

Enmascaramiento de nubes en la Regiones de "Madre de Dios", "Cuzco" y "Puno" (Perú) mediante el paquete landsat del lenguaje de programación R y la comparación con el enmascaramiento de CLASLite 3.1

Masking clouds in the Regions "Mother of God," "Cuzco" and "Puno" (Peru) by landsat package programming language R and comparison with masking CLASLite 3.1

Yonatan Tarazona Coronel

Universidad Nacional Mayor de San Marcos, Calle Germán Amézaga N° 375 - Edificio Jorge Basadre, Ciudad Universitaria, Lima 1.

Resumen

Las nubes y sombras de las nubes son características comunes en las imágenes visibles e infrarrojas, que son detectadas a distancia en muchas partes del mundo, particularmente en las regiones húmedas y tropicales como es el caso del ámbito de estudio del presente artículo. Muchos nos preguntaremos, cuál es la razón principal del enmascaramiento de nubes. Pues bien, la razón principal del enmascaramiento es eliminar cubiertas que no serán parte del análisis de deforestación lo cual es necesario apartar de tal manera que la interpretación final sea más rápida y más sencilla de realizar. Para ello se ha de usar algunas librerías de R como: *rgdal*, *sp*, *mgcv* y *landsat* que enmascare de manera bastante satisfactoria la totalidad de las nubes, sombras de las nubes, cuerpos de agua y las sombras topográficas.

Filtrar los píxeles de nubes en las imágenes ópticas, conlleva una tarea dificultosa si es que no se cuenta con algoritmos automáticos que están disponibles en programas como CLASLite o cualquier otro. Por ello este trabajo muestra una metodología bastante sencilla y semiautomática que involucra en gran porcentaje la programación en R-project, el manejo de QGIS (Software libre y de código abierto) y la utilización de las bandas 1, 2, 5 y 6 del satélite Landsat TM.

Como parte de la metodología se convertirá todas las bandas a utilizar en valores de reflectancia, es decir, se hará una calibración utilizando la propia metadata de la escena. Hay que señalar que para pasar la banda termal de valores de niveles digitales a valores de temperatura (grados Kelvin) se hará uso del paquete *landsat* de R-project, que nos permitirá lograr la calibración.

Culminado esta conversión a valores de reflectancia para todas las bandas a excepción de la banda termal y utilizando nuevamente la librería *Landsat*, crearemos una máscara de nubes con la banda 1 y 6. Este paso permitirá seleccionar aquellos píxeles que tienen un umbral característico de las nubes. La dificultad radica en que para un umbral de 0.00042 no solo seleccionará píxeles que son nubes, sino también otros píxeles que no lo son. Para minimizar este error utilizaremos algunos índices como el NDVI, NDSI, que nos permita, finalmente enmascarar las nubes, sombras de las nubes, cuerpos de agua y gran parte de las sombras topográficas.

Como punto final del presente estudio. Se hará una comparación del enmascaramiento obtenido con el enmascaramiento de CLASLite 3.1. CLASLite es un paquete de software creado por el equipo del Departamento de Ecología Global de la Institución Carnegie para la Ciencia. Está diseñado para la identificación altamente automatizada de la deforestación y degradación de bosques usando imágenes de sensores remotos.

Descriptor. *CLASLite, Librería rgdal, Deforestación, imágenes ópticas, Carnegui.*

Abstract

Clouds and cloud shadows are common features in the visible and infrared images, which are detected in many distant parts of the world, particularly in humid and tropical regions such as the field of study of this article. Many ask us, what is the main reason for the cloud masking. Well, the main reason is to eliminate masking covers that will not be part of the analysis of deforestation which is necessary to remove such that the final interpretation faster and easier to perform. For this we must use some R libraries as *rgdal*, *sp*, *mgcv* and *landsat* to mask quite satisfactorily all the clouds, cloud shadows, water bodies and topographic shading.

Filter the pixels of clouds in the optical images, entails a difficult task if you do not have automatic algorithms that are available in programs like CLASlite or otherwise. Therefore, this study shows a fairly simple and semiautomatic methodology that involves a large percentage programming in R -project, managing QGIS (Free Software and Open Source) and the use of the bands 1, 2, 5 and 6 satellite Landsat TM.

