

Introducción al procesamiento de imágenes usando GRASS GIS 7.0 - Version 0.2

Ivan Lizarazo

10 de noviembre de 2012



Attribution-ShareAlike
CC BY-SA

<http://creativecommons.org/licenses/by/3.0/>

Resumen

Este documento es un tutorial de procesamiento de imágenes de sensores remotos usando GRASS GIS versión 7.0. Su objetivo es mostrar que este programa de software libre permite realizar de manera eficiente un amplio rango de procesos digitales útiles para obtener información relacionada con las propiedades de la superficie terrestre. El autor explica paso a paso cómo realizar diversas tareas incluyendo correcciones radiométricas, índices de vegetación, análisis de componentes principales y clasificación no supervisada y supervisada de la cobertura del suelo. Igualmente, este documento explora la interacción entre GRASS GIS y el software estadístico R. El autor espera motivar a algunas personas a reemplazar el software propietario por herramientas de software libre.

1. Introducción

El lector puede utilizar cualquier imagen Landsat para realizar las operaciones que se indican en este tutorial. Sin embargo, si el usuario no está familiarizado con GRASS puede ser más conveniente que descargue la imagen utilizada por el autor, la cual corresponde a una pequeña ventana de Landsat 7 ETM+ que cubre una zona en el norte de Bogotá (Colombia). La imagen original es una escena completa, identificada con path 08 y row 57, tomada en enero de 2003. Dicha imagen fue descargada del sitio Global Land Cover Facility (GLCF) de la Universidad de Maryland. La ventana está en coordenadas geográficas WGS84. El identificador de dicha referencia espacial es el código EPSG 4326. La imagen de trabajo tiene 1091 columnas y 1093 filas y comprende siete bandas (tres en el espectro visible, dos en el infrarrojo cercano y una en el infrarrojo térmico). La resolución espacial de todas las bandas, excepto la correspondiente al infrarrojo térmico, es de 30 metros. La resolución radiométrica de la imagen es de 8 bits, es decir sus niveles digitales pueden ocupar el rango entre 0 y 255. El formato original es TIFF. El archivo de la imagen se puede descargar del siguiente enlace:

<https://docs.google.com/open?id=0BzEwvK1H17qecEViN0t1bmpfWk0>

Para ejecutar las instrucciones de este tutorial hay que tener instalado el software GRASS GIS el cual se puede descargar desde el sitio <http://grass.osgeo.org>. Existen versiones de GRASS para los sistemas operativos MS Windows, GNU Linux y MAC OSX. La versión de GRASS GIS utilizada para realizar este tutorial es la *7.0-svn* la cual se instaló en Ubuntu 12.04.1 LTS. La versión estable más reciente es *6.4.2* la cual es probable que pueda ser utilizada para desarrollar estos ejercicios.

Este tutorial asume que el lector tiene un conocimiento básico de percepción remota y de procesamiento digital de imágenes. Por ello, no se explicarán conceptos relacionados con esos temas. El objetivo es simplemente mostrar que el software GRASS GIS ofrece funcionalidades apropiadas para diferentes proyectos y que representa una alternativa libre que es muy competitiva en comparación con programas comerciales.

En esta versión del tutorial, los temas desarrollados son:

- creación de un espacio de trabajo
- importación de imágenes
- creación de composición a color
- histogramas y diagramas de dispersión
- cálculo de factores de índice óptimo (OIF)
- calibración radiométrica
- cálculo de reflectancia absoluta
- importación de datos de elevación

- análisis de componentes principales (*nuevo*)
- índices de vegetación (*nuevo*)
- clasificación no supervisada (*nuevo*)
- clasificación supervisada (*nuevo*)
- integración entre R y GRASS (*nuevo*)

2. Organización de datos en GRASS

Los datos de GRASS se guardan en un directorio que internamente se denomina *GISDATABASE*. Este directorio se llama usualmente *grassdata* y se debe crear antes de empezar a trabajar con GRASS. Este directorio puede ser creado usando el comando *mkdir* o bien un gestor de archivos en el directorio de inicio o en una carpeta de red compartida (por ejemplo, en una red sistema de archivos NSF). Dentro de dicho directorio, los datos de GRASS están organizados por proyectos que se almacenan en subdirectorios llamados *LOCATION* como se puede ver en la figura 1.

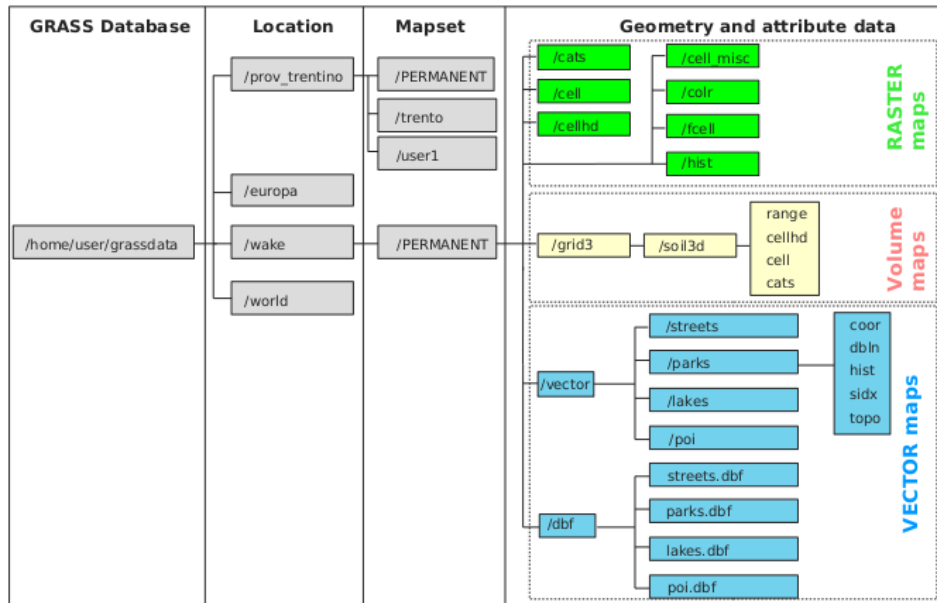


Figura 1. Organización de datos en GRASS

Cada *LOCATION* está definida por un sistema de coordenadas, una proyección cartográfica y unos límites geográficos. Los subdirectorios y archivos que definen una *LOCATION* se crean automáticamente cuando GRASS se inicia por

primera vez con una nueva LOCATION. Cada LOCATION puede tener varios MAPSETs (ver figura 1) que se usan para subdividir el proyecto en diferentes temas, sub-regiones o espacios de trabajo para integrantes individuales del grupo. Además de poder acceder a su propio MAPSET, cada usuario puede leer datos de otros MAPSETs pero únicamente puede modificar o remover únicamente los datos en su propio MAPSET. Todos los MAPSETs incluyen un archivo WIND que almacena los valores de coordenadas límite y la resolución raster que está actualmente seleccionada.

Cuando se crea una nueva LOCATION, GRASS crea automáticamente un archivo especial que se llama PERMANENT MAPSET. Ese MAPSET permite almacenar los datos centrales del proyecto, su extensión espacial default (dentro del archivo *DEFAULT_WIND*) y las definiciones del sistema de coordenadas. Únicamente el propietario de ese MAPSET puede modificar o remover ese archivo.

El archivo PERMANENT MAPSET es útil para proveer a un grupo de trabajo datos espaciales de referencia tales como elevación, vías y ríos al mismo tiempo que asegura su protección. Para importar datos en PERMANENT, simplemente hay que iniciar GRASS con la LOCATION relevante y el PERMANENT MAPSET. La organización interna y el manejo de LOCATION, MAPSETs y mapas se le deben dejar a GRASS. Es conveniente que cualquier operación que el usuario requiera tales como renombrar o copiar datos (que GRASS entiende como mapas) se debe realizar siempre usando los comandos de GRASS y no las funcionalidades del sistema operativo.

3. El concepto de región en GRASS

Una región es un concepto clave en GRASS. Una región define el área geográfica en la cual el programa debe trabajar. Una región se caracteriza por los siguientes parámetros:

- proyección geográfica (por ejemplo UTM, latitud-longitud, Gauss-Krueger, etc)
- extensión geográfica, es decir los límites North/South/East/West de la zona de estudio
- número de columnas y número de filas de los datos raster
- resolución espacial, es decir la extensión dividida entre el número de filas (resolución N-S) o de columnas (resolución E-W).

Los valores default de dichos parámetros para una *location* determinada se almacenan en el archivo *DEFAULT_WIND* dentro del mapset PERMANENT. Las definiciones de la región actual se almacenan en el archivo WIND dentro del *mapset* actual. Los valores almacenados permanecen válidos incluso si el usuario sale del programa y lo reinicia.

