

**UNIVERSIDAD NACIONAL AGRARIA
LA MOLINA**

FACULTAD DE CIENCIAS FORESTALES



**ANÁLISIS DE LA COBERTURA Y USO DE
LA TIERRA UTILIZANDO IMÁGENES DE
RESOLUCIÓN ESPACIAL MEDIA PARA EL
DISTRITO DE SAN RAMÓN –
CHANCHAMAYO – JUNÍN – PERÚ**

Presentado por:

Sheila Elena Gamarra Agama

TESIS PARA OPTAR EL TÍTULO DE
INGENIERO FORESTAL

Lima - Perú
2017

ACTA DE SUSTENTACIÓN DE TESIS

Los Miembros del Jurado que suscriben, reunidos para calificar la sustentación del Trabajo de Tesis, presentado por la ex-alumna de la Facultad de Ciencias Forestales, Bach.

SHEILA ELENA GAMARRA AGAMA, intitulado “ANÁLISIS DE LA COBERTURA Y USO DE LA TIERRA UTILIZANDO IMÁGENES DE RESOLUCIÓN ESPACIAL MEDIA PARA EL DISTRITO DE SAN RAMÓN -CHANCHAMAYO - JUNÍN - PERÚ”.

Oídas las respuestas a las observaciones formuladas, lo declaramos:

.....

con el calificativo de

En consecuencia queda en condición de ser considerada APTA y recibir el título de INGENIERO FORESTAL.

La Molina, 13 de Julio de 2017

.....
Mg. Sc. Victor Manuel Barrena Arroyo
Presidente

.....
Ing. Carlos Rafael Vargas Salas
Miembro

.....
Ing. Wilfredo Salvino Ojeda Ojeda
Miembro

Ing. Ethel Rubin De Celis Llanos
Asesor

DEDICATORIA

A Dios por derramar sus bendiciones sobre mí y llenarme de su fuerza.

A mi familia por todo el esfuerzo y sacrificio para brindarme todo el amor, la comprensión, el apoyo incondicional y la confianza en cada momento de mi vida y sobre todo en mis estudios universitarios.

AGRADECIMIENTOS

El presente documento de tesis ha sido desarrollado con apoyo, de las siguientes personas y organizaciones a quienes se les reconoce.

Al Laboratorio de Teledetección Aplicada, por facilitarme las licencias de los software utilizados.

A la Sala de Observación – OTCA, por el uso de sus instalaciones.

A mi Asesora Ethel Rubin De Celis Llanos, que me ha orientado, apoyado y corregido en el proceso de elaboración de tesis.

Al profesor Pedro Vásquez Ruesta, por facilitarme los equipos de medición.

A Lizette Gamarra Agama y Miguel Noa Villanueva, por apoyarme en la fase de campo.

A Alejandra Zamora Ríos, por brindarme su apoyo en la redacción del documento.

RESUMEN

En la presente investigación se muestra una contribución a la planificación y ordenamiento del territorio del distrito de San Ramón, mediante el conocimiento de la cobertura y uso de la tierra, utilizando imágenes Landsat 8, sensor pasivo de resolución espacial media del año 2016, con el propósito de determinar, cuantificar y visualizar, las coberturas y usos de la tierra. La metodología empleada muestra resultados eficientes con un sobresaliente aporte del algoritmo Random Forest en el proceso de clasificación, con un 97,2 por ciento de confiabilidad. Se usó como base la leyenda CORINE Land Cover, alcanzando 3 niveles de detalle y 13 clases; así mismo se empleó el software ENVI EX para la segmentación que más se ajuste a la imagen. La consistencia del mapa muestra resultados eficientes con un 95,33 por ciento de precisión global del mapa y un valor del índice Kappa de 0,922. De acuerdo a los resultados de la investigación, la clasificación supervisada utilizando el algoritmo Random Forest puede ser una alternativa eficiente y precisa a implementar para clasificar imágenes satelitales de sensores pasivos, cuando se requiere delimitar la cobertura y uso de la tierra en la Amazonía peruana, con posibilidad de aplicación en otras zonas del país.

Palabras Clave: Cobertura de la tierra, Landsat 8, Clasificación supervisada, Random Forest, Segmentación de imágenes.

ÍNDICE GENERAL

	Página
I. Introducción	1
II. Revisión de literatura	2
1. Cobertura de la tierra	2
2. Uso de la tierra.....	2
3. Teledetección	2
3.1. Tipos de sensores.....	2
3.1.1. Sensores Pasivos.....	3
3.1.2. Sensores activos	3
3.2. Resolución de un Sensor.....	4
3.2.1. Resolución espacial.....	4
3.2.2. Resolución temporal	4
3.2.3. Resolución espectral	4
4. Imágenes de satélite.....	4
5. El programa Landsat.....	5
5.1. Landsat 8.....	6
6. Combinación de bandas	8
7. Pre- procesamiento de las imágenes.....	9
7.1. Corrección de imágenes	9
7.2. Corrección radiométrica.....	9
7.2.1. Corrección atmosférica.....	9
7.2.2. Corrección geométrica.....	10
7.3. Unión de bandas.....	10
7.3.1. Imágenes pancromáticas y multiespectrales	11
7.4. Sistema de proyección	11
7.5. Mosaico.....	12
7.6. Ventanas de corte.....	12
8. Procesamiento de imágenes.....	13
8.1. Segmentación	13
8.2. Áreas de referencia.....	15
8.3. Información generada para la clasificación	15
8.4. Clasificación	17
8.4.1. Árboles de decisión.....	18
8.4.2. Algoritmos de clasificación.....	19
9. Indicadores de validación.....	20
9.1. Matriz de confusión.....	21
9.2. Índice Kappa (IK).....	21
10. Leyenda.....	23
10.1. Leyenda de CORINE Land Cover	23
10.2. Salida de campo.....	24
11. Metadatos	24
12. Estudios relacionados a la cobertura y uso de la tierra.....	25
III. Materiales y métodos.....	26
1. Zona de estudio	26
1.1. Fisiografía y suelos	27
1.2. Hidrografía	27

1.3.	Clima.....	28
1.4.	Contexto biológico	28
1.5.	Clasificación ecológica.....	29
1.6.	Clasificación por tipos de bosque	29
1.7.	Uso de la tierra.....	30
1.8.	Marco socioeconómico	30
1.9.	Accesibilidad.....	31
2.	Materiales y equipos	32
2.1.	Materiales.....	32
2.2.	Herramientas de software	32
2.3.	Equipos.....	33
3.	Metodología.....	35
3.1.	Adquisición de las imágenes satelitales.....	35
4.	Pre-procesamiento de las imágenes satelitales	35
4.1.1.	Unión de bandas.....	35
4.1.2.	Corrección geométrica de las imágenes.....	36
4.1.3.	Generación del mosaico.....	36
4.1.4.	Corte del mosaico	36
4.2.	Procesamiento	38
4.2.1.	Generación de las clases de la clave de clasificación	39
4.2.2.	Clasificación.....	39
4.3.	Salida de campo.....	41
4.4.	Edición.....	42
4.5.	Validación de los resultados.....	42
IV.	Resultados y discusión	44
1.	Unión de bandas.....	44
2.	Corrección geométrica de la imagen	44
3.	Generación del mosaico.....	45
4.	Corte del mosaico	46
5.	Generación de las clases de la clave de clasificación	46
6.	Segmentación de la imagen	49
7.	Generación de las áreas de entrenamiento	49
8.	Información generada para la clasificación	50
9.	Salida de campo.....	51
10.	Caracterización de las coberturas presentes en el distrito de San Ramón.....	52
11.	Clasificación supervisada	59
12.	Cobertura y uso de la tierra del distrito de San Ramón.....	60
13.	Proceso de clasificación supervisada	62
14.	Validación del mapa.....	63
V.	Conclusiones	67
VI.	Recomendaciones	68
VII.	Referencias bibliográficas	69
VIII.	Anexos.....	76

Índice de tablas

	Página
Tabla 1:	Fechas de Lanzamientos de los satélites Landsat..... 6
Tabla 2:	Distribución de las bandas OLI y TIRS en el espectro electromagnético. 7
Tabla 3:	Especificaciones técnicas de las imágenes Landsat 8 7
Tabla 4:	Combinaciones de bandas más comunes aplicados para Landsat 8 8
Tabla 5:	Valoración del índice Kappa..... 22
Tabla 6:	Límites del distrito de San Ramón..... 26
Tabla 7:	Trayectos terrestres de Lima a San Ramón 31
Tabla 8:	Imágenes Landsat 8..... 35
Tabla 9:	Clave de clasificación de la cobertura y uso de la tierra para el distrito de San Ramón..... 47
Tabla 10:	Clave de clasificación utilizada para calibrar el modelo de la clasificación de la cobertura y uso de la tierra para el distrito de San Ramón..... 48
Tabla 11:	Superficie de los tipos de cobertura y uso de la tierra del distrito de San Ramón para el año 2016..... 60
Tabla 12:	Resultados de los indicadores del proceso de clasificación supervisada para la IT2..... 62
Tabla 13:	Resultados de los indicadores de precisión del mapa. 63
Tabla 14:	Matriz de confusión correspondiente a la clasificación del análisis de cobertura y uso de la tierra del distrito de San Ramón con puntos de validación de campo..... 65

Índice de figuras

	Página
Figura 1: Proceso de Unión de Bandas.....	11
Figura 2: Zonas UTM en el Perú.....	12
Figura 3: Vista de la evaluación de las segmentaciones en una escena.	14
Figura 4: Estructura de un árbol de decisión.....	19
Figura 5: Mapa distrital de la provincia de Chanchamayo.	26
Figura 6: Diagrama de la metodología.....	34
Figura 7: Diagrama del procesamiento.....	37
Figura 8: Diagrama detallado del procesamiento.....	38
Figura 9: Mapa con los puntos al azar distribuidos en el área de estudio.	43
Figura 10: Vista de una imagen Landsat 8, antes y después de la unión de bandas	44
Figura 11: Imagen Landsat 8 re-proyectada a UTM/WGS-84/Zona 18S	45
Figura 12: Mosaico generado para el área de estudio	45
Figura 13: Corte del mosaico para el área de estudio	46
Figura 14: Segmentación adecuada para el corte del mosaico (seg_2030.shp)	49
Figura 15: Área de entrenamientos distribuidos en la imagen.....	50
Figura 16: Lista de la información generada para la clasificación.	50
Figura 17: Mapa de la ruta y puntos colectados en campo.	51
Figura 18: Vistas de las imágenes de los centros poblados	52
Figura 19: Vistas de las imágenes de la Red vial.....	52
Figura 20: Vistas de las imágenes del mosaico agropecuario	53
Figura 21: Vistas de las imágenes de los bosques húmedos.....	54
Figura 22: Vistas de las imágenes de arbustos y matorrales.....	54
Figura 23: Vistas de las imágenes de la puna.....	55
Figura 24: Vistas de las imágenes de las áreas pedregosas naturales.....	55
Figura 25: Vistas de las imágenes de tierras desnudas o con poca vegetación.	56
Figura 26: Vistas de las imágenes de las zonas quemadas.	56
Figura 27: Vistas de las imágenes de tierras altoandinas sin vegetación.	57
Figura 28: Vistas de las imágenes de los bofedales.	57
Figura 29: Vistas de las imágenes de ríos y quebradas.....	58
Figura 30: Vistas de las imágenes de las lagunas.....	58

Figura 31:	Clasificación supervisada denominada (it_2.shp).....	59
Figura 32:	Superficie de los tipos de cobertura de la tierra del distrito de San Ramón para el año 2016.	61
Figura 33:	Superficie de los tipos de uso de la tierra del distrito de San Ramón para el año 2016.	62
Figura 34:	Mapa de cobertura y uso de la tierra del distrito de San Ramón para el año 2016	66

Índice de anexos

	Página
Anexo 1 Adquisición de las imágenes satelitales	76
Anexo 2 Pre-procesamiento de las imágenes Landsat 8.....	77
Anexo 3 Procesamiento	80
Anexo 4 Salida de campo	89
Anexo 5 Edición	98
Anexo 6 Verificación de los resultados	100
Anexo 7 Caracterización de las coberturas presentes en el distrito de San Ramón.....	102

I. INTRODUCCIÓN

En el valle de Chanchamayo se encuentra el distrito de San Ramón, caracterizado por un paisaje montañoso y una topografía compleja; también presenta bosques de neblina de la vertiente oriental de la cordillera de los Andes que forma parte de un extenso sistema de cabeceras de cuenca, cuya vegetación actúa como captadora natural de humedad procedente de la llanura amazónica. Esta interrelación vincula los bosques montanos de neblina con la llanura amazónica, aportadora de gran cantidad de humedad. Sin embargo, es uno de los ecosistemas más amenazados por la deforestación y el cambio climático. (Reynel 2012).

Dentro del valle el potencial de uso mayoritario es el forestal; sin embargo, las áreas de bosque se ven afectadas por el crecimiento demográfico, que favorece la deforestación debido a la conversión de tierras forestales a usos agrícolas. (Reynel 2012)

En el distrito de San Ramón, la presión de la deforestación es alta; sin embargo, aún se observan fragmentos o relictos de bosque intactos o casi intactos, resultado de la presencia de menores niveles de impacto antropogénico sobre la vegetación conforme se incrementa la distancia a los centros poblados y en áreas con limitaciones de tránsito. (Antón y Reynel 2004).

El conocimiento de la cobertura y uso de la tierra es un paso fundamental para la gestión sustentable de los recursos naturales, la conservación de la biodiversidad, así como para mitigar los impactos en ecosistemas y el cambio climático. (FAO 2015)

El estudio de investigación tiene como objetivo general contribuir a la planificación y gestión informada, mediante el conocimiento de la cobertura y uso de la tierra, información básica y esencial en la toma de decisiones en el distrito de San Ramón, como objetivo específico se realizó una clasificación supervisada, utilizando el algoritmo *Random Forest* e imágenes de resolución espacial media (Landsat 8)

II. REVISIÓN DE LITERATURA

1. COBERTURA DE LA TIERRA

La cobertura de la tierra es la cobertura biofísica tal cual se observa sobre la superficie de la tierra (Di Gregorio 2005), a través de sensores remotos, independiente de su origen; incluyen la vegetación natural o cultivada, afloramientos rocosos, depósitos de arena, hielo, agua, construcciones hechas por el hombre, etc. (Lara *et al.* 2002).

El conocimiento de la cobertura de la tierra es un paso fundamental para la gestión sustentable de los recursos naturales, la conservación de la biodiversidad, mitigación de los impactos en las dinámicas territoriales y en el cambio climático. (FAO 2015).

2. USO DE LA TIERRA

El uso de la tierra está caracterizado por las funciones que se desarrollan sobre la cobertura del suelo, actividades realizadas por el hombre para producir, cambiarla o mantenerla (Herrera *et al.* 2009), con la finalidad de obtener beneficios mediante la producción de bienes y servicios. (Lara *et al.* 2002).

3. TELEDETECCIÓN

La teledetección es una técnica a través de la cual se obtiene información de un objeto sin tener contacto directo con él. Esto es posible debido a la relación sensor – cobertura (Martínez y Díaz 2005), basado en las propiedades de la radiación electromagnética y en su interacción con la superficie terrestre. (Suárez 2014)

3.1. TIPOS DE SENSORES

Los sensores de Teledetección son instrumentos que transforman la radiación electromagnética en información perceptible y analizable. Existen sensores pasivos y activos (Sánchez 2012), como se explica a continuación:

3.1.1. SENSORES PASIVOS

Los sensores pasivos detectan la radiación electromagnética emitida o reflejada de fuentes naturales. Muchos sensores pasivos tienen sensibilidad multispectral, lo cual significa que recogen simultáneamente datos de diferentes bandas espectrales. Las mediciones sobre diferentes zonas del espectro electromagnético aportan una amplia información sobre distintos aspectos del medio ambiente. Por ejemplo, la radiación ultravioleta se utiliza para monitorizar los niveles de ozono en las capas altas de la atmósfera; las bandas visible e infrarrojas permiten determinar la salud de la cubierta vegetal (cultivos, bosques); los captadores sensibles a la radiación infrarroja térmica sirven para determinar la temperatura del suelo, de las nubes, y de la superficie de los mares. (Sánchez 2012)

3.1.2. SENSORES ACTIVOS

Los sensores activos poseen fuentes internas que generan artificialmente la radiación. Entre los sensores activos, distinguimos el radar y el lidar. El radar trabaja en una banda del espectro comprendida entre 1 mm y 1 m. Se basa en el principio de que las microondas artificiales enviadas en una dirección determinada chocan con los objetos y son dispersadas. La energía dispersada se recibe, se amplifica y se analiza para determinar la localización y las propiedades de los objetos. Puesto que puede medirse también el tiempo que tarda un pulso de radiación en ir y volver, puede conocerse la distancia recorrida y generar así modelos digitales de elevaciones. El radar es insustituible en zonas con cobertura nubosa persistente, debido a su capacidad para atravesar las capas nubosas. El uso de longitudes de onda mayores, que no se encuentran en los espectros solar o terrestre, implica la necesidad de utilizar sensores activos, es decir aquellos que generan su propio flujo de radiación. Su mayor ventaja es que debido a su elevada longitud de onda (muy superior al tamaño de las gotas de agua en la atmósfera) no resulta absorbida por esta; además tratándose de un haz artificial puede manipularse la forma en que se emite para, así, obtener el máximo de información. (Sánchez 2012)

Las señales de radar se usan, entre otras aplicaciones, para el seguimiento de la retracción de los casquetes polares, como indicador del calentamiento global. El Lidar (Light Detection And Ranging) es un captador activo, análogo al radar pero con tecnología láser. Se utiliza para aplicaciones de topografía de precisión desde aviones. (Sánchez 2012)

3.2. RESOLUCIÓN DE UN SENSOR

El número de píxeles que posee un sensor de satélite define la resolución del mismo; es decir, la capacidad de apreciar detalles de un determinado tamaño en una imagen captada por dicho sensor. (Sánchez 2012)

Una imagen satelital se caracteriza por las siguientes resoluciones:

3.2.1. RESOLUCIÓN ESPACIAL

La resolución espacial hace referencia al objeto más pequeño que se puede distinguir en la imagen. Está constituida por el tamaño del píxel que se puede observar en la imagen medida en metros sobre el terreno, esto depende de la altura del sensor con respecto a la tierra, el ángulo de visión, la velocidad de escaneado y las características ópticas del sensor. (Sánchez 2012).

De acuerdo a la finalidad, precisión del análisis y disponibilidad de recursos con que se cuenta existen sensores: de alta resolución (0.5 – 1 m.), media resolución (15 – 30 m.) y baja resolución (250 – 1000 m.) (REDD – GTZ 2010).

3.2.2. RESOLUCIÓN TEMPORAL

La resolución temporal indica cuál es la frecuencia de revisita del satélite; es decir, cuánto tiempo debe pasar para que el satélite pueda obtener una nueva imagen sobre un área particular. (Chuvieco, citado por Sánchez 2009)

3.2.3. RESOLUCIÓN ESPECTRAL

La resolución espectral indica el número y el ancho de las bandas espectrales que son registradas por un sensor. (Sánchez 2009) La banda es una región del espectro electromagnético en el que un conjunto de detectores del sensor satelital es sensible; cuanto más estrechas sean estas bandas mayor será la resolución espectral. (Núñez *et al.* s.f.)

4. IMÁGENES DE SATÉLITE

Las imágenes satelitales son representaciones visuales de la información capturada por un sensor montado en un satélite artificial (Duarte s.f.).

Estos sensores recogen información valiosa, dentro de rangos específicos de longitud de onda del espectro electromagnético, que es registrada en distintos canales o bandas espectrales, de las características de la superficie de la Tierra. (Canada Center for Remote Sensing, citado por Pérez 2012).

Estas imágenes son matrices de celdas llamadas píxeles, formadas por un determinado número de filas y columnas. Cada una de esas celdas representa un área geográfica indivisible, determinando así el detalle espacial mínimo que se puede distinguir dentro de la imagen. El tamaño del pixel varía dependiendo del satélite y de los sensores que tomen la imagen. Cada pixel contiene un valor numérico, que representa en promedio la cantidad de energía solar que esa superficie emite o refleja. Como la energía que refleja la superficie depende de lo que haya sobre ella, los sensores ubicados en los satélites captan distintos niveles y calidades de energía, que luego pasan a ser distintos colores en una imagen ya formada. (Chuvienco, citado por Pérez 2012). Para visualizar estas imágenes satelitales se combina las distintas bandas digitales mediante los tres colores primarios (azul, verde y rojo) que capta el ojo humano. (Canada Center for Remote Sensing, citado por Pérez 2012)

5. EL PROGRAMA LANDSAT

Antes de 1972, la idea de utilizar datos de satélite para la vigilancia terrestre, la cartografía o la exploración era un concepto visionario. Esto se convierte en realidad con la aparición del Programa Landsat, el cual se constituye en una serie de misiones de observación de la Tierra por satélite gestionadas conjuntamente por la NASA y el Servicio Geológico de Estados Unidos (USGS). El programa Landsat ha revolucionado la forma de ver y estudiar nuestro planeta. Esta serie de datos, que se inició en 1972, es la más larga de la historia y continua registrando los cambios en la superficie terrestre desde el espacio. Landsat ha sido el único sistema de satélite diseñado y operado para observar repetidas veces la cubierta de la tierra con una resolución moderada. En la actualidad el programa se encuentra en su octava versión denominada: “Landsat Data Continuity Mission” (LDCM). Se trata del octavo satélite de observación de la serie Landsat y continuará el legado de archivo de los anteriores satélites, convirtiéndose de esta manera en el futuro de los satélites de observación de la tierra de mediana resolución con más historia. Este programa amplía, mejora y avanza el registro de imágenes multiespectrales, manteniendo la misma calidad de sus siete predecesores. (IGAC 2013)

Tabla 1: Fechas de Lanzamientos de los satélites Landsat

Satélite	Lanzamiento	Condición	Sensores
Landsat 1	23 Julio 1972	Dado de baja el 06 Enero 1978	MSS/RBV
Landsat 2	22 Enero 1975	Dado de baja el 27 Julio 1983	MSS/RBV
Landsat 3	05 Marzo 1978	Dado de baja el 07 Setiembre 1983	MSS/RBV
Landsat 4	16 Julio 1982	Dado de baja el 15 Junio 2001	MSS/TM
Landsat 5	01 Marzo 1984	Dado de baja el 2013	MSS/TM
Landsat 6	05 Octubre 1993	No entró en órbita. Perdido	ETM
Landsat 7	15 Abril 1999	Operacional con defectos en sensores	ETM+
Landsat 8	11 Febrero 2013	Operacional	OLI/TIRS

FUENTE: Fernández y Gilbes s.f.

5.1. LANDSAT 8

El satélite Landsat 8 incorpora dos instrumentos de recolección de datos: el Operational Land Imager (OLI), y un sensor térmico infrarrojo llamado Thermal Infrared Sensor (TIRS).

Las bandas espectrales del sensor OLI proporcionan una mejora de los instrumentos, debido a la incorporación de dos nuevas bandas espectrales: un canal profundo en el azul visible (banda 1), diseñado específicamente para los recursos hídricos e investigación en zonas costeras, y un nuevo canal infrarrojo (banda 9) para la detección de nubes cirrus. Adicionalmente una nueva banda de control de calidad se incluye con cada producto de datos generado. Esto proporciona información más detallada sobre la presencia de características tales como las nubes, agua y nieve. Por otra parte el sensor TIRS recoge dos bandas espectrales en longitudes de onda incluidas por la misma banda en los sensores TM y ETM+. (IGAC 2013)

Tabla 2: Distribución de las bandas OLI y TIRS en el espectro electromagnético.

Landsat 8 Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS)	Bandas	Longitud de Onda (micrómetros)	Resolución (metros)
	Banda 1 - Aerosol costero	0.43 - 0.45	30
	Banda 2 – Azul	0.45 -0.51	30
	Banda 3 – Verde	0.53 - 0.59	30
	Banda 4 – Rojo	0.64 - 0.67	30
	Banda 5 - Infrarrojo cercano (NIR)	0.85 - 0.88	30
	Banda 6 - SWIR 1	1.57 - 1.65	30
	Banda 7 - SWIR 2	2.11 - 2.29	30
	Banda 8 – Pancromático	0.50 - 0.68	15
	Banda 9 – Cirrus	1.36 - 1..38	30
	Banda 10 - Infrarrojo térmico (TIRS) 1	10.60 - 11.19	100
	Banda 11 - Infrarrojo térmico (TIRS) 2	11.50 - 12.51	100

FUENTE: USGS, citado por IGAC 2013

En la Tabla 2 se observa la distribución de las bandas en las imágenes Landsat 8 obtenidas por el sensor OLI y TIRS. Se trata de imágenes de resolución media que constan de nueve bandas espectrales con una resolución espacial de 30 metros para las bandas de 1 a 7 y 9. La banda 8 (pancromática) tiene una resolución de 15 metros. Las bandas térmicas 10 y 11 son útiles para proporcionar temperaturas más precisas de la superficie y se toman a 100 metros de resolución, pero se vuelven a remuestrear a 30 metros para que coincida con las bandas multiespectrales de OLI. (IGAC 2013)

Tabla 3: Especificaciones técnicas de las imágenes Landsat 8

	Características Técnicas
Nivel de tratamiento	1T (Imágenes Ortorectificadas)
Sensor	Multiespectral (pushbroom imager)
Resolución Espacial	15 metros (Panchromatic) 30 metros (VIS, SWIR) 100 metros (TIRS)
Resolución radiométrica	12 bits
Resolución temporal	16 días
Formato de entrega	GeoTIFF
Sistema de coordenadas	UTM/ WGS-84/ Zona 17
Metadatos	Formato TXT
Tipo de Licencia	Gratuito
Tamaño de escena	180km * 180km

FUENTE: NASA, citado por MAE – MAGAP 2015

En la Tabla 3 se observan las especificaciones de las imágenes Landsat 8, cuyo tamaño aproximado de escena es 180 km de norte a sur por 180 kilómetros de este a oeste. Se recogen aproximadamente 400 escenas al día, las cuales son cargadas en el servidor de USGS con el fin de que se encuentren disponibles para su descarga 24 horas después de la adquisición. (IGAC 2013)

6. COMBINACIÓN DE BANDAS

Para la visualización de los datos de teledetección es mejor tener una representación en colores, ya que el ojo humano capta mejor el color que los niveles de gris. (Melh y Peinado s.f.)

Según BIODAMAZ (2004), con el conocimiento del comportamiento espectral de los materiales es posible su identificación a través de las composiciones de color. En una composición de color se utilizan los tres colores primarios (rojo, verde y azul). Cuando estos tres colores son combinados en diversas proporciones producen diferentes coloraciones en la región visible del espectro. Asociando cada banda espectral (no necesariamente del visible) a un color primario por separado, para que finalmente resulte una imagen con combinaciones de color.

Banda: Grupo de datos que representan valores de intensidad de la luz reflejada o del calor emitido en una zona o fracción específica del espectro electromagnético (rojo, verde, azul, infrarrojo cercano e infrarrojo térmico). También se le conoce como “canal”.

Tabla 4: Combinaciones de bandas más comunes aplicados para Landsat 8

<i>Descripción</i>	<i>Combinación (RGB)</i>
Color natural	432
Falso color (urbano)	764
Color infrarrojo (vegetación)	543
Agricultura	652
Penetración atmosférica	765
Vegetación saludable	562
Tierra / agua	564
Natural con remoción atmosférica	753
Infrarrojo de onda corta	754
Análisis de vegetación	654

FUENTE: Puentes 2010

En la Tabla 4 se observa las combinaciones posibles para obtener diferentes efectos con las bandas del satélite Landsat 8.

7. PRE- PROCESAMIENTO DE LAS IMÁGENES

El pre-procesamiento es un conjunto de técnicas que se aplican a las imágenes satelitales con el objetivo de mejorar la calidad de los resultados. (BIODAMAZ 2004)

7.1. CORRECCIÓN DE IMÁGENES

Los procesos de corrección de imágenes buscan la eliminación de anomalías detectadas en la imagen y se corrigen con operaciones de corrección geométrica o radiométrica. (REDD – GTZ 2010)

7.2. CORRECCIÓN RADIOMÉTRICA

Las anomalías radiométricas son distorsiones que afectan a los Niveles Digitales (ND) de los píxeles. Estos Niveles Digitales (ND), que son la traducción digital de un flujo energético recibido por el sensor para una determinada banda del espectro (Suárez 2014), pueden evidenciar fallos en los sensores, que generan píxeles incorrectos (Pérez 2012). Estas anomalías deben ser corregidas para que la información que nos proporcionen estos ND sea lo más fiel posible a la realidad que se estudia. (Suárez 2014)

Las correcciones radiométricas son un conjunto de operaciones destinadas a modificar los ND para aproximarse a los valores obtenidos en condiciones ideales (Chuvieco, citado por REDD – GTZ 2010). Comprenden la restauración de líneas o píxeles perdidos y la corrección del bandeo de la imagen. (Lozano y Sierra 2005)

7.2.1. CORRECCIÓN ATMOSFÉRICA

Las interferencias de la atmosfera alteran de forma sistemática los valores de los píxeles (Sánchez 2012), causando un efecto de dispersión que incremento la radiancia detectada por el sensor. Las correcciones atmosféricas eliminan los efectos atmosféricos, como el vapor de agua (humedad), aerosoles (de polvo, volcanes, etc.) y otros factores. (Carnegie Institution for Science 2013)

7.2.2. CORRECCIÓN GEOMÉTRICA

Las distorsiones que producen las principales fuentes de error en el momento de la adquisición de la imagen son: el movimiento en el sistema de escaneo y de la plataforma que sustenta al sensor, perspectiva de los elementos ópticos del sensor, cambios en la altitud de la plataforma, relieve del terreno, curvatura y rotación de la Tierra. Estas distorsiones ocasionadas por las variaciones en la geometría del sensor y de la Tierra son corregidas geoméricamente (Sánchez 2012).

Este proceso de georreferenciación consiste en el cambio de posición de los píxeles respecto a un sistema de coordenadas previamente definido. Comprende una transformación geométrica que relaciona las coordenadas de una imagen, con las coordenadas de un sistema de referencia. La relación entre los dos sistemas de coordenadas (mapa e imagen) puede ser calculada a través de la definición de puntos de control en el terreno. (Chuvieco, citado por Lozano y Sierra 2005)

7.3. UNIÓN DE BANDAS

La unión de bandas es una técnica que pretende mejorar una imagen multiespectral y así potenciar su uso en diversas aplicaciones (Medina, citado por Navarro *et al.* 2007), mediante un proceso Pan-sharpening. (Arenas *et al.* 1998).

Pan-sharpening utiliza una banda pancromática de alta resolución de una imagen multiespectral para afinar las bandas de menor resolución (Barton s.f.). Para ello se combinan los datos de una imagen pancromática (Banda 8) con los datos de una imagen multiespectral (Bandas 1 a 7 y 9) (Medina, citado por Navarro *et al.* 2007), la imagen unida aumenta su resolución espacial y al mismo tiempo preserva la información espectral en la imagen multiespectral. (Caloca, citado por Navarro *et al.* 2007). Con esta técnica se pretende mejorar la interpretabilidad de los datos, ya sea mejorando su calidad visual; facilitando la discriminación de ciertas categorías; aumentando la fiabilidad en la estimación de una determinada variable. (Chuvieco, citado por Cánovas y Alonso 2014). Convirtiéndose así, en una de las tareas más útiles y necesarias en aplicaciones de percepción remota, puesto que el mantener la calidad de la información espectral es de vital importancia para el análisis de firmas espectrales, tales como: análisis de vegetación, suelo, etc. (Caloca, citado por Navarro *et al.* 2007).

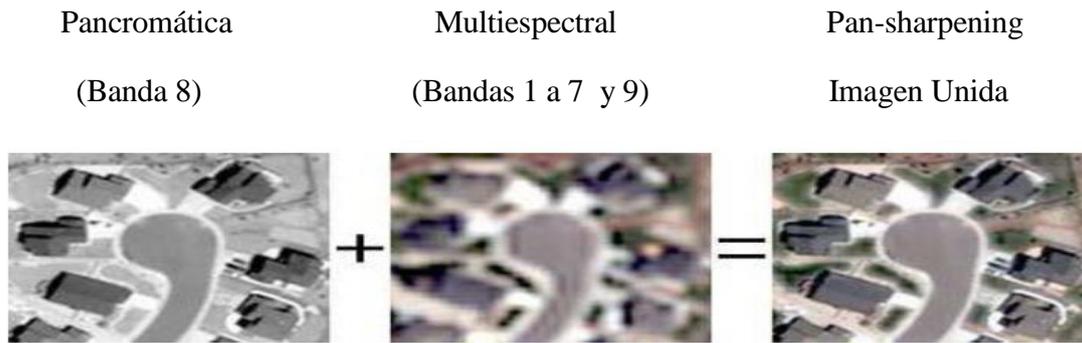


Figura 1: Proceso de Unión de Bandas

FUENTE: Barton s.f.

7.3.1. IMÁGENES PANCROMÁTICAS Y MULTIESPECTRALES

Una imagen pancromática consiste de una sola banda. Generalmente se exhibe como una imagen en escala de grises, es decir, que el brillo exhibido por un píxel en particular es proporcional al número digital del píxel, el cual está relacionado con la intensidad de la radiación solar reflejada por los objetos en el píxel y registrados por el detector. De modo que una imagen pancromática puede ser interpretada como una fotografía aérea en blanco y negro. (BIODAMAZ 2004)

Una imagen multiespectral consta de varias bandas. Cada banda puede ser exhibida como una imagen en escala de grises o, en combinación de tres bandas a la vez, como una composición de color. El análisis de una imagen multiespectral requiere del conocimiento de la firma espectral de los objetos en la escena. (BIODAMAZ 2004)

7.4. SISTEMA DE PROYECCIÓN

Según BIODAMAZ (2004) es el sistema que facilita la representación del mundo real (tridimensional), sobre una superficie plana (bidimensional). Es recomendable elegir el sistema de proyección oficial para la zona de trabajo, con el objeto de compatibilizar la información resultante con la información generada en el país. Por otro lado, deben definirse con claridad las necesidades de re-proyección de las imágenes a fin de garantizar la continuidad espacial de las imágenes.

MC: Meridiano Central

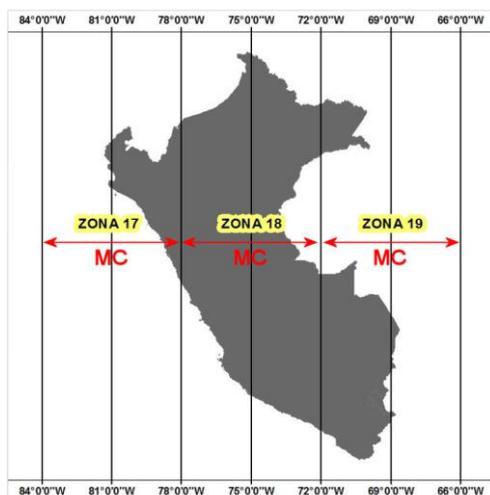


Figura 2: Zonas UTM en el Perú

FUENTE: BIODAMAZ 2004

Cada zona UTM tiene una extensión de 6°. En el Perú tenemos 3 Zonas UTM: Zona 17, Zona 18 y Zona 19, como se muestra en la Figura 2.