As part of the methodology will turn all of the bands used in reflectance values, i.e., a calibration will use the metadata of the scene itself. Note that to pass the thermal range with digital temperature levels to values (degrees Kelvin) will make use of landsat package R -project, allowing us to achieve calibration.

Completed this conversion to reflectance values for all bands except the thermal band using Landsat librereria again, we will create a cloud mask with band 1 and 6. This step will select those pixels having a characteristic threshold clouds. The difficulty is that for a threshold of 0.00042 not only select pixels which are clouds, but other pixels that are not. To minimize this error use some indices such as NDVI, NDSI, allowing us to finally masking clouds, cloud shadows, water bodies and much of topographic shading.

As a final point of this study. A comparison of masking obtained with masking will CLASlite 3.1. CLASlite is a software package created by the team of the Department of Global Ecology at the Carnegie Institution for Science. It is designed for highly automated identification of deforestation and forest degradation using remote sensing images.

Keywords. *CLASlite, Library rgdal, Deforestation, optical images.*

1. Introducción

En las regiones espectrales del visible e infrarrojo cercano de las imágenes Landsat TM, las nubes se caracterizan generalmente por tener una alta reflectancia que la superficie (tierra o mar) subyacente. Sin embargo, bajo ciertas condiciones de la superficie, en las zonas urbanas y urbanizadas, la diferencia de reflectancia entre la nube y la superficie subyacente puede ser muy baja (Li et al. 1999, Schröder et al. 2002). Las imágenes ópticas de teledetección también tienen el problema de las sombras de las nubes y sombras topográficas que también es objeto del presente estudio [2,3].

imágenes ópticas de teledetección también tienen el problema de las sombras de las nubes y sombras topográficas que también es objeto del presente estudio.

La presencia de la banda térmica en las imágenes Landsat 5 (banda 6) permite el análisis de la superficie en términos de las temperaturas superficiales o subterráneas canopy (Melesse y Jordan 2002). En esta banda, las nubes, el follaje

de árboles y la tierra húmeda aparecen en tonos oscuros (más frías), mientras que los pastizales y

vegetación de baja altitud, tierra seca, concreto y el asfalto aparecen en tonos más brillantes (más caliente) (Jensen 2000).

El desarrollo de algoritmos complejos que permita enmascarar nubes implica tanto la identificación de las características de las nubes, como también diferenciar las sombras de las nubes de las sombras topográficas. Varias técnicas han sido desarrolladas para minimizar el efecto de las nubes y las sombras de las nubes (Ackerman et al. 1998, Choi y Milton 1999, Helmer y Ruefenacht 2005, Li et al. 1999, Liew et al. 1998, Milton y Choi 1999, Schröder y col. 2002, Song y Civco 2002). Nos centraremos en enmascarar las nubes, cuerpos de agua, sombras de las nubes y las sombras topográficas, para esto utilizamos características espectrales ópticas y térmicas [2].

Mediante el Análisis de reflectancia espectral en las diferentes bandas nos permitirá identificar la

característica de las nubes que hay en las imágenes. Para ello utilizaremos el brillo en la banda 1(azul) y la temperatura de la banda térmica.

Teniendo tanto la banda azul, roja, infrarrojo cercano, la banda térmica y utilizando el paquete landsat de R-project podremos enmascarar las nubes utilizando la metadata de la propia imagen landsat.

Esto es muy importante mencionar, ya que muchos investigadores coinciden en que el enmascaramiento tiene que hacerse con las imágenes crudas y no con las imágenes procesadas.

2. Metodología

2.1 Área de estudio

El área de estudio está ubicado entre las regiones de Madre de Dios, Cuzco y Puno con una extensión de 5567km2 de la escena PATH/ROW 003-069. Se tomó parte de esta escena porque es un área donde la presencia de nubes es muy característica, y esto nos permitirá lograr el objetivo de este estudio, que es enmascarar las nubes, sombras de las nubes, cuerpos de agua y las sombras topográficas.