Cuál es la importancia de definir apropiadamente una región? Como se indicó anteriormente, la región define la extensión y la resolución de los datos dentro de la cual deben trabajar los comandos de GRASS. Eso significa, por ejemplo, que si la región esta definida para ocupar un área menor que el area ocupada por los datos (o mapas como los entiende GRASS) con los cuales está trabajando el usuario, el efecto del comando que permite desplegar esos datos (por ejemplo el comando `d.rast`) unicamente mostrará la porción del mapa que está contenida dentro de esa región.

Muchos comandos de GRASS trabajan únicamente dentro del área que está definida en la región, por ejemplo, los comandos de exportación o los módulos de procesamiento raster. Esta característica de GRASS permite que el usuario trabaje solamente en la región de interés y que no utilice recursos de computador en zonas geográficas que no son relevantes para sus propósitos.

De una manera similar, el usuario puede reducir la resolución de la región con el objeto de utilizar menos recursos de máquina. Por ejemplo, si el usuario requiere convertir datos vectoriales en datos raster utilizando una resolución menor a la definida en la región, simplemente puede cambiar la definición de la región, de manera que no consuma recursos innecesarios.

En general, es recomendable que el usuario revise las características de la región antes de realizar cualquier procesamiento. Un problema común en los usuarios que se inician en GRASS ocurre cuando intentan importar unos datos y luego, cuando quieren desplegarlos, solamente ven una ventana vacía. Ello se debe, casi siempre, a que se han definido los parámetros de la región de manera incorrecta, de manera tal que los datos importados cubren un área que está por fuera de la región.

Idealmente, la región default de una 'location' debería incluir el área que eventualmente va a ser ocupada por los datos geográficos que se van a trabajar en dicha 'location'. De esta manera, el usuario puede trabajar rutinariamente en el area que es su interes y, eventualmente, modificar los parámetros de la región para importar datos que cubran un área mayor.

4. Creación de una nueva LOCATION

Para ejecutar GRASS GIS en Linux se debe abrir una terminal. Teclee la expresión `grass70` como se indica en la figura 2:

En la ventana que se despliega, seleccione el directorio en el cual se almacenarán los datos, por ejemplo `grassdata`. Luego, haga clic en el botón *Location wizard* como se indica en la figura 3.

GRASS responde desplegando una serie de ventanas que permiten que el usuario defina una nueva LOCATION. En este caso, se puede utilizar la opción de indicar la referencia espacial utilizando el código EPSG. Como se indicó al principio de este documento, el código EPSG de la referencia espacial de nuestra imagen es 4326. Si el usuario desconoce cual es el código que identifica la referencia espacial de su imagen, puede buscarlo en el sitio <http://spatialreference.org>. Alternativamente, puede usar las otras opciones de

```
ivan@ubuntu:~$ grass70
WELCOME TO GRASS 7.0.svn

1) Have at your side all available GRASS tutorials

2) When working on your location, the following materials
are extremely useful:
- A topo map of your area
- Current catalog of available computer maps

3) Check the GRASS webpages for feedback mailinglists and more:
http://grass.osgeo.org
http://www.grass-gis.org

Hit RETURN to continue
```

Figura 2. Ejecucion de GRASS desde una terminal de Linux



Figura 3. Ventana de despliegue inicial de GRASS

creación de una LOCATION que ofrece GRASS. Al finalizar la creación de la nueva LOCATION, se puede proceder a ejecutar GRASS GIS haciendo clic en el botón *Start GRASS*.

5. Importación de la imagen al PERMANENT MAPSET

El programa inicia su ejecución mediante el despliegue de dos ventanas, una denominada *Layer Manager* y la otra *Map Display*. En la segunda fila de la ventana *Layer Manager*, en el tercer botón de izquierda a derecha, se encuentra la opción *Import raster or vector data*. Ese botón se puede utilizar para importar la imagen TIF descargada para trabajar con este tutorial. Al hacer clic en ese

botón, el programa despliega una ventana denominada *Import raster data* en donde el usuario debe ingresar el formato y el nombre del archivo que se desea importar tal como se muestra en la figura 4.

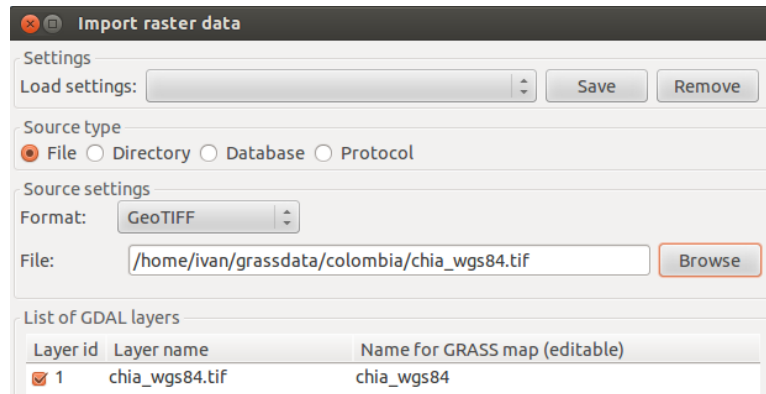


Figura 4. Importación de la imagen

```
(Sat Sep 22 11:44:17 2012)
r.in.gdal input=/home/ivan/grassdata/colombia/chia_wgs84.tif output=chia_wgs84
Datum <unknown> not recognised by GRASS and no parameters found
Projection of input dataset and current location appear to match
Raster map <chia_wgs84.1> created.
Raster map <chia_wgs84.2> created.
Raster map <chia_wgs84.3> created.
Raster map <chia_wgs84.4> created.
Raster map <chia_wgs84.5> created.
Raster map <chia_wgs84.6> created.
Raster map <chia_wgs84.7> created.
r.in.gdal complete.
(Sat Sep 22 11:44:19 2012) Command finished (1 sec)
```

Figura 5. Resultados de importar la imagen

En la figura 5 se pueden ver los resultados del proceso de importación. Se pueden observar dos cosas: (i) se han importado exitosamente cada una de las bandas de la imagen y se han creado siete mapas raster; y (ii) el programa GRASS no reconoce correctamente algunos parámetros asociados al datum de la imagen. Observe que todos los mapas raster tienen un prefijo común que es *wgs_84*. Esto significa que, para GRASS, las imágenes forman parte de un mismo ‘grupo’.

Para resolver el problema de indefinición de los parámetros de la referencia espacial, es conveniente revisar las características de la región creada por GRASS. Para ello, se debe usar la pestaña Search module del Layer Manager. En el árbol de menús, haga clic en *Settings/Region/Display Region*. El sistema muestra las definiciones que existen actualmente. Observe que el número de filas (rows) y de columnas (cols), entre otros parámetros, es incorrecto. Para cambiar esa definición, vuelva a la pestaña *Search module* y seleccione la opción *Set region*. El sistema responde desplegando la ventana *g.region* que tiene varias pestañas, entre ellas *Existing*, *Bounds* y *Resolution*. En la ventana *Existing*, en

la opción *[multiple]Set region to match raster map*, seleccione cualquiera de siete bandas almacenadas en el PERMANENT MAPSET. Luego, en la ventana *Bounds*, active la opción *Align region to resolution*. Enseguida, en la pestaña *Resolution*, ingrese el número de filas y de columnas que tiene la imagen importada.

```
(Sat Sep 22 11:59:33 2012) Command finished (0 sec)
(Sat Sep 22 12:07:56 2012)
g.region -p
projection: 3 (Latitude-Longitude)
zone: 0
datum: wgs84
ellipsoid: wgs84
north: 5:05:39.292578N
south: 4:47:51.515295N
west: 74:08:46.459894W
east: 73:51:00.636458W
nsres: 0:00:00.976923
ewres: 0:00:00.976923
rows: 1093
cols: 1091
cells: 1192463
(Sat Sep 22 12:07:57 2012) Command finished (0 sec)
```

Figura 6. Características de la región de trabajo corregida

Para verificar que la región tiene los parámetros apropiados, ejecute nuevamente el comando *Display Region*. Usted debería obtener resultados similares a los que se indican en la figura 6.

