encuentran en los rangos de 0 a 5% correspondientes a una zona plana.

**Figura 34.** Mapa de pendientes en la Cuenca del río Guachal.



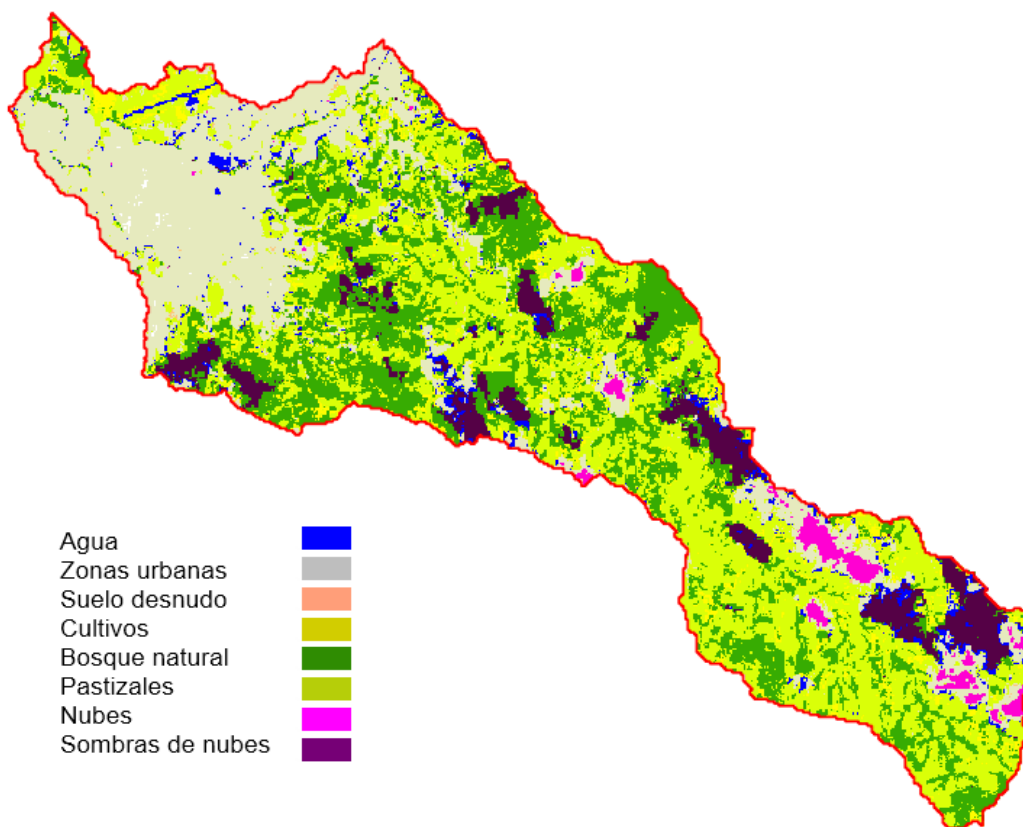
Basados en la información de la Tabla 13 aproximadamente 39000 hectáreas están dedicadas a actividades de siembra, manutención y cosecha de cultivos que representan la dinámica económica de la cuenca. Esta condición de uso y apropiación también denota la existencia de conflictos socio-ambientales por las actividades humanas y plantean una presión ambiental sobre los recursos naturales de estos ecosistemas específicamente en la invasión de terreno boscoso para agricultura y en la oferta y demanda de recursos hídricos para manutención de la misma. En este sentido, el territorio donde se cultiva la mayor parte de la caña de azúcar del país está ubicado en la (CARC), y específicamente en el Valle del Cauca, con una disponibilidad anual promedio de 467 m<sup>3</sup>/s, siendo aportados 272 m<sup>3</sup>/s por los diferentes ríos del departamento del Valle del Cauca. Esta zona se caracteriza por tener una elevada demanda de agua, la caña de azúcar tiene un uso intensivo del agua de 10.300 m<sup>3</sup>/ha Vs. 3.600 m<sup>3</sup>/ha y 3.400 m<sup>3</sup>/ha respectivamente para el maíz y la soja [40], también se concentra en esta área la mayor parte de la población y la actividad económica del Departamento, así de los 4,6 millones de habitantes del Valle del Cauca para 2015, 3.6 millones vivían en la zona plana. La mezcla de estos dos factores: la expansión agrícola y el asentamiento poblacional generan por una parte el decrecimiento de las zonas boscosas impulsado por la creación de nuevos cultivos o aumento en área de los ya existentes y un incremento del consumo de agua debido especialmente a los cultivos de caña. Esta situación genera una creciente concentración del recurso hídrico (y del suelo) en manos de un solo sector productivo: la caña de azúcar, mientras en

1980 la caña consumía el 23,6% del agua usada por la agricultura en la región, para 2009 ya se apropiaba del 58% y para el 2016 este supera el 62%. Esta condición se refleja en la pérdida de coberturas de esta cuenca y se verifica con el uso intensivo del suelo, lo cual se puede apreciar a partir del análisis espacial realizado.

El uso intensivo del territorio y de sus recursos interrelacionados en esta zona, comienza a evidenciar importantes conflictos por el uso del agua entre el consumo agrícola y el consumo humano. Como lo señala la Corporación Autónoma regional del Valle del Cauca, sobre la vertiente del río Cauca existe un desequilibrio temporal entre la oferta y la demanda de agua en los períodos de invierno y verano, desequilibrios que son resaltados igualmente por el Estudio Nacional de Aguas del IDEAM (2000), al afirmar que 14 de las 42 cabeceras del departamento del Valle del Cauca, ubicadas en la zona plana, tienen un índice de escasez entre medio alto y alto.

Sumado al análisis detallado que se realizó sobre la cuenca del río Guachal también se pueden observar comportamientos similares para las 5 sub-cuencas restantes de las que se obtienen modelos de clasificación y áreas de ocupación como se muestran a continuación:

**Figura 35.** Clasificación de coberturas en la Cuenca del Río Molino.

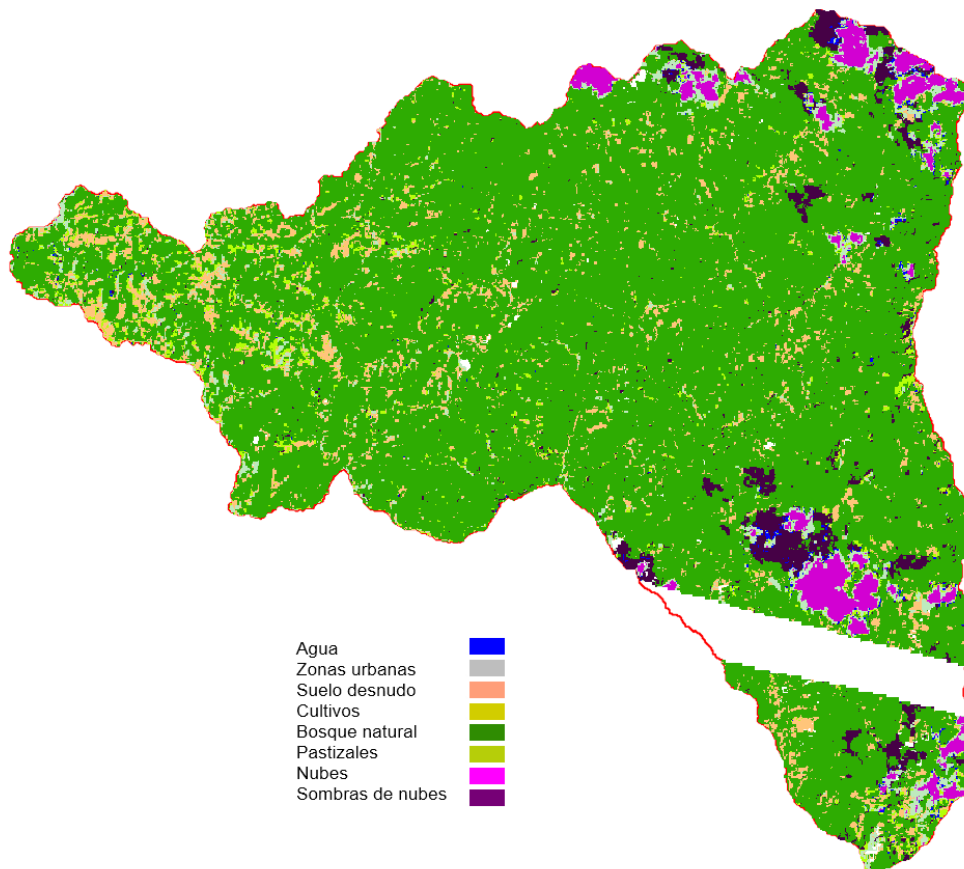


**Tabla 15.** Área de ocupación de las coberturas sobre la cuenca del río Molino

Cobertura	Área (ha)	Porcentaje de ocupación efectivo (%)
Cuerpos de agua	158,06	3,02
Zonas urbanas	1256,35	24,05
Suelo desnudo	7,15	0,13
Cultivos	332,75	6,36
Bosque natural	1285,58	24,60
Pastizales	1724,35	33,01
Nubes	369,40	7,07
Sombra de nubes	90,18	1,72
<b>Total</b>	<b>5223.85</b>	<b>100.00</b>
Área real	4926.70	



