

Detección de arboles de castaña (*Bertholletia excelsa*) en la amazonía boliviana mediante imágenes satelitales

Frank Paul de la Barra Martínez¹
Alberto Centellas Quezada²
Pieter Zuidema³
Foster Brown⁴

¹Universidad Amazónica de Pando - UAP/PROMAB
Av. 9 de Febrero #1, Cobija- Pando, Bolivia
siginvest@gmail.com

²Programa de Manejo de la Amazonía Boliviana - PROMAB
Av. 9 de Febrero #1, Cobija - Pando, Bolivia
alberto.centellas@gmail.com

³Institute of Environmental Biology, Utrecht University
PO Box 80084, 3508 TB Utrecht, The Netherlands
p.a.zuidema@uu.nl

⁴Woods Hole Research Center
Falmouth, MA 02540-1644, Massachusetts, USA
fbrown@whrc.org

Abstract. The investigation is on the detection of Brazil nut trees (*Bertholletia excelsa*), through satellite imagery. It is located in the town of Filadelfia, the department of Pando - Bolivia. The objective of this work is to detect the location of the forest species of interest and its spatial distribution, based on Landsat imagery combined with field data. The model uses the spectral bands of satellite imagery (Landsat-5/TM), not visual interpretation, analyzed from the perspective that the relation between sampling coordinates and the digital values of image, are sufficiently representative to detect the species of interest in all the satellite imagery. This has resulted in the development of the model, which uses as input the information field sampling of the kind of interest and satellite imagery, by processing based on spatial analysis, statistical and logical, the model detects forest mosaics of interest. The map is represented by location coordinates estimates for the trees of the species of interest. For the validation it was used the censuses of chestnut of 9 sites (2,035 Brazil nut trees), which resulted in the average success of detection of 77%.

Keywords: remote sensing, tropical forest, Brazil nut trees spectral bands, digital values, space distribution, sensoriamento remoto, floresta tropical, bandas espectrais, valores digitais, distribuição do espaço.

1. Introducción

El lanzamiento de la serie de satélites Landsat a partir de 1972, ha abierto una nueva forma para adquirir información de los recursos naturales y con ello la potencialidad de nuevas técnicas de análisis digital, para identificar y mapear los recursos forestales (Bauer et al, 1994; Wolter e al, 1995). Estos satélites, ofrecen una visión global de objetos y detalles de la superficie terrestre, además de facilitar la comprensión de las relaciones entre ellos, que es difícil lograrlo cuando se observan a ras de la tierra.

Para el mapeo de la vegetación, en un área determinada de la región amazónica, la diversidad, el difícil acceso y el bajo grado de exploración de grandes partes de la región, generan varias dificultades para realizar este trabajo. Por otro lado, la interpretación de imágenes de satélite de estas regiones (NDVI y combinación de bandas) y clasificación normal (supervisada y no supervisada), lleva a resultados poco exactos, debido a la complejidad del terreno y la sombra topográfica (Muller et al, 2002). Bajo este contexto, el

planteamiento de una nueva perspectiva de interpretación y clasificación de imágenes, con el fin de detectar especies forestales de interés para una región, se hace indispensable el uso de información satelital. La relación entre el bosque y el suelo; y la forma como se interrelacionan, constituyen los aspectos más importantes en la detección de especies forestales desde el espacio. El uso de bandas espectrales de la imagen satelital, no como una combinación visual, sino como valores medidos desde el espacio de la cobertura terrestre y analizados desde la perspectiva de la ocurrencia de una determinada especie, se debe a condiciones mínimas que lo propician y/o provocada por la misma ocurrencia.

El objetivo de este trabajo, es detectar la ubicación de la especie forestal de interés y su distribución espacial, basándose en valores digitales de la imagen del sensor Landsat, junto a información del muestreo de campo de la especie forestal de interés e imágenes de satélite. El procesamiento está basado en el análisis espacial, estadístico y lógico; y el modelo planteado en la “Clasificación en árbol” que puede considerarse como un ejemplo sencillo de sistema experto (Skidmore, 1989).

2. Metodología

2.1. Área de estudio.

Se ubica entre los paralelos: latitud sur $11^{\circ} 25' 34''$ y longitud oeste $68^{\circ} 53' 59''$; latitud sur $11^{\circ} 28' 15''$ y longitud oeste $68^{\circ} 51' 26''$, que corresponde al municipio de Filadelfia del departamento de Pando, Bolivia. El área estudio esta a una distancia de 54 km de la ciudad de Cobija (Figura 1) y pertenece a la región amazónica boliviana.