6. Creación de una composición a color

Para visualizar la imagen en una composición a color, haga clic en el botón situado en el séptimo lugar de la primera fila y que se llama *Add various raster map layers*. En el menú que se despliega seleccione *Add RGB map layer*. El sistema despliega la ventana de ejecución del comando *d.rgb*. En esa ventana, usted debe seleccionar cada una de las bandas (raster map, en el lenguaje de GRASS) asignadas a cada canal de color, *red*, *green* y *blue*. Como resultado, GRASS adiciona esa composición a color en la ventana *Layer Manager*. Allí, haga clic derecho en esa composición y seleccione la opción *Zoom to selected map(s)*. Luego, vaya a la ventana *Map Display* y haga clic en el primer botón denominado *Display map* para desplegar la composición a color. En la figura 7 se muestra parcialmente la imagen usando una composición a color RGB541.

En caso que la composición a color no produzca los resultados de visualización esperados (es decir que no se obtenga una imagen atractiva y fácil de interpretar), el usuario puede modificar la asignación de bandas espectrales a canales de color haciendo clic derecho sobre el layer que representa dicha composición en la ventana *Layer Manager* y usando la opción *Properties*.



Figura 7. Composición a color RGB541

7. Obtención de histogramas y diagramas de dispersión

Un histograma es un gráfico que muestra la frecuencia asociada a cada uno de los niveles digitales presentes en una banda individual. El histograma permite inferir, entre otras cosas, el tono predominante al visualizar una banda específica, lo mismo que el nivel de contraste correspondiente. Un diagrama de dispersión, también conocido como ploteo bidimensional, es un gráfico que muestra los niveles digitales de todos los píxeles de la imagen en dos bandas específicas. El diagrama de dispersión permite observar la correlación existente entre las bandas en cuestión.

En la ventana *Map Display*, el cuarto botón de derecha a izquierda, denominado *Analyze map*, permite obtener histogramas y diagramas de dispersión, entre otras cosas. Haga clic en ese botón y luego en *Create histogram of raster map*. En la ventana que se despliega, seleccione la opción *Histogram imagery group* y luego, en la opción *Select image group* indique que el grupo de bandas se llama *chia_wgs84*. Haga clic en *OK* para obtener los histogramas de todas las bandas de la imagen tal como se ve en la figura 8.

Para obtener ploteos de dispersión, haga clic en el botón *Analyze map* y use la opción *Create bivariate scatterplot of raster maps*. En la ventana que se abre,

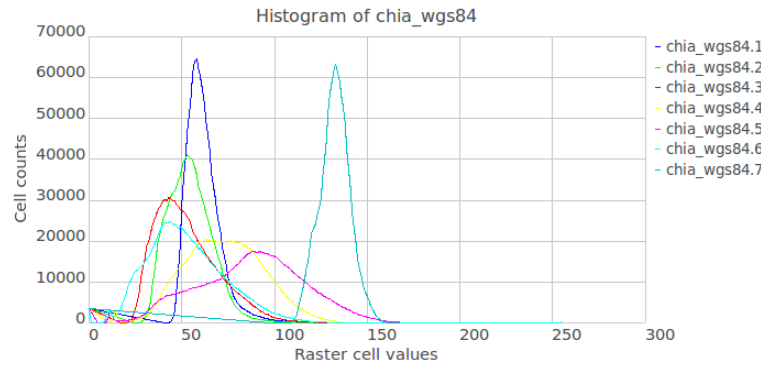


Figura 8. Histograma de todas las bandas de la imagen

seleccione las dos bandas de interés, por ejemplo las bandas 1 y 4 de *chia_wgs84*, y haga clic en OK. El resultado debe ser similar al que se muestra en la figura 9.

8. Obtención de valores OIF

El Factor de Índice Optimo (OIF, por sus siglas en inglés) es un valor que expresa la cantidad de información no redundante asociada a una combinación de tres bandas. En general, entre más alto es el valor de OIF para una combinación específica de tres bandas, mejor es la composición a color obtenida usando esa combinación de bandas.

En la ventana *Layer Manager* seleccione el tabulador *Search Module*. En el árbol de menús, haga clic en *Imagery*. Luego, haga clic en *Reports and Statistics* y posteriormente en la opción *OIF for Landsat TM*. En la ventana que se despliega, denominada *i.oif* indique el nombre del mapa raster que representa cada una de las bandas Landsat-TM, de una manera similar a la indicada en la figura 10.

Observe que, en este caso particular, el mapa raster *chia_wgs84.6* corresponde a la banda 7 de Landsat-TM. Esto se debe al orden específico que el autor utilizó para unir en un solo archivo las bandas individuales que se descargaron del sitio GLCF (este orden puede ser útil para interpretar correctamente perfiles espectrales pero puede traer complicaciones para realizae procesos posteriores). Luego de realizar la asignación requerida, haga clic en *Run*. Los resultados se muestran parcialmente en la figura 11.

El siguiente enlace proporciona algunos ‘tips’ que pueden ser útiles para definir combinaciones de bandas apropiadas para discriminar diversos tipos de cobertura del suelo: <http://web.pdx.edu/~emch/ip1/bandcombinations.html>

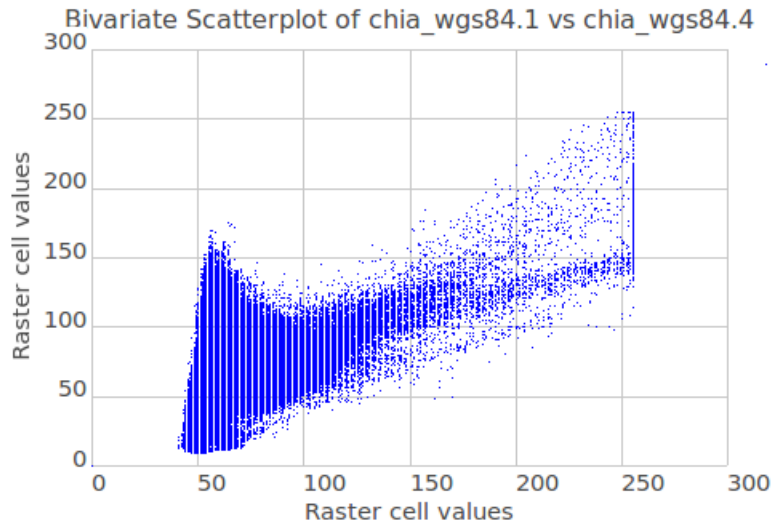


Figura 9. Gráfico de dispersión entre las bandas 1 y 4

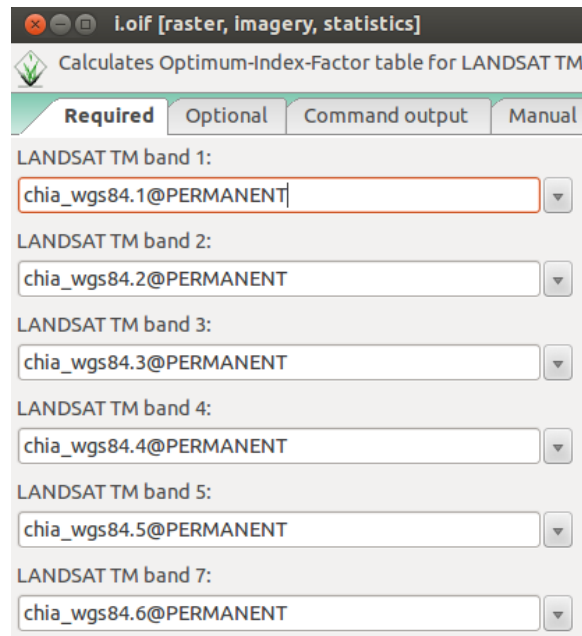


Figura 10. Asignación de mapas raster a cada una de las bandas Landsat

```
(Sat Sep 22 14:49:47 2012)
i.oif image1=chia_wgs84.1@PERMANENT image2=chia_wgs84.2@PERMANENT
Calculating Standard deviations for all bands...
Calculating Correlation Matrix...
r.covar: complete ...
Calculating OIF for the 20 band combinations...
347: 42.5344
345: 42.0629
145: 41.5852
457: 41.3590
147: 41.2375
247: 37.3572
245: 36.7323
134: 34.0029
234: 31.6122
157: 28.7791
357: 28.2516
124: 27.6301
257: 27.5799
135: 26.9503
```

Figura 11. Valores OIF correspondientes a la imagen *chia_wgs84*

9. Calibración radiométrica

La calibración radiométrica consiste en la conversión de los niveles digitales de cada banda de una imagen Landsat en radiancia espectral en el sensor. Este proceso se puede realizar en GRASS utilizando el comando *i.landsat.toar*. Usted puede acceder a las funcionalidades de este comando utilizando la pestaña *Search module* de *Layer Manager*. Para ello, vaya al árbol de menús, haga clic en *Imagery* y luego en *Satellite images tools*. Luego, haga clic en *Landsat DN to radiance/reflectance* para obtener la ventana de ejecución del comando *i.landsat.toar*.