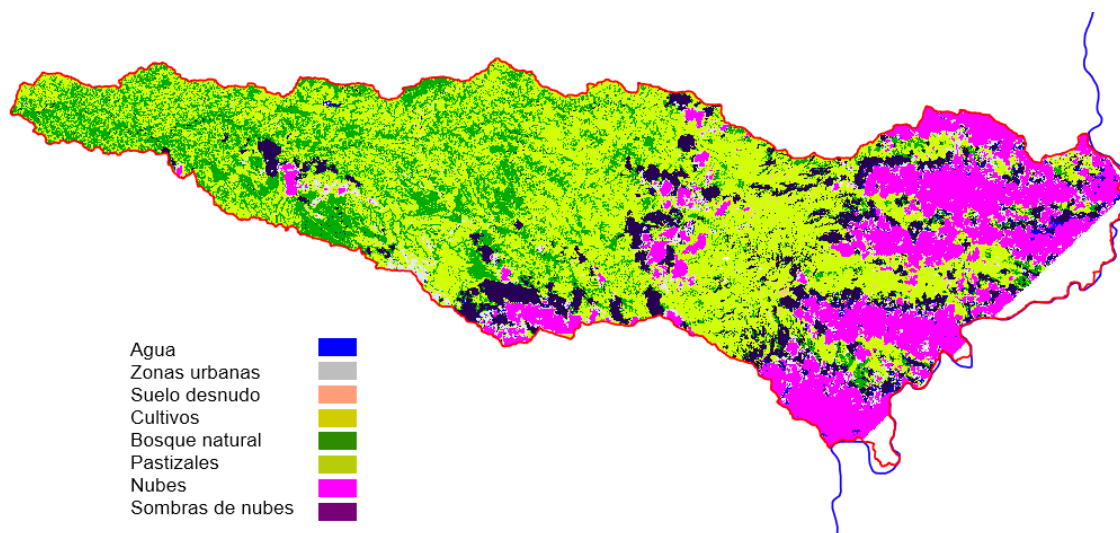
**Figura 36.** Clasificación de coberturas en la Cuenca del Río Mondomo.



**Tabla 16.** Área de ocupación de las coberturas sobre la cuenca del río Mondomo

Cobertura	Área (ha)	Porcentaje de ocupación efectivo (%)
Cuerpos de agua	90,27	0,40
Zonas urbanas	778,95	3,50
Suelo desnudo	1550,68	6,97
Cultivos	14,26	0,06
Bosque natural	17977,30	80,90
Pastizales	522,74	2,35
Nubes	699,16	3,14
Sombra de nubes	586,12	2,63
Total	22219.40	100.00
Área real	21808.80	

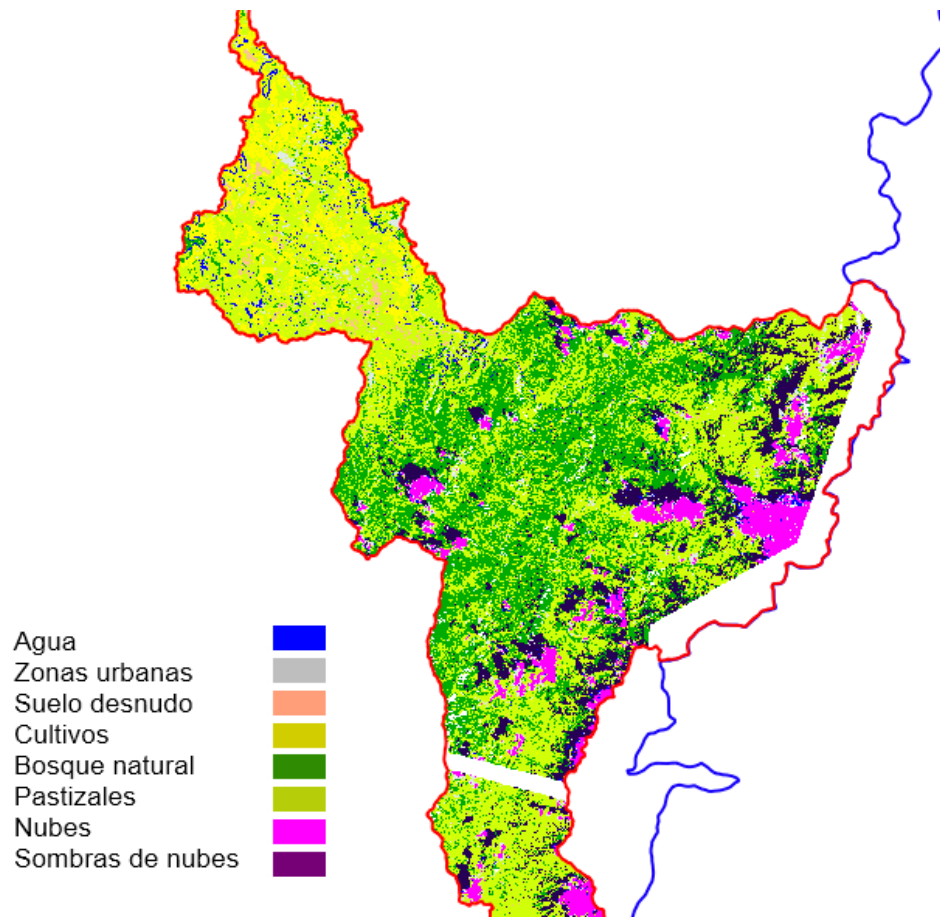
**Figura 37.** Clasificación de coberturas en la Cuenca del Río Palacé.



**Tabla 17.** Área de ocupación de las coberturas sobre la cuenca del río Palacé

Cobertura	Área (ha)	Porcentaje de ocupación efectivo (%)
Cuerpos de agua	790,85	1,19
Zonas urbanas	3202,63	4,83
Suelo desnudo	327,62	0,49
Cultivos	1091,25	1,64
Bosque natural	13342,50	20,13
Pastizales	29058,10	43,84
Nubes	6490,33	9,79
Sombra de nubes	11967,50	18,05
Total	66270,80	100,00
Área real	64295,30	