Existe otro valor que siempre tenemos que tener en cuenta al hablar de proyección: el datum es un valor de referencia que se utiliza cuando se transforma la esfera terrestre en un plano, es decir cuando se hace una proyección. Es un modelo de la tierra que define el tamaño y forma de la tierra.

A lo largo de los años se han usado diversos valores de Datum. En el Perú los más utilizados por el organismo oficialmente responsable de generar la cartografía base oficial es el Instituto Geográfico Nacional del Perú (IGN) son:

WGS84: World Geodesic System 1984

PSAD56: Provisional para América del Sur 1956

7.5. MOSAICO

Un mosaico es un conjunto de dos o más imágenes enlazadas en forma contigua; tal ensamblaje es necesario cuando el área de interés supera el área cubierta por una sola imagen. (BIODAMAZ 2004)

7.6. VENTANAS DE CORTE

Cuando el volumen de información es grande, es importante definir estrategias que permitan manipular el producto final con facilidad y rapidez, para lo cual es necesaria la generación de ventanas de corte del área de interés (BIODAMAZ 2004).

8. PROCESAMIENTO DE IMÁGENES

Conjunto de procedimientos para manipular las imágenes digitales a través de la ejecución de diversas técnicas y operaciones de clasificación, facilitando su análisis e interpretación (BIODAMAZ 2004). Existen muchos programas de software desarrollado para este fin, el cual incorpora diversas aproximaciones (entre las cuales se hallan presentadas en éste documento).

8.1. SEGMENTACIÓN

La segmentación es un proceso que considera la relación entre la respuesta espectral y la distribución espacial de los píxeles (Suárez 2014), generando polígonos que identifica conjuntos de píxeles que definen áreas geográficas homogéneas que representan algún tipo de unidad espacial (Castillo *et al.* 2015). Esta técnica intenta reproducir la identificación de objetos homogéneos según como la percibiría el ojo humano. (Suárez 2014)

Para segmentar las imágenes satelitales en el programa ENVI EX, se utiliza dos parámetros: Scale (tamaño de los segmentos) y Merge (grado de unión de los segmentos), los cuales se eligen iterativamente generando pruebas, hasta alcanzar la segmentación deseada, en base a la combinación adecuada de ambos parámetros. (MINAM 2014)

Scale: Parámetro de entrada requerido para el algoritmo de segmentación basado en bordes que utiliza ENVI EX. Este parámetro, junto con una función de densidad y el gradiente original del mapa, son necesarios para que el algoritmo de segmentación pueda calcular un mapa con gradientes modificados, y luego lo segmenta. Al elegir un valor de “Scale” alto, se definen pocos segmentos, mientras que la elección de un valor de “Scale” bajo dará lugar a que se generen más segmentos. Idealmente, uno debe elegir el valor de “Scale” más alto que delinee los límites de los polígonos lo mejor posible. Una buena segmentación asegura que los resultados de la clasificación sean más precisos. Si se escoge un valor de “Scale” demasiado alto, los límites entre los segmentos no estarán debidamente delineados y se perderán polígonos de interés. Los valores de “Scale” van desde 0.0 (la segmentación más fina) hasta 100.0 (la segmentación más gruesa), siendo 50.0 el valor por defecto. (ITT 2009)

Merge: Parámetro usado de manera opcional para unir pequeños segmentos dentro de áreas más grandes con textura, como árboles, nubes o cultivos, donde la sobre-segmentación puede ser un problema. (ITT 2009)

El parámetro representa el valor umbral “lambda”, el cual va de 0.0 a 100.0. Por ejemplo, si la imagen muestra una vegetación abundante, aumentar el valor de “Scale” (en el paso previo de segmentación) puede no delinear suficientemente las áreas cubiertas de árboles. Los resultados pueden ser sobre-segmentados si se elige bajo el valor de “Scale” o sub-segmentado si se pone alto el valor de “Scale”. El utilizar “Merge” sería útil en este caso para mejorar la delimitación de los árboles. (ITT 2009)

Segmentación con valores de los parámetros:

Menores valores de Scale —————> Segmentos más pequeños

Mayores valores de Scale —————> Segmentos más grandes

Menores valores de Merge —————> Segmentos menos fusionados

Mayores valores de Merge —————> Segmentos más fusionados

Según MINAM (2014) se considera una segmentación incorrecta cuando los segmentos no reconocen adecuadamente el límite entre dos clases. Sin embargo se aceptan los casos, donde la segmentación incluye más bordes de los que el intérprete distingue.

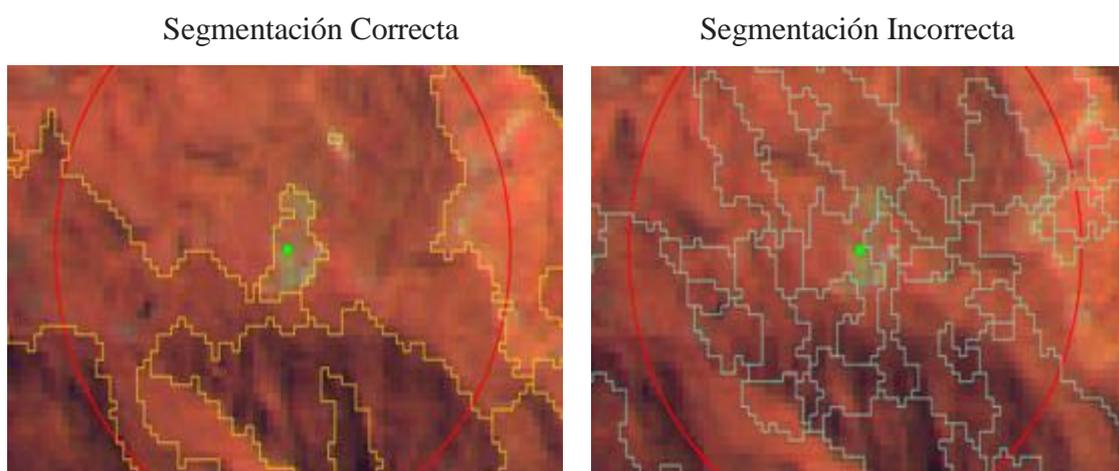


Figura 3: Vista de la evaluación de las segmentaciones en una escena.

FUENTE: Science for a changing world (USGS).

8.2. ÁREAS DE REFERENCIA

Son polígonos que constituyen una muestra representativa de cada una de las clases presentes en la escena. Son áreas homogéneas que representan a las clases de la leyenda y que servirán de insumo para entrenar el modelo de la clasificación (MINAM 2014).

Para generar las áreas de referencia se emplean datos recopilados en la salida de campo, con datos construidos previamente en gabinete. (Martínez *et al.* 2015)

Según MINAM (2014) se distinguen tres tipos:

Tipo 1. Áreas de entrenamiento, aquellas áreas que representan las clases significativas en la escena y que han sido reconocidas correctamente por la segmentación. Estas áreas posteriormente serán empleadas para seleccionar los segmentos que contribuirán con información para realizar la clasificación supervisada. Por ejemplo: Vegetación natural terrestre arbórea cerrada, cultivos, entre otros.

Tipo 2. Áreas de asignación manual, aquellas áreas que pueden estar bien identificadas por el proceso de segmentación, pero que representan clases poco recurrentes en la escena, no son posibles de clasificar con éxito por el clasificador y generan en el modelo un conflicto con las demás clases. Por ejemplo: áreas urbanas, piscigranjas, entre otros.

Tipo 3. Áreas de superposición, aquellas áreas en donde el proceso de segmentación no logró reconocer adecuadamente los bordes de las coberturas que se quiere representar. Por ejemplo: ríos, caminos, entre otros.

8.3. INFORMACIÓN GENERADA PARA LA CLASIFICACIÓN

La información generada para la clasificación tiene como objetivo incrementar la capacidad de clasificación, brindando al algoritmo de clasificación información que contribuye a aumentar las variables a tener en cuenta, para evaluar la información digital de los píxeles en el proceso de clasificación. Se emplea como información generada para la clasificación la información derivada del DEM (modelo digital de elevación), la curvatura y la pendiente. La selección de las capas a emplear está sujeta a la compatibilidad con el sensor empleado, la disponibilidad y accesibilidad a las capas, y el grado de contribución con el análisis. Es muy importante que todas las capas sean generadas con la misma extensión, tamaño de celda y proyección espacial, así mismo, deben estar almacenados en un mismo directorio. (MINAM 2014)

DEM (Modelo Digital de Elevación)

Un modelo digital de elevación es una representación visual y matemática de los valores de altura con respecto al nivel del mar, que permite caracterizar las formas del relieve y los elementos u objetos presentes en el mismo. Estos valores están contenidos en un archivo de tipo raster con estructura regular, el cual se genera utilizando un equipo de cómputo y software especializados. En los modelos digitales de elevación existen dos cualidades esenciales que son la exactitud y la resolución horizontal o grado de detalle digital de representación en formato digital, las cuales varían dependiendo del método que se emplea para generarlos. (INEGI s.f.)

CURVATURE (Curvatura)

Según ESRI (2008) corresponde a la curvatura de una superficie de raster e incluye, opcionalmente, la curvatura del perfil y el plano. A continuación se explica estos conceptos.

- La salida principal del algoritmo es la curvatura de la superficie celda por celda, calculada usando los valores del DEM de esa celda y de sus vecinos circundantes. La curvatura es la derivada segunda de la superficie, o la pendiente de la pendiente.
- Hay dos tipos de curvatura de salida opcionales: la curvatura del perfil está en la dirección de la pendiente máxima y la curvatura del plano es perpendicular a la dirección de la pendiente máxima.
- Una curvatura positiva indica que la superficie es convexa hacia arriba en esa celda. Una curvatura negativa indica que la superficie es cóncava hacia arriba en esa celda. Un valor 0 indica que la superficie es plana.
- En la salida del perfil, un valor negativo indica que la superficie es convexa hacia arriba en esa celda. Un perfil positivo indica que la superficie es cóncava hacia arriba en esa celda. Un valor 0 indica que la superficie es plana.
- En la salida del plano, un valor positivo indica que la superficie es convexa hacia arriba en esa celda. Un plano negativo indica que la superficie es cóncava hacia arriba en esa celda. Un valor 0 indica que la superficie es plana.

- Las unidades del raster de curvatura de salida, así como las unidades del raster de curvatura de perfil de salida opcional y el raster de curvatura plana de salida son una centésima parte (1/100) de una unidad z. Los valores esperados de los tres rasters de salida de un área accidentada pueden variar de -0.5 a 0.5; mientras que para montañas empinadas accidentadas, los valores pueden variar entre -4 y 4. Tenga en cuenta que es posible exceder este rango para ciertas superficies de raster.

SLOPE (Pendiente)

Según ESRI (2008) identifica la pendiente (gradiente o tasa de cambio máximo en el valor z) desde cada celda de una superficie de raster. Los conceptos esenciales asociados al cálculo de esta característica del relieve son:

- Pendiente es la tasa de cambio máximo en el valor z desde cada celda.
- El uso de un factor z es esencial para corregir los cálculos de pendiente cuando las unidades z de la superficie se expresan en unidades diferentes de las unidades x,y del terreno.
- El rango de valores en la salida depende del tipo de unidades de medición.
- Para grados, el rango de valores de pendiente es de 0 a 90.
- Para aumentar el porcentaje, el rango es de 0 a esencialmente infinito. Una superficie plana es 0 por ciento, una superficie de 45 grados es 100 por ciento y, a medida que la superficie se vuelve más vertical, la elevación en porcentaje se vuelve cada vez mayor.
- Si las celdas vecinas son No_Data, se les asigna el valor de la celda central; entonces se computa la pendiente.

8.4. CLASIFICACIÓN

La clasificación es un proceso de ordenar los píxeles en números de clases o categorías, basado en sus valores digitales; es decir si un pixel satisface un conjunto de características, este será asignado a esa clase o categoría (Ruíz y Salcedo 2005). Esta clasificación tiene como propósito convertir las imágenes que contiene los datos en mapas temáticos. (Núñez *et al.* s.f.)

Las técnicas de clasificación se dividen en:

- **Clasificación Supervisada**

La clasificación supervisada parte de un cierto conocimiento de la zona de estudio, adquirido por experiencia previa o por trabajos de campo. Esta mayor familiaridad con la zona, permite al intérprete delimitar sobre la imagen unas áreas piloto, que se consideran suficientemente representativas de las categorías que componen la leyenda. Estas áreas se denominan campos de entrenamiento. El término indica que tales áreas sirven para “entrenar” al ordenador en el reconocimiento de las distintas categorías. En otras palabras, a partir de ellas el ordenador calcula los Niveles Digitales (ND) que definen cada una de las clases, para luego asignar el resto de los píxeles de la imagen a una de esas categorías en función de sus Niveles Digitales (ND). (Chuvienco 1995)

- **Clasificación No Supervisada**

La clasificación no supervisada consiste en definir las clases espectrales presentes en la imagen. No implica ningún conocimiento previo de la zona de estudio por lo que la intervención humana se centra en la interpretación de los resultados (Sánchez 2009).

Esta clasificación asume que los Niveles Digitales (ND) de la imagen se agrupan en una serie de conglomerados que corresponde a un grupo de píxeles con un comportamiento espectral homogéneo. (Núñez *et al.* s.f.)

8.4.1. ÁRBOLES DE DECISIÓN

Los árboles de decisión constituyen uno de los métodos más utilizados en la clasificación de la cobertura y uso de la tierra a partir de imágenes de satélite, debido a su simplicidad e interpretabilidad y la posibilidad de ser representados gráficamente. (Rodríguez 2011).

Un árbol de decisión es un modelo de predicción de forma gráfica y analítica para poder llevar a cabo la clasificación de los datos utilizados, mediante diferentes caminos posibles. (Alcácer 2013). Mediante estos datos se construye una estructura arbórea que sirve para representar y categorizar una serie de condiciones que ocurren de forma sucesiva, para la resolución de un problema (Corso s.f.). Esta serie de condiciones se organiza de forma jerárquica, y se aplica sucesivamente desde una raíz hasta llegar a un nodo terminal u hoja del árbol. (Breiman, citado por Rodríguez y Chica 2012)

En la Figura 4 observamos la estructura de un árbol de decisión; al nodo raíz le pertenecen todos los casos de la muestra que se quiere clasificar, a los nodos internos les corresponde una condición elegida en el conjunto de atributos, y cada ramificación de un nodo interno representa un valor del dominio de la condición del atributo y los nodos terminales u hojas contienen una única etiqueta. Desde la raíz del árbol hasta una determinada hoja, el camino conforma una regla de clasificación. (Rodríguez y Chica 2012).

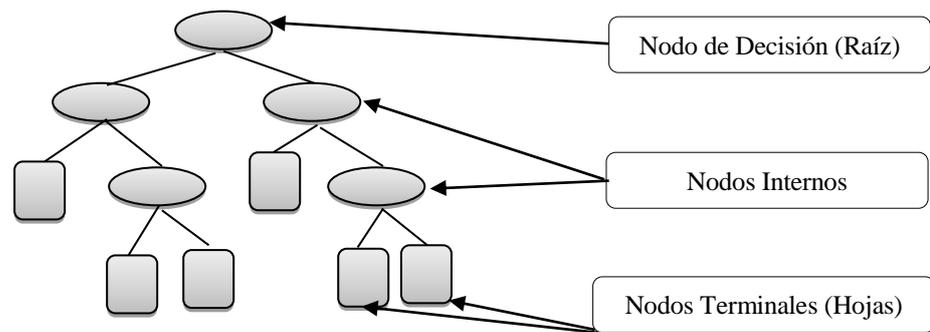


Figura 4: Estructura de un árbol de decisión.

FUENTE: Rodríguez y Chica 2012

8.4.2. ALGORITMOS DE CLASIFICACIÓN

Un Algoritmo es un conjunto de pasos lógicos, secuenciales y metódicamente aplicados para dar solución a una cierta clase de problema. (Balderrama 2008)

a. *Random Forest*

Random Forest es un algoritmo de clasificación no lineal que permite identificar patrones en los datos que se incorporan como entrada de información. Este algoritmo emplea un sistema de “votos” entre el conjunto de árboles para “decidir” la clase. Cada árbol calibrado es un voto, y se asigna la clase en función de cuántos votos obtiene cada una de las unidades a ser clasificadas, la clase más popular (Breiman 2001). *Random Forest* incrementa la diversidad de los árboles de decisión haciéndolos crecer a partir de diferentes subconjuntos de datos creados mediante un procedimiento denominado *bagging*. Ésta es una técnica diseñada para la creación de datos de entrenamiento que remuestrea aleatoriamente el conjunto de datos originales con reemplazamiento, es decir, sin eliminar el conjunto de datos seleccionados de forma previa a la elección del siguiente subconjunto. Por tanto, algunos datos pueden ser usados más de una vez en el entrenamiento de los clasificadores individuales. (Breiman, citado por Rodríguez y Chica 2012)

El algoritmo *Random Forest* implementado en el software Rproject dispone de un conjunto de árboles de decisión que fueron ajustados en base a re-muestreo de las observaciones y variables. (García *et al.* 2011)

b. Algoritmo See5/C5

El algoritmo See5/C5 es la versión más moderna del algoritmo ID3 y el algoritmo C4.5, a lo largo de las dos últimas décadas. (Quinlan, citado por Bonsón *et al.* 1999)

El algoritmo See5/C5 permite construir automáticamente a partir de un conjunto de datos o entrenamiento un árbol de clasificación que representa la relación que existe entre la decisión y sus atributos o variables. El algoritmo realiza particiones binarias sucesivas en el espacio de las variables explicativas. Para realizar cada partición se escoge la variable que aporta más información en función de una medida de entropía o cantidad de información (Bonsón *et al.* 1999). A partir de este árbol se puede elaborar reglas de clasificación fácilmente interpretables. (Alcácer 2013). Para ello ERDAS IMAGINE creó una interface para extrapolar los modelos de árboles de decisión derivados en pixeles clasificados. (MINAM 2014)

9. INDICADORES DE VALIDACIÓN

La validación de la exactitud temática permite definir el grado de fidelidad de los valores obtenidos en el análisis de cobertura y uso de la tierra, producido por la metodología aplicada con respecto a su verdadera característica encontrada en el mundo real y su clasificación correcta. (Llactayo *et al.* 2013)

Para ésta validación se emplea un conjunto de datos de campo y datos al azar para evaluar la consistencia del mapa, a través de métricas cuantitativas de evaluación: la matriz de confusión, el índice Kappa (IK) y el porcentaje correctamente clasificado (PCC). (MINAM 2014)

9.1. MATRIZ DE CONFUSIÓN

La matriz de confusión, también llamada matriz de error o de contingencia es una herramienta que permite evaluar la precisión de una clasificación, contrastando los resultados de la clasificación con información de referencia. Está compuesta por filas y columnas. Las filas corresponden a las clases reinterpretadas a partir de las imágenes base; mientras que las columnas corresponden a las clases provenientes de la clasificación (mapa). En la diagonal de la matriz se computa el número de puntos de verificación donde concuerdan ambas fuentes (reinterpretación y mapas); los puntos ubicados fuera de la diagonal suponen errores de asignación que se computan en los márgenes de la matriz, ya sea como errores de omisión y como errores de comisión. Los residuales en filas indican tipos de cubierta reinterpretadas que no se incluyeron en el mapa; los residuales en columnas representan las categorías del mapa que no se ajustan con la reinterpretación. A partir de la matriz de confusión se puede calcular un indicador de precisión global del mapa, que considera la relación entre el número de puntos correctamente asignados (sumatoria de la diagonal) y el total de los lugares de referencia. (Molinero 2001)

9.2. ÍNDICE KAPPA (IK)

El índice Kappa es otra forma de medir la exactitud de un mapa, ya que incluye dentro del cálculo todos los valores de la matriz y no solamente de los extremos. (Sánchez 2009)

A continuación, según Congalton, citado por Sánchez (2009), se muestra como se calcula el índice Kappa:

$$K = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} * x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} * x_{+i})}$$

Donde:

r: número de filas

x_{ii}: número de observaciones en una fila i y una columna i

x_{i+} y x_{+i}: son los totales marginales de una fila i y de una columna i respectivamente.

N: total de número de observaciones.

Según Cerda y Villarroel (2008), el índice Kappa es otro estimador que refleja la concordancia entre las dos observaciones (mapa y reinterpretación). Puede tomar valores entre 0 y 1: mientras más cercano a 1, mayor es el grado de concordancia entre las observaciones; por el contrario, mientras más cercano a 0, mayor es el grado de discordancia entre ellas. Este coeficiente toma en cuenta los errores de comisión y de omisión.

Según López de Ullibarri y Pita (2001), la máxima concordancia posible corresponde a $\kappa = 1$. El valor $\kappa = 0$ se obtiene cuando la concordancia observada es precisamente la que se espera a causa exclusivamente del azar. Si la concordancia es mayor que la esperada simplemente a causa del azar, $\kappa > 0$, mientras que si es menor, $\kappa < 0$. El mínimo valor de κ depende de las distribuciones marginales.

Tabla 5: Valoración del índice Kappa.

<i>Índice Kappa</i>	<i>Fuerza de Concordancia</i>
0.00	Pobre
0.01 – 0.20	Leve
0.21 – 0.40	Aceptable
0.41 – 0.60	Moderada
0.61 – 0.80	Considerable
0.81 – 1.00	Casi perfecta

FUENTE: Landis y Koch, citado por Cerda y Villarroel (2008)

<i>Valor K</i>	<i>Fuerza de Concordancia</i>
< 0.20	Pobre
0.21 – 0.40	Débil
0.41 – 0.60	Moderada
0.61 – 0.80	Buena
0.81 – 1.00	Muy buena

FUENTE: López de Ullibarri y Pita (2001)

En la Tabla 5 se presenta la valoración del índice Kappa según dos autores diferentes.

Según Zamora *et al.* (2016), se presenta los siguientes conceptos:

- **Error de Comisión:** Es el error que resulta de la clasificación, en la cual un pixel correspondiendo a una clase de cobertura no es clasificada de esa manera.
- **Error de Omisión:** Es el error que resulta de la clasificación, en la cual un pixel es clasificado como una clase de cobertura a la que no pertenece.
- **Exactitud:** Es el nivel de concordancia entre el valor real y el resultado de las observaciones o estimaciones de una característica.
- **Precisión:** Es el nivel de concordancia entre medidas repetidas de la misma característica. Se representa como una estrecha agrupación de resultados de los puntos de muestreo o parcelas. La precisión es inversamente proporcional al error.

10. LEYENDA

La leyenda es la aplicación de una clasificación en un área específica utilizando una escala definida de mapeo. (Di Gregorio 2005). Está estructurada de manera jerárquica en base a criterios fisonómicos de altura y densidad, claramente definidos y aplicables a todas las unidades consideradas para un grupo de coberturas del mismo tipo. (IDEAM 2010)

10.1. LEYENDA DE CORINE LAND COVER

CORINE Land Cover es una metodología para la construcción de mapas de cobertura y uso de la tierra. Esta metodología emplea una leyenda jerárquica que permite describir, caracterizar, clasificar y comparar las características de la cobertura y uso de la tierra, interpretadas a partir de la utilización de imágenes de satélite. Ésta leyenda vincula distintos niveles de detalle espacial (escala espacial) con distintos niveles de detalle temático (niveles de la leyenda jerárquica). (IDEAM *et al.* 2008)

10.2. SALIDA DE CAMPO

El levantamiento de datos de campo es fundamental tanto para la calibración de la clasificación como para la verificación de los resultados. En el proceso de calibración del clasificador, permite al intérprete conceptualizar mejor las coberturas presentes en el área de trabajo y definir mejor las áreas de entrenamiento de las mismas. (MINAM 2014)

La salida de campo consiste en el recorrido de rutas trazadas alrededor de un conjunto de puntos, sobre las coberturas que se quiere hacer el reconocimiento y en las que sea posible ejecutar el recorrido, de acuerdo a las condiciones del terreno. Este trabajo tiene como propósito realizar la verificación de las clases de cobertura interpretadas. El empleo de un GPS permite la incorporación de las rutas para facilitar la captura de los datos a lo largo del itinerario y las correspondientes fotografías terrestres, que sirven como respaldo de información digital. Los datos recopilados en campo son sistematizados para facilitar su uso en el proceso de interpretación. Una vez sistematizados, los datos de campo sirven como insumos fundamentales para la clasificación supervisada (para definir áreas de entrenamiento), permitiendo la integración de las bases de datos y asegurándose que el registro fotográfico corresponda a la información levantada en campo. (Martínez *et al.* 2015)

11. METADATOS

Los metadatos son datos acerca de los datos y sirven para suministrar información sobre el origen, calidad y condición de estos. Consisten en información que caracteriza a los datos, describen el contenido, calidad, condiciones, historia, disponibilidad. Estos metadatos permiten a una persona ubicar y entender los datos. Incluyen: información requerida para determinar qué conjuntos de datos existen para una localización geográfica particular, información necesaria para determinar si un conjunto de datos es apropiado para fines específicos, información requerida para recuperar o conseguir un conjunto ya identificado de datos, información requerida para procesarlos y utilizarlos. (GEOIDEP 2003)

12. ESTUDIOS RELACIONADOS A LA COBERTURA Y USO DE LA TIERRA

Duarte (s.f), realizó un estudio sobre el “Mapa Forestal de Honduras derivado con imágenes de alta resolución, una herramienta innovadora y útil para la toma de decisiones nacionales y sub-nacionales”. Se describen las principales actividades realizadas en el marco del diseño y construcción del mapa forestal de Honduras, estableciendo una secuencia de procedimientos que permiten obtener información cartográfica sobre la extensión y distribución de la cobertura boscosa a nivel nacional, utilizando imágenes de alta resolución RapidEye. Se realizó una clasificación con criterios espectrales, espaciales y estadísticos a través de un algoritmo de árboles de decisión *Random Forest* implementado en el software Rproject.

García *et al.* (2011), realizaron un estudio sobre “Cambio en la cobertura del suelo en la cuenca media y alta de Lules durante los últimos 20 años”. Se analizó el cambio en la cobertura de la tierra durante los últimos 20 años, mediante el procesamiento de imágenes satelitales Landsat 2 MSS y Landsat 5 TM. Se realizó una clasificación supervisada a través del algoritmo *Random Forest* implementado en el software Rproject. Este método dispone de un conjunto de árboles de decisión que fueron ajustados en base a re-muestreo de las observaciones y variables. Dicho algoritmo fue señalado como el más adecuado para este estudio.

ATECMA S.L. (2007), realizó un estudio sobre “Identificación y diagnóstico de la red de corredores ecológicos de la región de Murcia”. Se utilizó el algoritmo *Random Forest* implementado en el programa Rproject. Con ésta técnica se obtuvieron mapas predictivos continuos en los que cada celda contiene la probabilidad de que un determinado hábitat esté presente en función de un conjunto de predictores climáticos, topográficos, edáficos, y usos del suelo. El algoritmo de *Random Forest* utilizado para generar el mapa de idoneidad ofrece además la posibilidad de interpretar las variables que determinan la distribución del hábitat, proporcionando un gráfico en el que se asignan valores a cada variable en función de la importancia o peso que éstas adquieren en la modelización del hábitat.

III. MATERIALES Y MÉTODOS

1. ZONA DE ESTUDIO

El distrito de San Ramón se encuentra ubicado en la provincia de Chanchamayo, departamento de Junín. Se halla en selva alta, con una superficie de 591.67 km², equivalente al 12.53 por ciento de la provincia de Chanchamayo. Éste distrito es estratégicamente la puerta de ingreso de comunicación a la Selva Central de nuestro país, discurriendo en su territorio dos ríos principales: el río Tarma y el río Tulumayo. (Municipalidad Distrital de San Ramón 2009)

Tabla 6: Límites del distrito de San Ramón

Dirección	Límites
Norte	Distrito de Chanchamayo
Este	Distrito de Vitoc
Sur	Distrito de Palca
Oeste	Distrito de Huasahuasi

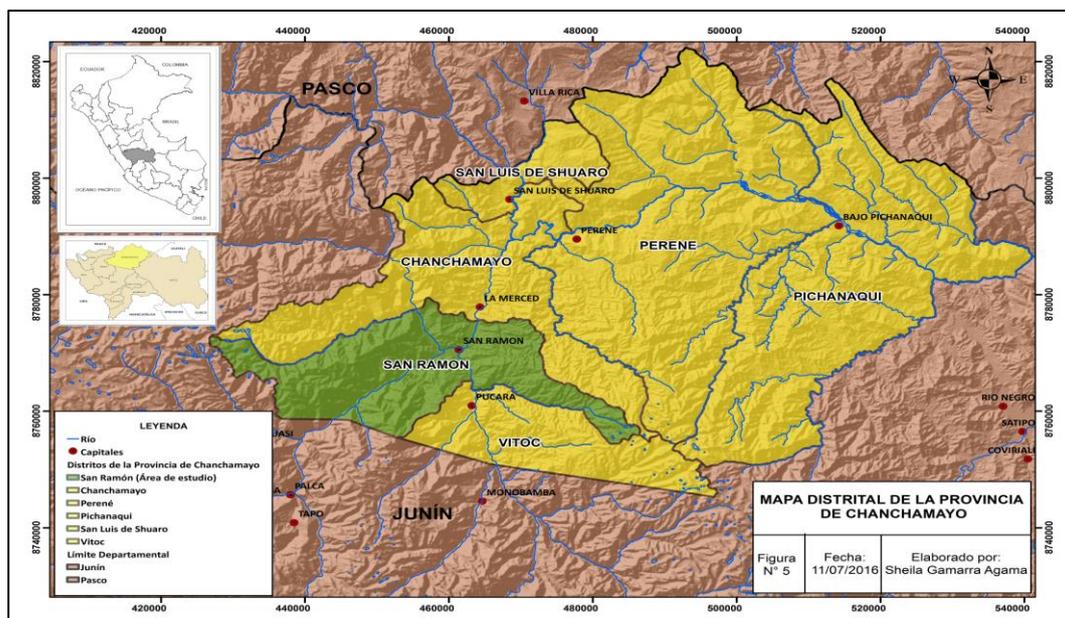


Figura 5: Mapa distrital de la provincia de Chanchamayo.

FUENTE: Elaboración propia

1.1. FISIOGRAFÍA Y SUELOS

La provincia de Chanchamayo tiene un paisaje montañoso y una topografía compleja originada por contrafuertes de la cordillera oriental andina, con pendientes marcadas a muy marcadas, frecuentemente de 60 a 100%. Las pendientes muy fuertes representan el 80% de la superficie total, los paisajes de colinas de menor altura representan el 18% y las terrazas de origen aluvial ubicadas en las márgenes de los ríos representan solamente el 2%. (Bullón, citado por Antón y Reynel 2004).

Los suelos de los espacios Montanos como los de San Ramón son Litosoles-Cambisoles dístricos y éutricos, de acuerdo al sistema de clasificación de FAO (IGN, citado por Reynel 2012). Los Litosoles son suelos superficiales cuya profundidad está limitada por masas o estratos de roca dura a partir de los 10 cm de profundidad. Los Cambisoles son suelos tropicales caracterizados por un horizonte B con conspicua presencia de hierro; los Cambisoles dístricos agrupan a aquellos con porcentaje de saturación de bases menores al 50% y los Cambisoles éutricos tienen similares características pero se diferencian por su porcentaje de saturación de bases por encima del 50%. (Sánchez, citado por Reynel 2012)

1.2. HIDROGRAFÍA

El distrito de San Ramón colinda con dos ríos principales: el río Tarma y el río Tulumayo, ambos de regular caudal y muy torrentosos. Estos ríos forman el río Chanchamayo que luego de un corto recorrido se une con el río Paucartambo para formar el río Perené. (Dancé y Kometter, citados por Cáceres y Reynel 2010). El río Perené es el principal colector de la zona, que nace con el nombre de río Tulumayo en las alturas de las provincias de Jauja y Concepción del departamento de Junín, perteneciente a la cuenca amazónica. El río Tulumayo, al llegar al distrito de San Ramón, recibe por la margen izquierda las aguas del río Palca y del río Ulcumayo, llamado también Oxabamba en su curso inferior. (Reynel 2012)

1.3. CLIMA

Las condiciones climáticas se encuentran documentadas por la existencia de una estación meteorológica asociada a la base aérea del distrito de San Ramón. La temperatura es alta, con algunos meses de temperatura templada. La temperatura máxima promedio corresponde a los meses de Octubre a Noviembre con 30,1 °C; la temperatura mínima corresponde al mes de Julio con 16,7 °C. La temperatura media anual 23,1 °C. (Galdos, citado por Antón y Reynel 2004).

La precipitación marca dos estaciones bien definidas, una con baja precipitación entre junio a agosto y otra con abundante precipitación entre diciembre a mayo. La precipitación total anual promedio se halla entre 1970 - 2104 mm, con un promedio alrededor de 2000 mm. (Antón y Reynel 2004)

1.4. CONTEXTO BIOLÓGICO

- Flora

Se perciben tres estratos principales en la estructura del bosque maduro: el estrato emergente del dosel, el estrato intermedio o arbóreo y un estrato oprimido o de sotobosque. Existe una gran cantidad de epífitas, tales como Orquídeas y Bromelias, helechos arbóreos y palmeras pertenecientes al género *Ceroxylum*. Así mismo, revelan la presencia de un número relativamente alto de especies endémicas, y algunas especies nuevas, aun no descritas para la ciencia. Dentro de las especies más representativas de flora se encuentran los árboles, como son: el Diablo Fuerte *Podocarpus* spp. (actualmente escasos), los Cedros *Cedrela montana* (muy escasos), especies de la familia Lauráceas *Ocotea* spp. y *Nectandra* spp. (relativamente abundantes) y especies de la familia Rubiáceas, a la cual pertenece el árbol de la Quina o Cascarilla *Cinchona pubescens*. (Reynel 2012)

- Fauna

La fauna silvestre es rica en el valle de Chanchamayo. La diversidad de avifauna es alta, y dentro de las especies conspicuas, se observa al gallito de las rocas *Rupicola peruviana*, el pájaro relojero *Momotus aequatorialis*, así como tucanes, loros, quetzales y gavilanes. (Reynel 2012).