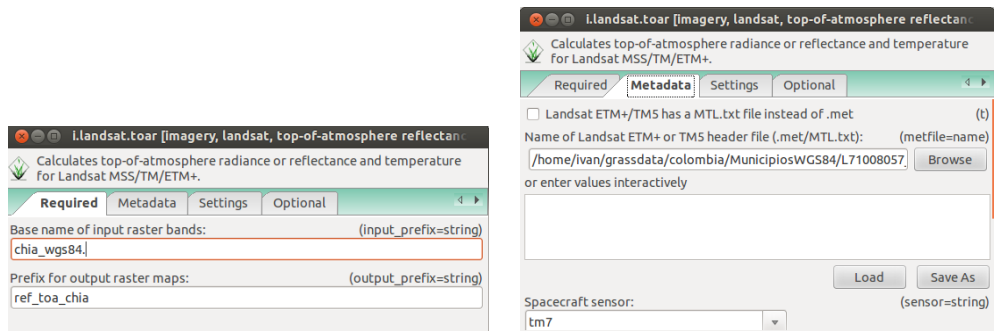


Figura 12. Parámetros suministrados para realizar calibración radiométrica

Para realizar la calibración radiométrica de una imagen se requiere conocer los datos del sensor y algunos metadatos, tales como fecha y hora de adquisición, y las funciones de calibración que relacionan los niveles digitales (ND) y los valores de radiancia espectral de cada banda. Dichos datos son usualmente proporcionados por los distribuidores de las imágenes. En este caso, se utilizó el archivo de metadatos obtenido del sitio GLCF el cual se puede descargar en este enlace <http://db.tt/IJXSfNqz>. En la figura 12 se muestran los parámetros suministrados para realizar la calibración radiométrica de la imagen que estamos trabajando. En la figura 13 se observa la banda 1 original y la banda 1 calibrada.

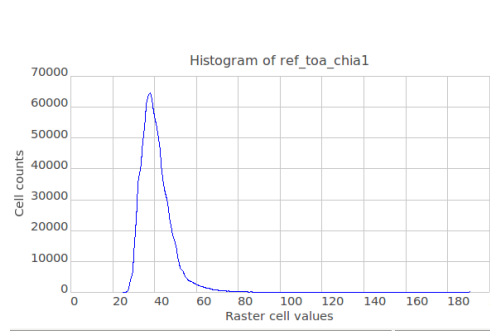
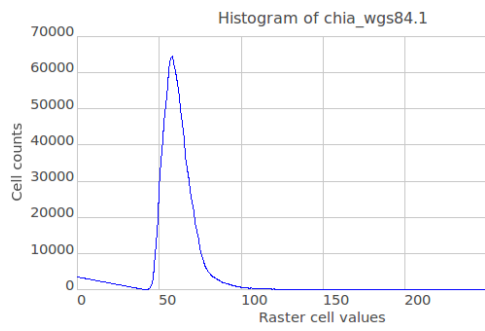


Figura 13. Comparación entre la banda 1 original y banda 1 calibrada

10. Corrección atmosférica

La corrección atmosférica de una imagen permite obtener valores de reflectancia de la superficie del terreno. La reflectancia es útil en estudios dirigidos a la obtención de variables biofísicas a partir de imágenes de sensores remotos.

La corrección atmosférica se puede realizar utilizando el comando *i.atcor* que utiliza el algoritmo 6S (Second Simulation of a Satellite Signal in the Solar Spectrum). Para ejecutar dicho comando, vaya al árbol de menús, haga clic en *Imagery* y luego en *Satellite images tools*. Luego, haga clic en *Atmospheric correction*. Enseguida, se despliega la ventana asociada al comando *i.atcor*.

Para ejecutar dicho comando se deben ingresar, a través de un archivo de texto o de manera interactiva en una caja de texto, algunos parámetros relacionados con el sensor, las condiciones de adquisición de la imagen (tiempo y localización), el modelo atmosférico, la altitud del sensor y del terreno y la banda de interés. En este caso, los parámetros utilizados para obtener reflectancia absoluta en la banda 1, a partir de la radiancia en el sensor en la misma banda, fueron los siguientes:

- 8 # indica que es una imagen ETM+
- 01 11 17.67 -74.000 5.000 # imagen tomada el 11 de enero, 17 : 42 GMT, long. 74W y lat. 5N
- 1 # modelo atmosférico tropical
- 1 # modelo de aerosoles continental
- 50 # visibilidad para el modelo de aerosoles [km]
- -2.60 # altura del terreno [km] * -1
- -1000 # imagen tomada desde un sensor satelital
- 61 # banda espectral 1

En la figura 14 se muestra la imagen de reflectancia absoluta de la banda 1 con su correspondiente histograma. Observe que los valores de reflectancia han sido rescalados de sus valores originales en el rango [0,1] a un rango de enteros sin signo de 8 bits [0,255].

En el siguiente enlace se pueden consultar cuáles son los parámetros requeridos para ejecutar el comando *i.atcor* con diferentes tipos de imágenes satelitales: http://grass.osgeo.org/gdp/html_grass70/i.atcorr.html

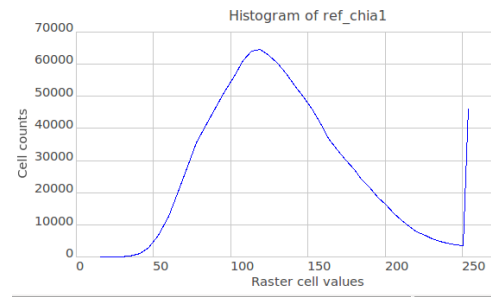


Figura 14. Imagen de reflectancia absoluta en la banda 1 con su correspondiente histograma

11. Importación de datos de elevación del terreno

En la siguiente dirección se pueden descargar datos raster de elevación de Suramérica: http://dds.cr.usgs.gov/srtm/version2_1/SRTM3/South_America/

Estos datos fueron obtenidos por NASA dentro del proyecto Shuttle Radar Topographic Mission (SRTM) realizado en Febrero de 2000. Estos datos tienen una resolución espacial de 90 metros y una exactitud de posición vertical de 15 metros para el 90% de los datos.

Los datos SRTM están organizados en hojas (*tiles* en inglés) de dimensión un grado por un grado. Los archivos con datos SRTM contienen las coordenadas del centro del pixel de la esquina inferior izquierda (por ejemplo, el archivo N51E010 tiene su celda inferior izquierda en longitud 10E y latitud 51N). Para saber cuál es el nombre de las hojas SRTM que cubren la región de interés, se puede usar la siguiente instrucción que permite visualizar una grilla de un grado por un grado: *d.grid size=1*

Al examinar la grilla desplegada se puede saber cuáles son los archivos SRTM que se deben descargar. En el caso de nuestra zona de estudio, los archivos a descargar son los siguientes: N04W074.hgt.zip, N04W075.hgt, N05W074.hgt y N05W075.hgt. Luego de realizar su descarga, los datos deben descomprimirse en el directorio de trabajo.

Para importar los datos SRTM HGT en GRASS se puede utilizar el comando *Import/link rasterorvectordata* (tercer botón desde la izquierda en la segunda fila de GRASS GIS Layer Manager). Haga clic en ese botón, luego en *Import raster data*. En la ventana que se despliega se debe indicar, entre otros parámetros, que el formato de los datos es *SRTM HGT*.

Luego de ejecutar el comando de importación, usted puede visualizar dichos datos. Observe que el área cubierta por cada *tile* importada es mucho mayor que el área cubierta por la imagen. Para cubrir solamente el área ocupada por la imagen se necesita realizar un mosaico que, primero una los cuatro *tiles* y, luego, extraiga únicamente los datos correspondientes a la región de trabajo. Para ello, se puede utilizar el comando *i.image.mosaic* que se ejecuta yendo al árbol de menús de la pestaña *Search module*, haciendo clic en *Imagery* y luego en *Mosaic images*.

En la figura 15 se observa el mosaico obtenido, el cual cubre únicamente la región de interés. El mosaico, que tiene un nivel de opacidad del 60%, está superpuesto sobre la imagen Landsat. Este ejemplo muestra la ventaja del concepto de *region* utilizado por el software GRASS GIS.

Los metadatos del mosaico de elevación se pueden obtener usando el comando *r.info* al cual se puede acceder en la pestaña *Map layers*, haciendo clic derecho en la capa *strm_chia*. Como se observa en la figura 16, el mosaico tiene el mismo número de columnas y filas (y, por ende, el mismo tamaño de pixel) que las bandas de la imagen *chia_wgs84*.