Aquí un mapa que muestra el área de estudio.

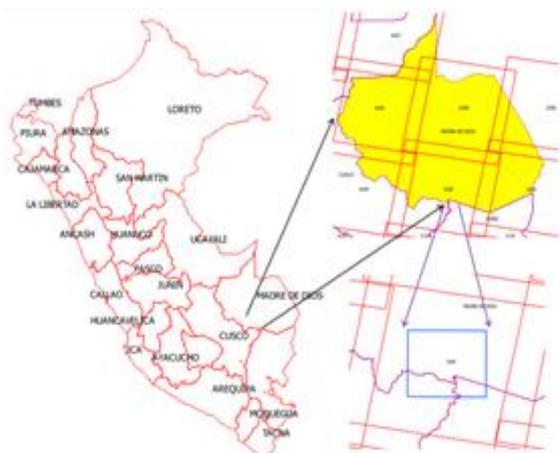


Figura1. Ubicación del área de estudio

2.2 Datos

Imágenes Utilizadas

Se empleó una pequeña extensión de una imagen Landsat TM de 30 m de resolución, fue adquirida el 16 de julio del 2005. Esta corresponde a la órbita PATH/ROW 003_069. La imagen fue descargada de la colección del USGS (U.S. Geological Survey) (<http://glovis.usgs.gov/>) de distribución libre y poseen un nivel de corrección radiométrico y geométrico L1T, es decir, tiene una buena precisión geométrica.

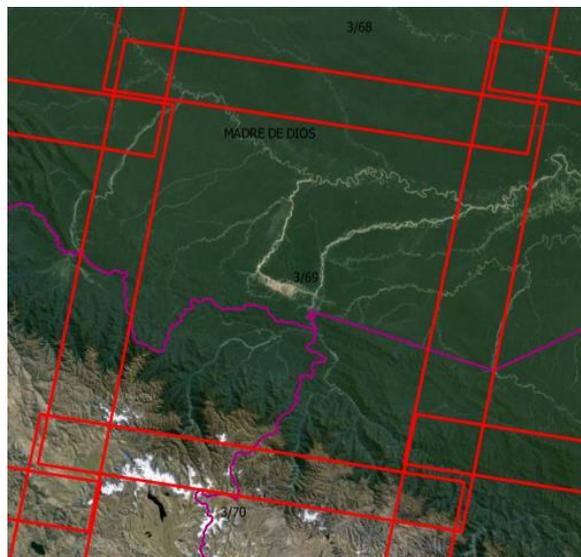


Figura2. Escena path 003, row 069

2.3 Máscara de nubes

2.3.1 Creando una máscara de nubes de una imagen Landsat usando la banda 1 y 6

Utilizando la banda 1 y la banda 6 de Landsat para identificar nubes y crear una máscara de nubes.

```
>clouds(band1, band6, level = 0.0014, buffer=5)
```

Argumentos

- band1 nombre del archivo o fichero de imagen (matriz, trama de datos, o SpatialGridDataFrame) de Landsat banda 1.
- band6 nombre del archivo o fichero de imagen (matriz, trama de datos, o SpatialGridDataFrame) de Landsat banda de 6.
- Level nivel de umbral para la toma de nube / no nube. El umbral predeterminado es adecuado para

valores de reflectancia y de temperatura, y se debe ajustar para su uso con DN.

Buffer Tamaño de la extensión de los píxeles alrededor.

Las nubes reflejan (alto) en la banda 1 y frío (bajo) en la banda de 6, por lo que la relación de las dos bandas es alto sobre las nubes.

La relación debe ser ajustada para el tipo de datos, ya sea de reflectancia o DN.

Línea de consolas en R.

```
>b1.ra<-radiocorr(b1,Grescale=0.7628,Brescale=1.52,sunelev=42.53582961, edist=ESdist("2005-07-16"), Esun= 1983, method= "apparentreflectance")
```

```
>b6.termal<thermalband(b6,band=6)
>cloud00042<clouds(b1.ra,b6.termal,level=0.00042,buffer= 1)
>writeGDAL(cloud00042,'cloud_00042.tif',drivername="GTiff")[1]
```

2.3.2 Analizando el Índice de Vegetación de Diferencia Normalizada (NDVI)

El índice de Vegetación de Diferencia Normalizada es, en algunos casos importante para filtrar las nubes, ya que el valor del NDVI es casi 0 para las nubes.