Figura 1. Ubicación referencial del área de estudio.

2.2. Imágenes de satélite utilizadas

La imagen Landsat-5/TM, con todas las bandas espectrales, menos la banda térmica, con los siguientes rangos espectrales: banda 1 (0.45-0.52 μm), banda 2 (0.52-0.60 μm), banda 3 (0.63-0.69 μm), banda 4 (0.76-0.90 μm), banda 5 (1.55-1.75 μm) y la banda 7 (2.08-2.35 μm). Subescena de 25 km^2 (Figura 2).

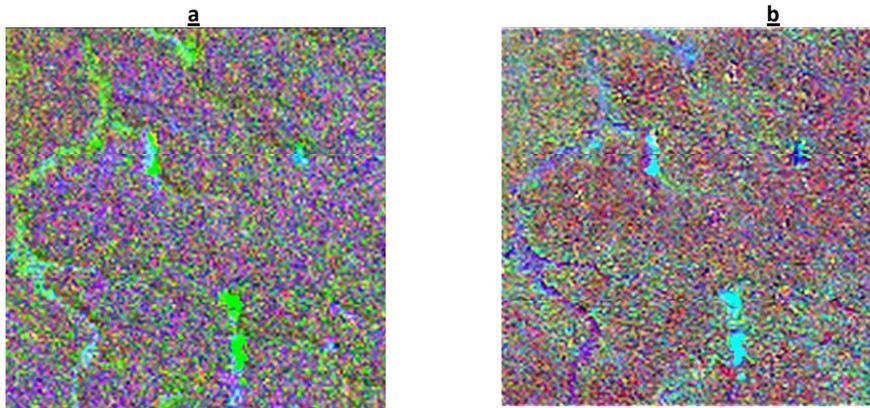


Figura 2 -. Imágenes Landsat-5/TM transformadas para obtener, dos imágenes híbridas basadas en las bandas espectrales de sensor landsat: (a) los cocientes de los canales 5/7, 3/1 y 4/3 y (b) los cocientes de los canales 5/7, 5/4 y 3/1, en las dos imágenes se les asigno respectivamente los colores rojo, verde y azul. Del área ubicada entre paralelos UTM ($x=510742$, $y=8736948$) y ($x=515742$, $y=8731949$), los mismos están en la zona 19 sur.

2.3. Censos forestales utilizados

Se ha utilizado datos de los censos forestales de plantas de castaña en producción con un DAP promedio de 127 cm y una altura promedio de 36 m de 10 sitios diferentes, del municipio de Filadelfia y provenientes de comunidades campesinas, como de predios privados “Barracas” (Tabla 1). Las coordenadas mencionadas en esta tabla, hacen referencia al centro del área del censo.

Tabla 1. Información referencial de los censos forestales de *Bertholletia excelsa*. El sitio 1, fue utilizado para fines de entrenamiento del modelo y los restantes sitios para fines de validación del modelo.

Sitio	Ubicación referencial (UTM/zona 19 sur, datum WGS 84)	Superficie del área de censo forestal (ha)	Abundancia por has (Numero de árboles ha ⁻¹)
1	X=513086, Y=8734919	475	1.72
2	X=517599, Y=8671150	11,096	1.68
3	X=596979, Y=8718797	243	2.35
4	X=621571, Y=8735820	2,357	2.07
5	X=542417, Y=8687978	2,226	3.12
6	X=631616, Y=8742584	1,678	2.45
7	X=611669, Y=8722104	11,715	1.87
8	X=548539, Y=8680944	20,601	1.78
9	X=543750, Y=8669157	10,760	1.61
10	X=539833, Y=8660686	16,934	1.70

2.4. Entrada de datos (input)

El input del modelo consistió en coordenadas de entrenamiento e imágenes de satélite (Figura 3). El modelo requiere de coordenadas de entrenamiento de 500 árboles y que tengan una precisión de 5 m. Las imágenes de la escena Landsat-5/TM, fueron transformadas a dos imágenes híbridas figura 2(a) y figura 2(b) con el programa ERDAS. Se lo hizo mediante el

comando “Índices”, mediante las funciones [Mineral Composite] y [Hydrothermal composite], especificando punto flotante como tipo de datos. Posteriormente ambas imágenes híbridas fueron desintegradas, para tener las bandas en archivos separados.