Algunos de los mamíferos importantes de la zona son: el Cutpate *Dasyprocta variegata*, el Zamaño *Agouti paca*, los Armadillos *Dasybus* sp., Puerco-espines *Coendou bicolor*; Sajinos *Pecari tajacu*, y el Oso de anteojos *Tremarctos ornatus*; especie categorizada en peligro por sus escasas poblaciones, así como varias especies de primates, dentro de ellas el Mono choro *Lagothrix cana*. También se observan: Coatíes *Nasua nasua* y varias especies de felinos, como el tigrino *Leopardus pardalis*, el margay *Leopardus tigrinus*, el puma *Puma concolor* y el otorongo *Panthera onca*. (Reynel 2012)

1.5. CLASIFICACIÓN ECOLÓGICA

De acuerdo a los criterios de clasificación ecológica basada en Zonas de Vida desarrollados por Holdridge (1978), el cual estratifica las áreas naturales sobre la base de parámetros de temperatura, precipitación, altitud y latitud. Ellas se plasman en el Mapa Ecológico elaborado por la ONERN (1976) y luego actualizado por INRENA (1995). Las 9 Zonas de Vida reconocidas en el ámbito de la provincia de Chanchamayo son: Bosque seco Tropical (bs-T), Bosque húmedo Premontano Tropical (bh-PT), Bosque muy húmedo Premontano Tropical (bmh-PT), Bosque muy húmedo Montano bajo Tropical (bmh-MBT), Bosque muy húmedo Montano Tropical (bmh-MT), Bosque pluvial Premontano tropical (bp-PT), Bosque pluvial Montano bajo Tropical (bp-MBT), Bosque pluvial Montano tropical (bp- MT) y Páramo pluvial subalpino tropical (pp-SAT). (Reynel 2012)

1.6. CLASIFICACIÓN POR TIPOS DE BOSQUE

En el valle de Chanchamayo, el bosque se encuentra constituido de los siguientes estratos: Bosque de Vigor Alto, Bosque de Vigor Medio, Bosque de Vigor Bajo, Bosque achaparrado, Bosque secundario temprano o vegetación arbustiva y Vegetación de pacal o bambú nativo. Esta estratificación de la vegetación se basa en variables que permiten reflejar los niveles de diversidad biológica presentes en cada área, tales como: el estadio de sucesión del bosque, el número de estratos, el nivel de apertura del dosel y la composición florística. (Reynel 2012)

1.7. USO DE LA TIERRA

Dentro del valle el potencial de uso mayoritario es el de tierras forestales; sin embargo, las áreas de bosque se ven afectadas por el crecimiento demográfico, que favorece la deforestación debido a la conversión de tierras forestales a usos agrícolas. La Selva Central es una de las regiones del país con mayor presión de agricultura por nuevas tierras. Es importante enfatizar que la agricultura migratoria, con rozo-tumba-quema y posterior abandono de tierras, constituye una de las principales amenazas a la conservación de los recursos forestales en la zona. (Reynel 2012)

1.8. MARCO SOCIOECONÓMICO

La provincia de Chanchamayo registra una población de 132 900 habitantes; es la segunda provincia más poblada del departamento de Junín, alcanza una densidad poblacional de alrededor de 28 habitantes por kilómetro cuadrado. La mayor parte de la población está conformada por migrantes procedentes de las zonas andinas. (INEI, citado por Reynel 2012)

Las actividades económicas de mayor importancia actual son la agrícola y forestal. Lamentablemente, la última entendida en un sentido de la extracción destructiva y desordenada de productos del bosque, aunque se observan iniciativas de reforestación en predios privados.

El turismo ha experimentado un sano incremento en los últimos años, pues la Selva Central está llena de sitios con belleza natural y ambientes atractivos para el turismo y ecoturismo.

Tradicionalmente, el valle de Chanchamayo ha sido productor de café y frutas tropicales. En los últimos años las variaciones en los precios han ocasionado oscilaciones en la producción cafetalera. La fruticultura creció con incidencia en el cultivo de los cítricos, por lo que constituye una de las despensas fundamentales de fruta para los mercados de Tarma, La Oroya, Huancayo y Lima. Los cultivos de maíz y plátano ocupan una proporción importante de las áreas dedicadas a cultivos en limpio en la zona. Todo ello ha sido favorecido por su cercanía a la ciudad de Lima y la reciente presencia de vías asfaltadas en buen estado. (Reynel 2012)

La extracción forestal constituyó uno de los rubros de actividad económica de mayor importancia en el pasado. Se inició hacia 1920, consumiendo rápidamente las maderas más valiosas de los bosques existentes, dentro de ellas: Cedro *Cedrela odorata*, Diablo Fuerte o Romerillo *Podocarpus* spp, actualmente muy escasos. Hacia los años 80, Chanchamayo todavía seguía constituyendo un área importante de producción maderera, aportando el 10% de la producción de madera aserrada al país. (Casas, citado por Reynel 2012). Para el año 1990, Junín había sido deforestado en un 28%. El mayor desembosque se ha producido en las localidades de San Ramón, La Merced y Satipo, que son las más accesibles. (INRENA, citado por Reynel 2012)

Los bosques remanentes que existen actualmente se encuentran localizados en colinas altas y montañas con inaccesibilidad natural y, de acuerdo a un cálculo reciente, abarcan una superficie de 84,688 ha en la provincia de Chanchamayo. (INRENA, citado por Reynel 2012)

1.9. ACCESIBILIDAD

El acceso al distrito de San Ramón es por vía terrestre. Desde Lima es relativamente fácil: Se debe viajar a través de la carretera central, con un recorrido total de 303 km (Aguilar, citado por Reynel 2012)

Tabla 7: Trayectos terrestres de Lima a San Ramón

<i>Origen</i>	<i>Destino</i>	<i>Distancia (km)</i>	<i>Vía</i>
Lima	La Oroya	181	Carretera asfaltada
La Oroya	Tarma	57	Carretera asfaltada
Tarma	San Ramón	65	Carretera asfaltada
Total		303	

FUENTE: (Aguilar, citado por Reynel 2012).

2. MATERIALES Y EQUIPOS

Para la recopilación y procesamiento de la información se emplearon los siguientes materiales y equipos.

2.1. MATERIALES

- Dos imágenes de resolución media para el año 2016 (Landsat 8)
- Cuatro imágenes de resolución alta para el año 2012 (RapidEye)
- Modelo Digital de Elevación (DEM) de 30 metros.
- Imágenes SPOT del servidor Google Earth
- Mapas del Perú: Mapa ecológico, Mapa de ecorregiones y el Mapa de cobertura vegetal.
- Scripts:“recogiendoInfoRaster_noTC.R”, “Filtrar.tabla.todo.rf.R” y “metricas.validacion.R”, elaborados por Carlos Alberto Arnillas en el año 2012.
- Fichas de campo
- Mapas de recorrido de campo

2.2. HERRAMIENTAS DE SOFTWARE

ARCGIS 10.1

Es una plataforma de sistema de información geográfica (SIG) que permite crear, analizar, almacenar y difundir datos geoespaciales, para resolver problemas complejos con fines geográficos. (MAE – MAGAP 2015)

ENVI EX

ENVI por sus siglas en inglés Environment por Visualizing Images (entorno para visualización de imágenes) (S-R-GIS 2005), es un software completo de visualización y análisis de las imágenes obtenidas en la teledetección. Posee un conjunto muy completo de funciones de análisis y tratamiento que permite trabajar con imágenes multiespectrales, hiperespectrales y de radar (Ruiz y Salcedo 2005).

ENVI EX está diseñado específicamente en torno a las necesidades de los usuarios de SIG y construido para interactuar y funcionar con los datos y capas de ArcGis. (S-R-GIS 2005)

ERDAS IMAGINE 9.1

Es un paquete de software basado en raster y diseñado específicamente para extraer información de imágenes remotamente adquiridas. El programa incluye un sistema comprensivo de herramientas para crear imágenes para la inclusión en un SIG. (Sánchez 2012)

R PROJECT 3.1.1

R es un sistema para análisis estadístico y gráfico creado por Ross Ihaka y Robert Gentleman en 1997. R tiene una naturaleza doble, de programa y lenguaje de programación estadística, tomando como base las características del software comercial S-PLUS. Se distribuye gratuitamente bajo los términos de General Public Licence. (Paradis 2003)

2.3. EQUIPOS

Computadora:

Procesador: Intel ® Core (TM) i3-4130 CPU @ 3.40GHz

Memoria instalada (RAM): 8.00 GB (7.88 GB utilizable)

Tipo de sistema: Sistema operativo de 64 bits, procesador x64

GPSmap 76S Garmín:

Tamaño de la pantalla: 4,1 x 5,6 cm

Resolución de la pantalla: 180 x 240 pixeles

Batería y duración: 2 pila AA, 16 horas

Memoria interna: 24 MB

Cámara Digital: Canon 60D

Batería: LP – EG, 18 megapíxeles.

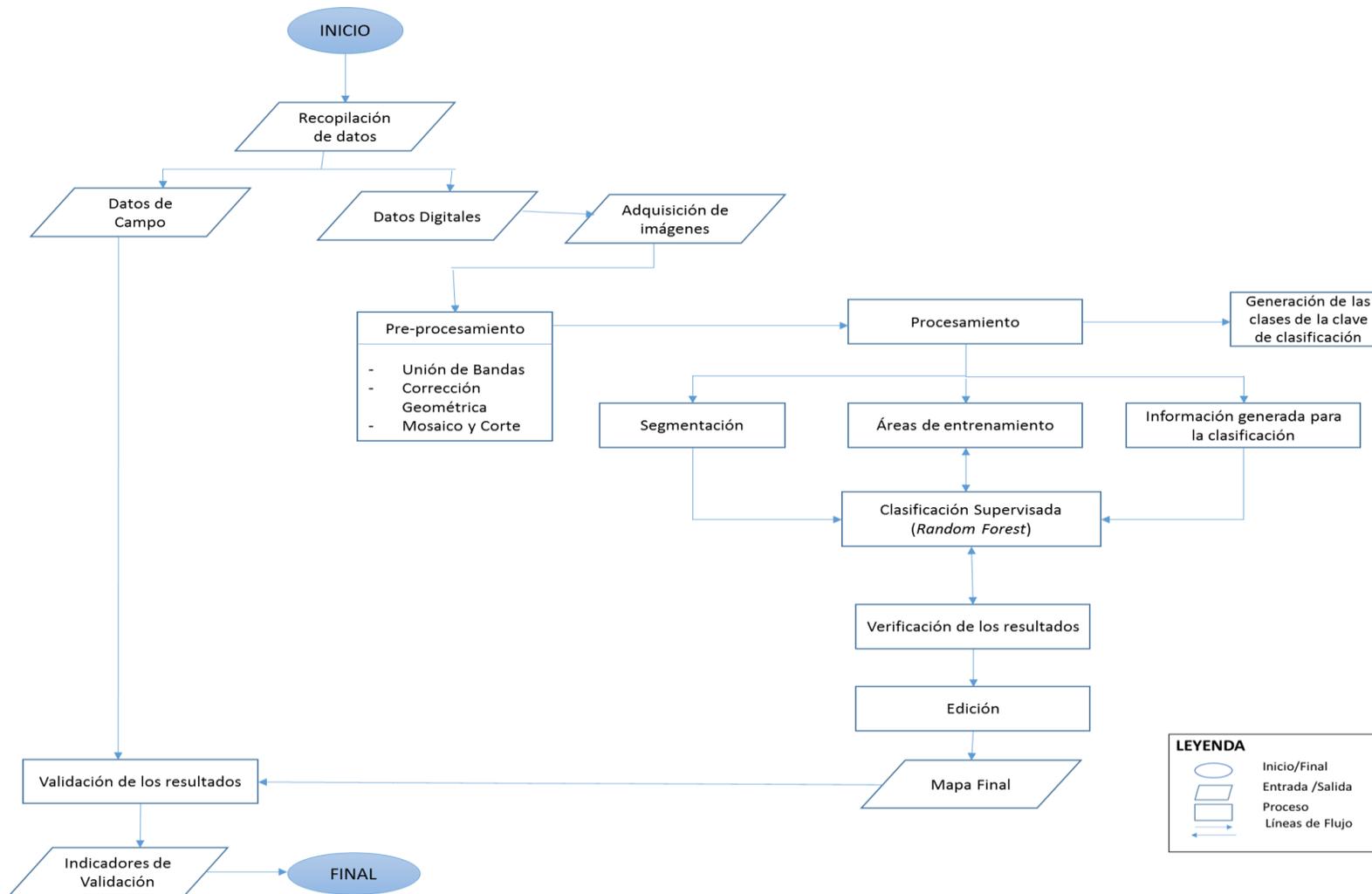


Figura 6: Diagrama de la metodología

FUENTE: Elaboración propia

3. METODOLOGÍA

La metodología empleada para la elaboración del Mapa de Cobertura y Uso de la Tierra del distrito de San Ramón se basó en la “Metodología de clasificación semiautomática, para el mapeo de cobertura y uso de la tierra, empleando *Random Forest*”, por el Centro de Datos para la Conservación (CDC-UNALM 2012).

El proceso metodológico está compuesto secuencialmente por los siguientes pasos:

3.1. ADQUISICIÓN DE LAS IMÁGENES SATELITALES

Las imágenes satelitales de resolución espacial media (Landsat 8) fueron descargadas gratuitamente en formato TIFF del servidor The USGS Global Visualization Viewer (GloVis), disponible en: <http://glovis.usgs.gov/> (Ver Anexo 1A)

Se descargó dos imágenes satelitales para cubrir la extensión del área de estudio. Se consideró la presencia de un 50% de nubes. (Ver Anexo 1B)

Tabla 8: Imágenes Landsat 8

<i>Sensor</i>	<i>Path - Row</i>	<i>Fecha</i>
Landsat 8	006 - 068	01/06/2016
Landsat 8	007 - 068	26/07/2016

FUENTE: Elaboración propia

4. PRE-PROCESAMIENTO DE LAS IMÁGENES SATELITALES

4.1.1. UNIÓN DE BANDAS

Se realizó la unión de bandas en el programa Erdas Imagine 9.1, mediante un proceso de mejoramiento utilizando la herramienta Pan-Sharpener, basado en el algoritmo “Pan-sharpening”, para lo cual se necesitó una imagen pancromática (banda 8) de 15 metros de resolución y una imagen multiespectral (bandas 1 a 7 y 9) de 30 metros de resolución. (Ver Anexo 2A)

4.1.2. CORRECCIÓN GEOMÉTRICA DE LAS IMÁGENES

Las imágenes Landsat 8 se encuentran ortorrectificadas en el sistema de coordenadas de UTM/WGS-84/Zona 18N por lo que fue necesario re-proyectarlas a UTM/WGS-84/Zona 18S.

Las dos imágenes Landsat 8 fueron re-proyectadas en el programa Erdas Imagine 9.1, para este proceso se utilizó la herramienta Rproject, para lo cual se necesitó la imagen unida a 15 metros de resolución. (Ver Anexo 2B)

4.1.3. GENERACIÓN DEL MOSAICO

La generación del mosaico se elaboró en el programa Erdas Imagine 9.1, donde se adicionaron las dos imágenes Landsat 8 y se empleó la herramienta MosaicPro. (Ver Anexo 2C)

4.1.4. CORTE DEL MOSAICO

Para facilitar los análisis posteriores fue necesario cortar el mosaico en el programa Erdas Imagine 9.1, donde se utilizó la herramienta Create Subset Image, mediante un AOI del área de interés. (Ver Anexo 2D)

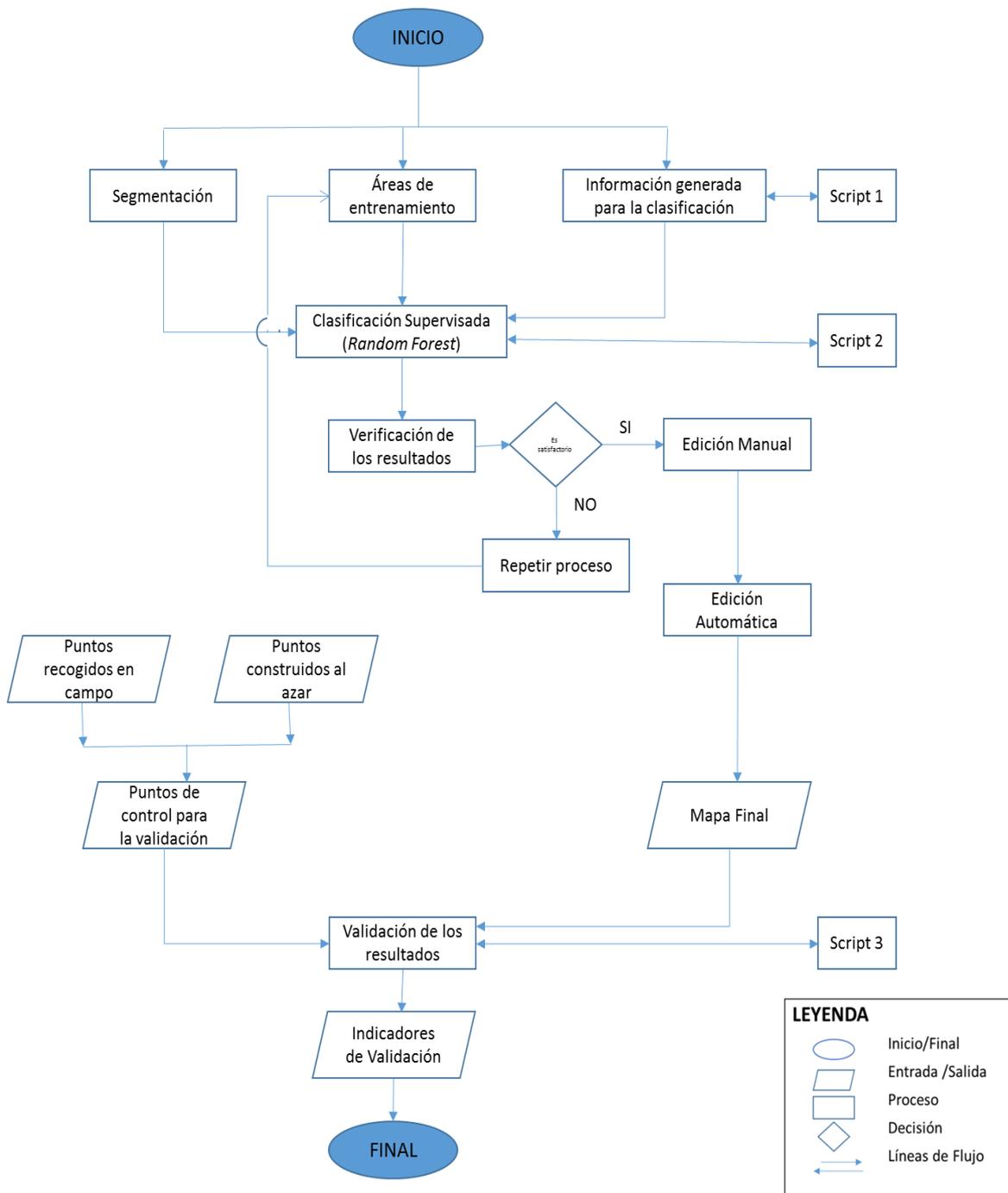


Figura 7: Diagrama del procesamiento

FUENTE: Elaboración propia

4.2. PROCESAMIENTO

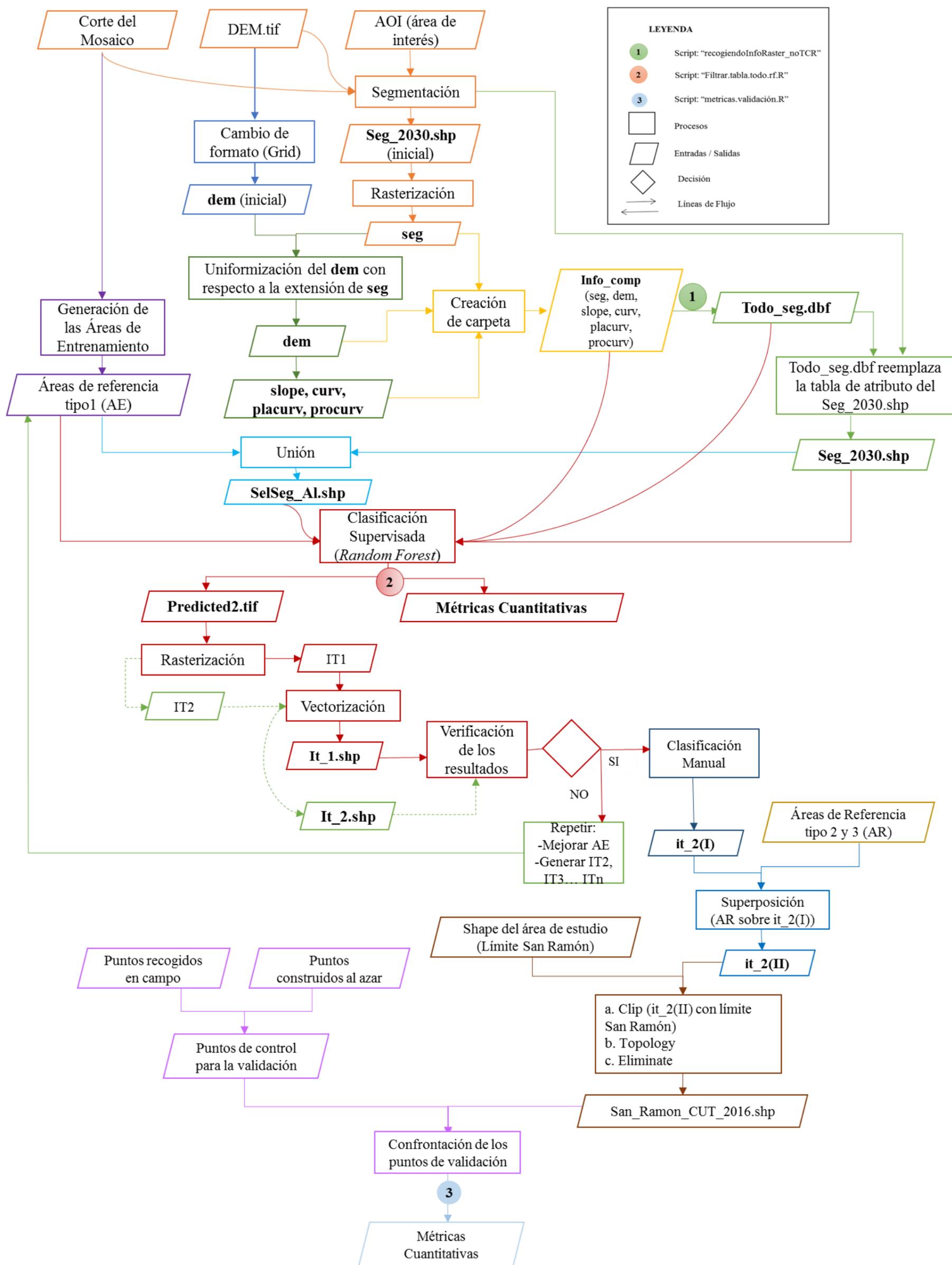


Figura 8: Diagrama detallado del procesamiento

FUENTE: Elaboración propia

4.2.1. GENERACIÓN DE LAS CLASES DE LA CLAVE DE CLASIFICACIÓN

El procesamiento se inició con el reconocimiento de las clases de la clave de clasificación en la imagen a interpretar. Estas clases fueron identificadas con el apoyo de datos digitales y datos obtenidos en la salida de campo. Como datos digitales se tienen: las imágenes Landsat 8 del año 2016; las imágenes RapidEye del año 2012; las imágenes Spot del servidor Google Earth; y la información vectorial del mapa ecológico, mapa de ecorregiones y el mapa de cobertura vegetal del Perú.

4.2.2. CLASIFICACIÓN

a. Segmentación

Este proceso se elaboró en el programa ENVI EX, para lo cual se requirió como insumos: el corte del mosaico, el AOI del área de interés y el modelo digital de elevación (DEM). Para definir la segmentación que más se ajuste a la escena se utilizó dos parámetros: Scale y Merge. Se realizaron pruebas iterativamente variando los valores de ambos parámetros hasta alcanzar la segmentación deseada.

b. Generación de la áreas de entrenamiento

Luego de definir la segmentación más adecuada para la imagen, se construyó las áreas de entrenamiento (áreas de referencia de tipo 1) en la plataforma de ArcGis 10.1. Se construyó 100 polígonos distribuidos de manera homogénea en todo el área de estudio. (Ver Anexo 3B)

Para generar las áreas de entrenamiento se emplearon datos recopilados de la salida de campo y la interpretación visual de los datos digitales.

c. Información generada para la clasificación

Se creó una carpeta denominada info_comp, donde fueron almacenadas todas las capas resultantes.

En la plataforma de ArcGis 10.1 se generó la información para la clasificación, para lo cual se requirió como insumos el shapefile de la segmentación (seg_2030) y el modelo digital de elevación (DEM).

Del shapefile de la segmentación (seg_2030) se generó el raster seg, utilizando la herramienta Feature to raster y del modelo digital de elevación (DEM) se generó el raster dem, utilizando la herramienta Raster Calculator.

Se uniformizó la capa del raster dem con respecto al raster seg, utilizando la opción Environment Settings (configuración del entorno), empleando la herramienta Resample para uniformizar el tamaño de celda, la extensión y la proyección espacial. (Ver Anexo 3C)

Finalmente se generaron las capas topográficas derivadas del raster dem, mediante la herramienta Surface, para la pendiente se utilizó la herramienta Slope y para la curvatura se utilizó la herramienta Curvature.

Después de generar la información para la clasificación, se construyó una base de datos (en formato dbf) denominada “Todo_seg”, mediante un script desarrollado para el software R 3.1.1 denominado “recogiendoInfoRaster_noTC.R”. En este software se instaló las librerías requeridas: raster, foreign y rgdal; para ejecutar satisfactoriamente el proceso.

Finalmente se verificó que la carpeta de info_comp tenga la siguiente información: dem, slop, curv, placurv, procurv, seg y seg_2030.

d. Clasificación Supervisada

Se elaboró una clasificación supervisada en base a segmentos y como clasificador se empleó el algoritmo de clasificación Random Forest, mediante un script desarrollado para el software R Project 3.1.1 denominado “Filtrar.tabla.todo.rf.R”. En este software se instaló las librerías requeridas: Random Forest, raster, irr y foreign; para ejecutar satisfactoriamente el proceso. (Ver Anexo 3G)

Se creó una carpeta denominada Clasificación, donde fueron almacenados los insumos requeridos: áreas de entrenamiento, la segmentación previamente modificada con el archivo dbf de la base de datos “Todo_seg” (seg_2030), la carpeta info_comp, el shapefile SelSeg_Al.shp y Todo_seg.

Se generó el shapefile SelSeg_Al.shp uniendo las áreas de entrenamiento con la segmentación (seg_2030), con la finalidad de asignar a los segmentos la clase del área de entrenamiento que se intercepta con cada uno de ellos, para este proceso se utilizó la herramienta Spatial Join.

El script empleado generó el raster predicted2.tiff. y las métricas cuantitativas de la clasificación.

En la plataforma de ArcGis 10.1 se generó el raster de la clasificación denominado iteración (IT1), que tuvo como insumo el raster predicted2.tif, para lo cual se utilizó la herramienta Raster Calculator; seguidamente este raster (IT1) se convirtió en un shapefile denominado (it_1), donde se utilizó la herramienta Raster to polygon. (Ver Anexo 3F).

Se revisó visualmente la clasificación y se realizó una segunda iteración (IT2), debido a la confusión de algunas coberturas, donde se modificó algunas áreas de entrenamiento y se generó nuevamente el shapefile del SelSeg_Al. Finalmente se elaboró nuevamente el procedimiento de clasificación generando un shapefile final denominado (it_2).

e. Verificación de Resultados

La verificación consistió en una edición manual sobre el resultado de la clasificación automática, elaborado en la plataforma de ArcGis 10.1, para facilitar este proceso se necesitó como apoyo una grilla de 50x10 km².

Inicialmente se asignaron manualmente las dos clases que no se reconocieron en la clasificación, que fueron las zonas quemadas y las tierras altoandinas sin vegetación.

Luego se hizo una revisión visual por todo el área de estudio para cambiar algunos polígonos incorrectamente asignados en la clasificación. Seguidamente se generó las áreas de referencia de tipo 2 que fueron los centros poblados y las áreas de referencia tipo 3 que fueron los ríos, quebradas y las redes viales, presentes en el área de estudio. Finalmente se superpuso las áreas de referencia tipo 2 y 3 sobre la imagen clasificada, donde se utilizó la herramienta Update.

4.3. SALIDA DE CAMPO

En la salida de campo se colectó 50 puntos y consistió en el recorrido de rutas trazadas a lo largo de la carretera del área de estudio, sujeto a la accesibilidad del lugar, registrando datos relevantes para este estudio. Estos datos de campo permitieron conceptualizar mejor las coberturas presentes en el área de estudio y definir mejor las áreas de entrenamiento para el proceso de clasificación.

Se empleó un GPS map 76S Garmín para facilitar la captura de datos a lo largo del itinerario y se tomaron las correspondientes fotografías terrestres que sirvieron como respaldo de los datos digitales. Para facilitar la sistematización de estos datos se trabajó con una ficha de

campo definida para el estudio y como datos digitales impresos se utilizó mapas de la ruta con datos de 50 puntos predefinidos con las clases de la leyenda previamente definida.

4.4. EDICIÓN

La edición se elaboró en la plataforma de ArcGis 10.1, después de la edición manual de la imagen clasificada, para éste proceso se empleó las siguientes herramientas: Clip, se cortó el shapefile de la clasificación con respecto a los límites del distrito de San Ramón; Dissolve, fusionó los elementos semejantes; Topology, corrigió los polígonos del shapefile final y Eliminate, eliminó los falsos polígonos que se crearon durante la superposición de las áreas de referencia (Ver Anexo 5).

Finalmente se obtuvo el shapefile del mapa final denominado San_Ramon_CUT_2016

4.5. VALIDACIÓN DE LOS RESULTADOS

En la plataforma ArcGis 10.1 se construyó un shapefile con 100 puntos al azar distribuidos al interior del área de estudio, estos puntos fueron interpretados visualmente a partir de la imagen satelital Landsat 8, imágenes RapidEye y Google Earth.

Se unió los puntos de campo y los puntos al azar para contrastarse con la información del mapa final (San_Ramon_CUT_2016), interpretado en dichos puntos, conformándose una capa de validación. (Ver Anexo 6)

Para la validación se empleó el script denominado “metricas.validacion.R”, desarrollado para el software R Project 3.1.1, donde se instaló las librerías requeridas: irr y foreign; para ejecutar satisfactoriamente el proceso. Finalmente se evaluó la consistencia del mapa, a través de métricas cuantitativas: la matriz de confusión, el índice Kappa (IK) y el porcentaje correctamente clasificados (PCC). (Ver Anexo 3G)

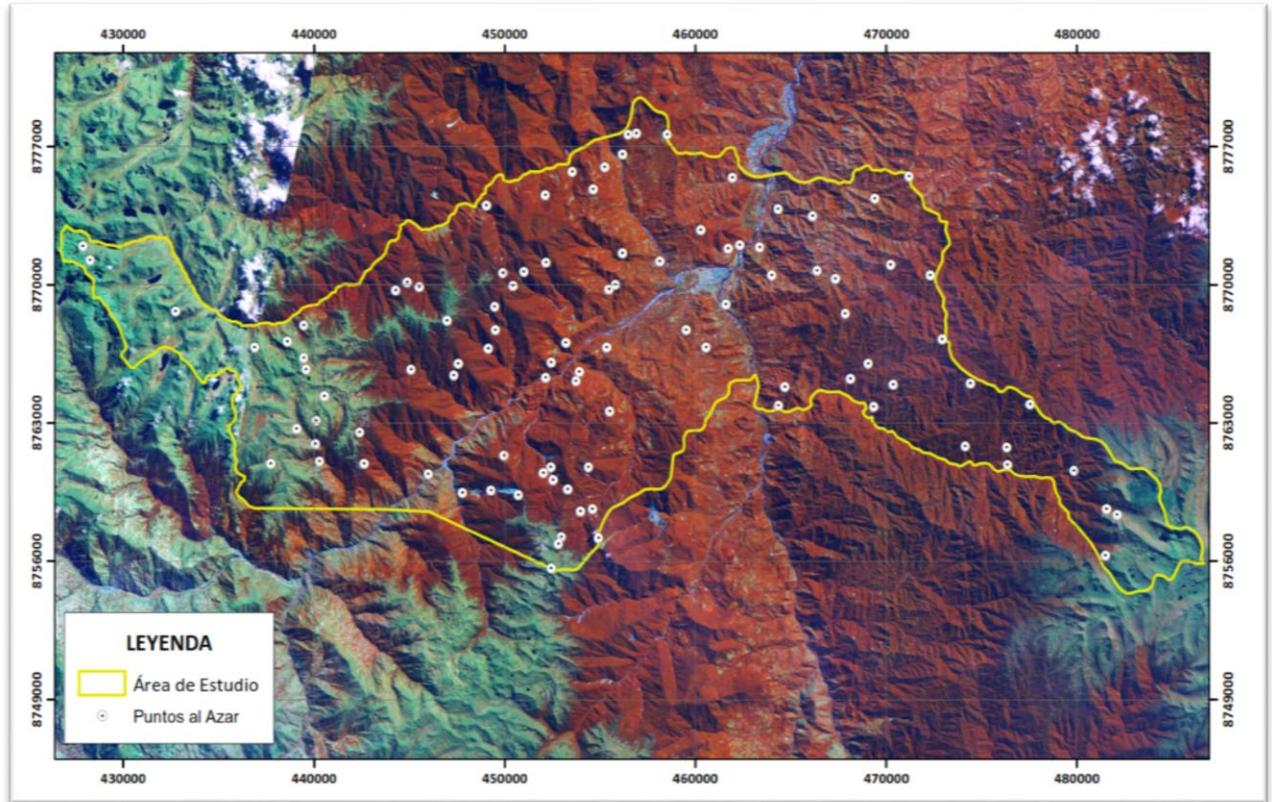


Figura 9: Mapa con los puntos al azar distribuidos en el área de estudio.

FUENTE: Science for a changing world (USGS)

IV. RESULTADOS Y DISCUSIÓN

1. UNIÓN DE BANDAS

La unión de bandas mejoró la resolución espacial de las imágenes pasando de 30 metros a 15 metros, de esta manera se mejoró la calidad visual de la imagen y así mismo preservó su información espectral. Al mejorar la calidad visual de la imagen se pudo discriminar mejor las diferentes coberturas presentes en ella.

Imagen Landsat 8 a 30 metros

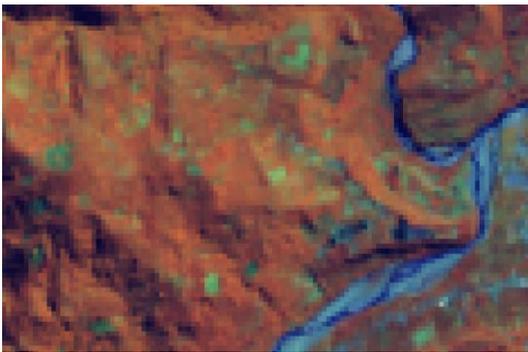


Imagen unida Landsat 8 a 15 metros

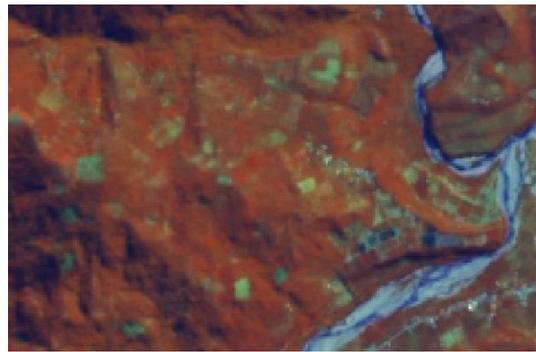


Figura 10: Vista de una imagen Landsat 8, antes y después de la unión de bandas

FUENTE: Science for a changing world (USGS)

En la Figura 10 se observa dos fragmentos de una misma imagen Landsat 8. A la izquierda, la imagen tiene 30 metros de resolución, antes del procesamiento; a la derecha, tiene 15 metros, luego de la unión de bandas.

2. CORRECCIÓN GEOMÉTRICA DE LA IMAGEN

Las imágenes Landsat 8 fueron re-proyectadas a UTM/WGS-84/Zona 18S. Si éstas imágenes satelitales no se re-proyectan no serán compatibles con el sistema de proyección oficial de la zona de estudio.