Analizando los valores del NDVI se obtiene que el umbral menor o igual a 0.4 se ajuste bastante a lo deseado.

Línea de consolas en R.

```
>ndvi < (b4_clip$band1 - b3$band1) / (b4_clip$band1 + b3$band1)
>ndvi04 <- ifelse(ndvi > 0.4, 1, 0)
>NDVI <- b4_clip
>NDVI$band1 <- ndvi04
>writeGDAL(NDVI, 'NDVI.tif', drivername = "GTiff")
```

2.3.3 Umbral de la Temperatura de la banda termal

Este filtro examina los valores de temperatura de la banda 6 para potenciales píxeles de nubes. Si un valor de píxel excede los 291 grados Kelvin, entonces no se considera como nube. Por el contrario si la temperatura es inferior o igual a 291

se considera nube. Es importante señalar que esta temperatura puede no ser características de las nubes en condiciones diferentes. Siendo un poco más preciso, si la temperatura está entre 291-300 K entonces se considera nube, pero en este caso fue menor o igual a 291 grados kelvin [3].

Línea de consolas en R.

```
>b6.termal<-thermalband(b6,band=6)
>b6.termal_1<-ifelse(b6.termal$band1<=291, 1, 0)
>b6.termal$band1<-b6.termal_1
>writeGDAL(cloud00042,'cloud_00042.tif',drivername="GTiff")[1]
```

Análisis vectorial

Los resultados originados en los puntos 2.3.1 y 2.3.2 tienen que pasar por un proceso de intersección en QGIS.

Este resultado se tendrá que fusionar con el resultado del punto 2.3.3 para finalmente obtener la máscara de nubes.

2.4 Máscara de cuerpos de agua, sombras de nubes y sombras topográficas

Uso de las bandas 2 y 5 de Landsat

Los valores de los píxeles de las bandas 2 y 5 se utilizan para formular el índice de la nieve de diferencia normalizada (NDSI). El filtro NDSI se expresa como [3]:

$$>NDSI = (banda\ 2 - banda\ de\ 5) / (banda\ 2 + banda\ de\ 5)$$

Este filtro es particularmente útil para la detección de los cuerpos de agua como ríos, lagos y lagunas. Con un umbral mayor o igual a 0 obtendremos buenos resultados.

```
>ndsi <- (b2_clip$band1 - b5$band1) / (b2_clip$band1 + b5$band1)
>ndsi_0<- ifelse(ndsi >= 0, 1, 0)
>NDSI_0<-b2
>NDSI_0$band1<-ndsi_0
>Image(NDSI_0)
>writeGDAL(NDSI_0,'NDSI_0.tif', drivername = "GTiff")
```

3. Resultados

3.1 Creando una máscara de nubes de una imagen Landsat usando la banda 1 y 6

Primero con un umbral 0.0014.

Este umbral nos da por defecto el paquete landsat de R, esto no indica que este umbral se adaptará adecuadamente a nuestra área de estudio, para ello hay que realizar varias pruebas.