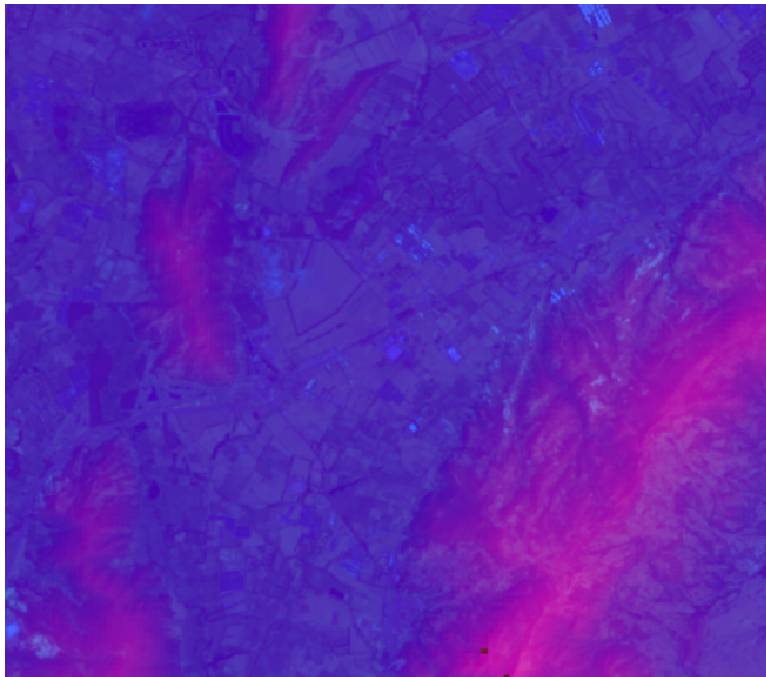


Figura 15. Mosaico de datos de elevación en la zona de estudio superpuesto sobre la imagen Landsat.

```
r.info map=strm_chia@PERMANENT
-----
| Layer:      strm_chia@PERMANENT          Date: Wed Oct  3 19:50:00 2012
| Mapset:    PERMANENT                    Login of Creator: ivan
| Location:  bogota
| DataBase:  /home/ivan/grassdata
| Title:     ( strm_chia )
| Timestamp: none
|-----
| Type of Map: raster                      Number of Categories: 0
| Data Type:  DCELL
| Rows:      1093
| Columns:   1091
| Total Cells: 1192463
| Projection: Latitude-Longitude
|             N: 5:05:39.292578N   S: 4:47:51.515295N   Res: 0:00:00.976923
|             E: 73:51:00.636458W   W: 74:08:46.459894W   Res: 0:00:00.97692
| Range of data:  min = 2533  max = 15879
|-----
```

Figura 16. Metadatos del mosaico de datos de elevación.

12. Análisis de componentes principales

El análisis de componentes principales (PCA, por sus siglas en inglés) tiene como propósito obtener nuevas bandas que contengan información no redundante (es decir, que no tengan ninguna relación entre sí). En general, es posible afirmar que los tres primeros componentes principales de una imagen representan aproximadamente el 95% de la información contenida en las bandas originales. PCA es una técnica muy útil para reducir el volumen de información original de una imagen. Igualmente, es una técnica muy popular en la realización de estudios multi-temporales.

La figura 17 muestra los parámetros utilizados para obtener los componentes principales de la imagen usada en este tutorial. La figura 18 muestra los autovalores (*eigenvalues*), autovectores (*eigenvalues*) y porcentajes de varianza asociados a cada uno de los siete componentes principales. Observe que los tres primeros componentes principales representan aproximadamente el 96.5% de la información original. la figura 19 muestra una composición a color con los tres primeros componentes principales. Se puede observar que los colores de la escena son diferentes a los obtenidos mediante composiciones basadas en las bandas originales.

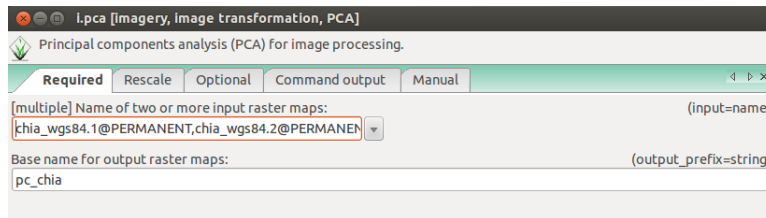


Figura 17. Parámetros de entrada para obtener componentes principales.

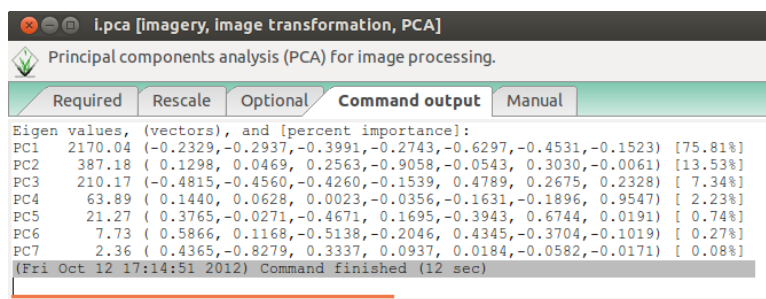


Figura 18. Resultados del análisis de componentes principales.

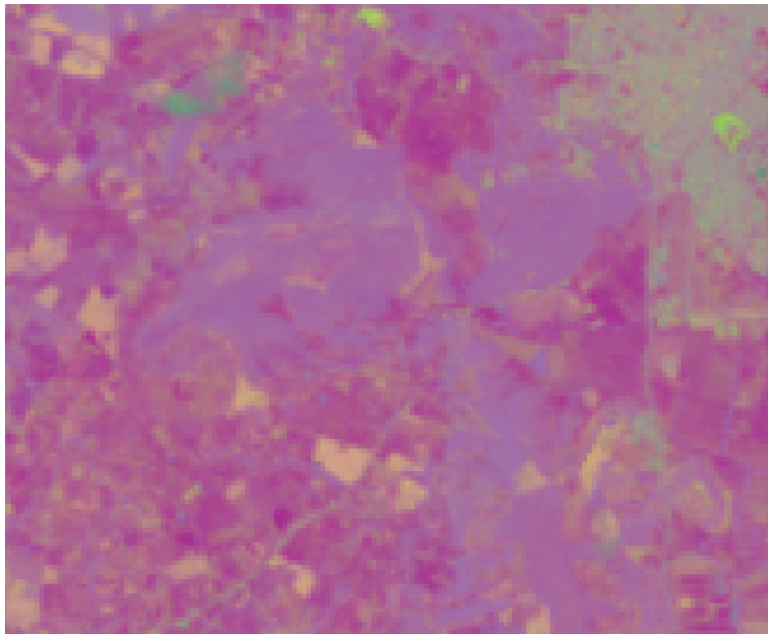


Figura 19. Composición a color RGB321 usando componentes principales.

13. Creación de una banda auxiliar de elevación

La disponibilidad de una capa raster de elevación que cubre la región de trabajo (el area de interés) con el mismo tamaño de pixel de las bandas originales puede ser aprovechada para crear un ‘grupo’ de bandas que agregue, junto a las bandas originales o a las bandas obtenidas mediante componentes principales, una banda auxiliar de elevación. Esta banda auxiliar, aunque no contiene información espectral, puede ser útil para discriminar coberturas del suelo que son espectralmente similares pero que ocurren en diferentes pisos altitudinales. En este tutorial, vamos a crear un grupo que incluya los componentes principales y la banda de elevación. Para el efecto, se usará el comando *i.group* el cual se encuentra en el árbol de menús, haciendo clic en *Imagery* y luego en *Create/edit group*. En la figura 20 se pueden ver los parámetros utilizados para crear un grupo denominado *chia8*. Este grupo puede emplearse posteriormente en la realización del proceso de clasificación de la cobertura del suelo.