**Figura 38.** Clasificación de coberturas en la Cuenca del Río Palo.

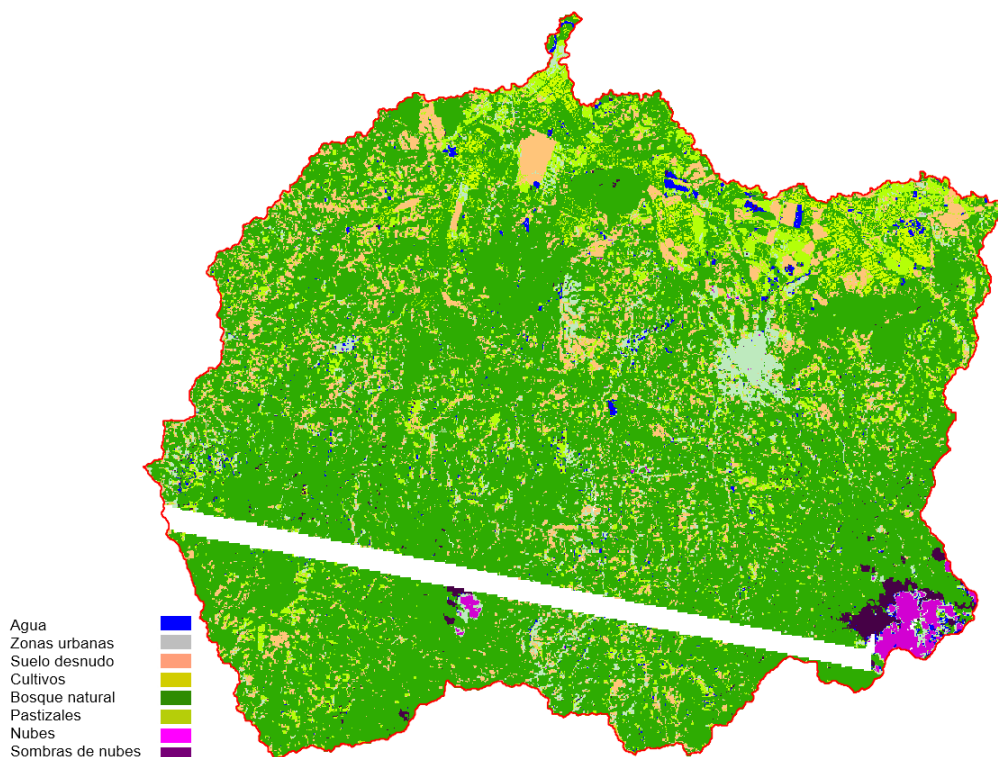


**Tabla 18.** Área de ocupación de las coberturas sobre la cuenca del río Palo

Cobertura	Área (ha)	Porcentaje de ocupación efectivo (%)
Cuerpos de agua	1894,48	1,65
Zonas urbanas	3899,61	3,40
Suelo desnudo	973,50	0,84
Cultivos	10129,00	8,83
Bosque natural	31857,30	27,79
Pastizales	47847,10	41,74
Nubes	11772,60	10,27
Sombra de nubes	6243,14	5,44

Total	114616.73	100.00
Área real	117279.00	

**Figura 39.** Clasificación de coberturas en la Cuenca del Río Quinamayó.



**Tabla 19.** Área de ocupación de las coberturas sobre la cuenca del río Quinamayó

Cobertura	Área (ha)	Porcentaje de ocupación efectivo (%)
Cuerpos de agua	468,76	0,95
Zonas urbanas	2345,47	4,76
Suelo desnudo	4811,69	9,76
Cultivos	1013,06	2,05
Bosque natural	35923,50	72,90
Pastizales	3906,81	7,92
Nubes	449,01	0,91

Sombra de nubes	354,82	0,72
Total	49273.12	100.00
Área real	48142.10	

Finalmente en la Tabla 20 se muestra el área de ocupación total de las coberturas sobre las seis sub-cuencas del río Cauca.

**Tabla 20.** Área de ocupación total de las coberturas sobre las seis sub-cuencas del río Cauca

Cobertura	Área (ha)	Porcentaje de ocupación efectivo (%)
Cuerpos de agua	4493,42	1,42
Zonas urbanas	14255,51	4,50
Suelo desnudo	10005,31	3,16
Cultivos	34692,92	10,96
Bosque natural	113998,78	36,03
Pastizales	95315,60	30,12
Nubes	22554,32	7,12
Sombra de nubes	21057,15	6,65
Total	316373,01	100.00
Área real	314850,80	

Usando el modelo de clasificación que surge de la base de datos incrementada en 18494 registros por cobertura, se logra extraer la información condensada en la Tabla 20, en la cual se evidencia que la instancia con más extensión territorial es la correspondiente a Bosques naturales con 113998,78 hectáreas seguida por pastizales con 95315,60 hectáreas, observando estas dos coberturas separadas de las demás no se percibe un conflicto ambiental ni mucho menos invasión a la zona arbórea pues el área de bosques sigue siendo mayor que la de pastizales por 18683,18 ha, pero no hay que ver de manera individual factores que son grupales, pues el uso de suelos para actividades agrícolas no solo abarca el producto

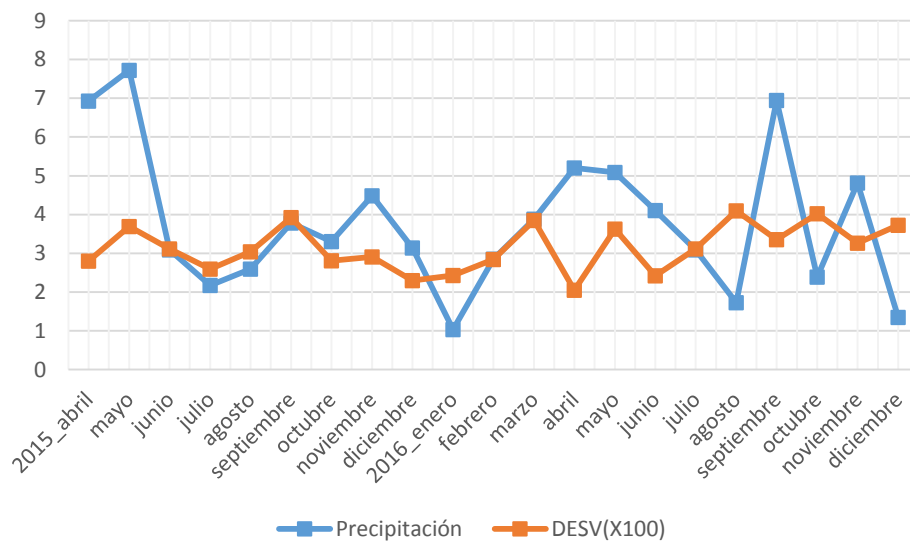
final que para este caso serían los cultivos sino también las coberturas que se generan camino a esa instancia, por lo tanto dos etapas se incluirían dentro de las dinámicas de rotación, la primera incluye los suelos desnudos con un área de 10005,31 ha y la segunda cultivos con una superficie igual a 34692,92 ha, sumadas las anteriores dos coberturas a el área de los cultivos se obtiene la superficie total de las instancias involucradas en los procesos agrícolas para las seis sub-cuencas seleccionadas, este valor es igual a 140013,83 hectáreas que representa el 44,24% del total de coberturas detectadas, estos nuevos valores calculados dan una aproximación más cercana a los fenómenos socio-ambientales que se presentan sobre parte de la CARC, donde el uso de suelo da más prioridad al desarrollo económico que al ambiental, otorgando 26015,05 hectáreas más de cultivos, pastizales y suelos desnudos que el área destinada a la cobertura de bosques, una condición no idónea para ninguna cuenca hidrográfica pues como se había mencionado anteriormente las zonas boscosas son un área importante para los diferentes procesos ambientales como la conservación de los recursos hídricos, la disponibilidad y conservación de hábitats para plantas y animales, el control de los niveles de erosión, humedad y calidad del aire, y la nutrición del caudal viviente de la naturaleza sirviendo como fuente continua de proteínas que deben ser tenidas en cuenta al momento de implementar otra clase de actividades que alteren el uso del suelo sobre esta cobertura y que por los resultados obtenidos en la clasificación no se están llevando acabo, además si se contrastan nuevamente el área destinada a la economía agrícola con la cantidad de recursos hídricos disponibles sobre las seis sub-cuencas detectados por el modelo de clasificación, la cobertura de cuerpos de agua con una área de 4493,42 no da abasto suficiente para cubrir con las necesidades humanas representada por las zonas urbanas ni mucho menos para continuar con la extensión tan prolongada de tierras dedicadas a la agricultura.