Figura 3. Enmascaramiento con un umbral de 0.0014 en una imagen en falso color

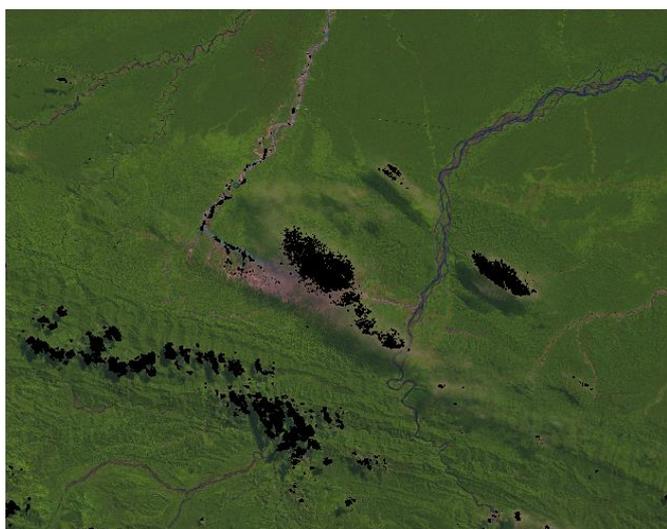


Figura 4. Con un umbral de 0.0008

Veamos los enmascaramientos en las imágenes en color real de tal manera que sean más notorias las nubes, ya que en un falso color no se aprecia con nitidez.

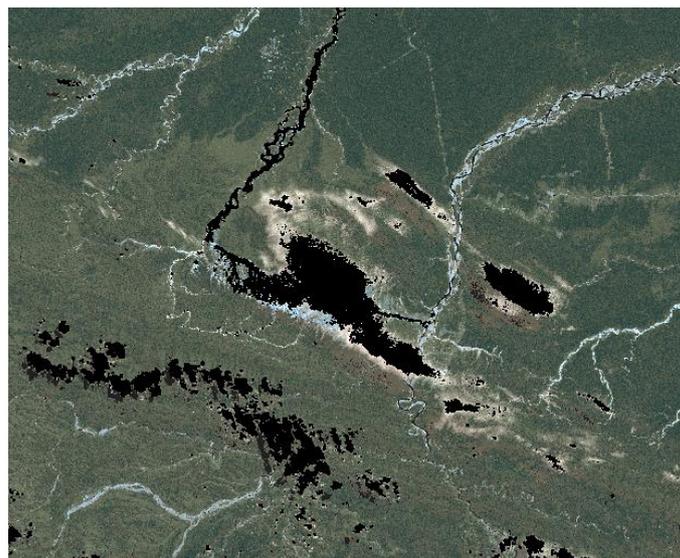


Figura 5. Con un umbral de 0.0006 en una imagen en color real



Figura 6. Con un umbral de 0.00042

Vemos que con este umbral enmascara casi la totalidad de las nubes, pero como todo algoritmo no todo lo que toma es lo correcto. Pues este umbral también toma píxeles que son ríos, o todo aquello que está expuesto a la evaporación produciendo la formación de pequeñas gotas de agua que no son otra cosa que "nube", por ello lo enmascara. Para separar únicamente las nubes debemos ayudarnos de la imagen termal.

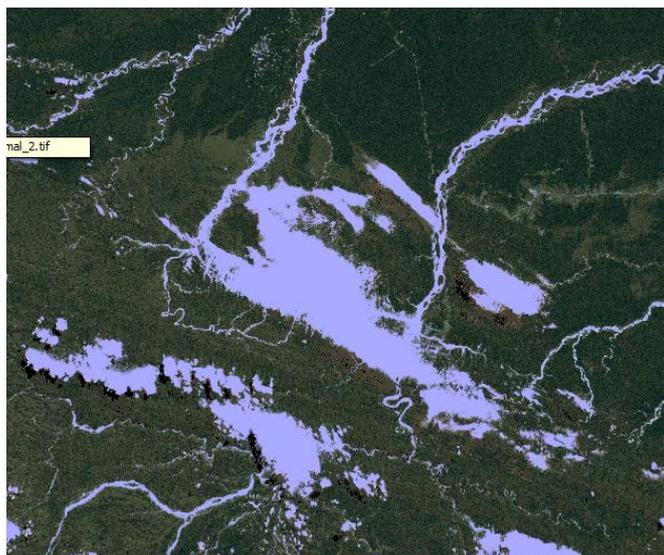
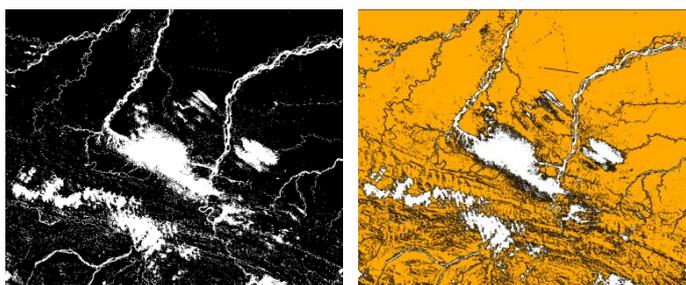


Figura 7. Enmascaramiento con un umbral de 0.00042

Analizando el Índice de Vegetación de Diferencia Normalizada (NDVI)

Con un umbral menor o igual a 0.4, será considerado como nube y lo contrario no será considerado como nube. Esto es lo que nos interesa para filtrar los píxeles [2].