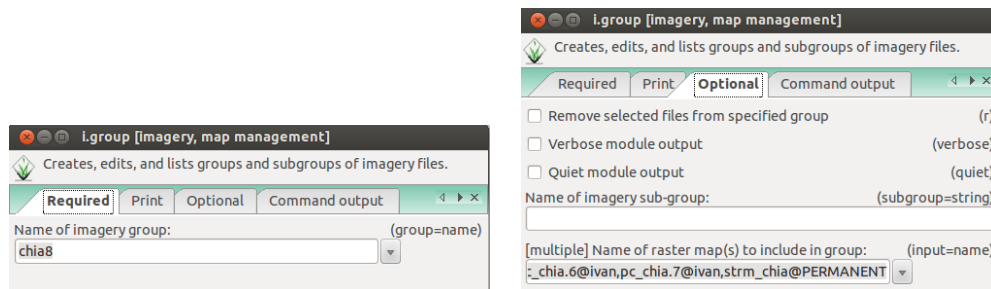


Figura 20. Parámetros utilizados para crear un grupo con siete componentes principales y una banda auxiliar de elevación

14. Índices de vegetación

Un índice de vegetación es una transformación que aprovecha las diferencias en reflectancia de la vegetación en las bandas roja e infrarroja cercana. Uno de los índices de vegetación más utilizado es el denominado índice de vegetación de diferencias normalizado (NDVI, por sus siglas en ingles). Para obtener un NDVI apropiado es conveniente utilizar una imagen que tenga corrección atmosférica. El módulo ‘map calculator’ de GRASS es una interface a la funcionalidad *r.mapcalc*, el cual permite construir fácilmente una variedad amplia de modelos computacionales basados en capas raster (incluyendo bandas de una imagen). Este módulo será usado en este tutorial para ingresar la fórmula correspondiente a NDVI: $NDVI = (NIR - Red)/(NIR + Red)$, en donde NIR es la reflectancia en el infrarrojo cercano (banda 4 de Landsat TM), y Red es la reflectancia en el rojo (banda 3 de Landsat TM).

En el menu Raster de Layer Manager, seleccione Raster map calculator. En la ventana resultante, ingrese *chia.ndvi* como el nombre de la capa raster que se va a crear, luego use los otros botones para crear una fórmula similar a la siguiente:

$$\text{float}(\text{ref_chia_4} - \text{ref_chia_3})/\text{float}(\text{ref_chia_4} + \text{ref_chia_3}) \quad (1)$$

Enseguida presione el botón Run para ejecutar el proceso. La figura 21 muestra el índice NDVI obtenido al igual que sus estadísticas (obtenidas haciendo clic derecho en dicha capa y seleccionando la opción Univariate raster statistics). Observe que entre más densidad o vigor de la vegetación existe, más alto es el valor de NDVI. Observe también que los valores NDVI se encuentran en el rango -0.32 a 0.925.

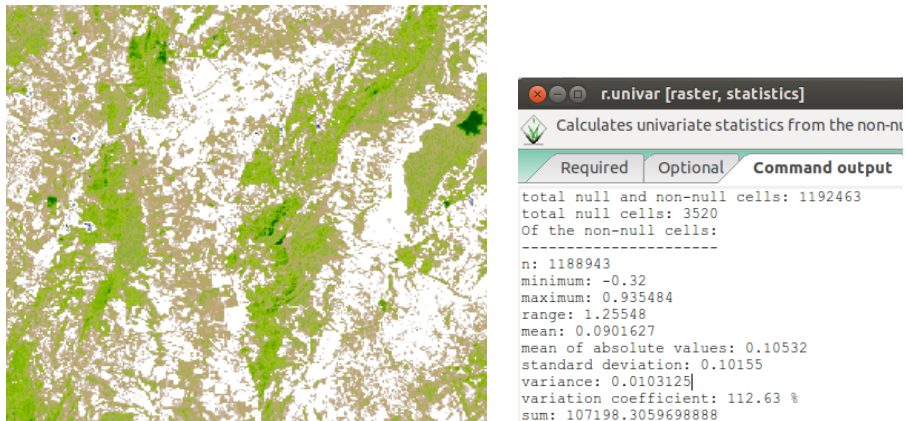


Figura 21. Índice de vegetación NDVI y sus correspondientes estadísticas.

15. Clasificación no supervisada

Una clasificación no supervisada es un proceso de agrupación de píxeles en ‘clusters’ o clases espectrales. Estos clusters no tienen significado real en sí mismos, pero el usuario puede establecer cuál es su relación con las categorías temáticas de la zona de estudio mediante un proceso de interpretación y análisis visual. En el siguiente ejercicio se realizará un ejemplo de obtención de clusters a partir de un subconjunto de las bandas de imagen, y luego, se utilizará el algoritmo de probabilidad máxima para asignar a cada píxel una clase de cobertura del suelo utilizando como base dichos clusters. Para el efecto, se deben realizar los siguientes pasos:

1. en el menú Imagery, seleccione Develop images y groups—Create/edit group. En la ventana Required ingrese el nombre del grupo de imágenes que se va a utilizar, por ejemplo *chia8*; luego, en el tabulador opcional, ingrese el nombre del subgrupo que se va a crear, por ejemplo *4bandas*; enseguida, ingrese los nombres de los tres primeros componentes principales y de la banda auxiliar de elevación. Finalmente, ejecute el proceso.
2. en el menú Imagery, seleccione el proceso unsupervised classification y luego, en la pestaña Classify image—Clustering input, indique que *chia8* es el nombre del imagery group, *4bandas* el nombre del subgrupo, *4bandas_no_sup* el nombre del archivo de firmas espectrales que obtendrá, y 10 el número de clusters deseado. Observe los parámetros default que se encuentran en la pestaña Settings y, en caso necesario, realice modificaciones. Finalmente, ejecute el proceso correspondiente. Observe que el módulo de GRASS que ejecuta ese proceso se llama *i.cluster*
3. en el menú Imagery, seleccione Classify image—Maximum likelihood classification (MLC). En la pestaña Required, indique el nombre del imagery group, por ejemplo *chia8*, el nombre del subgroup, por ejemplo *4bandas*, el nombre del archivo de firmas, por ejemplo *4bandas_no_sup*, y el nombre del mapa raster en el cual se van a almacenar los resultados de la clasificación, por ejemplo *10clusters*. Finalmente, ejecute el proceso correspondiente. Observe que el módulo de GRASS que ejecuta ese proceso se llama *i.maxlik*.

Ahora es tiempo de examinar los resultados de la clasificación, los cuales se muestran en la figura . Para ser justos, no se deben esperar resultados muy buenos, toda vez que en la ejecución de una clasificación no supervisada no se realiza ninguna selección de muestras de entrenamiento que representen categorías reales. Es bastante probable, entonces, que estos resultados apenas puedan ser calificados como aceptables.

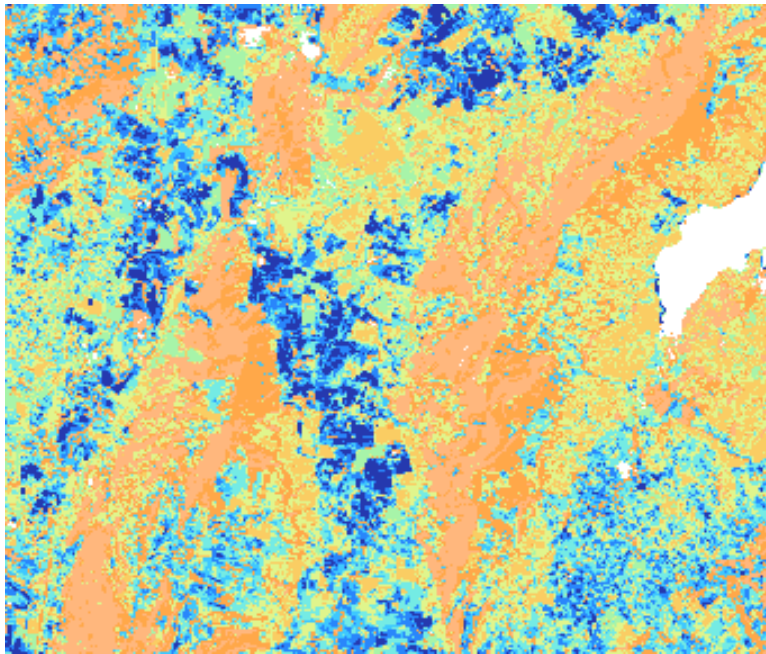


Figura 22. Resultados de ejecutar una clasificación no supervisada usando tres componentes principales y la banda auxiliar de elevacion.

16. Clasificación supervisada

Para realizar una clasificación supervisada, es necesario crear primero muestras de entrenamiento por cada categoría de interés, es decir, dibujar polígonos que representen las diferentes clases de cobertura, por ejemplo, bosques, pastos, cultivos, agua, construcciones, etc. Estos polígonos son utilizados para crear firmas espectrales de cada clase que posteriormente se utilizan para asignar a cada píxel de la imagen alguna de las categorías temáticas de interés.