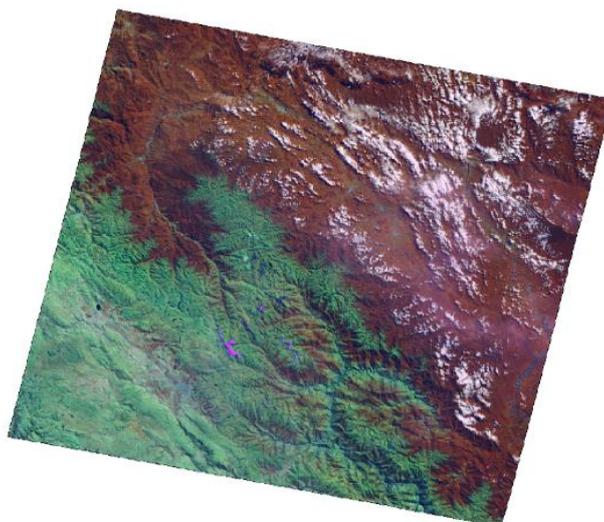


Figura 11: Imagen Landsat 8 re-proyectada a UTM/WGS-84/Zona 18S

FUENTE: Science for a changing world (USGS)

3. GENERACIÓN DEL MOSAICO

El área de estudio requirió parte de dos imágenes satelitales Landsat 8, por lo que fue necesario realizar un mosaico para cubrir la extensión del área.

Si las imágenes se hubieran trabajado independientemente los procesos posteriores tendrían que trabajarse para cada imagen por lo que tomaría mayor tiempo, debido a esto se decidió realizar un mosaico y se trabajó como una sola imagen.

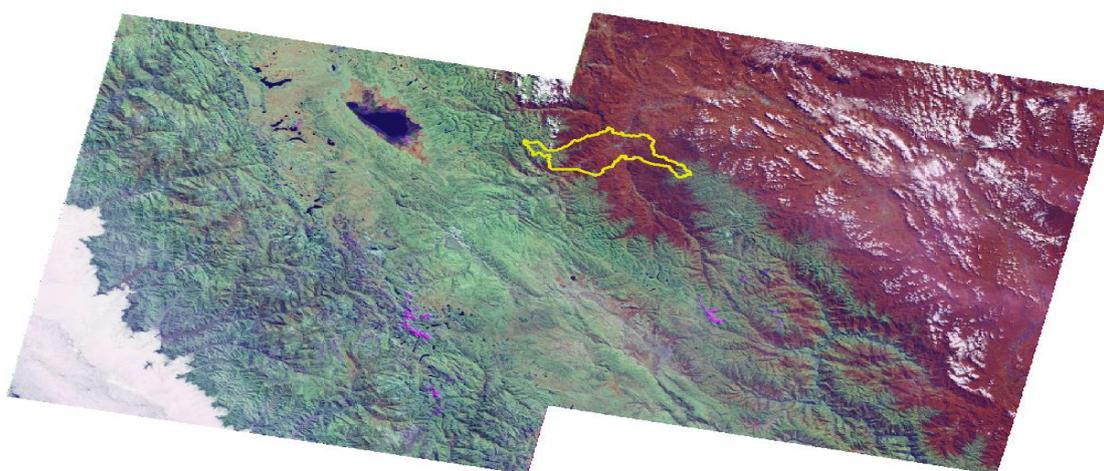


Figura 12: Mosaico generado para el área de estudio

FUENTE: Science for a changing world (USGS)

4. CORTE DEL MOSAICO

Se cortó el mosaico generado anteriormente, porque el volumen de información es grande y en el proceso de clasificación no genere información que no se necesita, facilitando así los análisis posteriores.

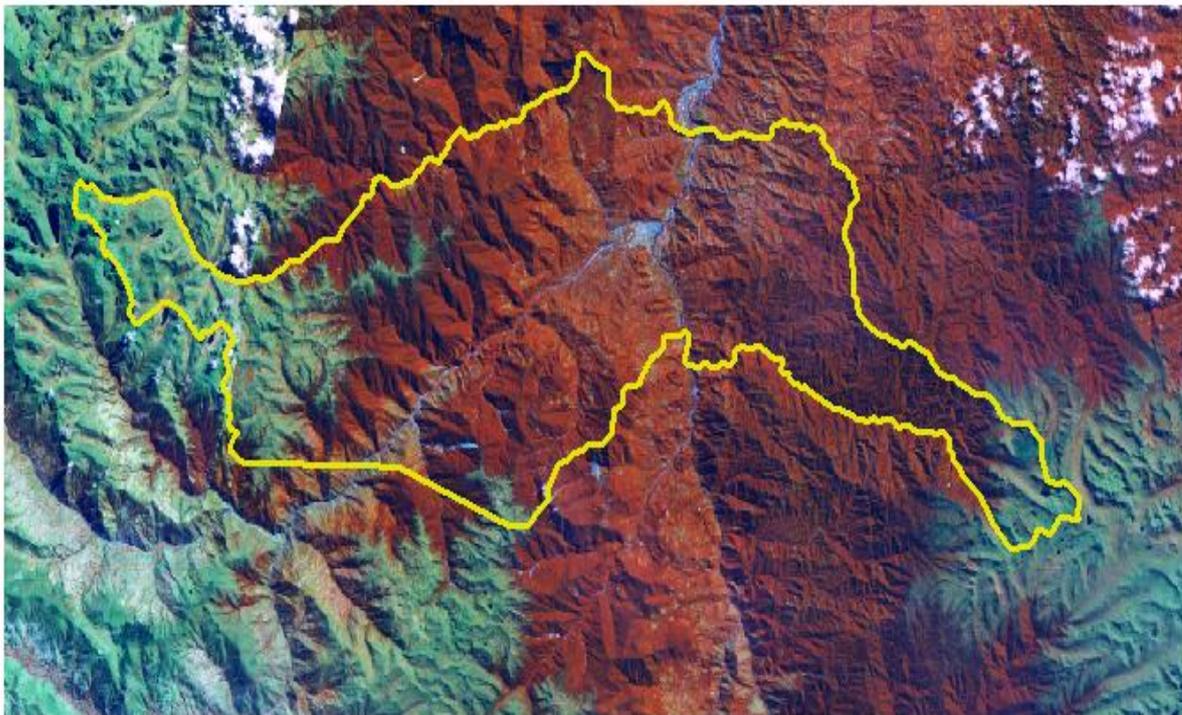


Figura 13: Corte del mosaico para el área de estudio

FUENTE: Science for a changing world (USGS)

5. GENERACIÓN DE LAS CLASES DE LA CLAVE DE CLASIFICACIÓN

El diseño de la clave de clasificación que se empleó está basado en la leyenda de CORINE Land Cover (ver Tabla 9). Para la definición de estas clases de la clave de clasificación fue muy importante el criterio del intérprete y los datos adquiridos en la salida de campo. Para el área de estudio se utilizaron 3 niveles de detalle y 13 clases.

Tabla 9: Clave de clasificación de la cobertura y uso de la tierra para el distrito de San Ramón.

1	TERRITORIOS ARTIFICIALIZADOS	
	1.1	Zonas urbanizadas
	1.1.1	Centros Poblados
	1.2	Zonas industriales o comerciales y redes de comunicación
	1.2.1	Red vial
2	TERRITORIOS AGRÍCOLAS	
	2.1	Mosaico de cultivos
	2.1.1	Mosaico agropecuario
3	BOSQUES Y ÁREAS SEMINATURALES	
	3.1	Bosques
	3.1.1	Bosque húmedo
	3.2	Áreas con vegetación herbácea y/o arbustiva
	3.2.1	Arbustos y matorrales
	3.2.2	Puna
	3.3	Áreas abiertas, sin o con poca vegetación
	3.3.1	Áreas pedregosas naturales
	3.3.2	Tierras desnudas o con poca vegetación
	3.3.3	Zonas quemadas
	3.3.4	Tierras altoandinas sin vegetación
4	ÁREAS HÚMEDAS	
	4.1	Áreas húmedas continentales
	4.1.1	Bofedales
5	SUPERFICIES DE AGUA	
	5.1	Aguas continentales
	5.1.1	Ríos y quebradas
	5.1.2	Lagunas

FUENTE: Tomado de la Leyenda de CORINE Land Cover.

Para obtener las 13 clases en la imagen se utilizaron tres tipos de áreas de referencia:

- Las áreas de referencia tipo 1 son también conocidas como áreas de entrenamiento (AE). Se usaron para clasificar las ocho clases (ver Tabla 10). Se construyeron en base a la segmentación y fueron empleadas para calibrar el modelo de clasificación.
- Las áreas de referencia tipo 2 son áreas asignadas manualmente. Corresponden a tres clases: las zonas quemadas (3.3.3), las tierras altoandinas sin vegetación (3.3.4) y los

centros poblados (1.1.1). Estas clases pueden estar bien reconocidas por la segmentación; pero si fueran asignadas mediante áreas de entrenamiento, se generaría un conflicto con las demás clases, por lo que se decidió asignarlas arbitrariamente.

- Las áreas de referencia tipo 3 son áreas que se superponen a las demás clases. Corresponden a los ríos (5.1.1) y red vial (1.2.1). En el proceso de segmentación no se logró reconocer adecuadamente por lo que fueron generadas independientemente.

Tabla 10: Clave de clasificación utilizada para calibrar el modelo de la clasificación de la cobertura y uso de la tierra para el distrito de San Ramón

2	TERRITORIOS AGRÍCOLAS	
	2.1	Mosaico de cultivos
	2.1.1	Mosaico agropecuario
3	BOSQUES Y ÁREAS SEMINATURALES	
	3.1	<i>Bosques</i>
	3.1.1	Bosque húmedo
	3.2	Áreas con vegetación herbácea y/o arbustiva
	3.2.1	Arbustos y matorrales
	3.2.2	Puna
	3.3	Áreas abiertas, sin o con poca vegetación
	3.3.1	Áreas pedregosas naturales
	3.3.2	Tierras desnudas o con poca vegetación
4	ÁREAS HÚMEDAS	
	4.1	Áreas húmedas continentales
	4.1.1	Bofedales
5	SUPERFICIES DE AGUA	
	5.1	Aguas continentales
	5.1.2	Lagunas

FUENTE: Tomado de la Leyenda de CORINE Land Cover.

6. SEGMENTACIÓN DE LA IMAGEN

Los valores óptimos obtenidos de Scale y Merge fueron de 20 y 30 respectivamente. Se definieron estos valores debido a que agrupa de manera correcta las diferentes clases presentes en la imagen. Se descartó los valores de 40 y 30 debido a que se consideró una segmentación incorrecta ya que los segmentos no reconocieron adecuadamente los límites entre diferentes clases. Sin embargo, se observó que la segmentación obtenida no identificó bien ciertas clases como: los ríos y las redes viales que posteriormente fueron editadas manualmente.

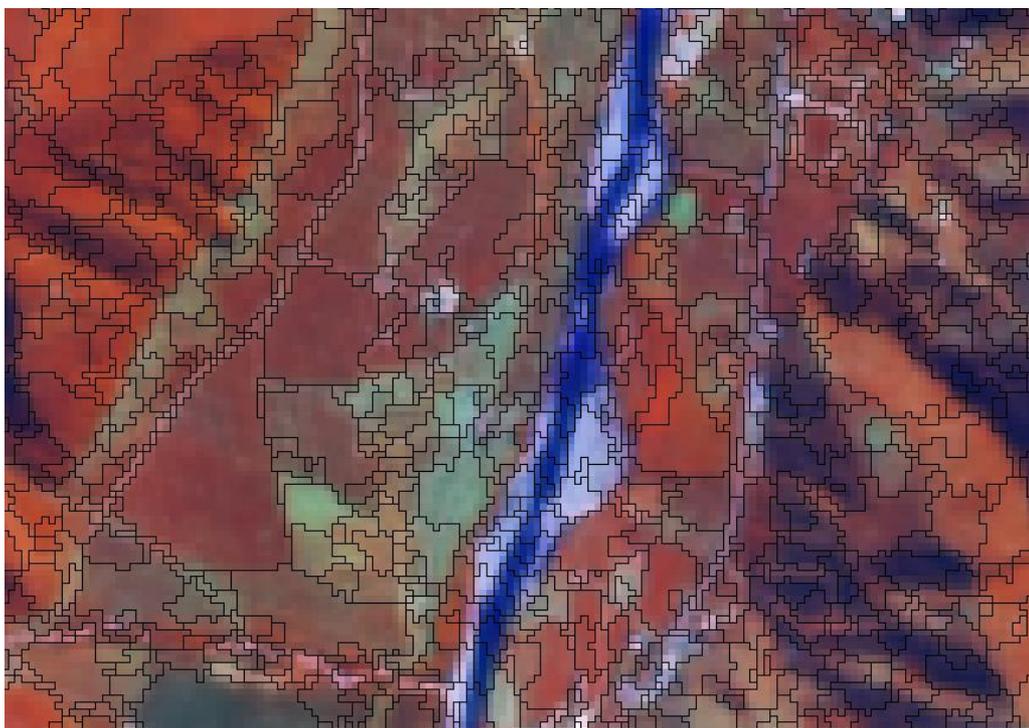


Figura 14: Segmentación adecuada para el corte del mosaico (seg_2030.shp)

FUENTE: Science for a changing world (USGS)

7. GENERACIÓN DE LAS ÁREAS DE ENTRENAMIENTO

Se generó 200 polígonos en total dentro del área de estudio para las ocho clases (Ver Tabla 10). Estos polígonos se construyeron en base a la segmentación y fueron empleadas para calibrar el modelo de clasificación. Cada polígono que se asignó se tuvo la certeza que era la clase correspondiente, debido a que la clasificación va a depender de la calidad de las mismas. Si un polígono no está bien identificado puede generar confusión en la clasificación obtenida.

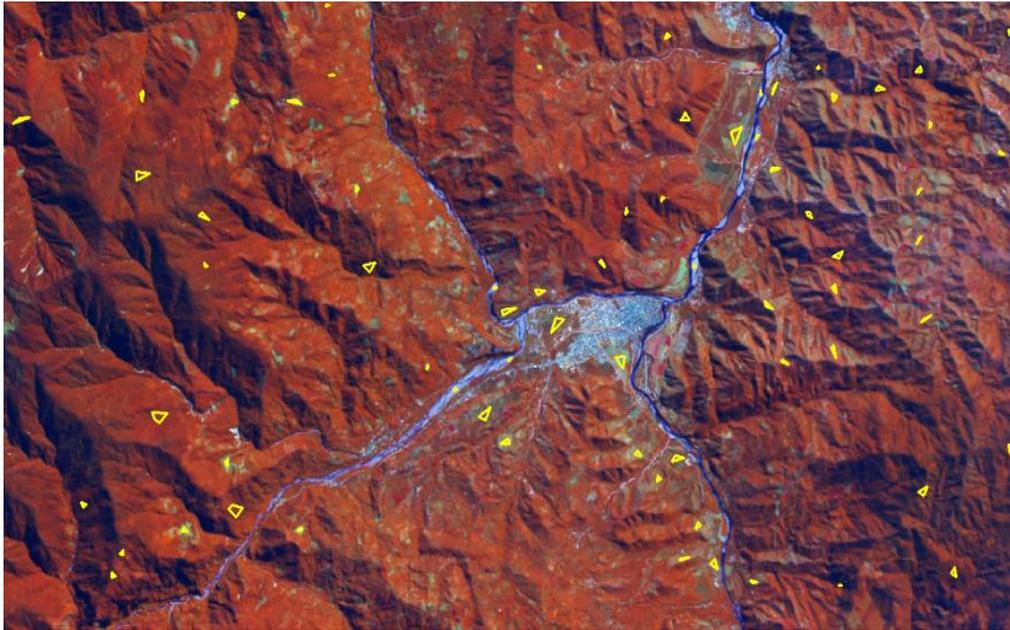


Figura 15: Área de entrenamientos distribuidos en la imagen

FUENTE: Science for a changing world (USGS)

8. INFORMACIÓN GENERADA PARA LA CLASIFICACIÓN

El resultado de las variables (capas topográficas) derivadas del raster DEM (Modelo Digital de Elevación) son los siguientes: curvatura, pendiente, forma y sombras. Estas variables obtenidas estuvieron sujetas a la compatibilidad con el sensor empleado, para evaluar la información digital de los píxeles en el proceso de clasificación e incrementar su capacidad.

Contents	
Name	Type
curv	Raster Dataset
dem	Raster Dataset
placurv	Raster Dataset
procurv	Raster Dataset
seg	Raster Dataset
slope	Raster Dataset

Figura 16: Lista de la información generada para la clasificación.

FUENTE: Elaboración propia

También se tuvo como resultado la base de datos “Todo_seg” (en formato .dbf), donde se intercambió por el archivo .dbf del shapefile de la segmentación (seg_2030), luego se cambió el nombre de Todo_seg.dbf por seg_2030, para que este lo reconozca como parte de su estructura. (Ver Anexo 3E)

9. SALIDA DE CAMPO

La salida de campo fue llevada a cabo del 28 al 31 de agosto del 2016. Como resultado se obtuvo 50 puntos colectados en todo el recorrido los cuales permitieron conceptualizar mejor las coberturas presentes dentro del área de estudio y así definir mejor las áreas de entrenamiento para el proceso de clasificación. Durante el recorrido se presentaron algunas limitaciones al momento de tomar los puntos, debido a que no se logró evaluar aquellos se encontraban en un predio privado.

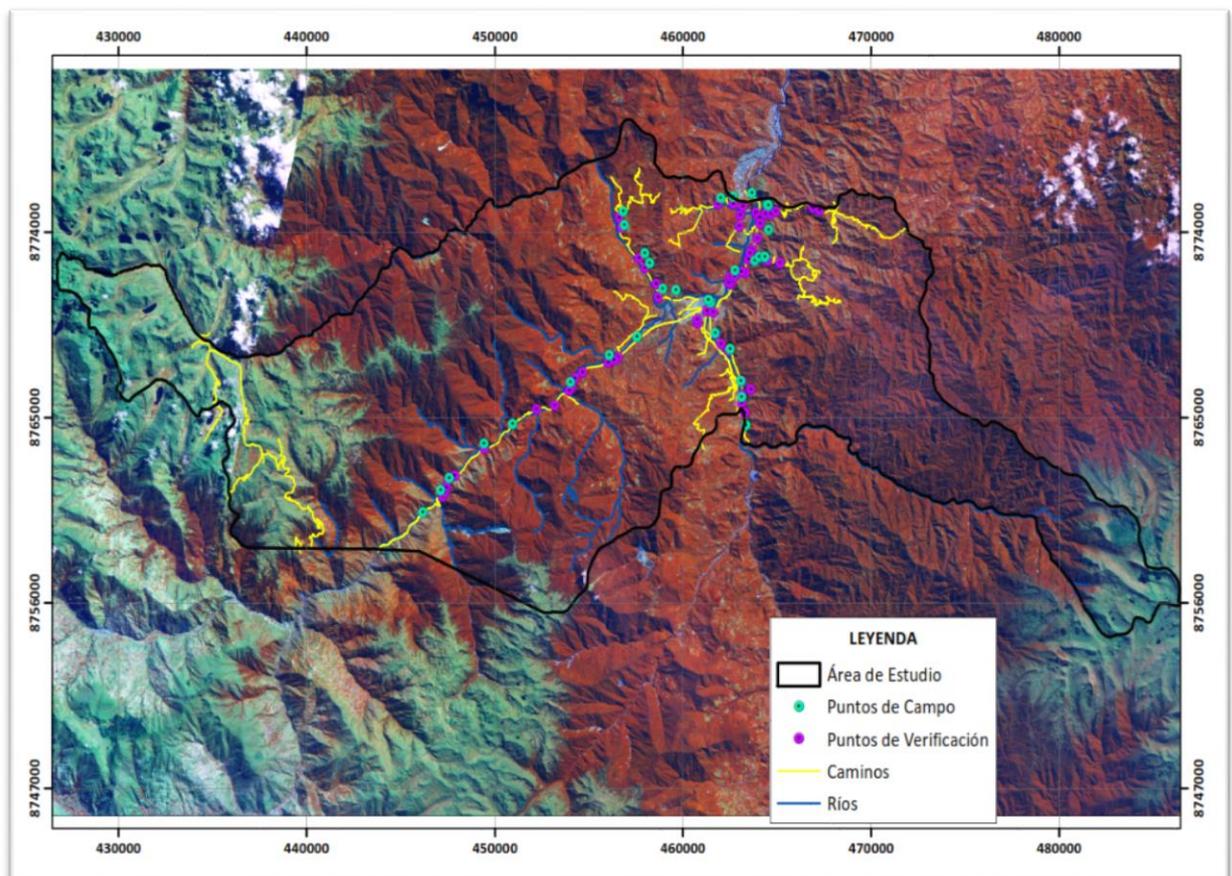


Figura 17: Mapa de la ruta y puntos colectados en campo.

FUENTE: Science for a changing world (USGS).

10. CARACTERIZACIÓN DE LAS COBERTURAS PRESENTES EN EL DISTRITO DE SAN RAMÓN.

- Centros poblados (1.1.1)

El distrito de San Ramón abarca un área de 155.07 hectáreas (0.25 por ciento), donde las viviendas se encuentran distribuidas de manera continua.

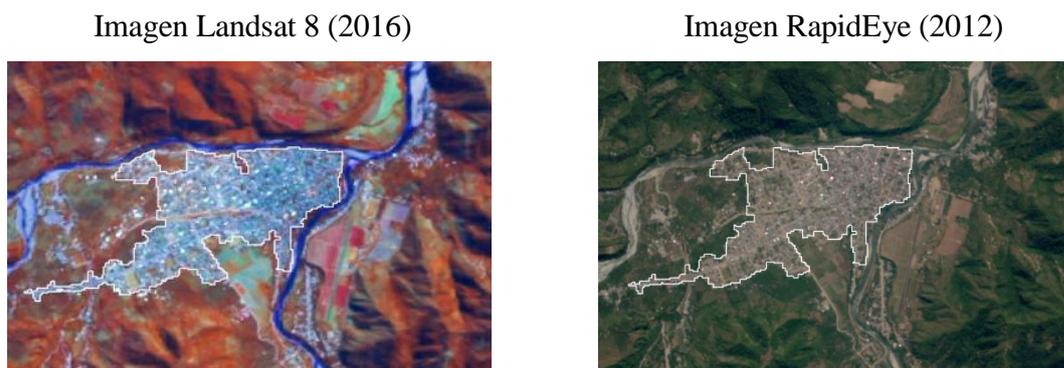


Figura 18: Vistas de las imágenes de los centros poblados

FUENTE: Science for a changing world (USGS).

- Red vial (1.2.1)

Dentro del distrito de San Ramón se observó carreteras asfaltadas y caminos afirmados, esta cobertura abarca 477.57 hectáreas (0.77 por ciento).

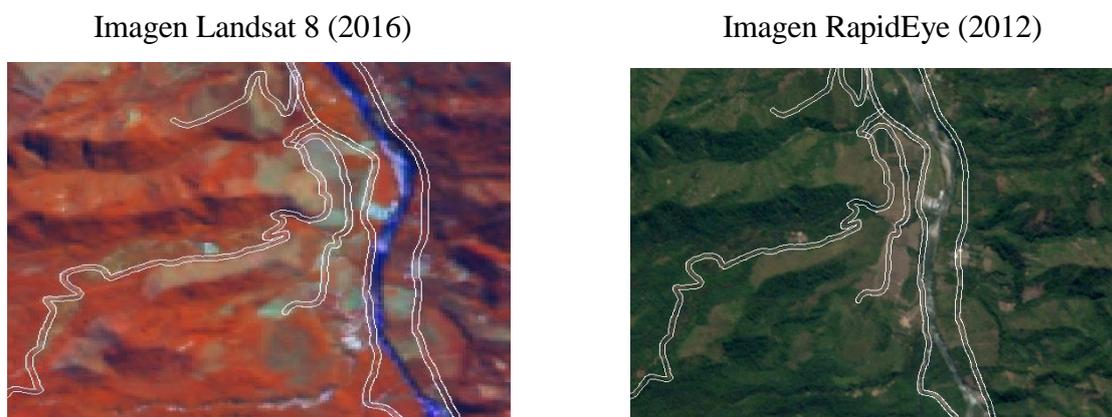


Figura 19: Vistas de las imágenes de la Red vial.

FUENTE: Science for a changing world (USGS)

- Mosaico agropecuario (2.1.1)

En el distrito de San Ramón el mosaico agropecuario comprende tierras ocupadas por cultivos, agroforestería y pastos (naturales o cultivados), donde se produce café, cacao, maíz, plátano, mandarina. En estos últimos años esta cobertura está avanzando debido al cambio de tierras forestales a usos agrícolas. Dentro de San Ramón la superficie de esta cobertura abarca 3,315.96 hectáreas (5.33 por ciento).

Imagen Landsat 8 (2016)



Imagen RapidEye (2012)

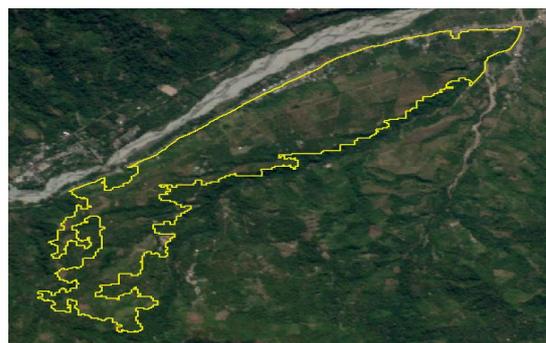


Figura 20: Vistas de las imágenes del mosaico agropecuario

FUENTE: Science for a changing world (USGS)

- Bosque húmedo (3.1.1)

En el valle de Chanchamayo se encuentra desde Bosques húmedos y muy húmedos hasta Bosques muy húmedo Montano desde los 1000 a 3000 m.s.n.m (Reynel 2012). De estos bosque se extrajeron madera muy valiosa como son el Cedro (*Cedrela montana*) y el Diablo Fuerte (*Podocarpus spp*); actualmente muy escasos. También se encuentran las familias de las Rubeaceas y Lauraceas. Dentro del área de estudio esta cobertura abarca 43,350.43 hectáreas (69.65 por ciento).

Imagen Landsat 8 (2016)



Imagen RapidEye (2012)

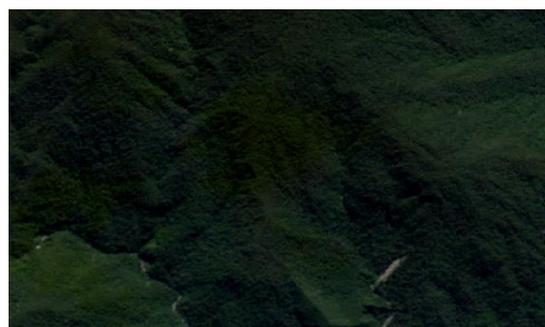


Figura 21: Vistas de las imágenes de los bosques húmedos.

FUENTE: Science for a changing world (USGS)

- Arbustos y matorrales (3.2.1)

Dentro del valle de Chanchamayo desde los 1500 hasta 2300 m.s.n.m. se encuentra los Bosques achaparrados y los Bosques secundarios tempranos o vegetación arbustiva de porte bajo (Reynel, 2012). En el distrito de San Ramón esta cobertura abarca 2,778.30 hectáreas (4.46 por ciento).

Imagen Landsat 8 (2016)

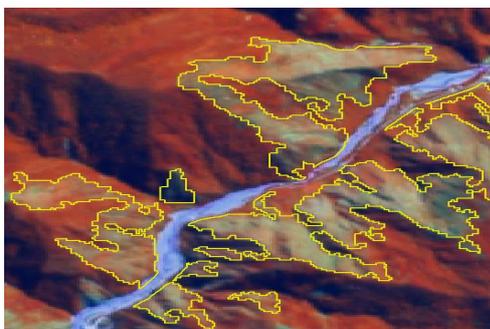


Imagen RapidEye (2012)



Figura 22: Vistas de las imágenes de arbustos y matorrales.

FUENTE: Science for a changing world (USGS)

- **Puna (3.2.2)**

En el distrito de San Ramón esta cobertura se encuentra desde 3500 hasta 4000 m.s.n.m., ubicado en la porción superior de la cordillera de los andes, está conformada mayormente por herbazales y abarca 9,809.38 hectáreas (15.76 por ciento).

Imagen Landsat 8 (2016)

Google Earth, Imagen SPOT (2014)



Figura 23: Vistas de las imágenes de la puna.

FUENTE: Science for a changing world (USGS)

- **Áreas pedregosas naturales (3.3.1)**

Dentro del área de estudio esta cobertura se encuentra en los bordes de los ríos conformando playas de piedras, generalmente desprovistos de vegetación, abarca 80.02 hectáreas (0.13 por ciento).

Imagen Landsat 8 (2016)

Imagen RapidEye (2012)

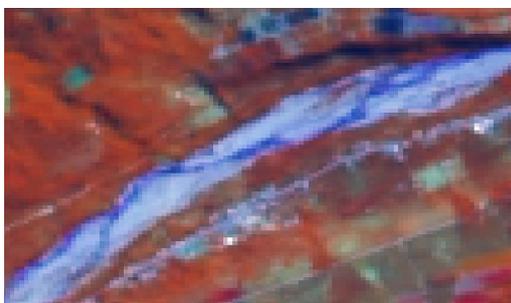


Figura 24: Vistas de las imágenes de las áreas pedregosas naturales.

FUENTE: Science for a changing world (USGS)

- **Tierras desnudas o con poca vegetación (3.3.2)**

Dentro del distrito de San Ramón se observó tierras desnudas por deslizamientos o por degradación extrema, esta cobertura abarca 77.79 hectáreas (0.12 por ciento).

Imagen Landsat 8 (2016)



Imagen RapidEye (2012)



Figura 25: Vistas de las imágenes de tierras desnudas o con poca vegetación.

FUENTE: Science for a changing world (USGS)

- **Zonas quemadas (3.3.3)**

Son zonas afectadas por incendios localizados debido al cambio de zonas de bosque a usos agrícolas favoreciendo a la deforestación. Otra de las actividades que favorece a estas zonas afectadas es la agricultura migratoria que son realizadas por quemas. Dentro del área de estudio esta cobertura abarca 58.49 hectáreas (0.09 por ciento).

Imagen Landsat 8 (2016)

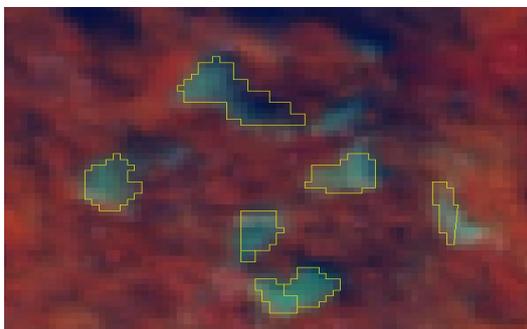


Imagen RapidEye (2012)



Figura 26: Vistas de las imágenes de las zonas quemadas.

FUENTE: Science for a changing world (USGS)

- Tierras altoandinas sin vegetación (3.3.4)

Este tipo de cobertura dentro del área de estudio se encuentra desde los 4200 m.s.n.m. en esta zona las herbáceas pierden cobertura debido a las condiciones extremas del clima, dejando áreas con suelos desnudos o afloramientos rocosos, esta cobertura abarca 534.02 hectáreas (0.86 por ciento).

Imagen Landsat 8 (2016)

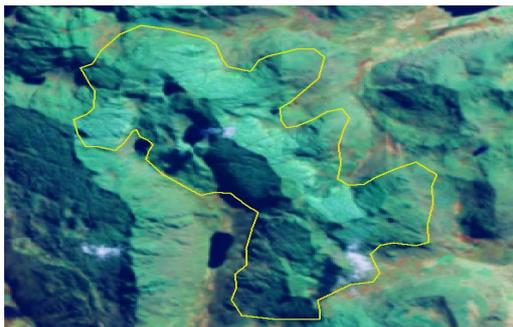


Imagen RapidEye (2012)

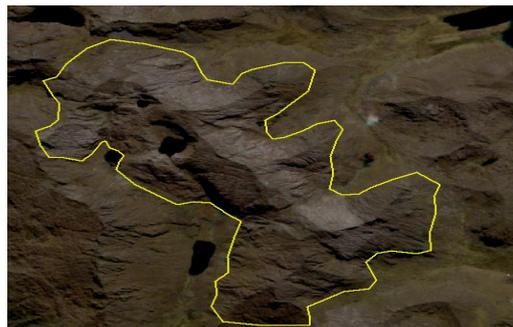


Figura 27: Vistas de las imágenes de tierras altoandinas sin vegetación.

FUENTE: Science for a changing world (USGS)

- Bofedales (4.1.1)

Dentro del área de estudio se encuentra cerca al pajonal altoandino o a las tierras sin vegetación, a partir de los 3800 m.s.n.m., esta cobertura abarca 900.30 hectáreas (1.45 por ciento).

Imagen Landsat 8 (2016)

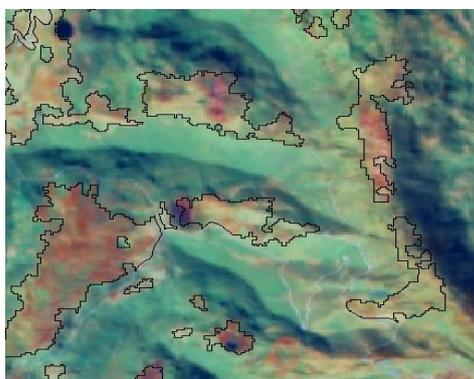


Imagen RapidEye (2012)



Figura 28: Vistas de las imágenes de los bofedales.

FUENTE: Science for a changing world (USGS)

- **Ríos y quebradas (5.1.1)**

Dentro del área de estudio pasan dos ríos principales el río Tarma y el río Tulumayo y abarcan 613.22 hectáreas (0.99 por ciento).

Imagen Landsat 8 (2016)



Google Earth, Imagen SPOT (2012)

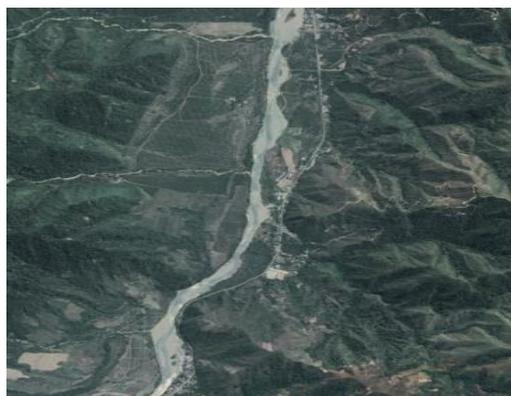


Figura 29: Vistas de las imágenes de ríos y quebradas.

FUENTE: Science for a changing world (USGS)

- **Lagunas (5.1.2)**

Dentro del área de estudio se encuentran en las zonas altas y abarcan 85.47 hectáreas (0.14 por ciento).

Imagen Landsat 8 (2016)



Imagen RapidEye (2012)

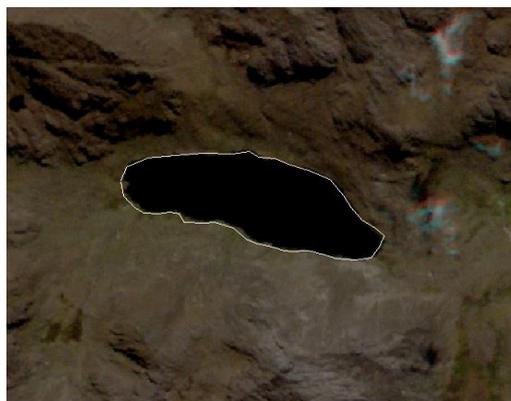


Figura 30: Vistas de las imágenes de las lagunas.