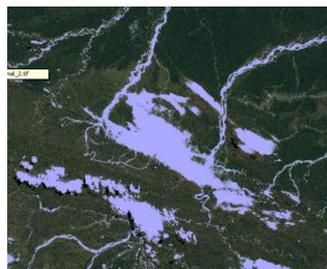
(a) Umbral ≤ 0.4

(b) Umbral > 0.4 en color naranja

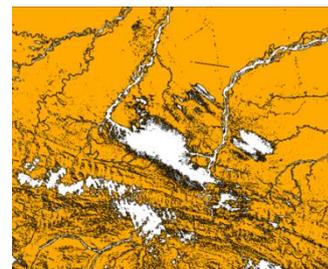
Figura 8. En la imagen izquierda (a), en escala de gris, notamos la nube con una tonalidad blanca y en la imagen derecha (b) notamos lo contrario, que es lo que nos interesa para filtrar algunos píxeles que no deben ser tomados.

Umbral de la Temperatura de la banda termal

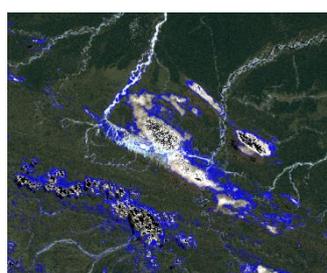
Mediante R haremos una condicional en donde los píxeles que tengan una temperatura menor o igual a 291 grados kelvin serán enmascarados como nubes.



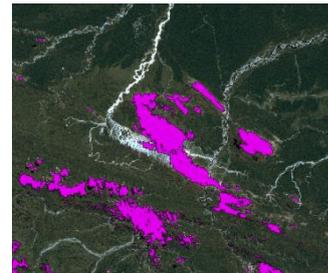
(a) Enmascaramiento con la banda 1 y 6 y con un umbral de 0.00042



(b) Enmascaramiento de nubes mediante el NDVI, umbral > 0.4



(d) Enmascaramiento Final de nubes.



(c) Enmascaramiento con una temperatura $\leq 291^{\circ}K$

Figura 9. En la imagen (d) puede mostrarse el resultado final del enmascaramiento de nubes. Esta imagen es el resultado de los tres pasos anteriores (a), (b) y (c).

3.2 Creando máscara de sombras de nubes, sombras topográficas y cuerpos de agua

Recordemos que el índice NDSI utiliza las bandas 2 y 5. En la banda 5 las nubes tiene reflexión casi cero, por lo que nos ayuda a determinar otros tipos de objetos como cuerpos de agua, sombras de las nubes y en menor medida sombras topográficas [2,3].

Es importante mencionar que el hecho de realizar el índice no indica que obtendremos de manera automática los objetos a enmascarar, sino que involucra un análisis minucioso para determinar el umbral característicos que nos permita determinar enmascarar, por un lado, las sombras de las nubes, las sombras topográficas y finalmente los cuerpos de agua que pueden ser ríos, lagunas u cualquier otro tipo con características espectrales similares.