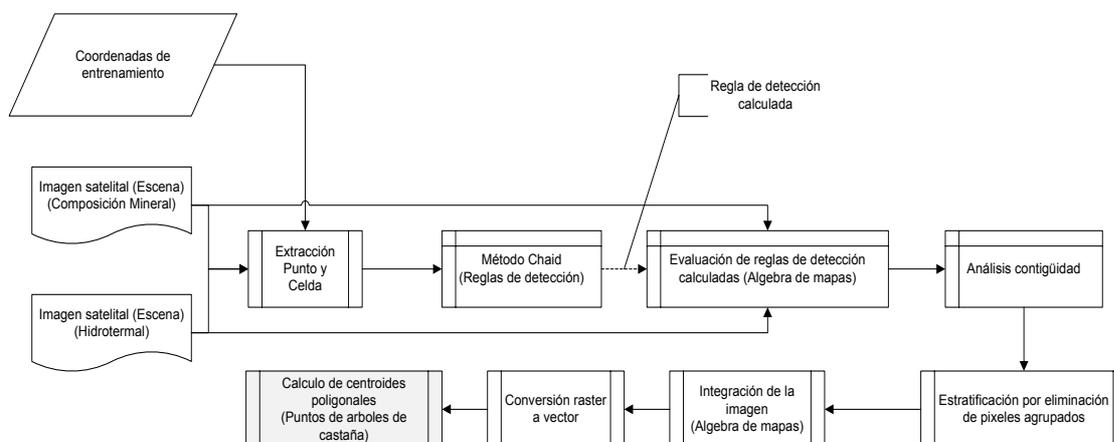


Figura 3. Modelo general de detección de especies forestales

2.5. Procesamiento

A las coordenadas de entrenamiento (muestreo), se les asignó el valor de cada banda correspondiente a las imágenes de composición mineral y alteración hidrotermal (ArcGIS, herramienta Extract values to points). Luego se aplicó el método de Chi-squared Automatic Interaction Detector (CHAID) (Morgan y Sonquist, 1963), que permite la creación de las reglas de detección automáticas de la especie forestal de interés, usando como variable dependiente la banda 2 (composición mineral), y como variables independientes las bandas restantes de composición mineral (excluyendo la banda 2) y de alteración hidrotermal. La ejecución de las reglas de detección, se realizó mediante el programa Arcgis Raster Calculator, que permite la ejecución de las mencionadas reglas, en cada una de las bandas de las dos imágenes (composición mineral y alteración hidrotermal). Para el análisis de contigüidad de píxeles en las imágenes resultado (reglas aplicadas), se utilizó el programa ERDAS Clump y el ERDAS Eliminate, para identificar las agrupaciones de píxeles (distancias mínimas entre píxeles) en cada una de las imágenes resultado (definidas en potencias de 5). Con el Arcgis Raster Calculator, se integró los resultados en una sola imagen, mediante la suma de todos los resultados, de esta manera se consiguió la imagen clasificada. Posteriormente, esta imagen fue transformada en formato vectorial (polígonos), para luego calcular los “centros poligonales” o “centroides”, que son puntos de ubicación estimativos de los árboles de la especie de interés.

2.6 Validación

Para evaluar la Correlación Bivariada entre lo detectado y lo observado (coordenadas de validación de la especie de interés), se realizó el cálculo de frecuencias de ambos datos (detectados y observados), por cuadrante, los cuales fueron de tamaños mayores a 2 has, pero no mayores al área que cubre las coordenadas de validación y se aplicó a todos los coeficientes de correlación (Spearman y Kendal), a excepción de Pearson. Finalmente para medir la precisión de la detección, se aplicó la precisión del producto. Este estadístico, indica la probabilidad de que un píxel de referencia sea correctamente clasificado o no. Es una medida del error de omisión (Jensen, 1981) que se produce, cuando un píxel posee en el terreno una cobertura determinada y no es asignado en el mapa de esa clase.

3. Resultados y discusión

El análisis de dispersión de lo detectado versus lo observado en campo, de arboles de la *Bertholletia excelsa*, en diferentes tamaños de cuadrantes, se presenta en la Figura 3. La correlación, va decreciendo de la más alta, que corresponde a la mayor superficie de observación (Fig. 3a), a la más baja que es de la menor superficie de observación (Fig.3c). Consiguientemente, a menor superficie de observación, mas es la diversidad de vegetación circundante, a la especie de interés que es detectada por el sensor del satélite y que trae como efecto que el modelo no pueda identificar con claridad la especie de interés, generando de esta manera un menor grado de correlación entre lo observado y lo detectado.