El modelo de clasificación obtenido así como los valores de área calculados para las sub-cuencas seleccionadas mediante la base de datos incrementada e igualada en registros da seguridad en los resultados alcanzados porque las estadísticas entregadas mediante la matriz de confusión así lo confirman, con un porcentaje

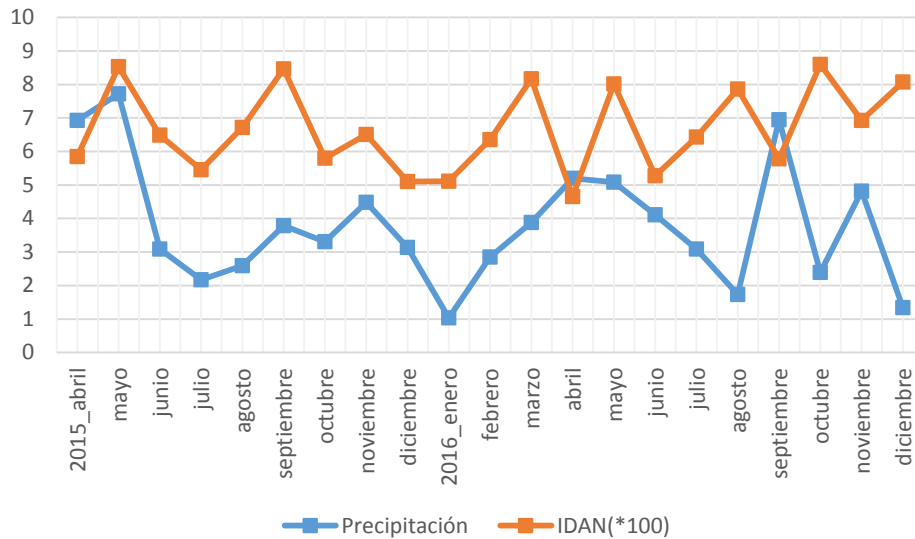
de precisión del 99% que otorga una cantidad de instancias correctamente clasificadas del 99.5% con un estadístico kappa igual a 0.99, además se valida mediante los porcentajes de estimación del suelo que factores como nubes y sombras de nubes que afectan a una base de datos compuesta por imágenes ópticas y de radar se vieron reducidos por la buena selección de escenas ocupando únicamente el 7% y 6% respectivamente del total de la imagen utilizada, dejando a disposición una mayor cantidad de área libre para la estimación del modelo de clasificación de coberturas.

Posteriormente se realiza sobre una zona de cultivos de caña un análisis temporal de los valores de precipitación mensual y las 25 características de radar utilizadas para este trabajo de grado, el periodo de estudio estuvo comprendido entre abril del 2015 y diciembre del 2016. A continuación se presentan algunos resultados representativos de los fenómenos observados.

**Figura 40.** Valores de precipitación mensual y característica desviación estándar a lo largo de los meses indicados.



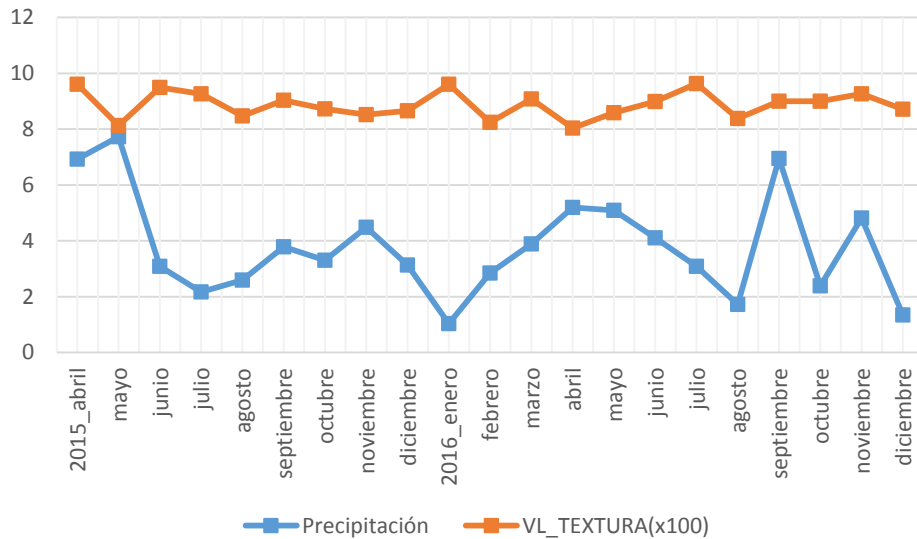
**Figura 41.** Valores de precipitación mensual y característica IDAN a lo largo de los meses indicados.



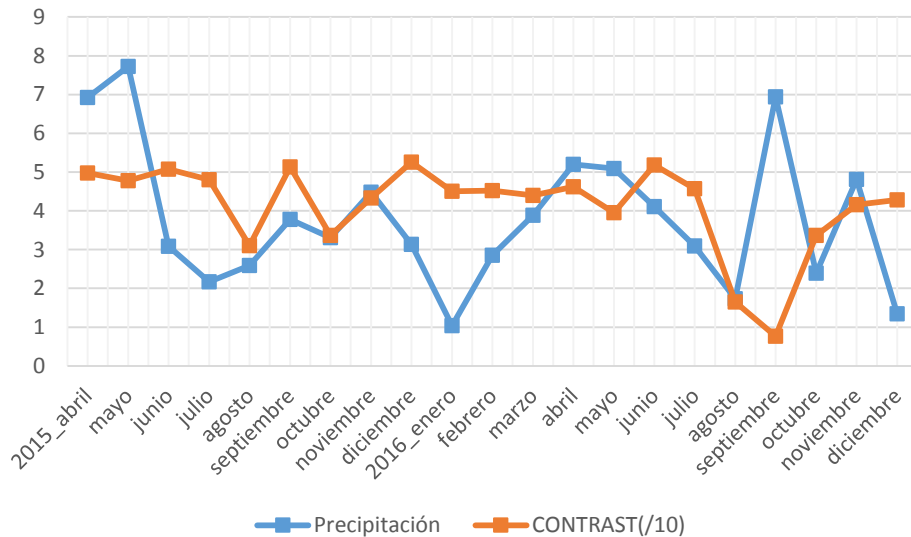
Para los meses comprendidos entre abril de 2015 y marzo de 2016 se evidencia de manera notoria la relación directa existente entre los valores de precipitación mensual, entregados por la estación meteorológica Hacienda Lucerna situada a 12,8 km del cultivo de caña seleccionado, y los valores de retrodispersión sobre la escena de radar correspondiente, de tal forma que, si los valores de precipitación incrementan o disminuyen sobre la zona de estudio así mismo lo harán los resultados de las características IDAN y desviación estándar, cabe aclarar que el anterior comportamiento no es descriptor de todos los meses en estudio, pues tanto en la Figura 40 y Figura 41 a partir de abril del 2016 hasta diciembre del mismo año las curvas de precipitación y características de radar no tienen una relación de proporcionalidad, lo anterior debido a lo apartada que se encontraba la estación meteorológica de la zona de estudio (cultivo de caña), así pues para ciertos días del mes, puede que el cultivo de caña seleccionado no se hubiera visto afectado por fenómenos de precipitación y sin embargo se reporte dentro de la gráfica un valor elevado producido principalmente por lluvias sectorizadas en zonas distintas a la seleccionada, situación que también se puede observar de manera más clara en la Figura 42 y Figura 43 donde en la mayor parte de la serie temporal no se evidencia correlación alguna entre la precipitación y los valores de retrodispersión para sus respectivas características de radar.



**Figura 42.** Valores de precipitación mensual y característica textura VL a lo largo de los meses indicados.



**Figura 43.** Valores de precipitación mensual y característica Contraste a lo largo de los meses indicados.

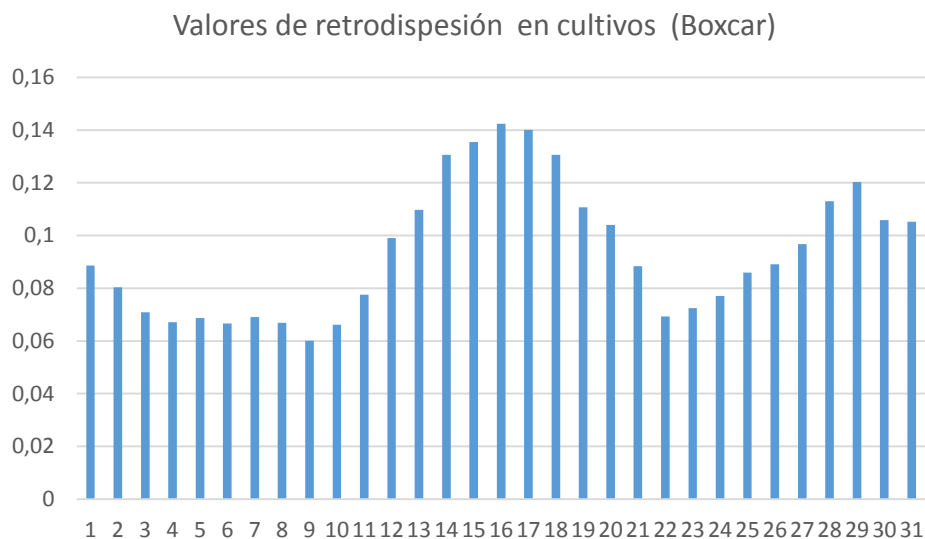


Como se manifestó anteriormente para la Figura 42 y Figura 43, la mayor cantidad de puntos e intervalos de tiempo no obedecen una relación de proporción directa entre los valores de precipitación y retrodispersión de radar, tanto para la características de contraste y textura VL, consecuencia de la falta de cercanía del cultivo de caña a la estación meteorológica, brindando esta última datos generales de precipitación de zonas que en ocasiones tenían fenómenos climáticos sec-