FUENTE: Science for a changing world (USGS)

11. CLASIFICACIÓN SUPERVISADA

El resultado de la clasificación supervisada utilizando el algoritmo *Random Forest* de la primera iteración fue revisado visualmente y se observó que había confusión en algunas coberturas, por lo que se decidió realizar una segunda iteración, para lo cual se modificó algunas áreas de entrenamiento. Finalmente la clasificación supervisada fue nuevamente revisada visualmente y tuvo un resultado satisfactorio, el cual se denominó “it_2.shp”.

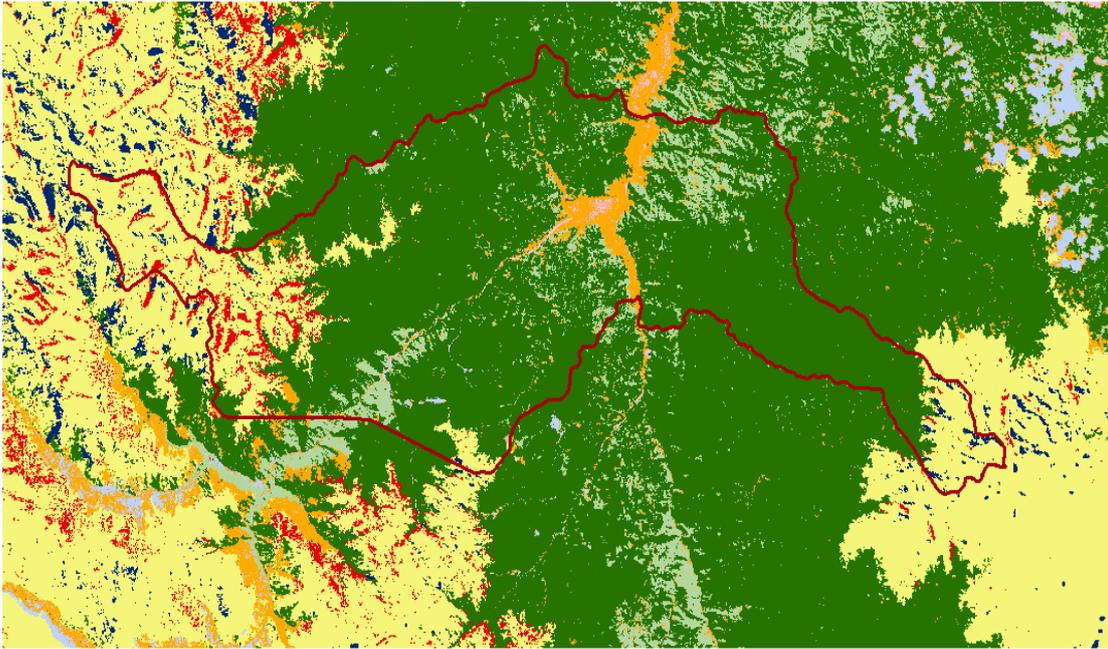


Figura 31: Clasificación supervisada denominada (it_2.shp)

FUENTE: Elaboración propia

12. COBERTURA Y USO DE LA TIERRA DEL DISTRITO DE SAN RAMÓN.

Tabla 11: Superficie de los tipos de cobertura y uso de la tierra del distrito de San Ramón para el año 2016.

<i>Tipos de Cobertura</i>	<i>Superficie (ha)</i>	<i>%</i>
Centros poblados	155.07	0.25
Red vial	477.57	0.77
Mosaico agropecuario	3,315.96	5.33
Bosque húmedo	43,350.43	69.65
Arbustos y matorrales	2,778.30	4.46
Puna	9,809.38	15.76
Áreas pedregosas naturales	80.02	0.13
Tierras desnudas o con poca vegetación	77.79	0.12
Zonas quemadas	58.49	0.09
Tierras altoandinas sin vegetación	534.02	0.86
Bofedales	900.30	1.45
Ríos y quebradas	613.22	0.99
Lagunas	85.47	0.14
TOTAL	62,236.01	100

FUENTE: Elaboración propia

La Tabla 11 muestra los resultados de la superficie de las 13 clases del mapa de cobertura y uso de la tierra del distrito de San Ramón identificado para el año 2016. La cobertura con mayor porcentaje son los bosques húmedos con 43,350.43 hectáreas (69.65 por ciento), seguido por la puna con 9,809.38 hectáreas (15,76 por ciento), mosaico agropecuario con 3,315.96 hectáreas (5.33 por ciento), arbustos y matorrales con 2,778.30 hectáreas (4.46 por ciento). Las demás coberturas ocupan un 4.8 por ciento.

En la Figura 32 se observa la superficie de los tipos de cobertura de la tierra del distrito de San Ramón para el año 2016. La cobertura de la tierra con mayor porcentaje son los bosques húmedos con un 69.65 por ciento, de estos bosques se han extraído maderas muy valiosas como son el Cedro (*Cedrela montana*) y el Diablo Fuerte (*Podocarpus spp*); actualmente muy escasos, seguido por la puna con un 15,76 por ciento, arbustos y matorrales con 4.46 por ciento y las demás coberturas de la tierra representan ocupan un 3.69 por ciento.

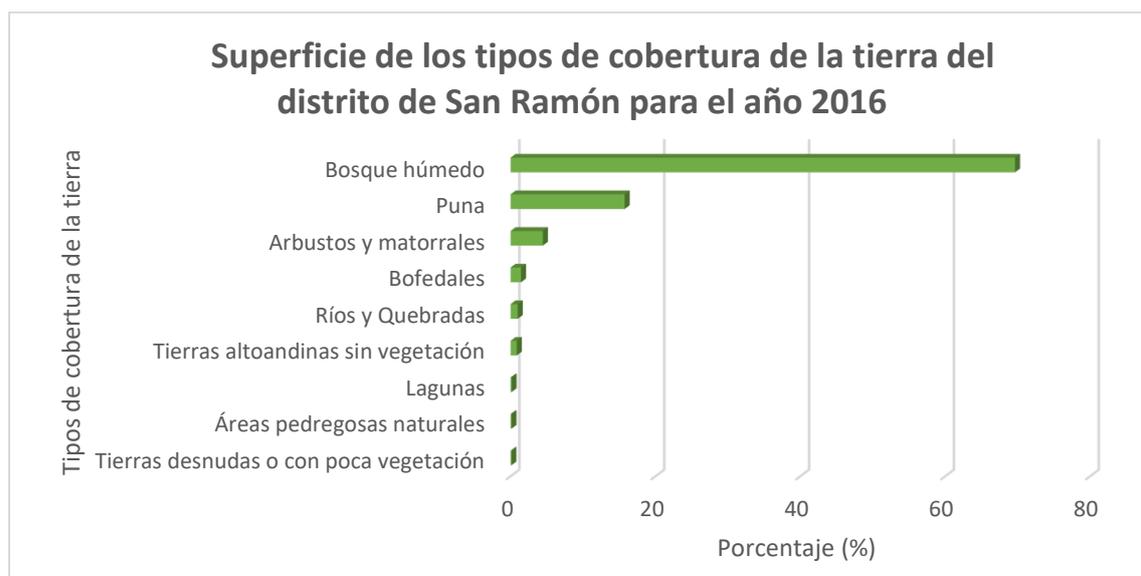


Figura 32: Superficie de los tipos de cobertura de la tierra del distrito de San Ramón para el año 2016.

FUENTE: Elaboración propia

En la Figura 33 se observa la superficie de los tipos de uso de la tierra del distrito de San Ramón para el año 2016. Los usos de la tierra con mayor porcentaje es el mosaico agropecuario con un 5, 33 por ciento; esta cobertura comprende tierras ocupadas por cultivos, agroforestería y pastos (naturales o cultivados), donde se produce café, cacao, maíz, plátano, mandarina; en estos últimos años esta cobertura está avanzando debido al cambio de tierras forestales a usos agrícolas, seguido por la red vial con un 0,77 por ciento, los centros poblados con un 0,25 por ciento y las zonas quemadas con un 0,09 por ciento, son zonas afectadas por incendios localizados debido al cambio de zonas de bosque a usos agrícolas favoreciendo a la deforestación y otra de las actividades que favorece a estas zonas afectadas es la agricultura migratoria que son realizadas por quemas.

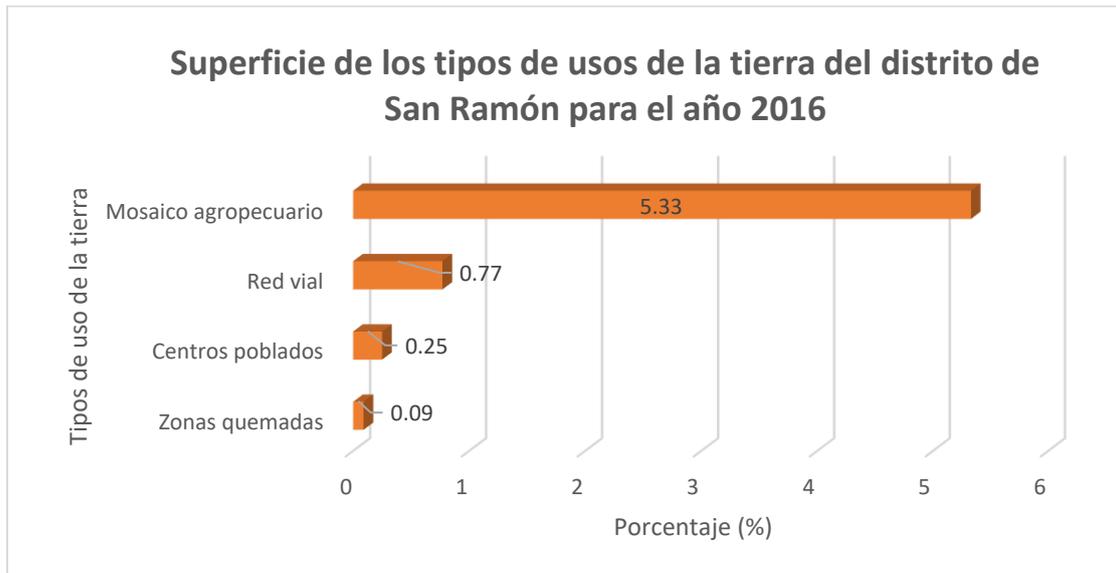


Figura 33: Superficie de los tipos de uso de la tierra del distrito de San Ramón para el año 2016.

FUENTE: *Elaboración propia*

13. PROCESO DE CLASIFICACIÓN SUPERVISADA

En el proceso de clasificación supervisada se observa un alto aporte del algoritmo de clasificación *Random Forest* para sus dos indicadores: el índice kappa y el porcentaje correctamente clasificado.

Tabla 12: Resultados de los indicadores del proceso de clasificación supervisada para la IT2.

<i>IK</i>	<i>PCC (%)</i>
0.972	0.977887

Donde: *IK*=Índice Kappa, *PCC*= Porcentaje Correctamente Clasificado

FUENTE: *Elaboración propia*

La Tabla 12 muestra los resultados de los dos indicadores del proceso de clasificación supervisada de la última iteración (IT2): el porcentaje correctamente clasificado que arrojó un valor de 0.977 y el índice Kappa que arrojó un valor de 0.972. Este último valor, según Landis y Koch, citado por Cerda y Villarroel (2008), se encuentra en el rango de “fuerza de concordancia casi perfecta”. Es decir, la clasificación supervisada generada utilizando el algoritmo *Random Forest* tiene un alto grado de correlación con la realidad.

14. VALIDACIÓN DEL MAPA

La validación del mapa ha sido calculada a través de la generación de los indicadores de precisión del mapa, que determina el grado de correlación del mapa de cobertura y uso de la tierra para el distrito de San Ramón con la realidad medida en campo y puntos al azar verificadas con imágenes de alta resolución RapidEye e imágenes de la plataforma Google Earth.

Tabla 13: Resultados de los indicadores de precisión del mapa.

<i>IK</i>	<i>PCC (%)</i>	<i>Precisión global del mapa</i>
0.922	0.9533	0.9533

Donde: IK=Índice Kappa, PCC= Porcentaje Correctamente Clasificado

FUENTE: Elaboración propia

La Tabla 13 muestra los resultados de los indicadores de precisión del mapa de cobertura y uso de la tierra para el distrito de San Ramón: la precisión global del mapa con un valor de 0.9533, el índice Kappa con un valor de 0.922, lo cual señala un alto grado de correlación con la realidad o exactitud temática del mapa producido; y el porcentaje correctamente clasificado que arrojó un valor de 0.953.

La Tabla 14 muestra los resultados de la matriz de confusión correspondiente a la clasificación del análisis de cobertura y uso de la tierra para el distrito de San Ramón, donde se contrastó el mapa final con los puntos de validación (al azar y campo).

La diagonal de la matriz muestra los puntos donde concuerdan ambas fuentes. Los puntos ubicados fuera de la diagonal suponen errores de omisión y comisión. A partir de esta matriz se calculó un indicador de precisión global del mapa que arrojó un valor de 0.9533, que considera la relación entre el número de puntos correctamente asignados (sumatoria de la diagonal) y el total de puntos de validación (al azar y campo).

La Figura 34 muestra el mapa de cobertura y uso de la tierra para el año 2016. Con este mapa se podrá tomar decisiones que contribuyan a la planificación del manejo sostenible del territorio y los recursos forestales, principalmente en la conservación de bosques y otros componentes que dependen de él. Así mismo facilitará a la gestión informada como son las autoridades del gobierno regional y local del departamento de Junín.

La metodología empleada que incluye la aplicación del algoritmo *Random Forest* permitió cuantificar las coberturas del área de estudio. Esto sumado a la experiencia del intérprete se obtuvo resultados más precisos.

Tabla 14: Matriz de confusión correspondiente a la clasificación del análisis de cobertura y uso de la tierra del distrito de San Ramón con puntos de validación de campo.

Mapa Final

Puntos de Validación (Al Azar Y Campo)

	111	121	211	311	321	322	331	332	333	334	411	511	512	Total	Exactitud	E. Omisión
111	2	0	0	0	0	0	0	0	0	0	0	0	0	2	100.00	0.00
121	0	1	0	0	0	0	0	0	0	0	0	0	0	1	100.00	0.00
211	0	1	19	1	0	0	0	0	0	0	0	0	0	21	90.48	9.52
311	0	2	1	89	0	0	0	0	0	0	0	0	0	92	96.74	3.26
321	0	0	0	0	14	0	0	0	0	0	0	1	0	15	93.33	6.67
322	0	0	0	0	0	10	0	0	0	0	0	0	0	10	100.00	0.00
331	0	0	0	0	0	0	1	0	0	0	0	0	0	1	100.00	0.00
332	0	0	0	0	0	0	0	1	0	0	0	0	0	1	100.00	0.00
333	0	0	0	0	0	0	0	0	1	0	0	0	0	1	100.00	0.00
334	0	0	0	0	0	0	0	0	0	1	0	0	0	1	100.00	0.00
411	0	0	0	0	0	1	0	0	0	0	2	0	0	3	66.67	33.33
511	0	0	0	0	0	0	0	0	0	0	0	1	0	1	100.00	0.00
512	0	0	0	0	0	0	0	0	0	0	0	0	1	1	100	0.00
Total	2	4	20	90	14	11	1	1	1	1	2	2	1	150		
Exactitud	100.00	25.00	95.00	98.89	100.00	90.91	100.00	100.00	100.00	100.00	100.00	50.00	100.00		Precisión global del mapa	0.9533
E. Comisión	0.00	75.00	5.00	1.11	0.00	9.09	0.00	0.00	0.00	0.00	0.00	50.00	0.00			

Fuente: Elaboración propia

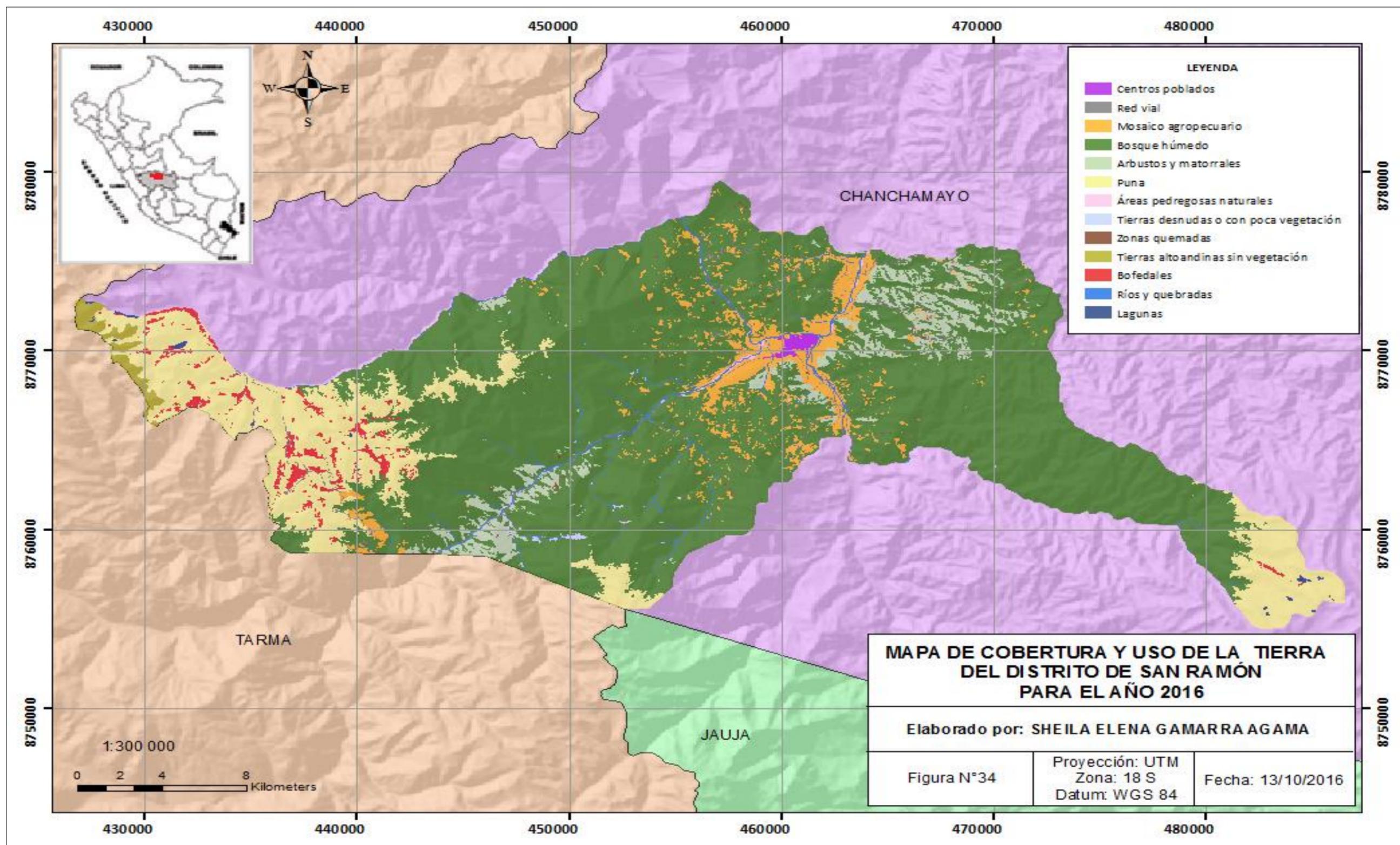


Figura 34: Mapa de cobertura y uso de la tierra del distrito de San Ramón para el año 2016

Fuente: Elaboración propia

V. CONCLUSIONES

- 1) El mapa de cobertura y uso de la tierra del distrito de San Ramón para el año 2016, es un insumo útil que contribuye a tomar decisiones acertadas en la planificación del manejo sostenible del territorio, los recursos forestales; principalmente en la conservación de los bosques y otros componentes que dependen de él, así mismo facilitará a la gestión informada como son las autoridades del gobierno regional y local del departamento de Junín; la cual responde al objetivo general.
- 2) El mapa de cobertura y uso de la tierra, ayudará a reconocer las actividades de conservación, recuperación y usos sostenibles de las fuentes de los servicios ecosistémico, que es uno de los objetivos de la Resolución Ejecutiva Regional N° 198-2017-GRJ/GR del gobierno regional del departamento de Junín.
- 3) La metodología empleada en el presente estudio muestra resultados eficientes para la clasificación de las imágenes junto con el algoritmo Random Forest.
- 4) El análisis de validación del mapa muestra un alto grado de correlación con la realidad.
- 5) El mapa de cobertura y uso de la tierra del distrito de San Ramón será base inicial para el monitoreo de la dinámica de los bosques de manera periódica.

VI. RECOMENDACIONES

- La información generada debería ser actualizada anualmente, a fin de verificar la evolución de los cambios en los diferentes tipos de coberturas y usos de la tierra del distrito de San Ramón.
- Se recomienda hacer un análisis multitemporal, tomando como base la información generada, permitiendo estimar la evolución de cambios teniendo en cuenta diferentes momentos del tiempo para hacer proyecciones sobre el futuro.
- Se sugiere la capacitación a los gobiernos regionales en la elaboración de mapas de cobertura y uso de la tierra y en la implementación de la metodología, de manera que pueda ser utilizada en la toma de decisiones orientadas a la planificación y el ordenamiento del territorio.
- Para estudios posteriores se recomienda trabajar con imágenes de alta resolución permitiendo un análisis más preciso de la cobertura y determinación del uso de la tierra.

VII. REFERENCIAS BIBLIOGRÁFICAS

- Alcácer, E. 2013. Análisis de riesgo crediticio de las empresas españolas mediante el uso de “*decisión trees*“. Tesis Mg. Sc. Valencia, ES, Universidad Politécnica de Valencia. 178 p.
- Antón, D; Reynel, C. 2004. Relictos de bosque de excepcional diversidad en los andes centrales del Perú. 1 ed. Lima, PE. 316 p.
- APN – SIB (Administración de Parques Nacionales – Sistema de Información de Biodiversidad, AR). 2005. Protocolo para el procesamiento de imágenes satelitales Landsat para aplicaciones de la administración de parques nacionales. Buenos Aires, AR. 21 p.
- Arenas, S; Haeger, JF; Jordano, D. 1998. Aplicación de técnicas de teledetección y GIS sobre imágenes Quickbird para identificar y mapear individuos de peral silvestre (*Pyrus bourgeana*) en bosque esclerófilo mediterráneo. Córdoba. ES. 16 p.
- ATECMA S.L. 2007. Identificación y diagnóstico de la red de corredores ecológicos de la región de Murcia. Metodología. ES. 83 p.
- Balderrama, C. 2008. Algoritmos. Metodología de la Programación I. 36 p.
- Barton, M. s.f. Manual Pansharpen. GRASS GIS (en línea). Arizona, USA. Consultado 20 oct. 2016. Disponible en <https://grass.osgeo.org/grass73/manuals/i.pansharpen.html>
- BIODAMAZ (Proyecto Diversidad Biológica de la Amazonía Peruana, PE). 2004. Manual para la elaboración de mosaicos de imágenes de satélite Landsat TM para selva baja peruana. Iquitos, PE. 120 p.
- Bonsón, E; Escobar, T; Martín, MP. 1999. Aplicación de los sistemas de inducción de árboles de decisión a la gestión empresarial: Toma de decisiones y control de tareas de decisión. Madrid, ES. 20 p.

- Breiman, L. 2001. Random Forest. Machine Learning (en línea). California, USA. Consultado 27 abr. 2016. Disponible en <http://www.springerlink.com/content/u0p06167n6173512/fulltext.pdf>
- Cabello, J. 2008. Bofedales en el Lago Chungara. Parque Nacional Lauca (en línea). Arica, CL. Consultado 22 ago. 2016. Disponible en <https://www.flickr.com/photos/jorgeleoncabello/5327464585>
- Cánovas, F; Alonso, F. 2014. Comparación de técnicas de fusión en imágenes de alta resolución espacial. Revista Internacional de Ciencia y Tecnología de la Información Geográfica. 18 p.
- Carnegie Institution for Science, USA. 2013. CLASlite, tecnología para el monitoreo de bosques. Guía del usuario (en línea). Stanford, USA. Consultado 5 mayo 2016. Disponible en <ftp://globalecology.stanford.edu/pub/clas/docs/Espa%F1ol/CLASlite%20Guia%20del%20Usuario%20ESPA%D1OL.pdf>
- Castillo, M; Samaniego, R; Kindgard, A. 2015. Mapa de cobertura y uso de la tierra 2012. Programa de colaboración de las Naciones Unidas para la reducción de emisiones de la deforestación y la degradación de bosques en los países en desarrollo (ONU-REDD). Panamá. 89 p.
- CDC – UNALM (Centro de Datos para la Conservación – Universidad Agraria la Molina, PE) 2012. Manual de la metodología de clasificación semiautomática, para el mapeo de cobertura y uso de la tierra, empleando *Random Forest*. Lima, PE. 70 p.
- Cerda, J; Villarroel, L. 2008. Evaluación de la concordancia inter-observador en investigación pediátrica: Coeficiente de Kappa. (en línea). Santiago, Chile. Consultado 12 abr. 2016. Disponible en <http://www.scielo.cl/pdf/rcp/v79n1/art08.pdf>
- Chuvieco, E. 1995. Fundamentos de Teledetección Espacial. 2 ed. Ediciones RIALP. 415 p.
- Corso, C. s.f. Aplicación de algoritmos de clasificación supervisada usando Weka. Tesis. Córdoba, AR, Universidad Tecnológica Nacional. 11 p.

- Di Gregorio, A. 2005. Sistema de clasificación de la cobertura de la tierra (LCCS), versión 2: Conceptos de clasificación y manual del usuario. Roma, IT. 208 p.
- Duarte, E. s.f. Mapa forestal de Honduras derivado con imágenes de alta resolución, una herramienta innovadora y útil para la toma de decisiones nacionales y sub-nacionales. Instituto Nacional de Conservación y Desarrollo Forestal, Áreas Protegidas y Vida Silvestre (ICF). Comayagüela, HO. 13 p.
- ESRI (Environmental Systems Research Institute) 2008. ArcGis Help 9.2 (en línea). Consultado 16 oct. 2016. Disponible en <http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=Curvature>
- FAO (Organización de Naciones Unidas para la Alimentación y la Agricultura, UY) 2015. Atlas de cobertura del suelo de Uruguay. Cobertura de suelo y cambios 2000 – 2011. Land Cover Classification System. Montevideo, UY. 52 p.
- Fernandez, W; Gilbes, F. s.f. Disponibilidad de datos de Landsat 8 y su procesamiento en arcgis. Geographic mapping technologies. 17 p.
- García, A; Piriz, V; Gasparri, N. 2011. Cambio en la cobertura del suelo en la cuenca media y alta del río Lules. Tucumán, AR. 13 p.
- GEOIDEP (Portal de la Infraestructura de Datos Espaciales del Perú, PE) 2003. Metadatos. (en línea). Lima, Perú. Consultado 18 ago. 2016. Disponible en <http://www.geoidep.gob.pe/metadatos>
- GloVis (Global Visualization Viewer, USA) 1879. USGS. Servicio Geológico de los Estados Unidos, por sus siglas en inglés (United States Geological Survey). (en línea), USA. Consultado 18 jul. 2016. Disponible en <http://glovis.usgs.gov/>
- Herrera, P; Correa da Silva, I; Soria, L; Quezada, B; Calderón, V; Calzadilla, M; Flores, M. 2009. Mapa de cobertura y uso de la tierra del departamento del Beni-2008. Memoria explicativa. Fundación Amigos del Museo NKM y Museo de Historia Natural Noel Kempff Mercado. Santa Cruz. 31 p.
- IDEAM (Instituto de Hidrología, Meteorología y Estudios Ambientales, CO) 2010. Leyenda nacional de coberturas de la tierra. Metodología CORINE Land Cover adaptada para Colombia Escala 1:100.000. Bogotá, CO. 72 p.

- IDEAM; IGAC; CORMAGDALENA (Instituto de Hidrología, Meteorología y Estudios Ambientales; Corporación Autónoma Regional del Río Grande de la Magdalena; Instituto Geográfico Agustín Codazzi, CO) 2008. Mapa de Cobertura de la Tierra Cuenca Magdalena-Cauca: Metodología CORINE Land Cover adaptada para Colombia a escala 1:100000. Bogotá. CO. 199 p.
- IGAC (Instituto Geográfico Agustín Codazzi, CO) 2013. Descripción y Corrección de Productos Landsat 8. Landsat Data Continuity Mission (LDCM) Versión 1.0, Grupo interno de trabajo de percepción remota y aplicaciones geográficas. Bogotá, CO. 46 p.
- INEGI (Instituto Nacional de Estadística y Geografía, MX) s.f. Modelo Digital de Elevación (MDE). Descripción. (en línea) Ciudad de México, MX. Consultado 15 ago. 2016. Disponible en <http://www.inegi.org.mx/geo/contenidos/datosrelieve/continental/queesmde.aspx>
- INEI (Instituto Nacional de Estadística e Informática, PE) 1998. Características del proceso de retorno y reasentamiento de la población desplazada. (en línea). Lima, PE. Consultado 11 ago. 2016. Disponible en <http://proyectos.inei.gob.pe/web/biblioineipub/bancopub/Est/Lib0017/indice.htm>
- ITT (Visual Information Solutions All Rights Reserved) 2009. ENVI EX User's Guide. (en línea). Consultado 28 oct. 2016. Disponible en http://www.exelisvis.com/portals/0/pdfs/enviex/ENVI_EX_User_Guide.pdf
- Lara, R; Lizeca, JL; Fukushima Y. 2002. Memoria del mapa de cobertura y uso de la tierra. Superintendencia Agraria. La Paz, BO. 23 p.
- Llactayo, W; Salcedo, K; Victoria, E. 2013. Memoria técnica de la cuantificación de cambios de la cobertura de bosque por deforestación en el ámbito de la Amazonía Peruana, periodo 2009-2010-2011. Ministerio del Medio Ambiente, Dirección General de Ordenamiento Territorial. Lima. PE. 50 p.
- López de Ullibarri, I y Pita, S. 2001. Medidas de concordancia: el índice Kappa. (en línea). Consultado 24 oct. 2016. Disponible en <http://www.fisterra.com/mbe/investiga/kappa/kappa2.pdf>

- Lozano, P; Sierra, P. 2005. Propuesta metodológica para delimitación de bosques de manglar a partir de datos satelitales utilizando SPRING versión 4.1, en la reserva de biósfera Ciénaga Grande de Santa Marta, Caribe Colombiano. CO. 4285 – 4293 p.
- MAE-MAGAP (Ministerio de Agricultura, Ganadería, Acuacultura y Pesca, EC) 2015. Protocolo metodológico para la elaboración del mapa de cobertura y uso de la tierra del Ecuador continental 2013 – 2014, escala 1:100.00. Quito, EC. 49 p.
- Martínez, J; Díaz, A. 2005. Percepción remota. Fundamentos de teledetección espacial. 62 p.
- Martínez, P; Palacios, JJ; Matínez, M; Zárate, R; García, J; Escobedo, R. 2015. Cobertura y uso de la tierra de las provincias de Alto Amazonas y Ramón Castilla. Instituto de investigaciones de la amazonia peruana. Iquitos, PE. 54 p.
- Melh, H; Peinado, O. s.f. Fundamentos del procesamiento digital de imágenes. Comisión Nacional de Actividades Espaciales (CONAE). Mendoza, AR. 24 p.
- MINAM (Ministerio del Ambiente, PE) 2014. Protocolo: Análisis de las dinámicas de cambio de cobertura de la tierra en la Comunidad Andina. Dirección General del Ordenamiento Territorial. Lima, PE. 72 p.
- MINAM (Ministerio del Ambiente, PE) 2015. Mapa nacional de cobertura vegetal: memoria descriptiva. Dirección General de Evaluación, Valoración y Financiamiento del Patrimonio Natural. Lima, PE. 105 p.
- Municipalidad Distrital de San Ramón, PE. 2009. Mejoramiento y ampliación de la gestión integral de los residuos sólidos municipales en la población urbana del distrito de San Ramón, provincia de Chanchamayo, región Junín. Perfil de pre-inversión SNIP. San Ramón, PE. 144 p.
- Molinero, L. 2001. Medidas de concordancia para variables cualitativas. Asociación de la Sociedad Española de Hipertensión. (en línea). España. Consultado 17 abr. 2016. Disponible en <http://www.seh-lelha.org/pdf/concor2.pdf>
- Navarro, C; Rivera, JP; Cepeda, J. 2007. Análisis de métodos de fusión de imágenes Landsat 8 sobre coberturas y uso del suelo en el municipio de Bahía de Banderas, Nayarit. Revista Educateconciencia. Tepic, MX. 14 p.

- Núñez, W; Vargas, E; Castro, E. s.f. Memoria descriptiva del mapa agrostológico – provincia de Paucartambo. Proyecto: Fortalecimiento del desarrollo de capacidades de ordenamiento territorial en la región de Cusco. Cusco, PE. 36 p.
- Paradis, E. 2003. R para principiantes. (en línea). Consultado 10 jun. 2016. Disponible en https://cran.r-project.org/doc/contrib/rdebuts_es.pdf
- Pérez, D. 2012. Identificación de ecosistemas en la provincia de Napo – Ecuador mediante análisis digital de imágenes satelitales. Tesis Mag. Sc. Quito, EC, Universidad San Francisco. 141 p.
- Puentes, C. 2010. Combinación de Bandas para Landsat. Cartografía CL. (en línea). Chile. Consultado 15 oct. 2016. Disponible en <http://www.cartografia.cl/beta/index.php/home/percepcion/830-2013-08-21-02-17-46?format=pdf>
- REDD GTZ (Programa de Reducción de Emisiones de Carbono causadas por la Deforestación y la Degradación de los Bosques, HO) 2010. Metodologías de análisis multitemporal de imágenes satelitales aplicable a nivel sub-nacional para los países miembros del CCAD: Belice, Costa Rica, El Salvador, Guatemala, Honduras, Nicaragua, Panamá, República Dominicana. Siguatepeque, HO. 61 p.
- Reynel, C. 2012. Flora y fauna del bosque montano nublado Puyu Sacha. Lima, PE. 382 p.
- Rodríguez, A. 2011. Metodología para detectar cambios en el uso de la tierra utilizando los principios de la clasificación orientada a objetos, estudio de caso piedemonte de Villavicencio, Meta. Tesis. Ing. Agr. Bogotá, CO, Universidad Nacional de Colombia. 94 p.
- Rodríguez, V; Chica, M. 2012. Clasificación de imágenes de satélite mediante software libre: Nuevas tendencias en algoritmos de inteligencia artificial. Madrid, ES. 7 p.
- Ruíz, D; Salcedo, J. 2005. Análisis comparativo de herramientas Software especializadas para el tratamiento digital de imágenes de Satélite. Caracterización de procesos básicos. Tesis. Ing. Civil. Bucaramanga, CO, Universidad Nacional de Santander. 183 p.