La detección de los objetos desde el espacio conlleva a considerar, la forma en que se representan los objetos (en términos informáticos) en la imagen de satélite. Considerando que la representación, es mediante formas cuadráticas llamados pixeles, que responde a una superficie constante, que está en función de la tecnología existente en el satélite. En consecuencia, las singularidades y detalles de la forma que puede tener la mancha de distribución de la especie de interés, están representados por pixeles que generan una cierta distorsión entre lo detectado de la superficie terrestre por el sensor y lo convertido por el satélite. Esto se hace evidente a medida que la escala se hace más pequeña, que trae en consecuencia la forma de la mancha, se vaya perdiendo su definición en la distribución de la especie de interés.

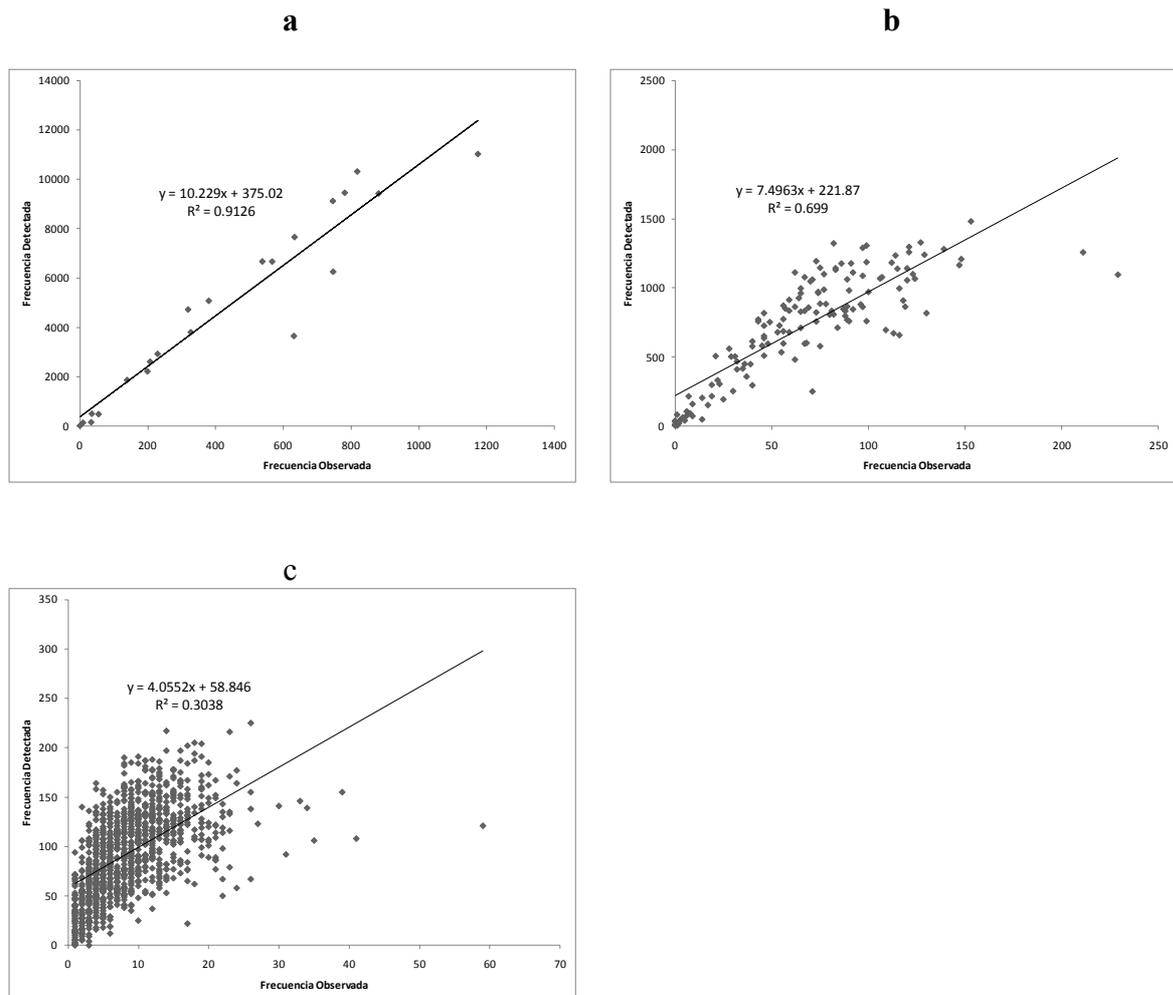


Figura 3. Relación de la frecuencia detectada por el modelo con referencia a la frecuencia observada de arboles de *Bertholletia excelsa*, en diferentes tamaños de cuadrantes: (a) 1000 has; (b) 100 has; (c) 10 has.