En GRASS, la clasificación supervisada se puede realizar utilizando alguna de las siguientes opciones:

- utilización de los comandos *i.gensig* e *i.maxlik* (utilizando un archivo raster con muestras de entrenamiento)
- utilización de los comandos *i.class* e *i.maxlik* (modo interactivo: creación de firmas espectrales por el usuario)

Enseguida se muestra como realizar el proceso de clasificación supervisada utilizando un archivo raster con muestras de entrenamiento. Para el efecto se deben ejecutar los siguientes pasos:

1. descargue un archivo con categorías de cobertura del suelo de la zona central andina de Colombia de la siguiente dirección: <https://docs.google.com/open?id=0BzEwvK1H17qeejNOZzhiVWxEcm8>
2. importe el archivo descargado en GRASS (este archivo se muestra en la figura 23).
3. utilice el comando *i.gensig* para obtener las muestras de entrenamiento a partir del archivo raster (este comando se puede ejecutar desde la consola de comandos del Layer Manager)
4. utilice el comando *i.maxlik* para realizar la clasificación de cobertura. Los resultados de la clasificación realizada se muestran en la figura 24)

Para información del lector, el archivo raster utilizado para obtener las muestras de entrenamiento corresponde a un subconjunto de los datos globales de cobertura del suelo denominados *GlobCover 2009*, los cuales fueron producidos por la agencia espacial europea (ESA). Dichos archivos se pueden descargar del siguiente enlace: <http://ionia1.esrin.esa.int/>

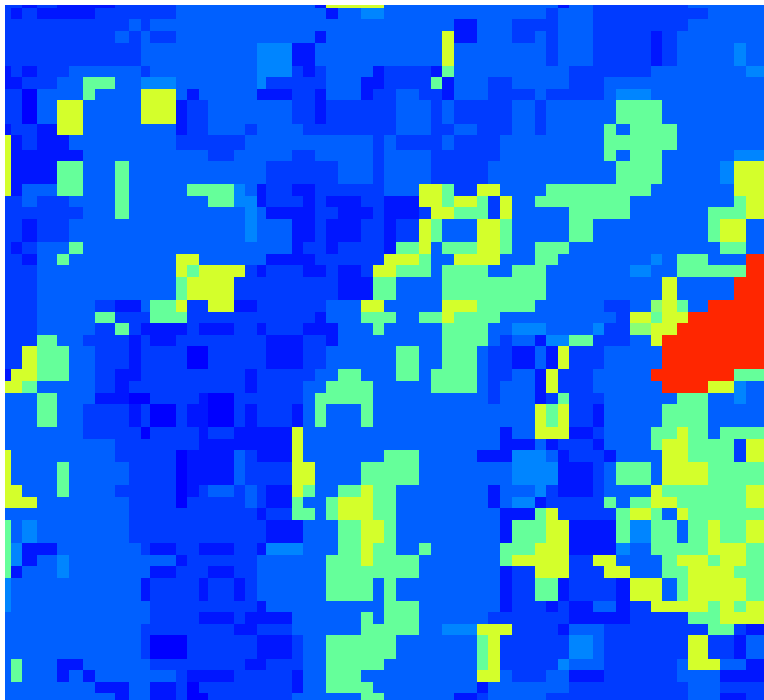


Figura 23. Vista parcial del archivo raster con categorías de cobertura.

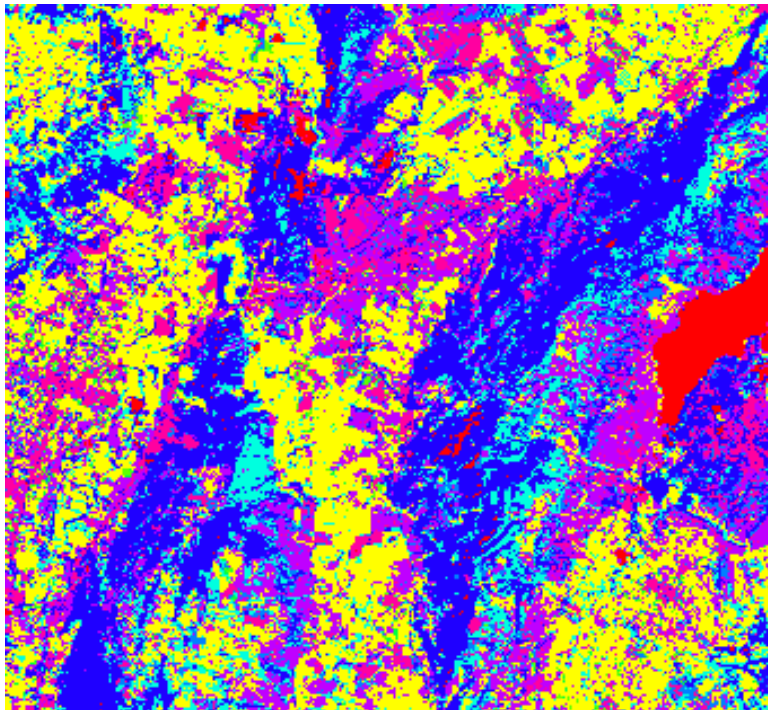


Figura 24. Clasificación de la cobertura del suelo usando un archivo raster de entrenamiento.

Enseguida se muestra como realizar el proceso de clasificación supervisada utilizando el modo interactivo. Para el efecto se deben ejecutar los siguientes pasos (tenga en cuenta que, debido a que este proceso no funcionó en GRASS 7.0, el autor lo realizó utilizando GRASS 6.4.2):

1. en el terminal de GRASS, ingrese la instrucción *d.mon X0* para desplegar un monitor X de ‘viejo estilo’ (este es un visor diferente del visor wxGUI de GRASS 7.0).
2. en el terminal de GRASS, ingrese el comando *i.class* que permite crear las firmas espectrales de cada clase de una manera interactiva. El programa le pregunta, de manera sucesiva, el nombre del mapa raster que va a ser desplegado (ingrese el nombre de una composición a color, por ejemplo *rgb742*), el nombre del grupo de imágenes (ingrese, por ejemplo, *chia8*), el nombre del subgrupo que se utilizará como base (ingrese, por ejemplo, *4bandas*), el nombre del nombre del archivo de salida (ingrese, por ejemplo *4bandas_sup*, y, finalmente, el nombre del archivo con las ‘semillas’ de las firmas (no ingrese nada). El programa despliega luego el mapa raster sobre el cuál el usuario debe ingresar de manera interactiva cada uno de los polígonos de entrenamiento (ver figura 25).
3. en el menú Imagery, seleccione Classify image—Maximum likelihood classification (MLC). En la pestaña Required, indique el nombre del imagery group, por ejemplo *chia8*, el nombre del subgroup, por ejemplo *4bandas*, el nombre del archivo de firmas, por ejemplo *4bandas_sup*, y el nombre del mapa raster en el cual se van a almacenar los resultados de la clasificación, por ejemplo *chia_sup*. Finalmente, ejecute el proceso correspondiente. Los resultados de la clasificación se muestran en la figura 26)

Usted puede realizar una estimación visual de la calidad de una clasificación mediante la comparación del mapa categórico y de una composición a color de la zona de estudio. Sin embargo, es conveniente que la evaluación de la calidad se realice mediante una evaluación cuantitativa de la exactitud temática. Para el efecto, GRASS requiere un archivo con muestras de validación (obviamente, este archivo debe ser diferente de las muestras utilizadas para el entrenamiento). Una vez se tenga ese archivo, se puede evaluar la exactitud temática utilizando el módulo *r.kappa* que produce una matriz de error con sus correspondientes indicadores.