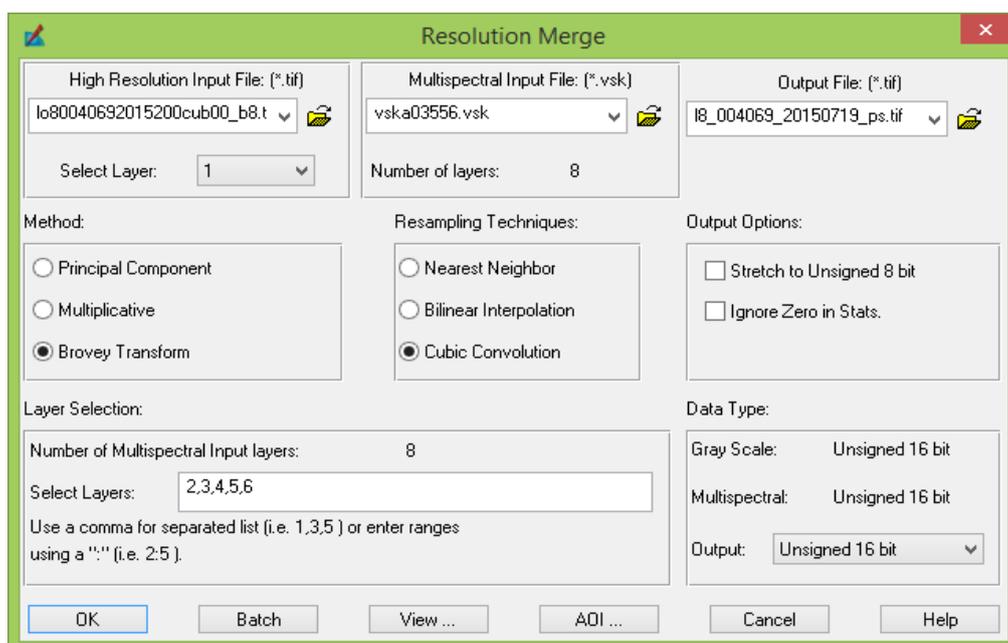
- Sánchez, A. 2009. Análisis multitemporal de la cobertura de la tierra en la cuenca del valle de Jesús de Otoro entre los años 2000 – 2006 aplicando técnicas de teledetección. Tesis Mag. Sc. Suyapa, HN, Universidad Nacional Autónoma de Honduras. 146 p.
- Sánchez, P. 2012. La teledetección enfocada a la obtención de mapas digitales. Tesis. Ing. Civil. Cuenca, EC, Universidad de Cuenca. 60 p.
- Suárez, A. 2014. Análisis y comparación de imágenes de diferentes sensores para verificar su eficiencia y utilidad en estudios multitemporales. Tesis Ing. Geógrafo. Sangolquí, EC, Universidad de las Fuerzas Armadas. 122 p.
- S-R-GIS (Sensores Remotos & Gis, CL) 2005. Geología y Geomática. (en línea). Santiago de Chile, CL. Consultado 28 oct. 2016. Disponible en <http://www.srgis.cl/software/>
- Zamora, S; López, T; Rodríguez J; Blanco, M. 2016. Metodología de validación de los mapas de usos de la tierra de Nicaragua. Ministerio del Ambiente y los Recursos Naturales de Marena. NI. 17 p.

ANEXO 2

PRE-PROCESAMIENTO DE LAS IMÁGENES LANDSAT 8

A. Unión de Bandas

Ventana de Resolution Merge de la herramienta Pan-Sharpen en el programa Erdas Imagine 9.1



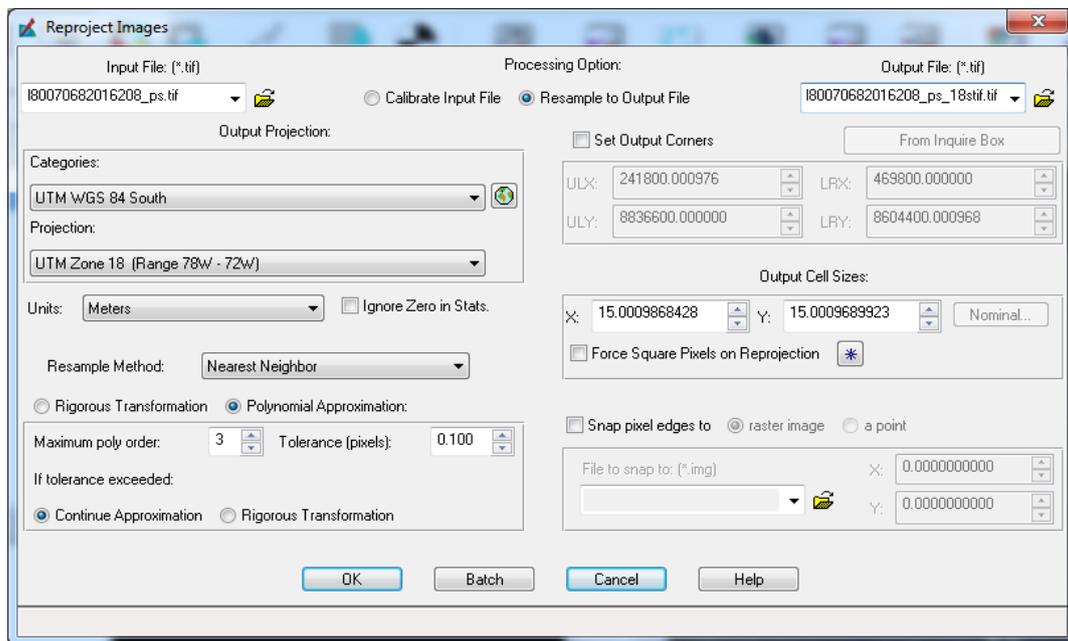
B. Re-proyección de la imagen

Especificaciones técnicas de la imagen Landsat 8

	Características Técnicas
Nivel de tratamiento	1T (Imágenes Ortorectificadas)
Sensor	Multiespectral (pushbroom imager)
Resolución Espacial	15 metros (Panchromatic) 30 metros (VIS, SWIR) 100 metros (TIRS)
Resolución radiométrica	12 bits
Resolución temporal	16 días
Formato de entrega	GeoTIFF
Sistema de coordenadas	UTM/ WGS-84/ Zona 17
Metadatos	Formato TXT
Tipo de Licencia	Gratis
Tamaño de escena	180km * 180km

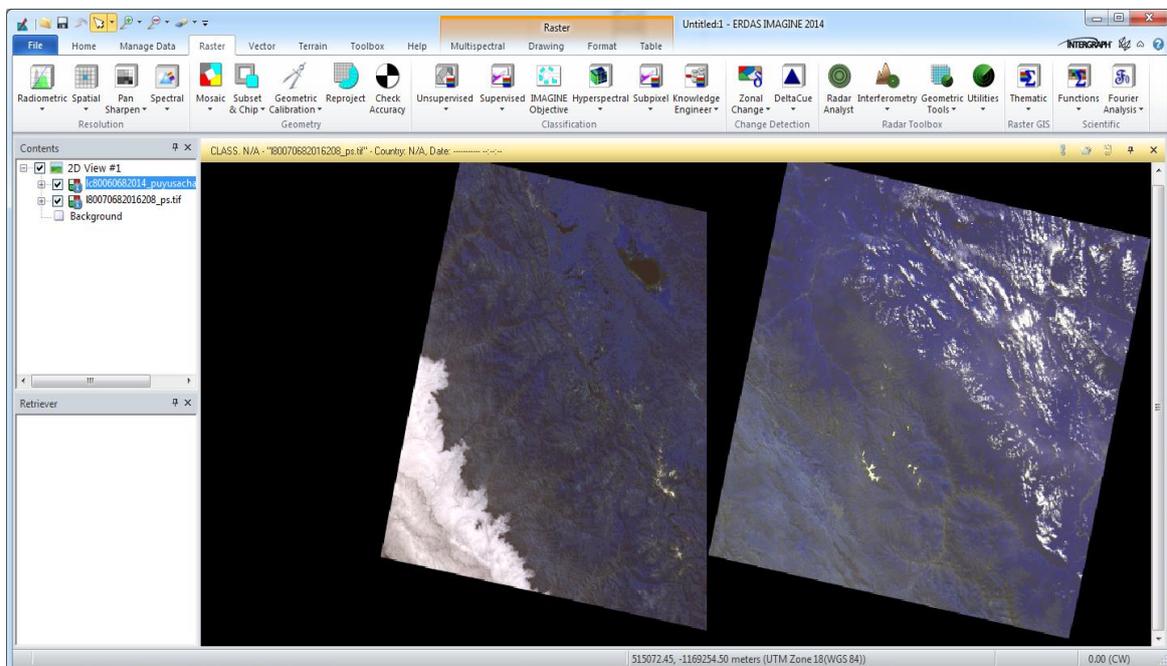
FUENTE: NASA, citado por MAE – MAGAP 2015

- Ventana de Reproject en el programa Erdas Imagine 9.1, donde se adicionó la imagen unida para reproyectarla al sistema de coordenada de UTM/WGS-84/Zona 18S.

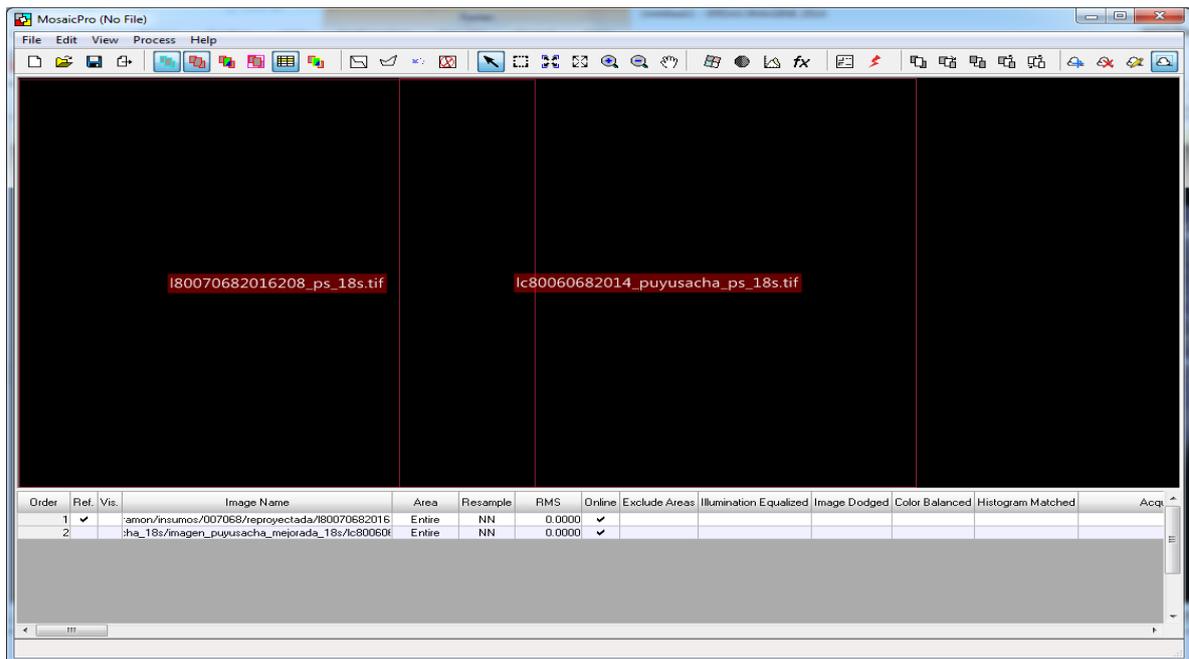


C. Generación del mosaico

- Ventana de las dos imágenes Landsat 8 adicionadas para el proceso en el programa Erdas Imagine 9.1

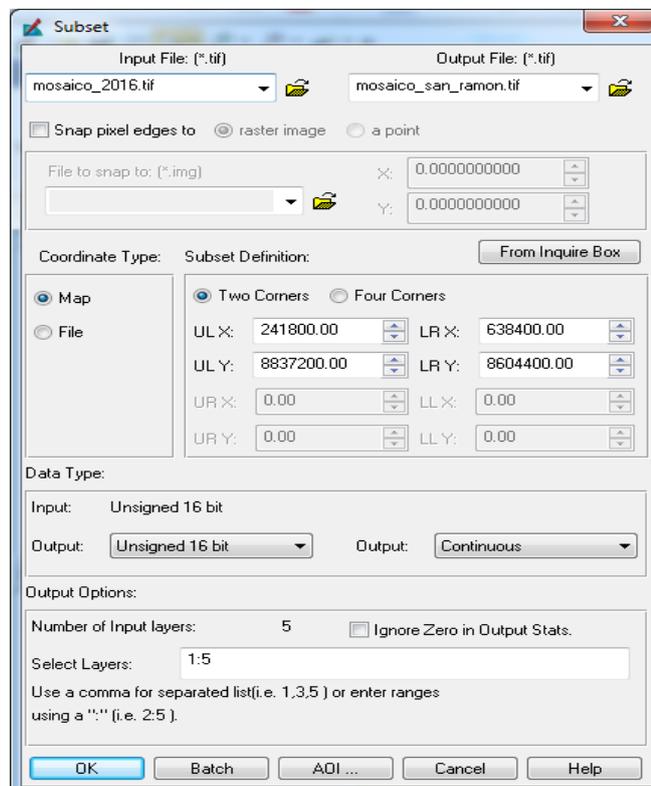


- Ventana de la herramienta MosaicPro en el programa Erdas Imagine 9.1



D. Corte del mosaico

- Ventana de la herramienta Create Subset Image en el programa Erdas Imagine 9.1, para elaborar el corte del mosaico.

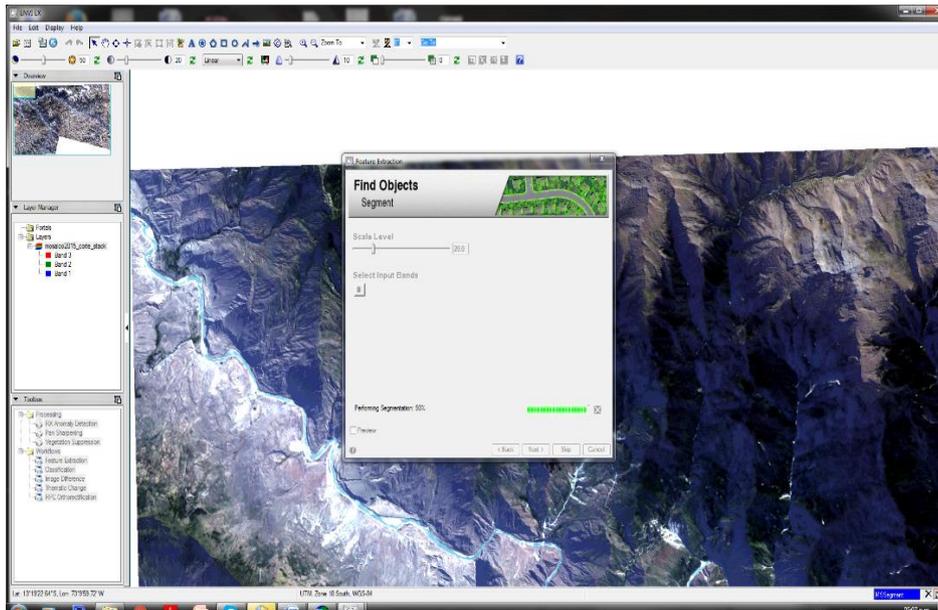


ANEXO 3

PROCESAMIENTO

A. Segmentación

- Ventana del proceso de segmentación en el programa ENVI EX



B. Generación de las áreas de entrenamiento

- Estructura de la tabla de atributos de las áreas de entrenamiento.

FID	Shape *	Id	Id_Leyenda	Certeza	Fuente	Comentario	Tipo_Uso
206	Polygon	0	321	1		matorral	1
207	Polygon	0	321	1		matorral	1
208	Polygon	0	321	1		matorral	1
209	Polygon	0	321	1		matorral	1
210	Polygon	0	321	1		matorral	1
211	Polygon	0	321	1		matorral	1
212	Polygon	0	321	1		matorral	1
213	Polygon	0	321	1		matorral	1
180	Polygon	0	512	1		Lagunas	1
181	Polygon	0	512	1		Lagunas	1
182	Polygon	0	512	1		Lagunas	1
183	Polygon	0	512	1		Lagunas	1
184	Polygon	0	512	1		Lagunas	1
185	Polygon	0	512	1		Lagunas	1
186	Polygon	0	512	1		Lagunas	1
0	Polygon	0	311	1		Bosque	1
1	Polygon	0	311	1		Bosque	1
2	Polygon	0	311	1		Bosque	1
3	Polygon	0	311	1		Bosque	1
4	Polygon	0	311	1		Bosque	1
5	Polygon	0	311	1		Bosque	1
6	Polygon	0	311	1		Bosque	1
7	Polygon	0	311	1		Bosque	1
8	Polygon	0	311	1		Bosque	1
9	Polygon	0	311	1		Bosque	1
10	Polygon	0	311	1		Bosque	1
11	Polygon	0	311	1		Bosque	1
12	Polygon	0	311	1		Bosque	1
13	Polygon	0	311	1		Bosque	1
14	Polygon	0	311	1		Bosque	1
15	Polygon	0	311	1		Bosque	1
16	Polygon	0	311	1		Bosque	1
17	Polygon	0	311	1		Bosque	1

C. Información generada para la clasificación

- Se configuró los siguientes campos en la opción *Environment Settings* (configuración del entorno) de las capas resultantes.

Campos a configurar	Descripción del campo	Opción a elegir	Descripción de la opción
Extent – General Settings	- Define el área de interés, mediante un rectángulo que identifica las coordenadas espaciales del área de estudio.	Same as Layer	El mismo extent de un layer seleccionado
Snap raster – General Setting	- Permite la configuración de un ajuste completo, asegurando que la alineación de las celdas del área de análisis coincida.	Directorio de un raster seleccionado	Indica el directorio de un raster seleccionado
Cartography Coordinate System – Cartography Setting	- Se especifica el Sistema de coordenadas.	Same as Layer	El mismo sistema de coordenadas de un layer seleccionado
Cell Size – Raster Analysis Setting	- Se especifica la resolución o tamaño de celda.	Same as Layer	El mismo tamaño de celda de un layer seleccionado.

FUENTE: CDC-UNALM 2012

D. Aplicación con el software R

- Script denominado “recogiendoInfoRaster_noTC.R”

```
#datos para R
#- convertir las regiones (todas) a raster
#- reproyectar DEM para que coincida con extent y cellsize de área de estudio (decidir qué dem se usa)
#- acomodar cada capa que se quiera trabajar para que coincida con extent y cellsize de área de estudio
#- En R:

library(raster)
library(foreign)

##### configuración Ejercicio_training_RandomForest

ruta <- "D:/SHEILA/Clasificacion/"
carpetaIndices <- "info_comp/"

# registrar archivo de regiones
archRegiones <- "info_comp/seg_2030.dbf"
rasterRegiones <- "info_comp/seg"

# registrar los archivos donde están cada una de las variables a evaluar
opc <- data.frame(variables = c("dem", "slope", "placurv", "procurv"),
                  archivos = c("dem", "slope", "placurv", "procurv"), stringsAsFactors = FALSE)
opc <- edit(opc)

funcs <- data.frame(funciones = c("mean", "sd", "min", "max"), titulos = c("MEAN", "STD", "MIN", "MAX"), :

#####

### abrir regiones
regiones <- raster(paste(ruta, rasterRegiones, sep=""))

# construir tabla de regiones de la segmentación (en formato raster)
tabreg <- as.data.frame(freq(regiones))

# y buscar las regiones con un solo pixel
solitarios <- which(tabreg[,2] == 1)

# tabla de las regiones de la segmentación (parte del formato shape)
tabla <- read.dbf(paste(ruta, archRegiones, sep=""))

# orden según region_id, guardando un correlativo
tabla$correlativo <- 1:nrow(tabla)
tabla <- tabla[order(tabla$REGION_ID),]

# verificar consistencia entre raster de regiones y tabla de regiones
if (any(tabla$REGION_ID != tabreg[,1])) {
  cat("Error en las tablas de las regiones. Revisar los parámetros de entrada.\n")
} else {
  # quitar variables inútiles
  lvarborrar <- c("AREA", "LENGTH", "COMPACT", "CONVEXITY", "SOLIDITY", "ROUNDNESS",
                 "FORMFACTOR", "ELONGATION", "RECT_FIT", "MAINDIR", "MAJAXISLEN", "MINAXISLEN",
                 "NUMHOLES", "HOLESOLRAT")
  tabla <- tabla[!(names(tabla) %in% lvarborrar)]
  rm(lvarborrar)

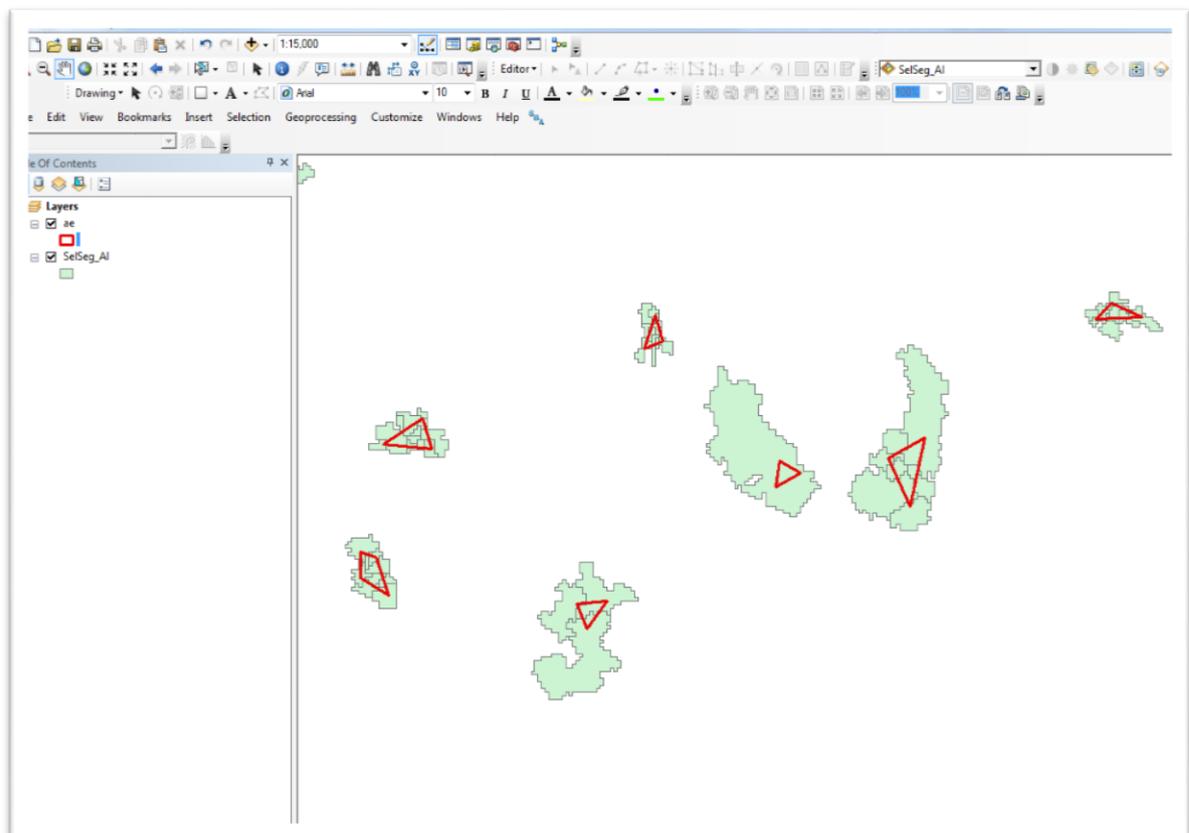
  # juntar info de número de píxeles
  tabla$numpixeles <- tabreg[,2]
```


F. Clasificación

- Insumos de la carpeta de Clasificación

Contents	
Name	Type
ae	Folder
info_comp	Folder
IT1	Folder
IT2	Folder
seg_2030.shp	Shapefile
seg_2030 - copia.dbf	dBASE Table
SelSeg_AI.shp	Shapefile
todo_seg - copia.dbf	dBASE Table

- Generación del shapefile denominado SelSeg_AI.shp



- Script denominado “Filtrar.tabla.todo.rf.R”.

```

# Código para juntar las tablas de varias fuentes.
# genera un reporte de qué se pudo hacer y qué no.

library(foreign)

# DEFINIR LA RUTA Ejercicio_training_RandomForest
ruta      <- "D:/SHEILA/Clasificacion/"
rutaSalida <- "D:/SHEILA/Clasificacion/IT2/"
# leer la tabla con todos los segmentos.
load(paste(ruta, "todo_seg.Rdata", sep=""))

# poner el nombre del dbf que tiene la información de los segmentos ya limpios POST spatial join
arch.seg <- "SelSeg_A1.dbf"

#####

# leer el archivo de la segmentación cruzado con áreas de entrenamiento ya limpio
tabseg <- read.dbf(paste(ruta, arch.seg, sep=""))

# filtrar los datos
tabla.sub <- tabla[match(tabseg$REGION_ID, tabla$REGION_ID),]
# verificar coincidencia
if (!(all(tabseg$REGION_ID == tabla.sub$REGION_ID)) ) cat("Error en las tablas")

# poner el orden de tabseg a tabla.sub y borrar la variable correlativo, y añadir ID_Leyenda
tabla.sub$Id_Leyenda <- tabseg$Id_Leyenda
# verificar si hay algún valor nulo.
if (any(is.na(tabla.sub))) apply(tabla.sub, 2, function(x) sum(is.na(x)))

#####

# preparar RandomForest
library(randomForest)
library(irr)
variables <- names(tabla.sub)
variables <- variables[!(variables %in% c("REGION_ID", "Id_Leyenda", "correlativo", "numpixeles", "predicte
variables.o <- variables # variables.o guarda la lista original de variables
cat("Revisar lista de variables:\n")
variables

#####

## este código es para filtrar la lista de variables y quedarse sólo con las más importantes
## cambiar el umbral según corresponda
# umbral <- 1; variables <- rownames(importance(tblcal.rfo))[importance(tblcal.rfo) > umbral]

rf.form <- as.formula(paste("Id_Leyenda ~ ", paste(variables, collapse = " + "), sep=""))
tblcal.rf <- randomForest (rf.form, data = tabla.sub, proximity=TRUE)

## guardar el rf original, con todas las variables
if (length(variables) == length(variables.o)) tblcal.rfo <- tblcal.rf

# generando la matriz de confusión y la tabla.sub de importancias

```

- Ventana del software Rproject 3.1.1 en el cual se generó los productos de la clasificación.

```

R Console
slope_MEAN      3.4843905
slope_STD       11.1791972
slope_MIN       0.2631188
slope_MAX       8.7763140
placurv_MEAN    0.9814983
placurv_STD     1.4677971
placurv_MIN     1.2198036
placurv_MAX     1.3555765
procurv_MEAN    1.6496113
procurv_STD     6.8768405
procurv_MIN     7.0880110
procurv_MAX     8.0264490
> tblcalmatriz <- tblcal.rf$confusion
> #tblcalmatriz <- edit(tblcal.rf$confusion)
> # generando estadísticos de prueba
> cat("\n-- Porcentaje correctamente clasificado:",
+     sum(diag(tblcalmatriz[,-ncol(tblcalmatriz)]))/sum(tblcalmatriz[,-ncol(tblc$
-- Porcentaje correctamente clasificado: 0.977887
> #####Kappa2(tabla.sub[,c("Id_Leyenda", "predicted")])
>
>
> ## antes de guardar el modelo, regresar a la línea 48
> ## y hacer la segunda corrida filtrando las variables poco importantes.
> #####

```

```

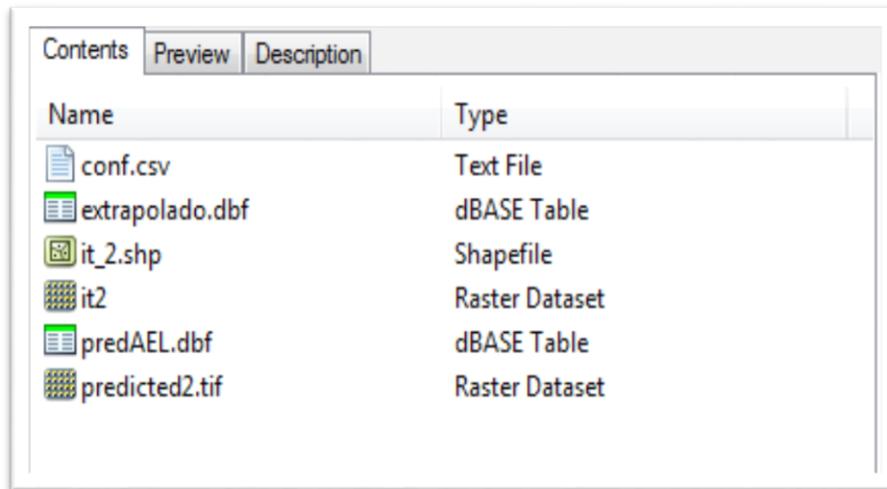
R Console
Error in save(tblcal.rf, umbral, file = paste(rutaSalida, "modrf.RData", :
  object 'umbral' not found
> write.csv(tblcal.rf$confusion, file=paste(rutaSalida,"conf.csv", sep=""))
>
> ### se corre cuando ya se tenga predAEL.dbf
> kappa2(tabla.sub[,c("Id_Leyenda", "predicted")])
Cohen's Kappa for 2 Raters (Weights: unweighted)

Subjects = 1221
Raters = 2
Kappa = 0.972

      z = 72
p-value = 0
>
> #####$
> ## Extrapolando
>
> # proyectando a otras áreas
> tablapred <- tabla["REGION_ID"]
> tablapred$predicted <- predict(tblcal.rf, newdata=tabla)
> # exportando la tabla
> write.dbf(tablapred, file=paste(rutaSalida, "extrapolado.dbf", sep=""))
>
>

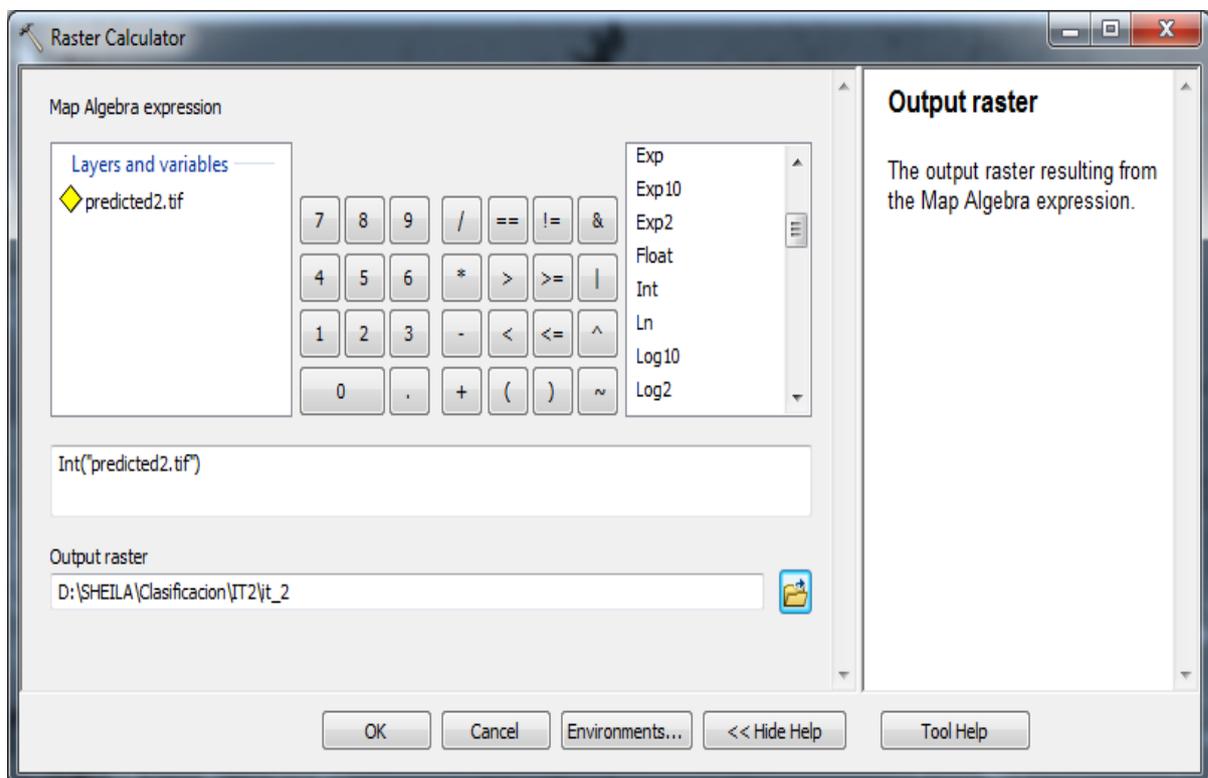
```

- Productos del proceso de clasificación



Name	Type
conf.csv	Text File
extrapolado.dbf	dBASE Table
it_2.shp	Shapefile
it2	Raster Dataset
predAEL.dbf	dBASE Table
predicted2.tif	Raster Dataset

- Ventana de la herramienta Raster Calculator, donde se generó el raster de la iteración.



G. Script

Script o archivo de procesamiento por lotes, desarrollado por Carlos Alberto Arnillas en el año 2012.

SCRIPT	FUNCIONES	LIBRERIAS	INSUMOS	PRODUCTOS
RecogiendolInfoRaster.R	<ul style="list-style-type: none"> - Convierte todas las regiones raster. - Re-proyecta el DEM para que coincida con el extent y cellsize del área de estudio. - Acomoda las capas que se quiera trabajar para que coincida con el extent y cellsize del área de estudio. 	library (raster) library (foreign)	info_com seg20_30.dbf	todo_seg.dbf todo_seg.Rdata
Filtrar.tabla.todo.R	<ul style="list-style-type: none"> - Código para juntar las tablas de varias fuentes. - Genera un reporte de qué se pudo hacer y qué no. 	library (foreign) library(random Forest) library (irr) library (raster)	todo_seg.Rdata SelSeg_Al.dbf	Tabla de Importancia de las variables: Cof.csv - matriz de Confusión. Porcentaje Correctamente Clasificado. Índice Kappa, Modelo de Clasificación predicted2.tif
Metricas.validacion.R	<ul style="list-style-type: none"> - Estima las métricas de la validación: Porcentaje Correctamente Clasificado (PCC), Índice Kappa (IK) y Matriz de confusión. 	library (irr) library (foreign)	puntos_val.dbf	PCC, Índice Kappa y matriz de confusión matrizconf2.cvs.

FUENTE: CDC-UNALM 2012

ANEXO 4

SALIDA DE CAMPO

- Materiales utilizados en la salida de campo.