En la Tabla 3, se observa, que existen valores altos de correlación entre lo detectado y lo observado y esto mismos, se refleja en cada uno de los sitios de validación realizada. El promedio de precisión global del modelo es 77%. Asimismo, cuando se compara los valores detectados con los observados, los detectados son menores, Esto se debe a que el modelo, da como resultado las coordenadas de detección de la especie de interés y que son representaciones de los centroides o centros de cada polígono detectado; es decir, es una coordenada por polígono detectado. Es posible que algunos polígonos puedan tener más de un individuo censado. Se optó por esta estrategia de determinar los centroides de los polígonos detectados, para disminuir el efecto de sobredimensión causado por el mismo pixel.

Cabe aclarar que el sitio 1, no es mostrado en la Tabla 3, porque ha sido considerado como un sitio de muestreo o de entrenamiento para el modelo de detección. Debe resaltarse el hecho, que la distancia entre el área de muestreo y validación es de más de 20 km, lo que denota que la muestra (sitio de entrenamiento), es representativa de la escena (182 x 182km) de la imagen de satélite.

Tabla 3. Grado de precisión por sitio de validación

Sitio	Nombre de referencia	Resumen			
		Detectado	No detectado	Observado	Precisión Global (R ²)
2	Arroyo Malecon	2,156	477	2,633	0.82
3	El Escondido	104	129	233	0.45
4	Nagasaki	580	126	706	0.82
5	San Silvestre	415	95	510	0.81
6	Santa Fe	569	128	697	0.82
7	Alianza	7,781	1,336	9,117	0.85
8	Curichon	7,986	1,989	9,975	0.80
9	San Antonio de Chive	4,661	1,203	5,864	0.79
10	Luz América	4,143	979	5,122	0.81

El grado de precisión del modelo, depende en gran medida de la tecnología existente, específicamente de la cantidad de decimales que soporte el sistema operativo y el procesador del equipo donde se realice el procesamiento de la información. La georeferenciación de la imagen también influye en la precisión, que por supuesto debe ser la misma que de las coordenadas de campo.

El modelo detecta estimativamente, la ubicación de la especie de interés y no debe considerarse como una sustitución del censo forestal, sino como una herramienta que permita a los profesionales o técnicos tomar la mejor estrategia para realizar el censo de la especie de interés.

En la identificación de los patrones de distribución de la *Bertholletia excelsa*, el modelo provee cinco grados de agrupación o pixeles adyacentes 5, 25, 125, 625 y 3125 pixeles definidos como categorías. Estos son determinados por la cercanía que hay entre pixeles, los mismos equivalen a las superficies de: 0.5, 2.9, 14, 72 y 361 ha (Figura 4). En el proceso de validación, se ha determinado la precisión global del modelo con todas las categorías;

En la Figura 4 (a, b, c), se presentan, las categorías que da como resultado el modelo, en función de 3 escenarios. La mayoría de las categorías, están presentes en el primer escenario (Fig.4a), que muestra todas las categorías que tiene un área específica. En el segundo escenario (Fig.4b), hay una significativa disminución de manchas menores a una hectárea y se

visualiza mejor las manchas de distribución de la especie de interés, considerando que son áreas mayores a una hectárea. Ya para el tercer escenario (Fig.4c), existe una notable disminución de las manchas, para estas áreas mayores a 5 hectáreas.

Por otro lado se observa también, que de las 5 categorías se mantienen 4, de las cuales la categoría 25 píxeles $\cong 2.9$ ha, tiene una mayor distribución y abarca más superficie, este escenario es especialmente útil, para establecer el plan de trabajo del censo, porque se puede tomar dos criterios, por una parte la superficie que abarca y por el otro lado la cercanía de caminos.