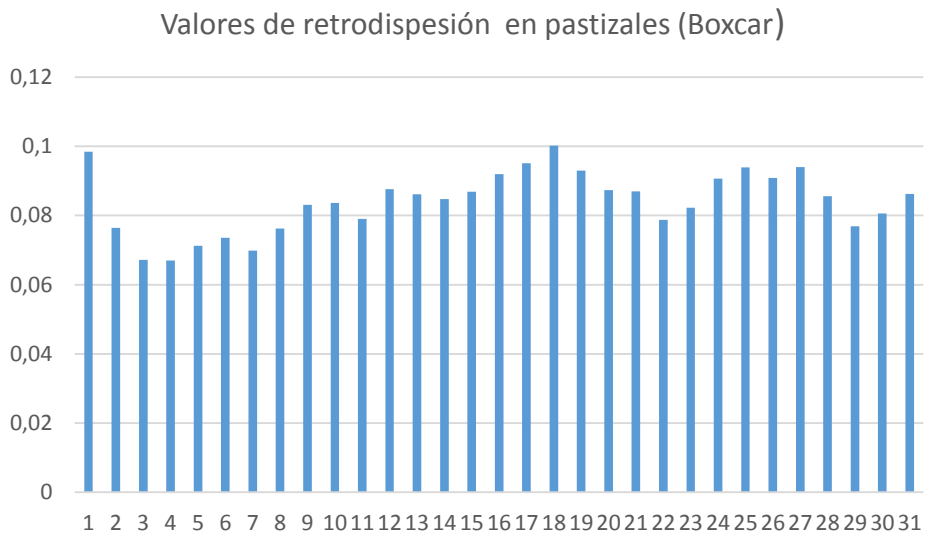
torizados que no representaban precisamente la dinámica de lo sucedido sobre el área de estudio.

Finalmente se realiza un análisis espacial sobre las siguientes áreas: cultivos, suelos desnudos, pastizales, zonas urbanas y bosques, con la finalidad de obtener un valor representativo para cada cobertura en cada característica de radar implementada, los resultados obtenidos a partir de la extracción de los valores de los pixeles sobre un perfil de línea y el posterior cálculo de promedio como valor representativo de cada zona se presentan a continuación:

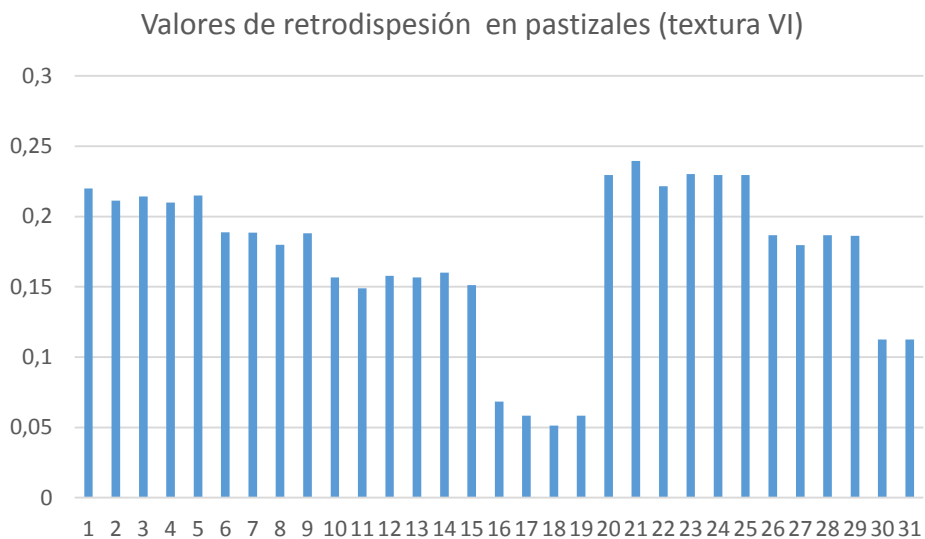
**Figura 44.** Valores de retrodispersión sobre la cobertura cultivos ubicada en la cuenca del río Guachal, para la característica Boxcar utilizando un perfil de línea.



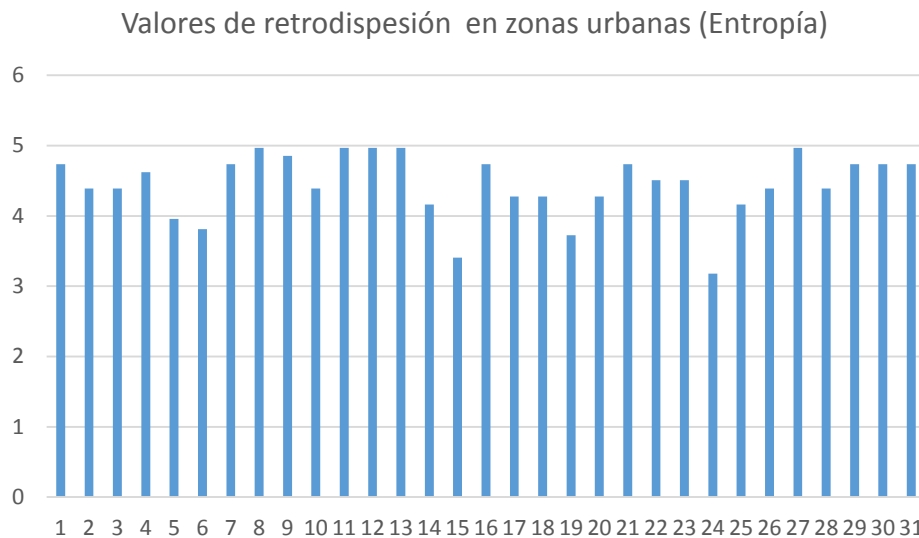
**Figura 45.** Valores de retrodispersión sobre la cobertura pastizales ubicada en la cuenca del río Guachal, para la característica Boxcar utilizando un perfil de línea.



**Figura 46.** Valores de retrodispersión sobre la cobertura pastizales ubicada en la cuenca del río Guachal, para la característica textura VI utilizando un perfil de línea.



**Figura 47.** Valores de retrodispersión sobre la cobertura zonas urbanas ubicada en la cuenca del río Guachal, para la característica Entropía utilizando un perfil de línea.



Los valores de retrodispersión obtenidos para las anteriores figuras surgen de un perfil de línea con longitud de 31 píxeles trazado sobre una imagen con resolución espacial igual a 15m, por ende la medida real de las coberturas biofísicas trabajadas sobre la cuenca del río Guachal fueron de 465 metros.