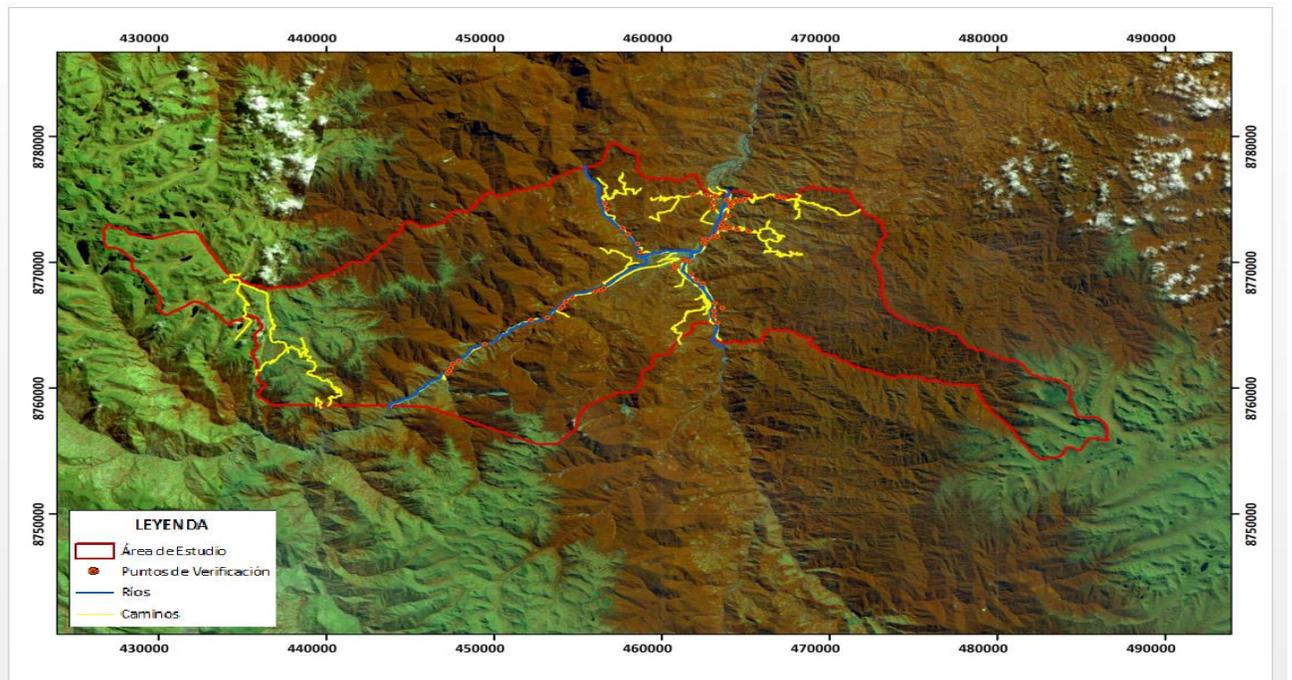


GPS Map 76S Garmin

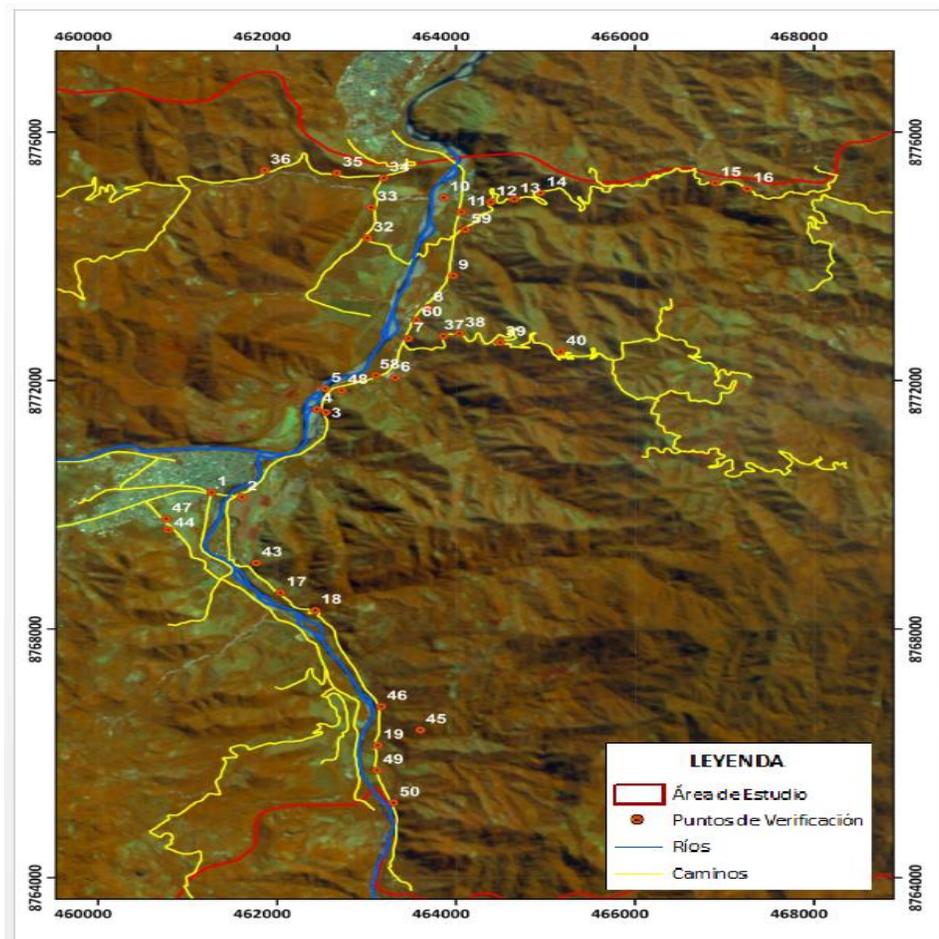
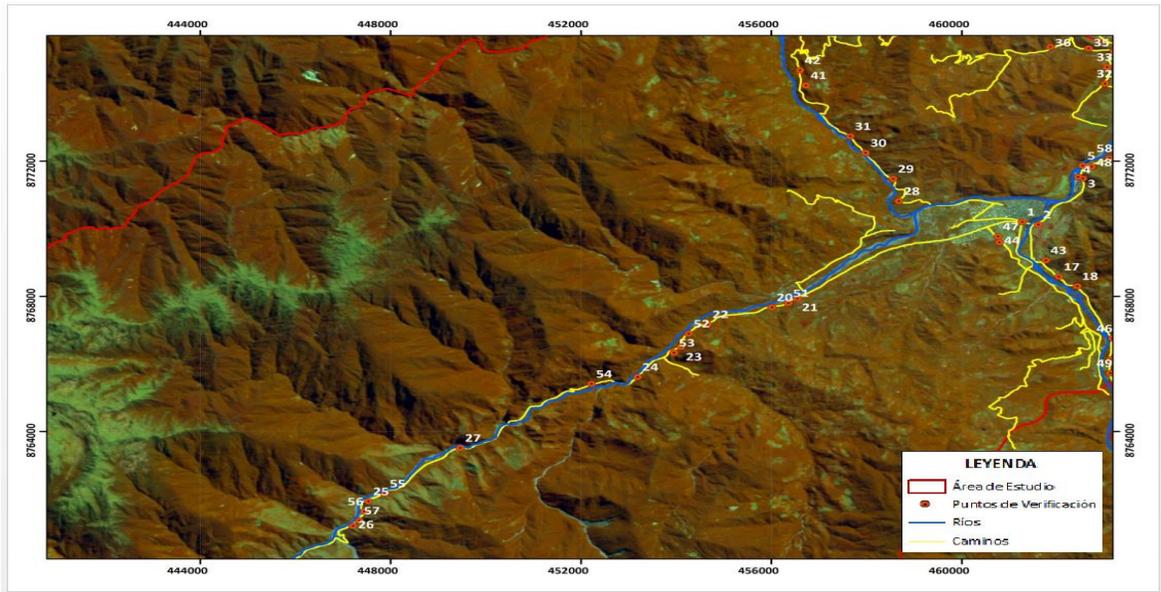


Cámara Fotográfica Canon 60D

- Mapas de la ruta y puntos de verificación que fueron recogidos en la salida de campo.



- Mapas de la ruta y puntos de verificación que fueron recogidos en la salida de campo.



- Fotos tomadas en la salida de campo.



- Ficha de campo definida para el estudio.

Ficha de Campo - Mapa de Cobertura y uso de la tierra			
			Nº: 08
Responsable	Sheila Gamarra	Fecha	28/08/2016
Asistente	Lizette Gamarra	Código WP	020
Código del GPS	S8	Error (m)	10.8
Coordenadas	0463997	Datum	WGS 84
	8773729	Zona UTM	18S
Altitud (msnm)	1005		
Departamento	Junín	Provincia	Chanchamayo
Lugar de referencia	-	Distrito	San Ramón
Cobertura	Arbustos y matorrales		
Observaciones (Fisonomía, Vegetación predominante, Actividad humana)			
La vegetación predominante que se observa son arbustos achaparrados y matorrales.			
Rumbo	Oeste	Distancia	30 m
Fotos			
Código Fotográfico	Orientación	Observaciones	
8179	Norte	Arbustos y matorrales	
8180	Sur	Cerca de la carretera	
8181	Este	Arbustos y matorrales	
8182	Oeste	Arbustos y matorrales	

- Datos colectados en la salida de campo

N°	Código GPS	Responsable	Asistente	Coord X	Coord Y	Altitud	Fecha	Código WP	Error (m)	Datum	Zona UTM	Departamento	Provincia	Distrito
1	S1	Sheila	Lizette	446174	8760420	1688	28/08/2016		10.8	WGS 84	18S	Junín	Chanchamayo	San Ramón
2	S2	Sheila	Lizette	447131	8761454	1656	28/08/2016	26	10.5	WGS 84	18S	Junín	Chanchamayo	San Ramón
3	S3	Sheila	Lizette	449431	8763740	1471	28/08/2016	27	10.5	WGS 84	18S	Junín	Chanchamayo	San Ramón
4	S4	Sheila	Lizette	464890	8775068	1231	28/08/2016		10	WGS 84	18S	Junín	Chanchamayo	San Ramón
5	S5	Sheila	Lizette	464656	8774853	1009	28/08/2016		10.1	WGS 84	18S	Junín	Chanchamayo	San Ramón
6	S6	Sheila	Lizette	464064	8774784	1007	28/08/2016		9.3	WGS 84	18S	Junín	Chanchamayo	San Ramón
7	S7	Sheila	Lizette	464107	8774334	931	28/08/2016		9.1	WGS 84	18S	Junín	Chanchamayo	San Ramón
8	S8	Sheila	Lizette	463997	8773729	1005	28/08/2016	20	10.8	WGS 84	18S	Junín	Chanchamayo	San Ramón
9	S9	Sheila	Lizette	463729	8773203	1235	28/08/2016		8.6	WGS 84	18S	Junín	Chanchamayo	San Ramón
10	S10	Sheila	Lizette	463508	8772694	1261	28/08/2016		9.2	WGS 84	18S	Junín	Chanchamayo	San Ramón
11	S11	Sheila	Lizette	462671	8771847	906	28/08/2016		9	WGS 84	18S	Junín	Chanchamayo	San Ramón
12	S12	Sheila	Lizette	462010	8768613	893	28/08/2016		8.6	WGS 84	18S	Junín	Chanchamayo	San Ramón
13	S13	Sheila	Lizette	460757	8769521	879	28/08/2016		8.5	WGS 84	18S	Junín	Chanchamayo	San Ramón
14	S14	Sheila	Lizette	461616	8775338	860	29/08/2016	35	9.7	WGS 84	18S	Junín	Chanchamayo	San Ramón
15	S15	Sheila	Lizette	462023	8775636	917	29/08/2016	36	8.4	WGS 84	18S	Junín	Chanchamayo	San Ramón
16	S16	Sheila	Lizette	463156	8772187	847	29/08/2016	34	9	WGS 84	18S	Junín	Chanchamayo	San Ramón
17	S17	Sheila	Lizette	464487	8775362	806	29/08/2016	12	9	WGS 84	18S	Junín	Chanchamayo	San Ramón
18	S18	Sheila	Lizette	464545	8775276	825	29/08/2016		8	WGS 84	18S	Junín	Chanchamayo	San Ramón
19	S19	Sheila	Lizette	464551	8774102	790	29/08/2016	9	10.1	WGS 84	18S	Junín	Chanchamayo	San Ramón
20	S20	Sheila	Lizette	463840	8772642	825	29/08/2016	37	10.2	WGS 84	18S	Junín	Chanchamayo	San Ramón
21	S21	Sheila	Lizette	464060	8772792	880	29/08/2016	38	8.5	WGS 84	18S	Junín	Chanchamayo	San Ramón
22	S22	Sheila	Lizette	464369	8772816	974	29/08/2016		10.9	WGS 84	18S	Junín	Chanchamayo	San Ramón
23	S23	Sheila	Lizette	462765	8772146	822	29/08/2016		9.1	WGS 84	18S	Junín	Chanchamayo	San Ramón
24	S24	Sheila	Lizette	461460	8770572	843	29/08/2016	1	8.9	WGS 84	18S	Junín	Chanchamayo	San Ramón
25	S25	Sheila	Lizette	461710	8769090	849	29/08/2016	43	7.5	WGS 84	18S	Junín	Chanchamayo	San Ramón
26	S26	Sheila	Miguel	462513	8768340	862	29/08/2016	18	9	WGS 84	18S	Junín	Chanchamayo	San Ramón
27	S27	Sheila	Miguel	463073	8766783	880	29/08/2016	46	9.4	WGS 84	18S	Junín	Chanchamayo	San Ramón

- Datos colectados en la salida de campo

N°	Código GPS	Responsable	Asistente	Coord X	Coord Y	Altitud	Fecha	Código WP	Error (m)	Datum	Zona UTM	Departamento	Provincia	Distrito
28	S28	Sheila	Miguel	463120	8766004	887	29/08/2016	49	9.8	WGS 84	18S	Junín	Chanchamayo	San Ramón
29	S29	Sheila	Miguel	463350	8764652	906	29/08/2016		9.5	WGS 84	18S	Junín	Chanchamayo	San Ramón
30	S30	Sheila	Miguel	459637	8771216	882	29/08/2016	28	10.8	WGS 84	18S	Junín	Chanchamayo	San Ramón
31	S31	Sheila	Miguel	458924	8771236	950	29/08/2016	28	9.5	WGS 84	18S	Junín	Chanchamayo	San Ramón
32	S32	Sheila	Miguel	458254	8772478	990	29/08/2016		10.1	WGS 84	18S	Junín	Chanchamayo	San Ramón
33	S33	Sheila	Miguel	457978	8772972	907	29/08/2016	31	8.5	WGS 84	18S	Junín	Chanchamayo	San Ramón
34	S34	Sheila	Miguel	456896	8774326	985	29/08/2016	41	10.1	WGS 84	18S	Junín	Chanchamayo	San Ramón
35	S35	Sheila	Miguel	456826	8774998	1006	29/08/2016	42	8.9	WGS 84	18S	Junín	Chanchamayo	San Ramón
36	S36	Sheila	Miguel	461357	8770706	830	30/08/2016		8.5	WGS 84	18S	Junín	Chanchamayo	San Ramón
37	S37	Sheila	Miguel	457568	8768926	959	30/08/2016		9.8	WGS 84	18S	Junín	Chanchamayo	San Ramón
38	S38	Sheila	Miguel	456088	8768024	1009	30/08/2016	51	9.5	WGS 84	18S	Junín	Chanchamayo	San Ramón
39	S39	Sheila	Miguel	454038	8766692	1091	30/08/2016		9	WGS 84	18S	Junín	Chanchamayo	San Ramón
40	S40	Sheila	Miguel	450946	8764716	1297	30/08/2016		8	WGS 84	18S	Junín	Chanchamayo	San Ramón
41	S41	Sheila	Miguel	447585	8762098	1582	30/08/2016		8	WGS 84	18S	Junín	Chanchamayo	San Ramón
42	S42	Sheila	Miguel	454730	8767166	1009	30/08/2016		8.1	WGS 84	18S	Junín	Chanchamayo	San Ramón
43	S43	Sheila	Miguel	453931	8766230	1002	30/08/2016		8.3	WGS 84	18S	Junín	Chanchamayo	San Ramón
44	S44	Sheila	Miguel	453328	8765663	908	30/08/2016		9	WGS 84	18S	Junín	Chanchamayo	San Ramón
45	S45	Sheila	Miguel	447354	8761267	1163	30/08/2016		9.9	WGS 84	18S	Junín	Chanchamayo	San Ramón
46	S46	Sheila	Miguel	446290	8760415	1236	30/08/2016		9	WGS 84	18S	Junín	Chanchamayo	San Ramón
47	S47	Sheila	Miguel	444689	8758994	1007	30/08/2016		9.2	WGS 84	18S	Junín	Chanchamayo	San Ramón
48	S48	Sheila	Miguel	458120	8769267	904	30/08/2016		8	WGS 84	18S	Junín	Chanchamayo	San Ramón
49	S49	Sheila	Miguel	459046	8769473	923	30/08/2016		8.3	WGS 84	18S	Junín	Chanchamayo	San Ramón
50	S50	Sheila	Miguel	461529	8769869	982	30/08/2016		9.3	WGS 84	18S	Junín	Chanchamayo	San Ramón

<i>N°</i>	<i>Cobertura</i>	<i>Rumbo</i>	<i>Distancia (m)</i>	<i>Observaciones</i>	<i>N</i>	<i>S</i>	<i>E</i>	<i>O</i>
1	Carretera	O	30	Carretera	7763	7764	7765	7766
2	Matorrales	O	20	Pocos arbustos	7779	7780	7781	7782
3	Bosque	N	15	Bosque a la ribera del río	7786	7787	7788	7789
4	Bosque	N	20	Bosque denso	8160	8161	8162	8163
5	Bosque	N	15	Bosque denso	8165	8166	8167	8168
6	Cultivos	N	20	Papaya, cacao	8170	8171	8172	8173
7	Cultivos	S	25	Papaya, cacao	8174	8175	8176	8177
8	Matorrales	S	30	Arbustos achaparrados	8179	8180	8181	8182
9	Matorrales	S	15	Arbustos achaparrados	8183	8184	8185	8186
10	Bosque	S	30	Bosque denso	8187	8188	8189	8190
11	Cultivos	O	30	Cacao, frutales	8191	8192	8193	8194
12	Cultivos	O	25	Cacao, frutales	8195	8196	8197	8198
13	Cultivos	O	10	Cacao, frutales	8199	8200	8201	8202
14	Bosque	N	10	Bosque	7800	7801	7802	7803
15	Bosque	N	10	Bosque denso	7810	7811	7812	7813
16	Cultivos	O	10	Cultivos de plátano, mandarina, cacao	7841	7842	7843	7844
17	Bosque	E	10	Bosque	7851	7852	7853	7854
18	Bosque	S	8	Bosque	7862	7863	7864	7865
19	Matorrales	N	15	Arbustos achaparrados	7870	7871	7872	7873
20	Bosque	O	20	Bosque	7875	7876	7877	7878
21	Matorrales	N	15	Arbustos dispersos y achaparrados	7881	7882	7883	7884
22	Matorrales	N	10	Arbustos dispersos y achaparrados	7887	7888	7889	7890
23	Bosque	N	10	Bosque a la ribera del río	7894	7895	7896	7897

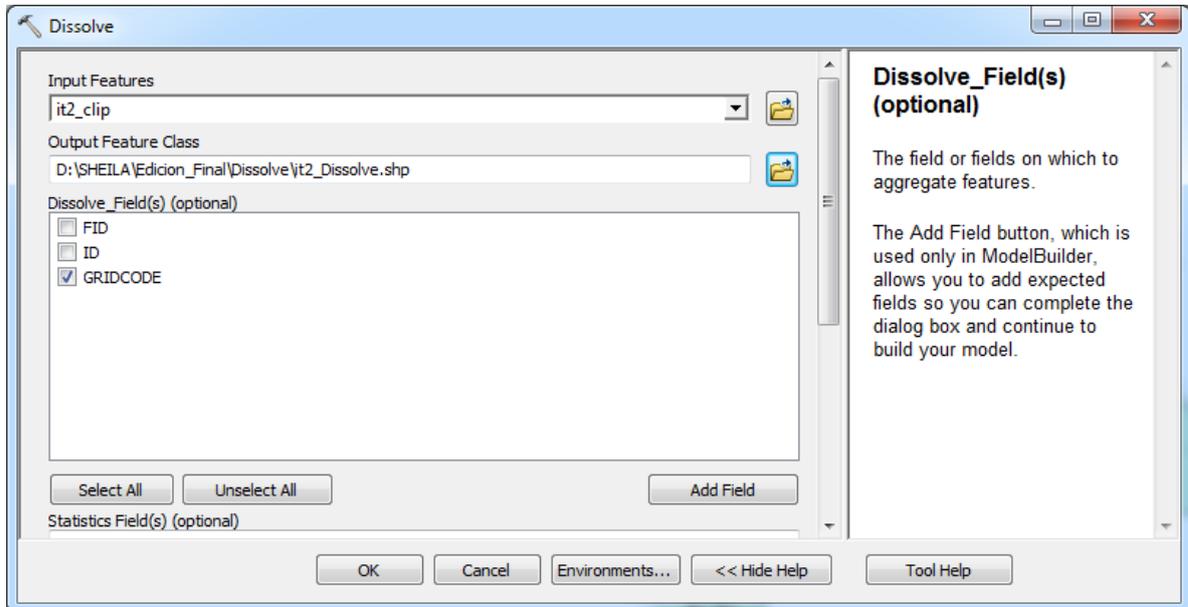
N°	Cobertura	Rumbo	Distancia (m)	Observaciones	N	S	E	O
24	Centro poblado			San Ramón	7898	7899	7901	7902
25	Cultivo	E	20	Cultivos de plátanos, cacao y palmeras	7906	7907	7908	7909
26	Bosque	E	25	Bosque hacia la carretera	7912	7913	7914	7915
27	Río	O	15	Río al costado de la ciudad	7921	7922	7923	7924
28	Cultivo	N	25	Plátanos, mandarinas, cacao	7926	7927	7928	7929
29	Bosque	O	20	Bosque a la ribera del río	7936	7937	7938	7939
30	Cultivo	N	20	Cacao, plátanos	7952	7953	7954	7955
31	Cultivo	S	15	Cacao, mandarina, plátano	7957	7958	7959	7960
32	Bosque	N	20	Bosque cerca de la Laguna la Auverna	7966	7967	7968	7969
33	Bosque	N	20	Bosque a la ribera del río, al Sur zona quemada	7991	7992	7993	7994
34	Bosque	E	15	Alrededor cubierta por bosque	8015	8016	8017	8018
35	Cultivos	E	10	Cultivos de papaya y plátanos	8027	8028	8029	8030
36	Centro Poblado	N		Distrito de San Ramón	8094	8095	8096	8097
37	Área pedregosas	O	50	Piedras redondeadas	8100	8101	8102	8103
38	Cultivo	E	10	Plátanos, mandarinas, cacao	8104	8105	8106	8107
39	Bosque	O	50	Bosque denso	8108	8109	8110	8111
40	Bosque	N	20	Bosque denso	8112	8113	8114	8115
41	Matorrales	E	30	Pocos arbustos, pasto seco	8156	8157	8158	8159
42	Bosque	N	15	Bosque denso	8203	8204	8205	8206
43	Bosque	E	20	Bosque denso	8207	8208	8209	8210
44	Bosque	N	30	Bosque denso	8211	8212	8213	8214
45	Matorrales	O	20	Matorrales y arbustos dispersos	8215	8216	8217	8218
46	Matorrales	O	25	Matorrales y arbustos dispersos	8219	8220	8221	8222
47	Matorrales	O	30	Matorrales y arbustos dispersos	8223	8224	8225	8226
48	Cultivos	O	20	Plátanos, mandarinas, cacao	8227	8228	8229	8230

<i>N°</i>	<i>Cobertura</i>	<i>Rumbo</i>	<i>Distancia (m)</i>	<i>Observaciones</i>	<i>N</i>	<i>S</i>	<i>E</i>	<i>O</i>
49	Cultivos	E	20	Plátanos, mandarinas, cacao	8231	8232	8233	8234
50	Cultivos	E	20	Plátanos, mandarinas, cacao	8235	8236	8237	8238

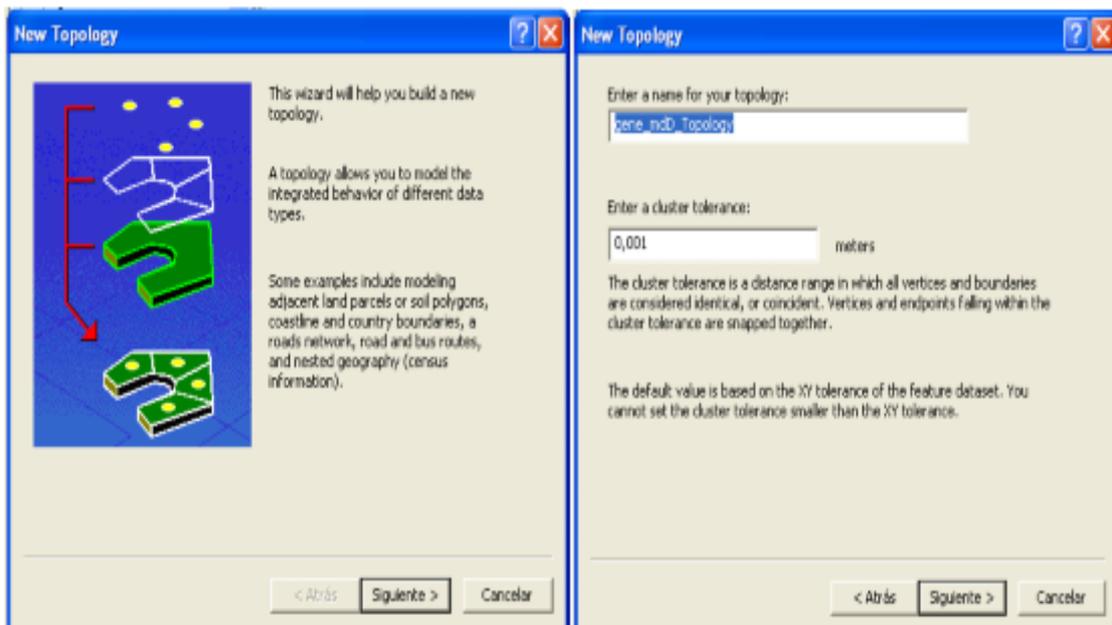
ANEXO 5

EDICIÓN

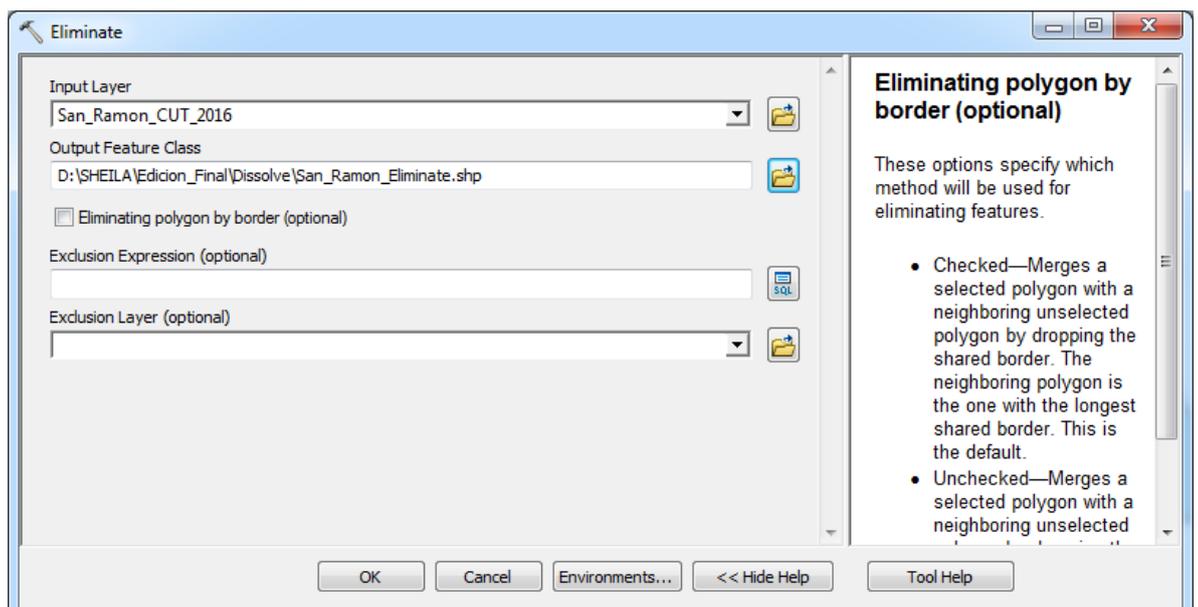
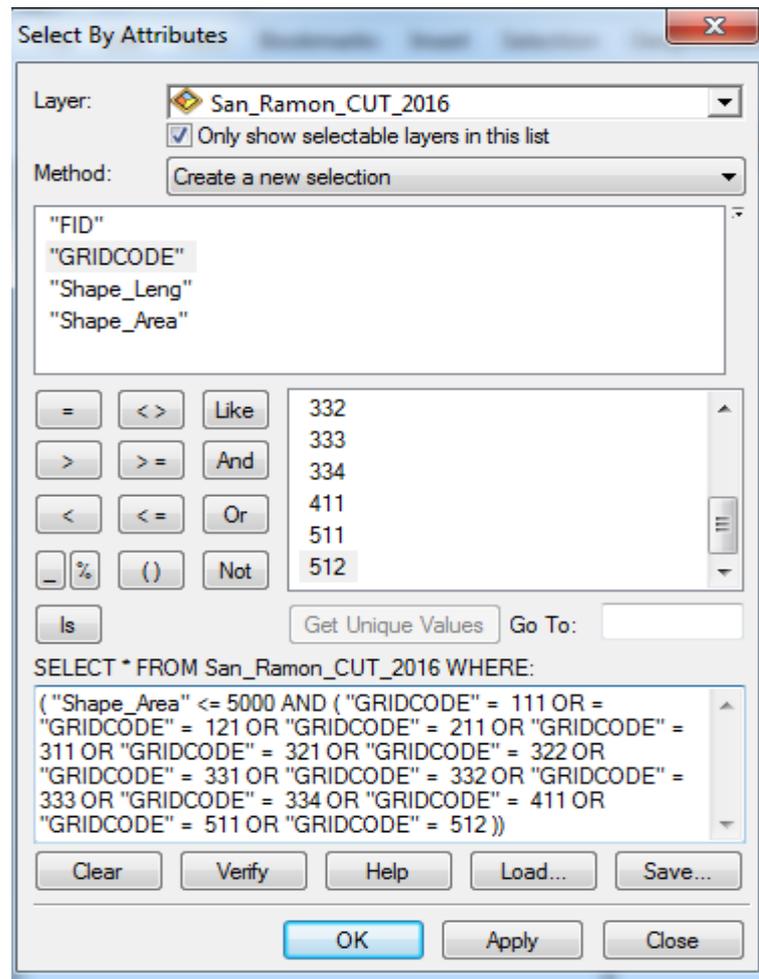
- Ventana de la herramienta Dissolve



- Ventana de la herramienta Topology



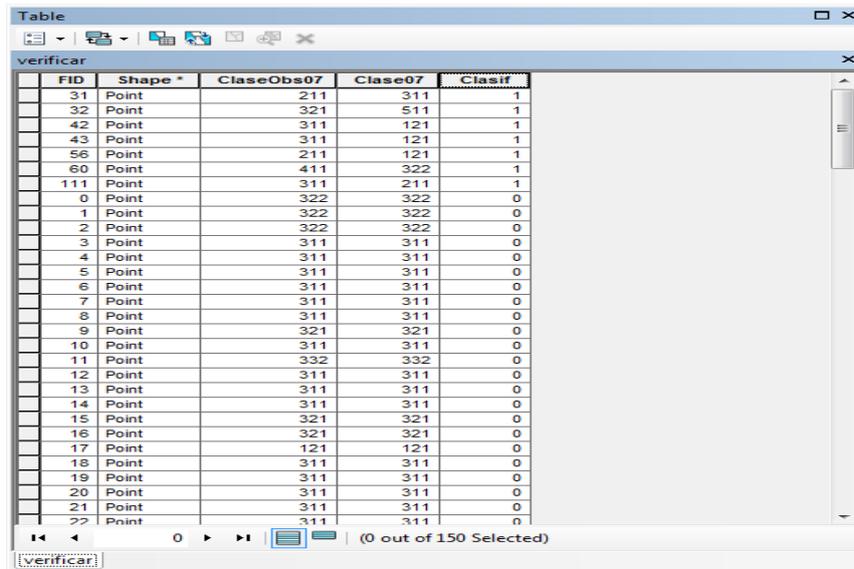
- Ventana de la herramienta Eliminate



ANEXO 6

VERIFICACIÓN DE LOS RESULTADOS

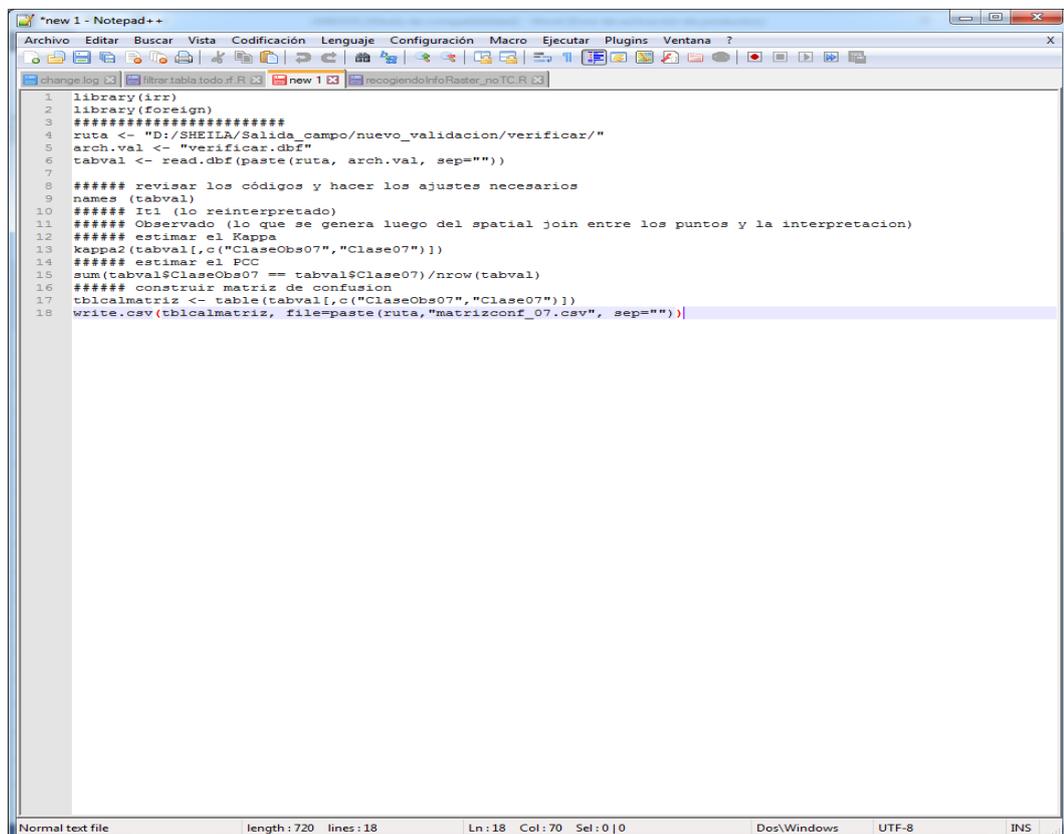
- Tabla de atributos del shapefile de verificación.