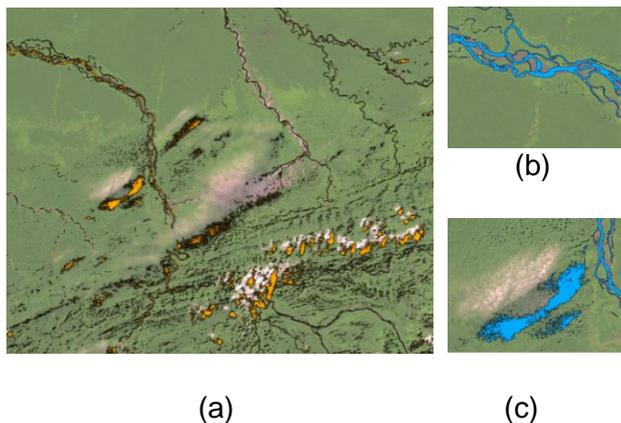


Figura 10. Con un umbral menor ≤ 0 del NDSI se obtiene buenos resultados. Y esto se aprecia en las pequeñas escenas (b) y (c) de la imagen (a) con mejor resolución espacial. En la imagen (b) se aprecia el enmascaramiento de los cuerpos de agua. En la imagen (c) de muestra el enmascaramiento de las sombras de las nubes.

3.3 Resultado final del enmascaramiento de nubes, cuerpos de agua, sombras de nubes y sombras topográficas.

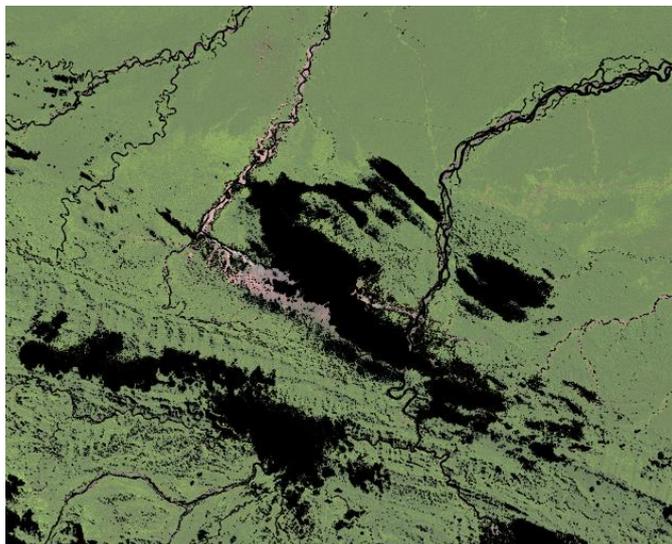


Figura 11. Este resultado final es la suma de todos los puntos tratados hasta ahora.

3.4 Comparación con el enmascaramiento de CLASlite 3.1

Una manera de corroborar nuestros resultados es una simple comparación visual con el enmascaramiento que realiza CLASlite 3.1. Es importante mencionar que CLASlite es un paquete de software creado por el equipo del Departamento de Ecología Global de la Institución Carnegie para

la Ciencia y que uno de sus procesos es enmascarar las nubes, cuerpos de agua, sombras de las nubes y sombras topográficas altamente automatizado, es decir, solamente se necesita como insumos de entrada: las imágenes crudas, algunos parámetros de su propia metadata y obtendremos el enmascaramiento como se muestra a continuación [4].