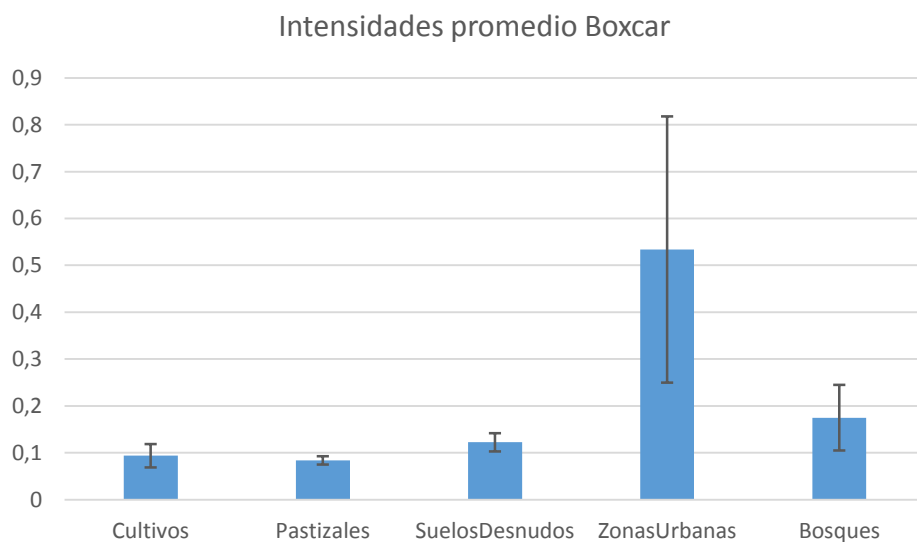
Se puede observar a través de las Figura 44 y Figura 45 el comportamiento de las coberturas cultivos y pastizales sobre la característica Boxcar, describiendo para la primera unos altos y bajos a lo largo de su extensión, esto debido a la presencia de una cobertura extra sobre lo que se creía ser únicamente cultivos, generando una mezcla espectral y por lo tanto diferentes valores de retrodispersión, particularmente es interesante observar en la Figura 45 el comportamiento sostenido de los valores de los píxeles para pastizales, los cuales se mantienen por encima de 0,06 hasta 0,1, este intervalo es el mismo que se presenta en la Figura 44 para las partes donde decrece la gráfica, esto permite inferir que la cobertura extra existente en el primer esquema es la correspondiente a pastizales, también es posible identificar un intervalo numérico que describa los cultivos a partir de la característica Boxcar, estos valores serían los comprendidos entre 0,11 y 0,14 para esta cobertura.

En la Figura 46 se evidencia una conducta similar a la descrita anteriormente, donde la gráfica no presenta un comportamiento homogéneo resaltando la existencia de más de una cobertura a lo largo del perfil de línea, esto se verifica también por la forma escalonada en la que van disminuyendo los valores de intensidad para los pastizales, pasando de una cifra máxima de 0,2 hasta llegar a 0,05, el intervalo de 0,2 a 0,15 será el que represente la cobertura pastizales para la característica de radar textura VI.

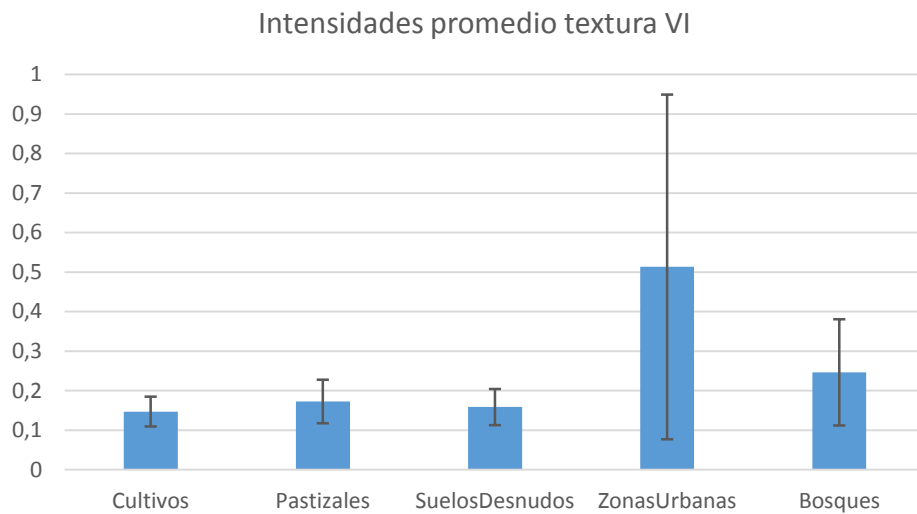
Los valores de intensidad del pixel entre los cuales se puede caracterizar la presencia de zonas urbanas en imágenes de radar para entropía son los comprendidos entre 3,3 y 5, ya que en este intervalo la Figura 47 presenta un comportamiento constante; cabe resaltar, que en esta imagen no hay presencia de coberturas diferentes a ciudades sobre el perfil de línea utilizado, dado que ninguna cifra de retrodispersión sale del rango especificado.

De los resultados obtenidos anteriormente para cada característica de radar se calcula un valor promedio que identifique a cada cobertura (ver Tabla 20); a continuación se observan aquellas figuras que representan la variedad de datos obtenidos para la separabilidad de capas naturales de vegetación:

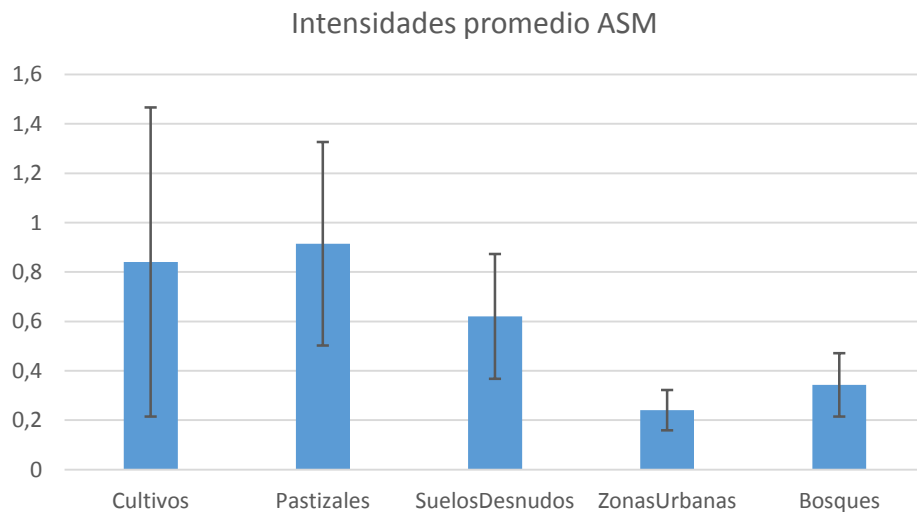
**Figura 48.** Valores de intensidad promedio de la característica de radar Boxcar para cinco coberturas biofísicas.



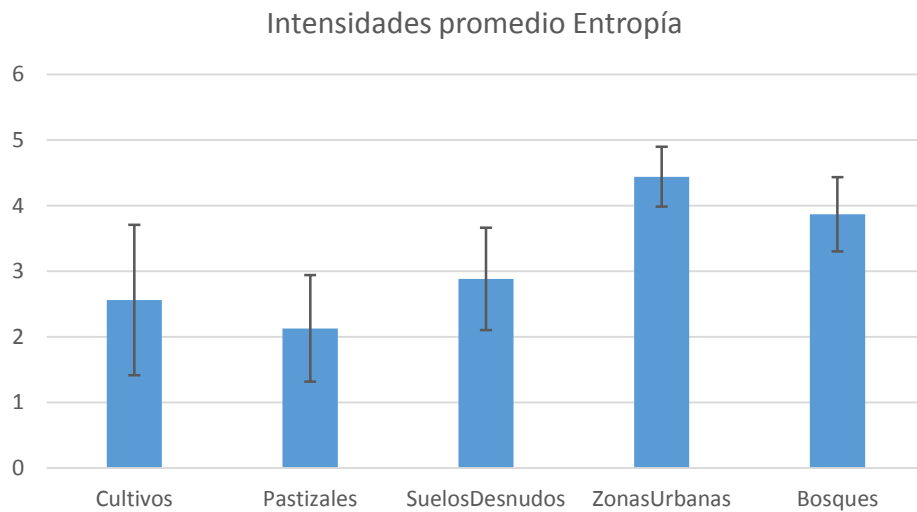
**Figura 49.** Valores de intensidad promedio de la característica de radar textura VI para cinco coberturas biofísicas.