The screenshot shows a window titled 'Table' with a sub-window 'verificar'. It displays a table with the following columns: FID, Shape *, ClaseObs07, Clase07, and Clasif. The data rows are as follows:

FID	Shape *	ClaseObs07	Clase07	Clasif
31	Point	211	311	1
32	Point	321	511	1
42	Point	311	121	1
43	Point	311	121	1
56	Point	211	121	1
60	Point	411	322	1
111	Point	311	211	1
0	Point	322	322	0
1	Point	322	322	0
2	Point	322	322	0
3	Point	311	311	0
4	Point	311	311	0
5	Point	311	311	0
6	Point	311	311	0
7	Point	311	311	0
8	Point	311	311	0
9	Point	321	321	0
10	Point	311	311	0
11	Point	332	332	0
12	Point	311	311	0
13	Point	311	311	0
14	Point	311	311	0
15	Point	321	321	0
16	Point	321	321	0
17	Point	121	121	0
18	Point	311	311	0
19	Point	311	311	0
20	Point	311	311	0
21	Point	311	311	0
22	Point	311	311	0

- Script denominado “metricas.validacion.R”



```
1 library(irr)
2 library(foreign)
3 #####
4 ruta <- "D:/SHEILA/Salida_campo/nuevo_validacion/verificar/"
5 arch.val <- "verificar.dbf"
6 tabval <- read.dbf(paste(ruta, arch.val, sep=""))
7
8 ##### revisar los códigos y hacer los ajustes necesarios
9 names(tabval)
10 ##### It1 (lo reinterpretado)
11 ##### Observado (lo que se genera luego del spatial join entre los puntos y la interpretacion)
12 ##### estimar el Kappa
13 kappa2(tabval[,c("ClaseObs07","Clase07")])
14 ##### estimar el PCC
15 sum(tabval$ClaseObs07 == tabval$Clase07)/nrow(tabval)
16 ##### construir matriz de confusion
17 tblcalmatriz <- table(tabval[,c("ClaseObs07","Clase07")])
18 write.csv(tblcalmatriz, file=paste(ruta,"matrizconf_07.csv", sep=""))
```

- Ventana del software Rproject 3.1.1 en el cual se generó las métricas de la verificación.

```

R Console
> ruta <- "D:/SHEILA/Salida_campo/nuevo_validacion/verificar/"
> arch.val <- "verificar.dbf"
> tabval <- read.dbf(paste(ruta, arch.val, sep=""))
>
> ##### revisar los códigos y hacer los ajustes necesarios
> names(tabval)
[1] "ClaseObs07" "Clase07" "Clasif"
> ##### It1 (lo reinterpretado)
> ##### Observado (lo que se genera luego del spatial join entre los puntos y $
> ##### estimar el Kappa
> kappa2(tabval[,c("ClaseObs07", "Clase07")])
Cohen's Kappa for 2 Raters (Weights: unweighted)

Subjects = 150
Raters = 2
Kappa = 0.922

z = 20.4
p-value = 0
> ##### estimar el PCC
> sum(tabval$ClaseObs07 == tabval$Clase07)/nrow(tabval)
[1] 0.9533333
> ##### construir matriz de confusion
> tblcalmatriz <- table(tabval[,c("ClaseObs07", "Clase07")])
> write.csv(tblcalmatriz, file=paste(ruta, "matrizconf_07.csv", sep=""))

```

A. Elaboración de la Metadata

Contents Preview Description

Print Edit Import

Mapa de cobertura y uso de la tierra del distrito de San Ramón para el año 2016

Shapefile

Thumbnail Not Available

Tags

Mapa de cobertura y uso de la tierra del distrito de San Ramón para el año 2016

Summary
 Análisis de la cobertura y uso de la tierra utilizando imágenes de resolución espacial media para el distrito de San Ramón - Chanchamayo - Junín - Perú.

Description
 El conocimiento de la cobertura y uso de la tierra es un paso fundamental para la gestión sustentable de los recursos naturales, la conservación de la biodiversidad, mitigación de los impactos en las dinámicas territoriales y en el cambio climático. (FAO 2015)

Credits
 There are no credits for this item.

Use limitations
 There are no access and use limitations for this item.

Extent
 West -75.671396 East -75.122569
 North -11.040392 South -11.268210

Scale Range
 Maximum (zoomed in) 1:5,000
 Minimum (zoomed out) 1:150,000,000

You are currently using the Item Description metadata style. Change your metadata style in the Options dialog box to see additional metadata content.

ANEXO 7

**CARACTERIZACIÓN DE LAS COBERTURAS PRESENTES EN EL DISTRITO
DE SAN RAMÓN**

Centros poblados

1.1.1

Descripción:

Áreas con centros poblados que presentan un mínimo de 100 viviendas agrupadas contiguamente. También se incluyen aquellos centros poblados que son capital de distrito, a pesar de no contar con el mínimo de 100 viviendas (INEI 1998). Además, la distribución de las viviendas se da de manera continua.



Área:

Hectáreas (Ha): 155.07

Porcentaje (%): 0.25



LANDSAT 8 (2016)



Imagen RapidEye (2012)

Red vial

1.2.1

Descripción:

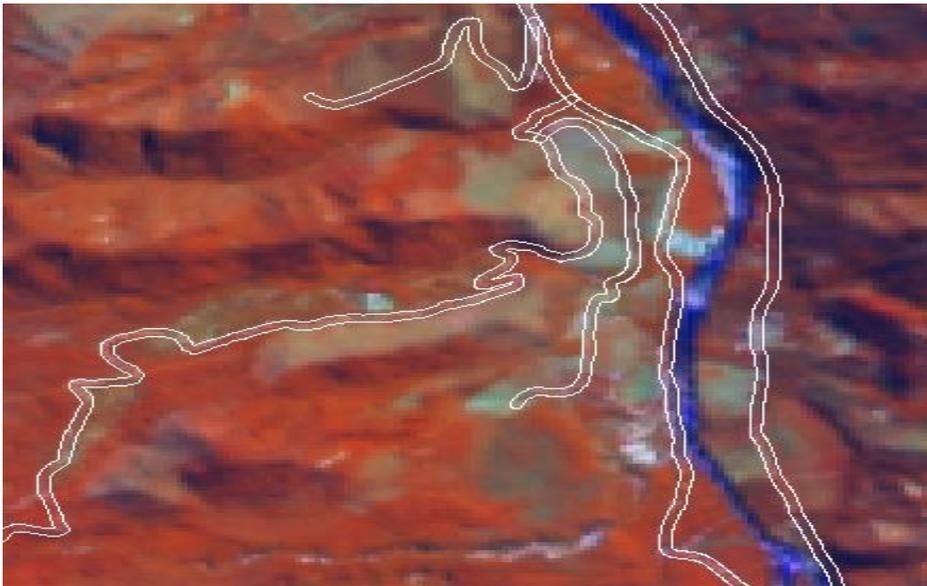
Espacios artificializados con infraestructuras de comunicaciones como autopistas, carreteras asfaltadas o afirmadas. (IDEAM *et al.* 2008)



Área:

Hectáreas (Ha): 477.57

Porcentaje (%): 0.77



LANDSAT 8 (2016)

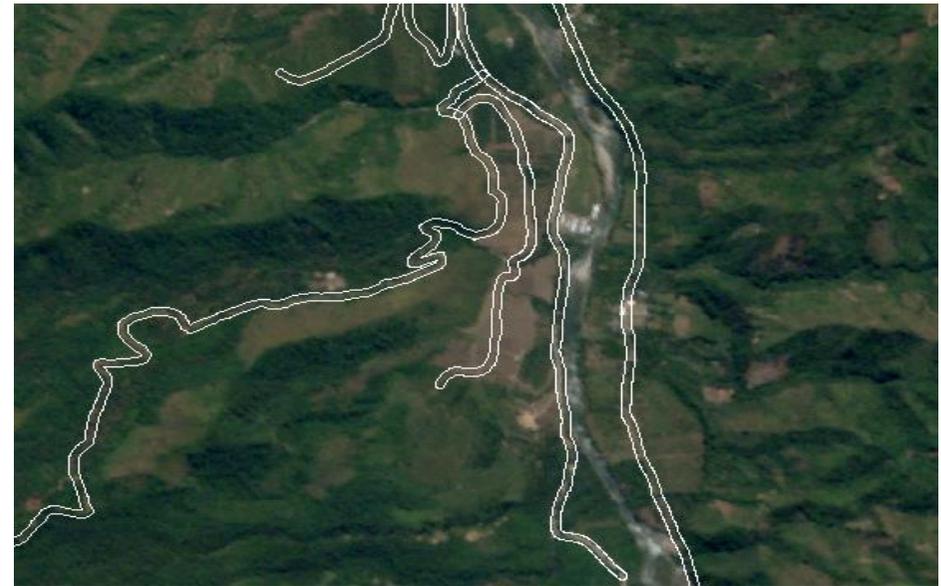


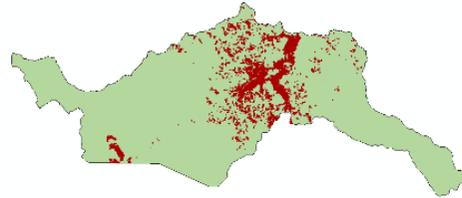
Imagen RapidEye (2012)

Mosaico agropecuario

2.1.1

Descripción:

Comprende tierras ocupadas por cultivos y pastos (naturales o cultivados), en los cuales el tamaño de las parcelas es muy pequeño y el patrón de distribución de lotes es demasiado intrincado para representarlos cartográficamente de manera individual. (IDEAM *et al.* 2008)



Área:

Hectáreas (Ha): 3,315.96

Porcentaje (%): 5.33



LANDSAT 8 (2016)

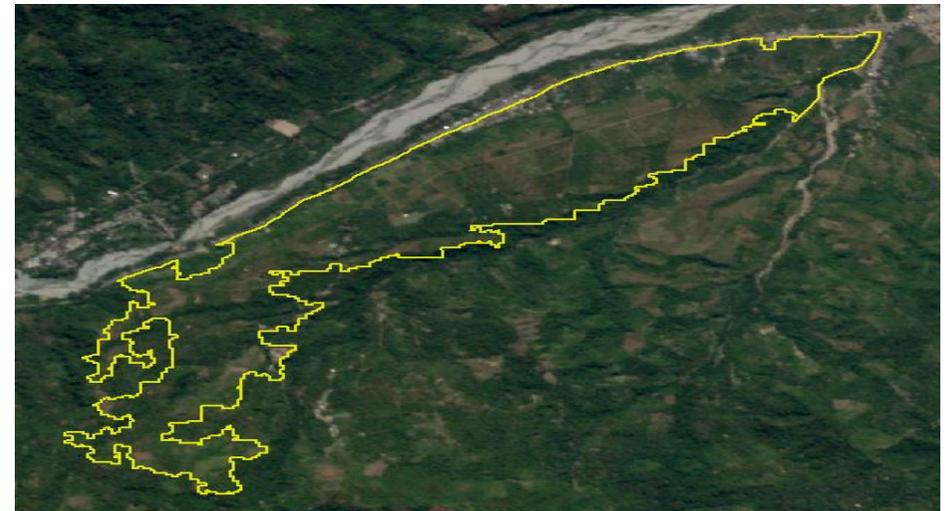


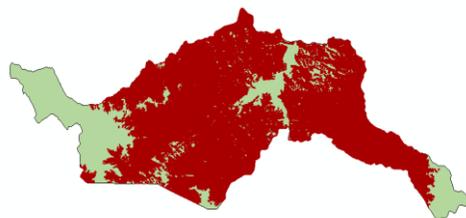
Imagen RapidEye (2012)

Bosque húmedo

3.1.1

Descripción:

Cobertura constituida por una comunidad vegetal dominada por elementos típicamente arbóreos, los que forman un estrato de copas (dosel) más o menos continuo, con altura superior a 5 metros. Estas formaciones no han sido intervenidas o su intervención ha sido selectiva y no ha alterado su estructura original. (IDEAM *et al.* 2008)



Área:

Hectáreas (Ha): 43,350.43

Porcentaje (%): 69.65



LANDSAT 8 (2016)

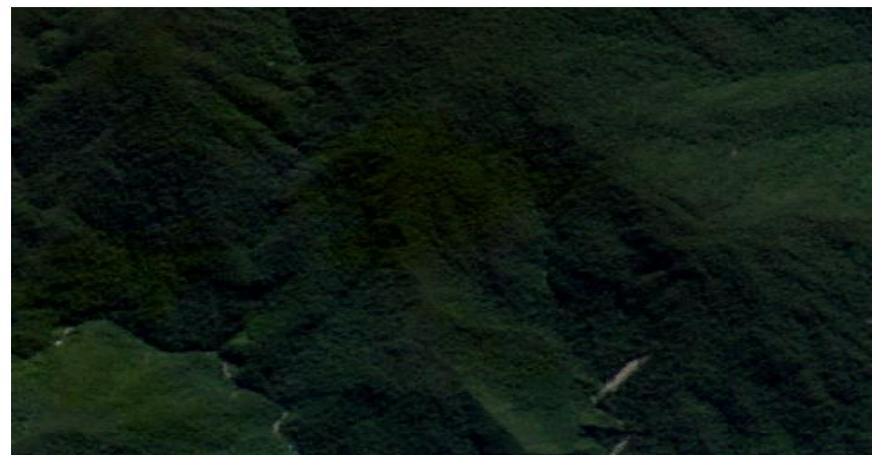


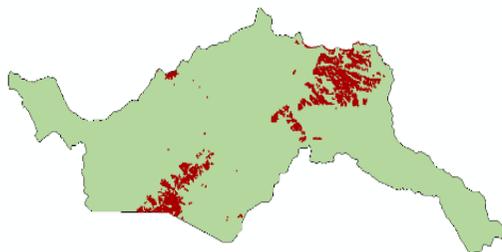
Imagen RapidEye (2012)

Arbustos y matorrales

3.2.1

Descripción

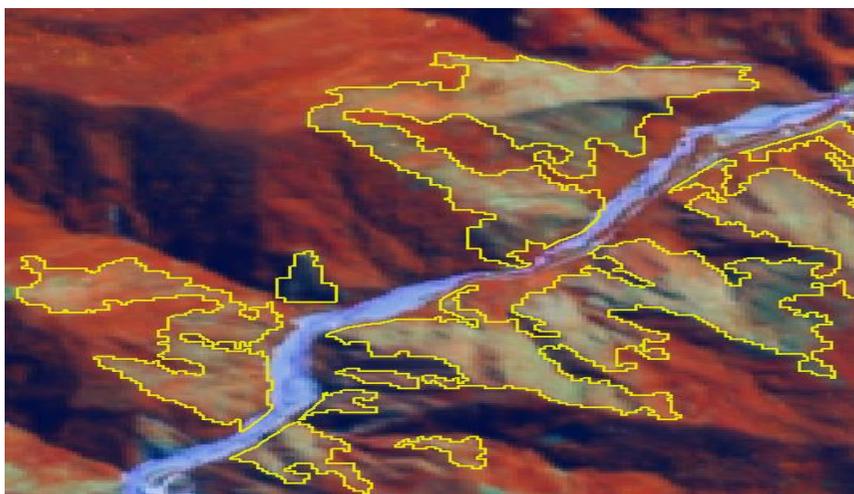
Coberturas constituidas por vegetación natural de porte arbustivo, mezclado con algunas especies arbóreas de porte muy bajo. Este tipo de coberturas se encuentran arriba de los 1500 m.s.n.m. (MINAM 2015)



Área:

Hectáreas (Ha): 2,778.30

Porcentaje (%): 4.46



LANDSAT 8, 2016



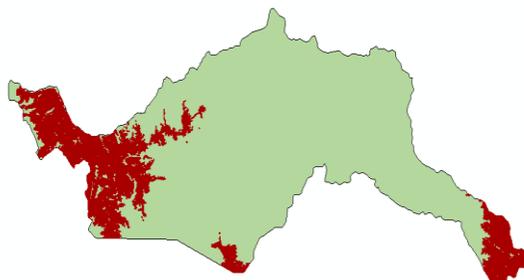
Imagen RapidEye (2012)

Puna

3.2.2

Descripción:

Esta cobertura vegetal está conformado mayormente por herbazales ubicado en la porción superior de la cordillera de los andes, aproximadamente entre 3800 y 4800 m.s.n.m. Se desarrolla sobre terrenos que van desde casi planos (planicies) hasta empinados o escarpados, en las depresiones y fondo de valles glaciares. (MINAM 2015)



Área:

Hectáreas (Ha): 9,809.38

Porcentaje (%): 15.76



LANDSAT 8, 2016



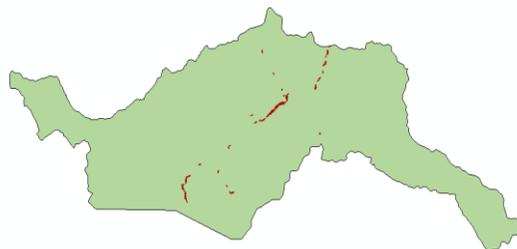
Google Earth. Imagen: SPOT (15/09/2014)

Áreas pedregosas naturales

3.3.1

Descripción:

Terrenos bajos y planos constituidos principalmente por suelos arenosos y pedregosos, generalmente desprovistos de vegetación o cubiertos por una vegetación de matorral ralo y bajo. Se encuentran principalmente conformando playas litorales, playas de ríos y bancos de arena de los ríos. (IDEAM *et al.* 2008)



Área:

Hectáreas (Ha): 80.02

Porcentaje (%): 0.13



LANDSAT 8 (2016)

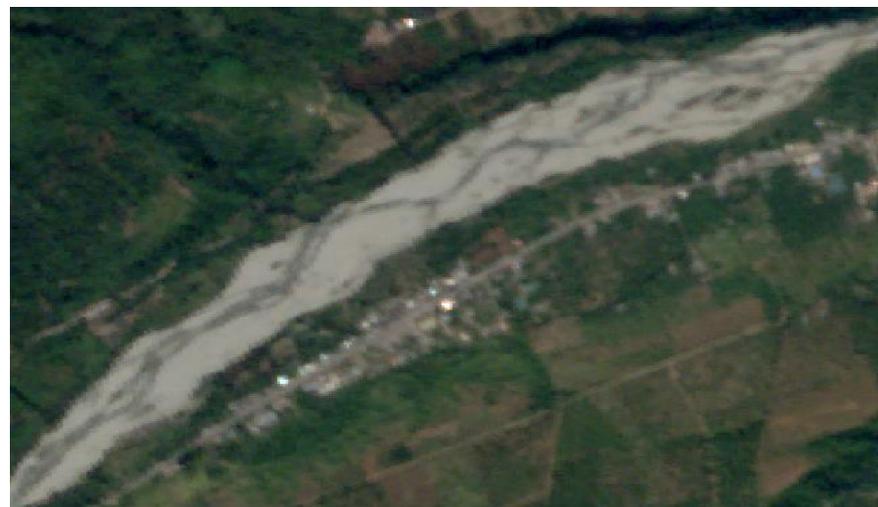


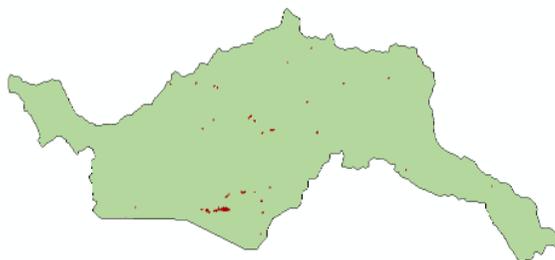
Imagen RapidEye (2012)

Tierra desnuda o con poca vegetación

3.3.2

Descripción:

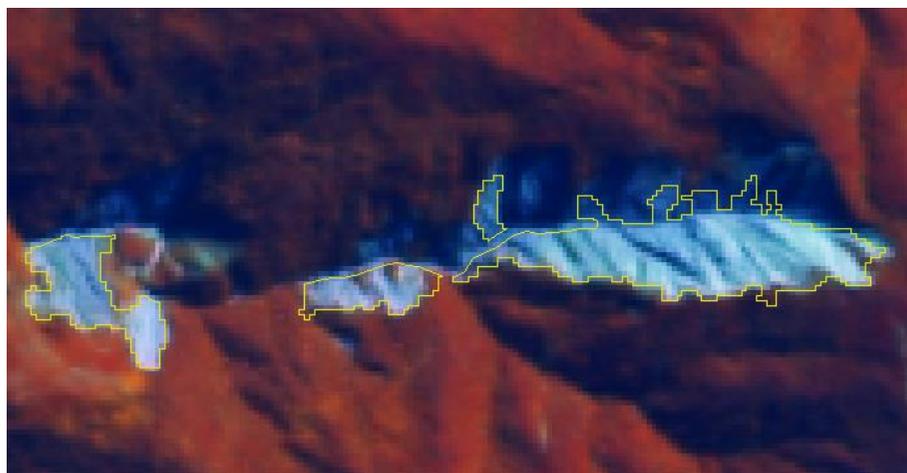
Superficie de terreno con poca o escasa cobertura vegetal, debido a la ocurrencia de procesos tanto naturales como antrópicos de erosión y degradación extrema. (IDEAM *et al.* 2008)



Área:

Hectáreas (Ha): 77.79

Porcentaje (%): 0.12



LANDSAT 8 (2016)



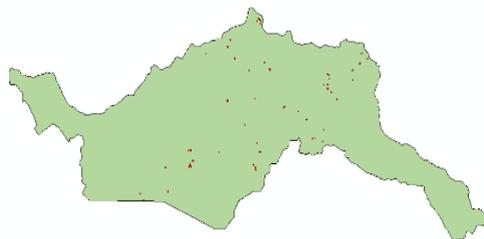
Imagen RapidEye (2012)

Zonas quemadas

3.3.3

Descripción:

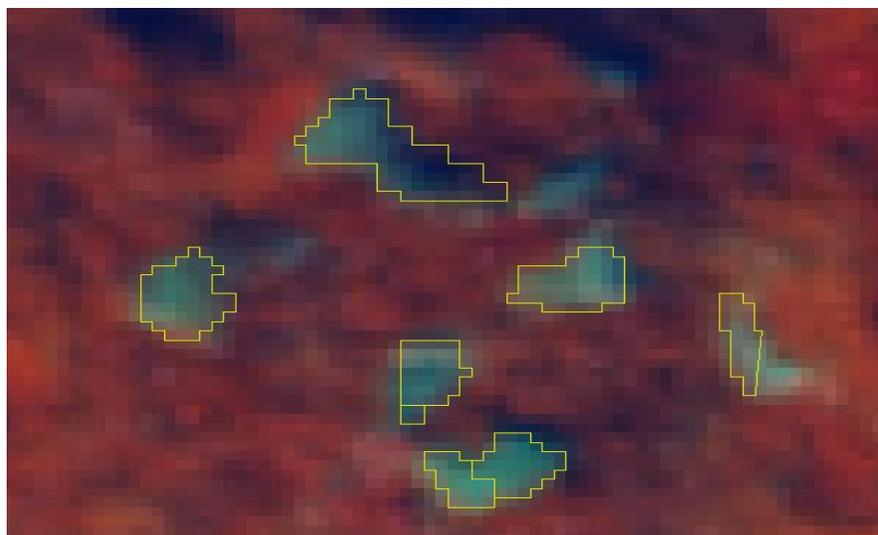
Zonas afectadas por incendios recientes, donde los materiales carbonizados todavía están presentes. Estas zonas hacen referencia a los territorios afectados por incendios localizados tanto en áreas naturales como semi-naturales, tales como bosques, cultivos y arbustos. (IDEAM *et al.* 2008)



Área:

Hectáreas (Ha): 58.49

Porcentaje (%): 0.09



LANDSAT 8 (2016)



Imagen RapidEye (2012)

Tierras altoandinas sin vegetación

3.3.4

Descripción:

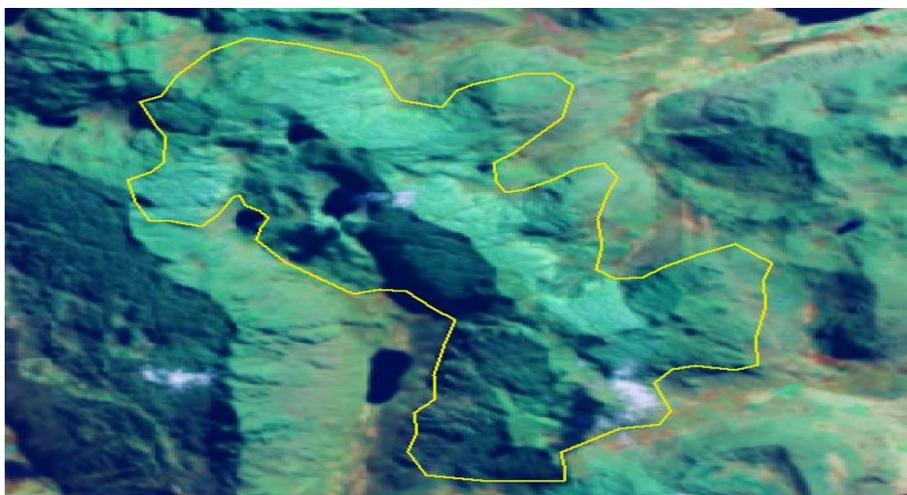
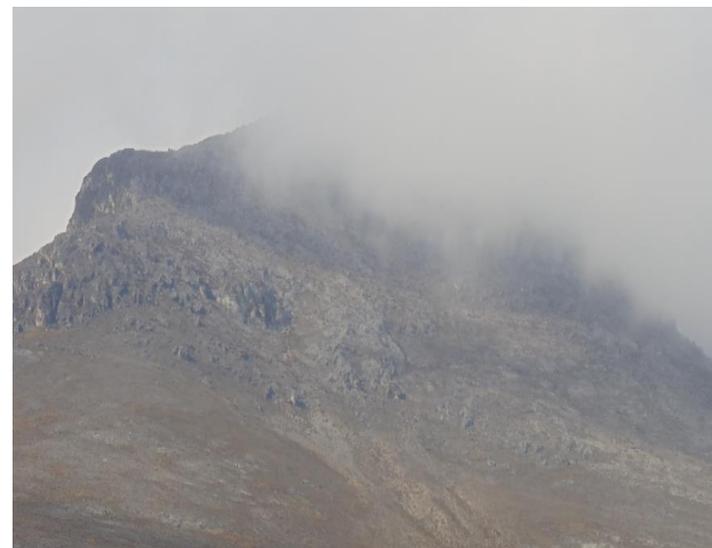
Situada a un nivel altitudinal superior a los 4500 m.s.n.m., en esta zona las herbáceas pierden cobertura debido a las condiciones extremas del clima, dejando áreas con suelos desnudos o afloramientos rocosos. Se hacen presentes especies de porte rastrero o almohadillado. (MINAM 2015)



Área:

Hectáreas (Ha): 534.02

Porcentaje (%): 0.86



LANDSAT 8 (2016)

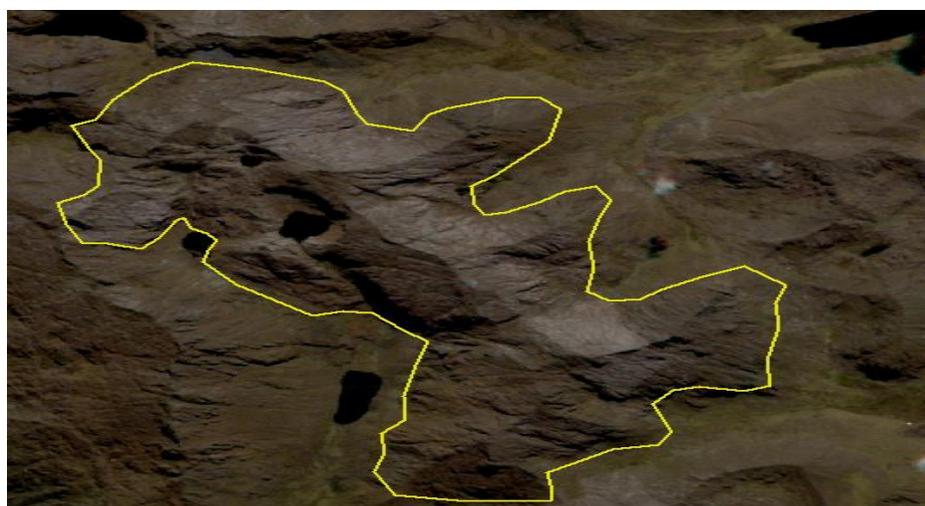


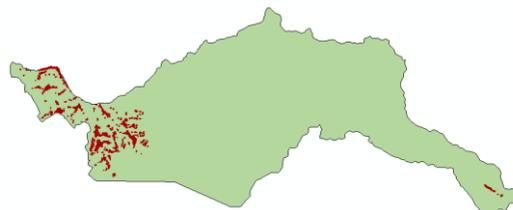
Imagen RapidEye (2012)

Bofedales

4.1.1

Descripción:

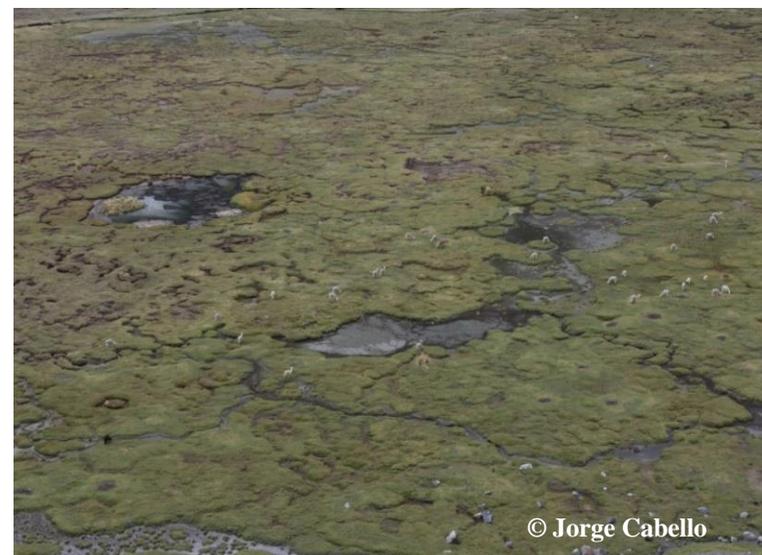
Llamados también “oconal” o “turbera”, constituye un ecosistema hidromórfico distribuido en la región altoandina. Este humedal altoandino se encuentra ubicado en los fondos de valle fluvio-glacial, planicies lacustres, piedemonte y terrazas fluviales. (MINAM 2015)



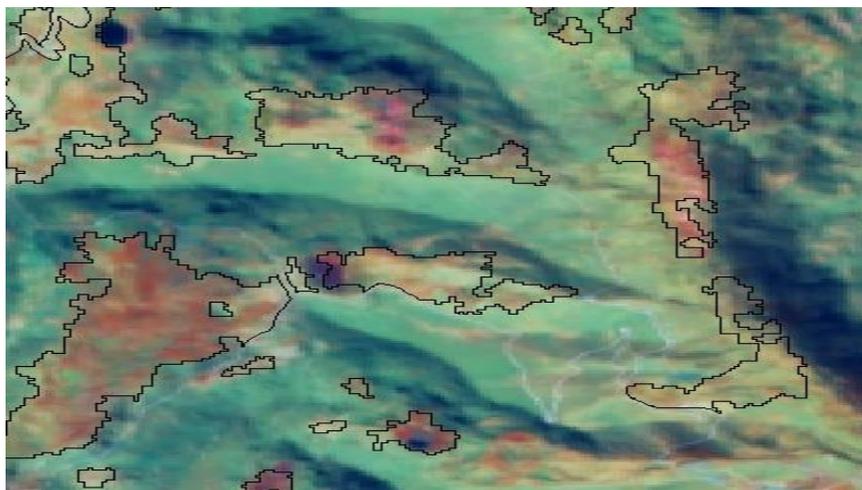
Área:

Hectáreas (Ha): 900.30

Porcentaje (%): 1.45



© Jorge Cabello



LANDSAT 8 (2016)



Imagen RapidEye (2012)

Ríos y quebradas

5.1.1

Descripción:

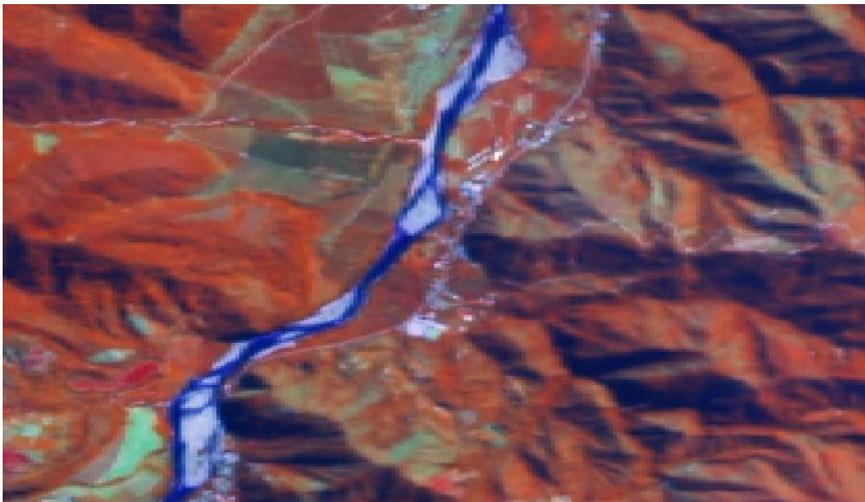
Corriente natural de agua que fluye con continuidad, posee un caudal considerable y desemboca en el mar, lago u otro río. (IDEAM *et al.* 2008)



Área:

Hectáreas (Ha): 613.22

Porcentaje (%): 0.99



LANDSAT 8 (2016)



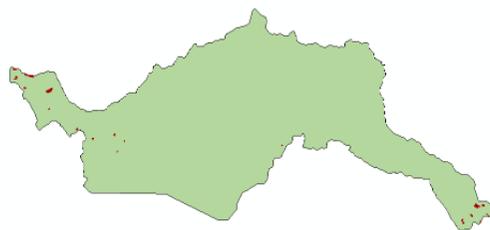
Google Earth. Imagen: SPOT (15/09/2012)

Lagunas

5.1.2

Descripción:

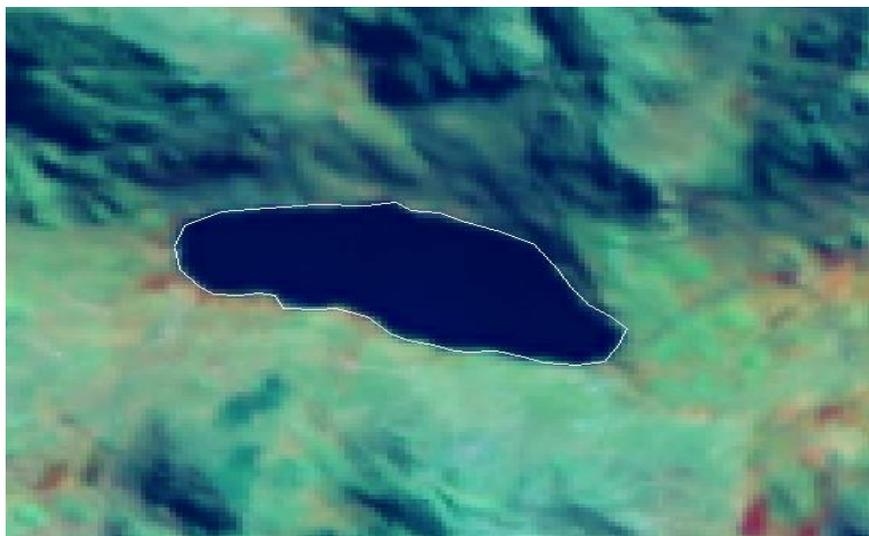
Superficies de agua naturales de carácter abierto o cerrado, dulce, que pueden estar conectadas o no con un río. (IDEAM *et al.* 2008)



Área:

Hectáreas (Ha): 85.47

Porcentaje (%): 0.14



LANDSAT 8 (2016)



Imagen RapidEye (2012)