Si el lector requiere entender a profundidad el tema de exactitud temática puede consultar el siguiente documento: http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CDAQFjAA&url=http%3A%2F%2Fwww1.usgs.gov%2Fvip%2Fstandards%2FNPSVI_Accuracy_Assessment_Guidelines_ver2.pdf&ei=HtyeUK2-M4bs9ASp-YGQBA&usg=AFQjCNGl5KgE_p_x8Ei2ZaAco_7L5ESq8w&cad=rja

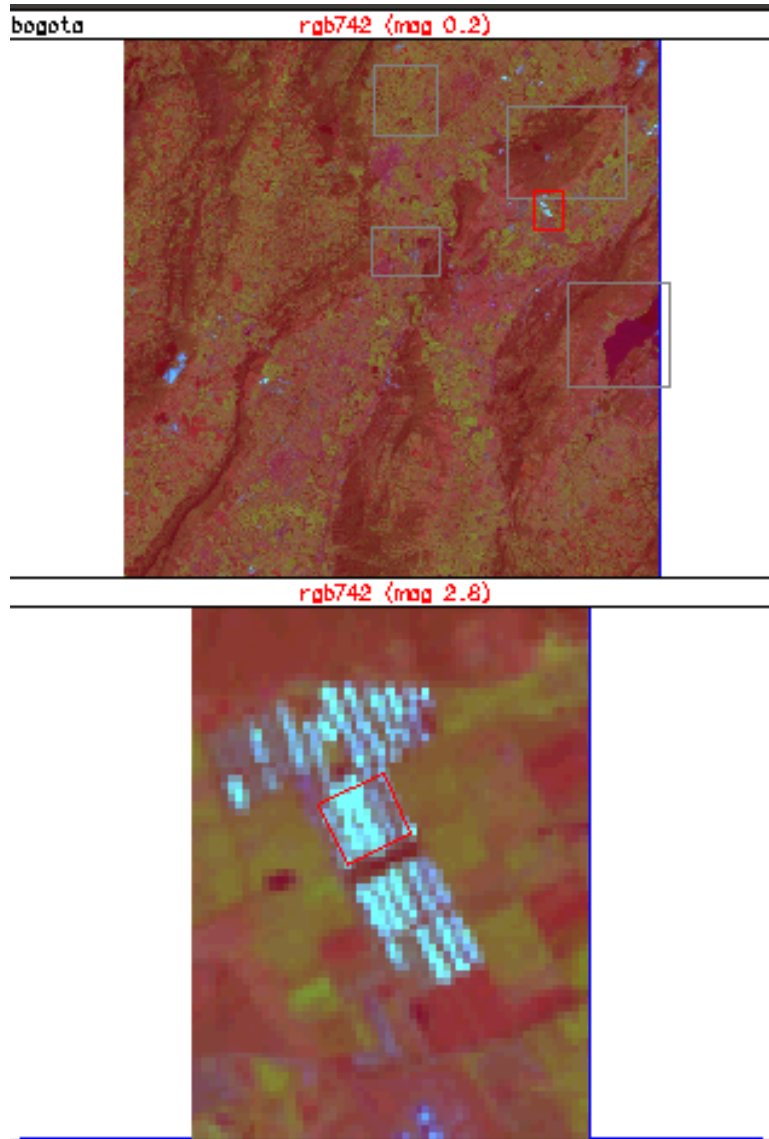


Figura 25. Monitor con la composición a color que sirve de referencia para ingresar los polígonos de entrenamiento.

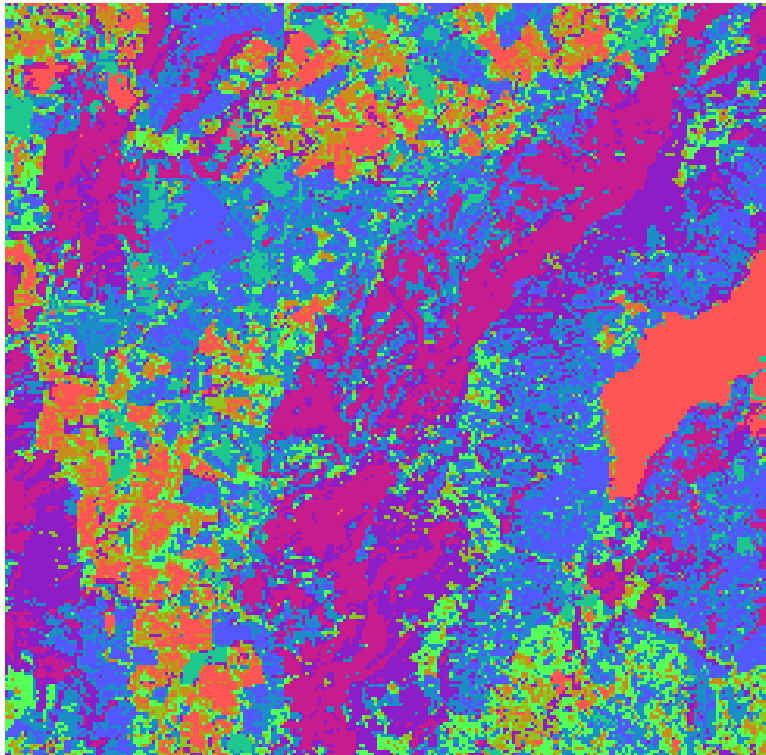


Figura 26. Clasificación de la cobertura del suelo a partir de polígonos dibujados por el usuario.

17. Integración entre R y GRASS

Las potencialidades del software estadístico *R* pueden ser utilizadas desde una sesión de GRASS. Para poder realizar la integración entre los dos programas es necesario instalar el paquete base de R, lo mismo que las librerías de R denominadas *sp* y *spgrass6*. Es posible que usted se pregunte cuál es la necesidad de utilizar el programa *GRASS* como intermediario del programa *R*. En opinión del autor, existen dos razones principales: (i) GRASS no tiene suficientes funcionalidades de análisis y modelamiento estadístico, y (ii) R tiene un problema grande de desempeño cuando se requiere procesar datos de gran volumen, como las imágenes de sensores remotos, debido a que dicho programa requiere almacenar los datos en memoria.

La instalación de R y demás librerías requeridas se describe en detalle en la siguiente dirección: http://grass.osgeo.org/wiki/R_statistics

Una vez que haya realizado dicha instalación, se puede ejecutar el programa GRASS y, luego, desde la ventana de comandos en la cual se esté ejecutando dicho programa, se debe iniciar R usando la instrucción *R*. Enseguida, dentro de la sesión de *R*, se debe cargar la librería que permite acceder a las funcionalidades de GRASS usando la instrucción *library(spgrass6)*.

La importación de las bandas de la imagen en R se realiza utilizando la siguiente instrucción:

```
> strm <- readRAST6(c("strm_ch1a", "ch1a_wgs84.7"), cat=c(TRUE,FALSE), ignore.stder  
r=TRUE, plugin=NULL)
```

Para plotear la banda de elevación en R se usa la siguiente instrucción:

```
> image(strm, attr=1, col=terrain.colors(20))  
> title("Mosaico de elevacion de la zona de estudio")
```

La figura 27 muestra un ploteo de la banda de elevación en R.

Para conocer las características del objeto *strm* se debe usar la instrucción *summary(strm)*. El resultado está desplegado en la figura 28. Observe que *strm* es un objeto de la clase *SpatialGridDataFrame*. Si el lector no está familiarizado con la librería *sp* de R, que permite manejar datos espaciales, se recomienda que revise los conceptos correspondientes en el siguiente enlace cran.r-project.org/web/packages/sp/sp.pdf

La librería *sp* proporciona clases y métodos que permiten manejar datos raster y vector desde R.

Mosaico de elevacion de la zona de estudio

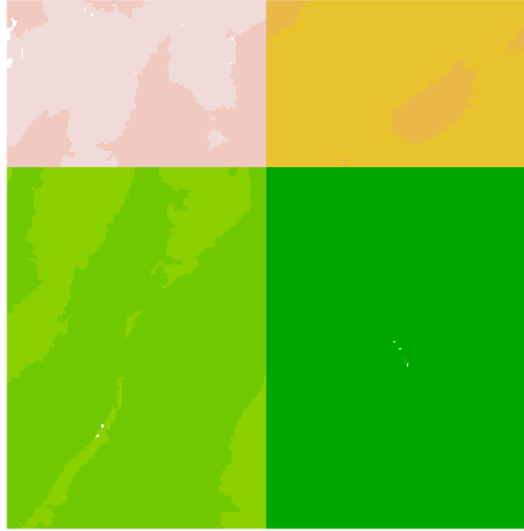


Figura 27. Ploteo de la banda de elevación en R.

```
> summary(strm)
Object of class SpatialGridDataFrame
Coordinates:
      min      max
coords.x1 -74.146239 -73.850177
coords.x2  4.797643  5.094248
Is projected: FALSE
proj4string :
[+proj=longlat +no_defs +a=6378137 +rf=298.257223563
+towgs84=0.000,0.000,0.000]
Number of points: 2
Grid attributes:
  cellcentre.offset  cellsize  cells.dim
1      -74.146103  0.0002713676    1091
2       4.797779  0.0002713676    1093
Data attributes:
  strm_chia  chia_wgs84.7
Min.   : 2533  Min.   : 0.0
1st Qu.: 2773  1st Qu.:127.0
Median : 6560  Median :133.0
Mean   : 7284  Mean   :132.1
3rd Qu.:10701 3rd Qu.:138.0
Max.   :15879  Max.   :168.0
NA's   : 780
```

Figura 28. Características del objeto *strm*.

El potencial de la integración de las funcionalidades de los programas R y GRASS es ilimitado. En la próxima versión de este tutorial, se explicará cómo realizar clasificación de cobertura del suelo usando algoritmos más ‘avanzados’ que los utilizados hasta el momento.