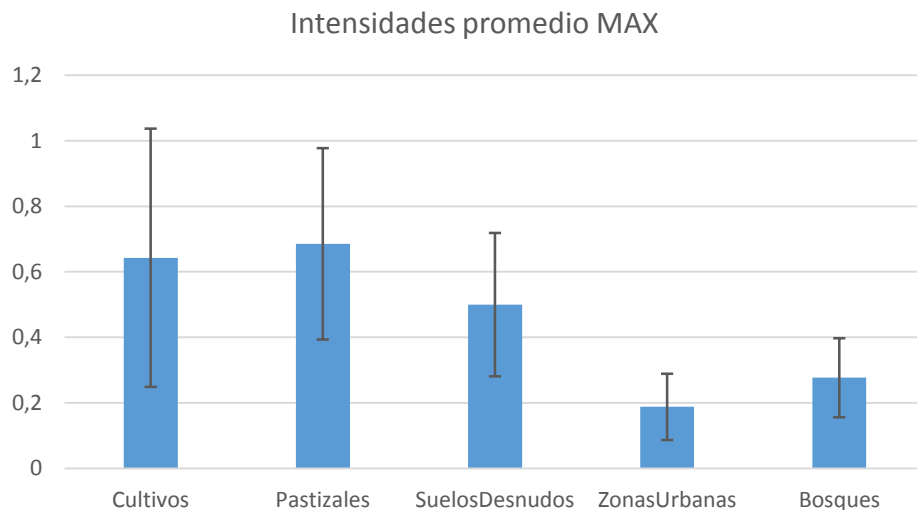
**Figura 50.** Valores de intensidad promedio de la característica de radar ASM para cinco coberturas biofísicas.



**Figura 51.** Valores de intensidad promedio de la característica de radar Entropía para cinco coberturas biofísicas.



**Figura 52.** Valores de intensidad promedio de la característica de radar MAX para cinco coberturas biofísicas.



**Tabla 21.** Consolidado del valor promedio de retrodispersión para cada tipo de cobertura en las características de radar más representativas

Tipo de cobertura	Característica de radar				
	Boxcar	Textura VI	ASM	Entropía	MAX
Cultivos	0,0937912	0,146995	0,840054	2,55994	0,642473
Pastizales	0,0840391	0,172543	0,913979	2,12765	0,685484
Suelos Desnudos	0,122363	0,158416	0,62052	2,88299	0,5

Zonas Urbanas	0,533792	0,513314	0,240591	4,44239	0,188172
Bosques	0,174926	0,246008	0,342742	3,86919	0,276882

Es una constante que en todas las características de radar seleccionadas la cobertura zonas urbanas sea aquella con mayor separabilidad, exceptuando los resultados correspondientes a la Figura 49 (Textura VI) en la que ninguna de las coberturas puede ser diferenciada, puesto que cada medida posee al menos un valor en el rango de incertidumbre de la otra. En la **Figura 48** se representan los valores promedios de Boxcar, para la cobertura zonas urbanas el rango de incertidumbre es el más alto pero a su vez el más distante del resto de valores, lo que hace posible identificar en la característica estudiada un rango aproximado sobre el cual tener certeza de la presencia de zonas urbanas sobre el resto de áreas vegetales sin tener en cuenta las áreas boscosas, que poseen un valor de incertidumbre superior con al menos un punto dentro del rango de zonas urbanas. Ahora bien, los resultados de las tres características de radar restantes: ASM, entropía y MAX correspondientes a las Figura 50, Figura 51 y Figura 52 respectivamente, agregan dos coberturas más al conjunto de las separables por medio de los valores de retrodispersión promedio obtenidos del perfil de línea trabajado para la zona de estudio, estas son bosques y pastizales, las cifras exactas para cada una de ellas se pueden apreciar en las Tabla 21. El anterior ejercicio se llevó a cabo con la finalidad de encontrar en las características de radar extraídas para este trabajo de grado aquellas en las que las coberturas detectadas fueran más separables numéricamente, a través del cálculo y posterior representación gráfica de los valores promedio de retrodispersión se logra establecer qué características favorecen más la separabilidad, los mejores resultados se obtienen de la característica entropía, se puede observar tanto en la Figura 51 como en los datos de la Tabla 21 que la diferencia entre cada valor promedio de retrodispersión para las capas de vegetación bosque, pastizales y zonas urbanas varían una de la otra en el orden de las décimas y con rangos de incertidumbre no cruzados entre sí, a diferencia del resto de características que también tienen valores distintos para cada cobertura pero con una proximidad numérica del orden de las centésimas junto con rangos de incertidumbre que se intersecan entre si impidiendo en algu-



nos casos distanciar todas las coberturas o como se había dicho anteriormente solo diferenciar de manera clara las zonas urbanas.

**Figura 53.** Imagen Falso color de la cuenca del río Palacé.



**Figura 54.** Mascara de sombras de radar presentes sobre la cuenca del río Palacé.



Sentinel-1 es inmune a la cobertura nubosa debido a la longitud de onda que maneja, lo que permite extraer información de las coberturas aún con presencia de nubosidad, esta extracción de información es posible mientras, en la imagen de radar, no haya pérdida de datos debido a las distorsiones geométricas. En la Figura 53 y Figura 54 se observan las zonas que no tienen información disponible sobre la cuenca del río Palacé para dos tipos de sensores: óptico y de radar, es posible detallar como la cobertura nubosa en la imagen óptica está concentrada so-

bre la zona de alta montaña de la cuenca, mientras que las distorsiones de radar se ubican en la parte baja de la región (zonas de color blanco en la Figura 54), por lo tanto al no existir una coincidencia espacial entre las dos zonas de pérdida de datos se puede recurrir al análisis espacial realizado de caracterización de coberturas biofísicas utilizando las características de radar y así lograr obtener información que en la imagen óptica está perdida. De forma similar, este mismo análisis puede ser llevado a cabo partiendo de las imágenes ópticas, encontrando información perdida debido a las distorsiones geométricas en las imágenes de radar, lo que conlleva a concluir que los dos tipos de información son complementarios y pueden obtener mejores resultados en procesos de análisis de coberturas y clasificación de imágenes satelitales.

## 5. CONCLUSIONES Y TRABAJOS A FUTURO

El pre-procesamiento de las imágenes de radar a través de la plataforma de aplicaciones Sentinel (SNAP) requiere un alto uso de memoria RAM para realizar los procesos de corrección de orbita, calibración, deburst, multilooking y corrección del terreno con efecto Doppler, en este trabajo se comprobó que antes de tener mayor frecuencia de procesador es necesario contar con un mínimo de 12 GB de memoria para realizar dichos procesos.

Obtenida la base de datos a través del satélite de radar Sentinel-1 se extrajeron 25 características biofísicas de las cuales cinco de ellas: boxcar, textura VI, ASM, entropía y MAX permiten diferenciar con su valor de retrodispersión las siguientes coberturas: zonas urbanas, pastizales, suelo desnudo, cultivos y bosques. Las cinco características establecen valores promedios con alta separabilidad entre zonas urbanas y bosques pero tres de ellas: ASM, entropía y MAX permiten diferenciar en su totalidad cada cobertura.

Observando la respuesta temporal de las características Sentinel-1 sobre un cultivo de caña se encuentra que existe una relación directa de proporción que afecta los valores de retrodispersión debido a la tasa de precipitación.

Un proceso de clasificación basado en el algoritmo de bosque aleatorio, utilizando imágenes satelitales ópticas y de radar fue realizado, uniendo las dos fuentes de información y conformando una base de datos de 18494 registros por cobertura se obtuvo una precisión general de 99,49%, el valor del estadístico kappa fue 0.99, obteniendo tasas de precisión para cada cobertura superiores al 98%. Con respecto a la selección de atributos se comprueba que las características superiores al 70% de importancia son las responsables del incremento en los estadísticos de precisión, aumentando en un 11,69% un proceso de clasificación realizado únicamente con características ópticas (Satélite Landsat 8), además, la selección de variables reduce el coste computacional en el nuevo conjunto de datos.

Basado en la clasificación de las seis sub-cuencas piloto seleccionadas para representar la CARC, se realizó una cuantificación espacial de las coberturas, don-

de se observó que las superficies clasificadas como suelos desnudos, pastizales y cultivos superan en 8,21% el área de los bosques, lo anterior evidencia un claro desequilibrio en el uso de suelos dando prioridad a las actividades agrícolas que a la conservación natural de la cuenca.