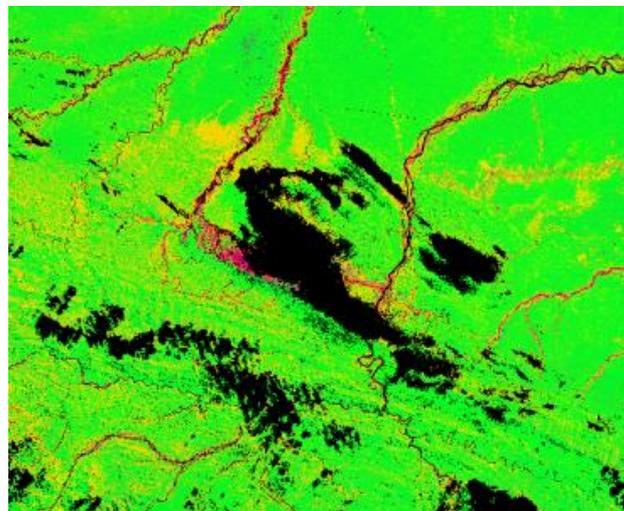


Figura 12. Enmascaramiento en CLASlite 3.1

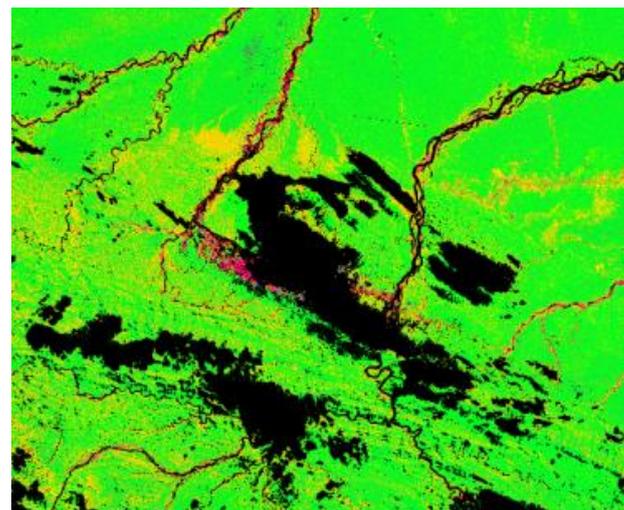


Figura 13. Enmascaramiento final del presente artículo con un ajuste de 2 píxeles.

4. Discusión

La metodología planteada, como todas en realidad, no son perfectas. Hay píxeles que no son nubes pero que son enmascarados. Esto sucede casi en toda la imagen, pero con menor fuerza en zonas con pendiente plana.

Lo cierto es, que con esta metodología los usuarios ahorrarían mucho tiempo dedicado a la verificación de la deforestación, ya que al tener un buen enmascaramiento se puede intersectar con el resultado de la deforestación y la información dentro del enmascaramiento se pueda eliminar sin ningún problema, de tal manera que solo se tendría que verificar una menor cantidad de píxeles.

5. Conclusiones

- Con un umbral de 0.00042 el enmascaramiento es casi total, pero lo perfecto no siempre se cumple, por lo que con este umbral también son enmascarados otros píxeles.
- El NDSI nos permitió filtrar algunos píxeles que no son nubes que fueron enmascarados con el umbral 0.00042.
- La imagen termal juega un papel muy importante en el enmascaramiento, porque nos permite discriminar si en realidad es nube o no, ya que las nubes tienen una temperatura por debajo de los 291 grados kelvin aproximadamente (para el presente estudio). Esta temperatura característica del presente artículo puede no caracterizar las nubes de otro ámbito de estudio, siempre será necesario realizar un análisis de los datos antes de enmascarar.
- Se podría decir que el enmascaramiento tiene una exactitud muy buena en zonas planas como la selva baja. Pero, en pendientes no planas la precisión del enmascaramiento se reduce.

6. Agradecimientos

Me gustaría agradecer al Dr. José Guerrero consultor ambiental en SIG y programador de Python, R y C++, por todas las enseñanzas brindadas a mi persona y hacerme cada vez mejor como profesional dedicado a la Teledetección y los SIG.

También agradecer a la **Dirección General de Ordenamiento Territorial del Ministerio del Ambiente Perú** por capacitarme en CLASlite, que hizo posible comparar esta pequeña metodología con el enmascaramiento que realiza CLASlite 3.1.

7. Referencias

- [1] Gyanesh Chander, Member, IEEE, Brian L. Markham, and Julia A. Barsi. **Revised Landsat-5 Thematic Mapper Radiometric Calibration**. Consultado el 16 de abril 2014.
- [2] Sebastián Martinuzzi, William A. Gould, and Olga M. Ramos González. **Creating Cloud-Free Landsat ETM+ Data Sets in Tropical Landscapes: Cloud and Cloud-Shadow Removal**. Consultado el 17 abril 2014.
- [3] Richard R. Irish•Science Systems and Applications, Inc. NASA's Goddard Space Flight Center, Greenbelt, Maryland. **Landsat 7 Automatic Cloud Cover Assessment**. Consultado el 18 de abril 2014.
- [4] Carnegie Institution For Science Department Of Global Ecology 260 Panama Street Stanford, Ca 94305 USA. **CLASlite Forest Monitoring Technology**. Consultado el 20 de abril 2014.