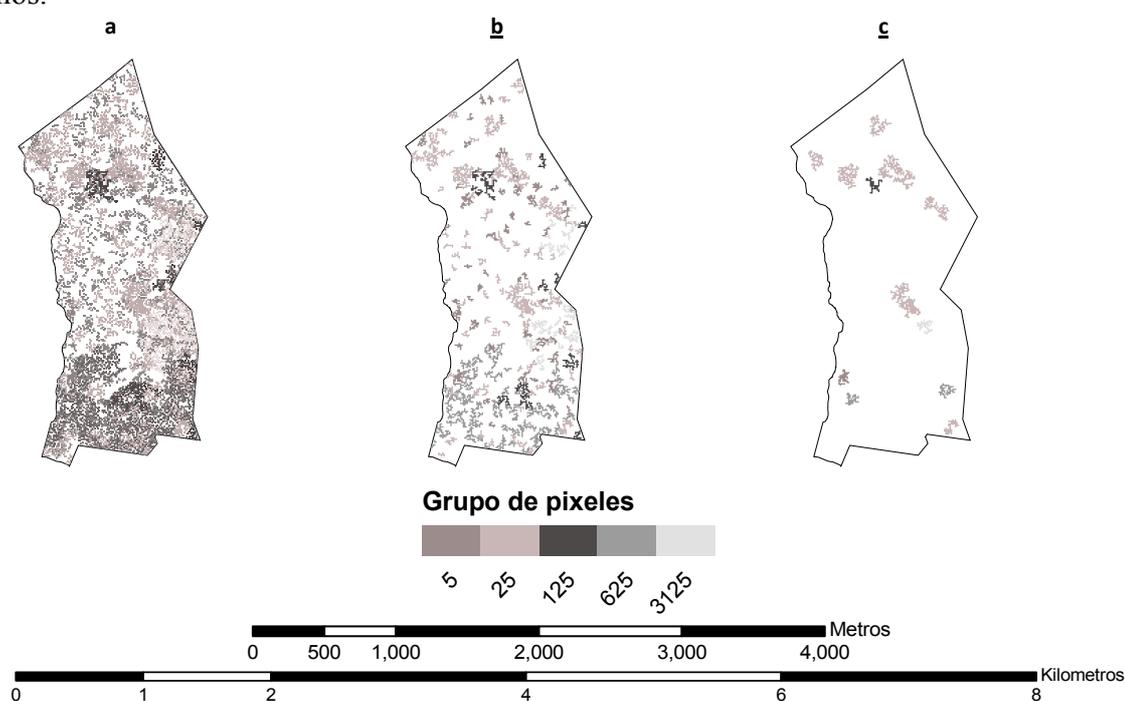


Figura 4. Escenarios y los patrones de distribución de la *Bertholletia excelsa* (escala 1:50000), detectados por el modelo: (a) áreas menores a una hectárea, (b) áreas mayores a una hectárea y (c) áreas mayores a 5 hectáreas, en la comunidad campesina de San Antonio de Filadelfia.

4. Conclusiones

El modelo estima la ubicación de manchas de castaña, con una precisión del 77%. El mismo se basa en índices relacionados con la composición del suelo y la capacidad de este, para mantener la humedad y la temperatura, en un cierto tiempo. El modelo plantea, que la intervención humana en los procesos de análisis y clasificación sea prácticamente nula y solo sea un medio de recolección de información de campo (coordenadas GPS).

5. Bibliografía

- Bauer, M.; Burk, A.; Ek, P.; Coppin, S., Lime, T., Walters, W. Satellite inventory of Minnesota forest resources. **Photogrammetric Engineering and Remote Sensing**, v. 60, n. 3, p. 287-298, 1994.
- Jensen, J. R. Urban Change detection mapping using Landsat digital data. **The American Cartographer**, v. 8, p. 127-147, 1981
- Morgan, J. N.; Sonquist, J. A. Problems in the Analysis of Survey Data, and a Proposal. **American Statistical Association**, v. 58, p. 415-434, 1963.

Mueller, R., Beck, S. G.; Lara,R. Vegetación potencial de los bosques de Yungas en Bolivia, basado en datos climáticos. **Ecología en Bolivia**, v.37, n.2, p. 5-14, 2002.

Wolter, P.T.; Mladenoff, D.J.; Host, G.E.; Crow, T.R. Improved forest classification in the Northern lake states using multi-temporal Landsat imagery. **Photogrammetric Engineering and Remote Sensing**, v.61, n.9, p. 1129-1143, 1995.

Skidmore, A. K. An expert system classifies eucalypt forest types using Thematic Mapper data and a digital terrain model. **Photogrammetric Engineering and Remote Sensing**, v. 55, n.10., - p. 1449-1464, 1989.