Como trabajos a futuro, se plantea agregar nuevas características al archivo de clasificación que involucren las dos polarizaciones del satélite sentinel-1, así mismo enriquecer el conjunto de datos de entrenamiento con una serie temporal más extensa y actualizada de la CARC con la finalidad de ver la evolución de las coberturas detectadas y mejorar los procesos de clasificación.

Por último, realizar la construcción de un índice de vegetación óptico a partir de las características extraídas de las imágenes de radar Sentinel-1 utilizando herramientas como la programación genética con el objetivo de detectar de manera más precisa las coberturas biofísicas de la región.

## 6. BIBLIOGRAFÍA

- [1] F. Capodici, G. Ciraolo, G. D'Urso, G. La Loggia, and A. Maltese, "Vegetation index retrieval by coupling optical and SAR images," 2012, p. 853118.
- [2] S. La Biodiversidad Y Asegurando Alimentación Para El Mundo, "LA BIODIVERSIDAD Y LA AGRICULTURA."
- [3] Lavado Angela Patricia, "Sector Agrícola de la Cuenca Alta del Río Cauca Altamente Vulnerable al Cambio Climático | DAPA." [Online]. Available: <http://dapa.ciat.cgiar.org/vulnerabilidad-del-sector-agricola-frente-al-cambio-climatico-en-la-cuenca-alta-del-rio-cauca/>. [Accessed: 31-Jul-2017].
- [4] L. V. Crowder, "Gramíneas y leguminosas forrajeras en Colombia.," *Boletín Técnico - Div. Investig. Agropecu.*, 1960.
- [5] M. Claudia Merino *et al.*, "Diagnóstico del estado de la Acuicultura en Colombia."
- [6] A. Ferro, "Advanced methods for the analysis of radar sounder and VHR SAR signals," no. December, 2011.
- [7] U.S. Department of the Interior and U.S. Geological Survey, "Landsat 8 | Landsat Missions." [Online]. Available: <https://landsat.usgs.gov/landsat-8>. [Accessed: 14-Mar-2018].
- [8] C. P. Ruiz, "Tipos de Sensores-Teledetección," vol. 8, p. 63.
- [9] J. L. Chavez, *Tratamiento digital de imágenes multiespectrales*. Lulu Com, 2012.
- [10] R. Touzi, "A review of speckle filtering in the context of estimation theory," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2392–2404, Nov. 2002.
- [11] A. K. Platónov, J. M. (José M. Redondo, and Universitat Politècnica de Catalunya. Laboratori d'Enginyeria Marítima., *Aplicación de imágenes de satélite SAR en los estudios de contaminación marina y de dinámica de las aguas en el mediterráneo noroccidental*. Universitat Politècnica de Catalunya, 2002.
- [12] European Space Agency, "Introducing Sentinel-1 / Sentinel-1 / Copernicus / Observing the Earth / Our Activities / ESA." [Online]. Available: [http://www.esa.int/Our\\_Activities/Observing\\_the\\_Earth/Copernicus/Sentinel-1/Introducing\\_Sentinel-1](http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Sentinel-1/Introducing_Sentinel-1). [Accessed: 15-Feb-2018].
- [13] Canada Centre for Remote Sensing, "Material educativo para tele-detección mediante radares."
- [14] B. Aldalur<sup>1</sup> and M. Santamaría, "Realce de imágenes: filtrado espacial," vol. 17, pp. 31–42, 2002.

- [15] U. de Murcia, "Técnicas de filtrado."
- [16] J. A. Sobrino, J. A. Sobrino, N. Raissouni, and N. Kerr, *Teledetección*. Universitat de València, 2000.
- [17] A. Lopes, R. Touzi, and E. Nezry, "Adaptive speckle filters and scene heterogeneity," *IEEE Trans. Geosci. Remote Sens.*, vol. 28, no. 6, pp. 992–1000, 1990.
- [18] R. P. Corredor, "Metodología para determinar la línea costera con imágenes radar en el Departamento del Magdalena," 2014.
- [19] Harris Geospatial Docs Center, "Adaptive Filters Documentation Center," 2018. [Online]. Available: <https://www.harrisgeospatial.com/docs/adaptivefilters.html#Enhanced2>. [Accessed: 07-Mar-2018].
- [20] G. Vasile, E. Trouve, Jong-Sen Lee, and V. Buzuloiu, "Intensity-driven adaptive-neighborhood technique for polarimetric and interferometric SAR parameters estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1609–1621, Jun. 2006.
- [21] Centro de Información Tecnológica (Chile), *Información tecnológica*. Centro de Información Tecnológica.
- [22] S. Putri Abdul Rahman Putra, S. Chong Keat, K. Abdullah, L. Hwee San, and M. Nawawi Mohd Nordin, "Texture Analysis of AIRSAR Images for Land Cover Classification."
- [23] Escuela superior de ingenieros, "Caracterización de Texturas. Métodos Estadísticos Basados en la Transformada Wavelet," vol. Capítulo 5.
- [24] M. Presutti, "La Matriz De Co-Ocurrencia En La Clasificación Multiespectral: Tutorial Para La Enseñanza De Medidas Texturales En Cursos De Grado Universitario. Tema do Trabalho: Recursos Didáticos para o Ensino de Sensoriamento Remoto," 2004.
- [25] M. Orradre Nazareno, "Metodos De Clasificación Digital De Imágenes Satelitales Para La Determinación De Usos De Suelo Agrícola En El Partido De Tandil."
- [26] E. Chuvieco, *Teledetección ambiental : la observación de la Tierra desde el espacio*. Ariel, 2010.
- [27] The Scientific blog of Ets Asset Management Factory, "Random forest vs simple tree | Quantdare," 2015. [Online]. Available: <https://quantdare.com/random-forest-vs-simple-tree/>. [Accessed: 15-Mar-2018].
- [28] L. Breiman, *Classification and regression trees*. Chapman & Hall/CRC, 1984.
- [29] D. R. Cutler *et al.*, "Random Forests For Classification In Ecology," *Ecology*,

vol. 88, no. 11, pp. 2783–2792, Nov. 2007.

- [30] L. Breiman, “Random Forests,” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
- [31] E. L. Pencue-Fierro, Y. T. Solano-Correa, J. C. Corrales-Munoz, and A. Figueroa-Casas, “A Semi-Supervised Hybrid Approach for Multitemporal Multi-Region Multisensor Landsat Data Classification,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 9, no. 12, pp. 5424–5435, Dec. 2016.
- [32] J. Ren, Z. Chen, Q. Zhou, and H. Tang, “Regional yield estimation for winter wheat with MODIS-NDVI data in Shandong, China,” *Int. J. Appl. Earth Obs. Geoinf.*, vol. 10, no. 4, pp. 403–413, Dec. 2008.
- [33] L. J. M. M. Sergio Andrés García J, “Método para identificación de cultivos de arroz (*Oryza sativa* L.) con base en imágenes de satélite.”
- [34] I. H. W. Eibe Frank, Mark A. Hall, “Data Mining: Practical Machine Learning Tools and Techniques,” 2016. [Online]. Available: <https://www.cs.waikato.ac.nz/ml/weka/citing.html>. [Accessed: 17-Apr-2018].
- [35] A. Liaw and M. Wiener, “Classification and Regression by randomForest,” vol. 23, 2002.
- [36] P. O. Gislason, J. A. Benediktsson, and J. R. Sveinsson, “Random Forests for land cover classification,” *Pattern Recognit. Lett.*, vol. 27, no. 4, pp. 294–300, Mar. 2006.
- [37] V. Urrea Gales, “Detección de interacciones genéticas asociadas a enfermedades complejas. Aplicación al cáncer de vejiga,” 2009.
- [38] N. Torbick *et al.*, “Monitoring Rice Agriculture across Myanmar Using Time Series Sentinel-1 Assisted by Landsat-8 and PALSAR-2,” 2017.
- [39] “Radar,” 2000. [Online]. Available: <http://www.radartutorial.eu/01.basics/Principio Radar.es.html>. [Accessed: 08-Feb-2018].
- [40] A. K. Chapagain and A. Y. Hoekstra, “Virtual water flows between nations in relation to trade in livestock and livestock products Value of Water,” 2003.