

Similitud de cobertura de suelo, mediante clasificación digital de imágenes satelitales, y su verificación en campo en la ciudad de Celendín, Cajamarca, 2022

Similarity of land cover, through digital classification of satellite images, and its field verification in the city of Celendin, Cajamarca, 2022

Agustín Medina Chávez^{1*}, Edgar Díaz Mori¹, Víctor Estela Becerra¹

¹Facultad de Ciencias Agrarias, Universidad Nacional de Cajamarca, Av. Atahualpa 1070, C.P. 06003, Cajamarca, Perú

* Autor de correspondencia: amedina@unc.edu.pe

Resumen

La investigación tuvo como objetivo calcular el porcentaje de similitud entre las coberturas identificadas en el terreno mediante puntos de muestreo y las categorías correspondientes en el mapa ráster derivado de la clasificación no supervisada, utilizando una imagen satelital Landsat 8 de la zona de interés. Se empleó una muestra que superaba los treinta y cinco puntos de muestreo, proporcionando un nivel de confianza del noventa por ciento y un margen de error del diez por ciento. Las imágenes de la zona de interés se descargaron de la USGS (United States Geological Survey), y se realizó una combinación de las primeras siete bandas. Luego, esta composición se fusionó con la banda número ocho para mejorar la resolución espacial, alcanzando una resolución de quince metros. Posteriormente, se llevaron a cabo correcciones de radiancia y reflectancia. Con las imágenes corregidas, se efectuó la combinación de bandas 543 para la categoría de vegetación, 764 para bosque y el índice (BSI) de suelo desnudo. Estos datos se representaron en un mapa ráster, y sobre cada categoría se superpusieron los puntos de muestreo correspondientes para determinar la cantidad de puntos que coincidían tanto en el terreno como en el mapa ráster. Los resultados revelaron una coincidencia del 93,5% en la categoría de vegetación, 95% en suelo desnudo y 91% en la categoría de bosque.

Palabras clave: banda, cobertura del suelo, imágenes satelitales, ráster, similitud de categorías

Abstract

The research aimed to calculate the percentage of similarity between land cover identified in the field through sampling points and the corresponding categories in the raster map derived from unsupervised classification, using a Landsat 8 satellite image of the area of interest. A sample exceeding thirty-five sampling points was utilized, providing a confidence level of ninety percent and a ten percent margin of error. Images of the area of interest were downloaded from the United States Geological Survey (USGS), and a combination of the first seven bands was performed. Subsequently, this composition was merged with band eight to enhance spatial resolution, achieving a resolution of fifteen meters. Radiance and reflectance corrections were then applied. With the corrected images, band combinations 543 for vegetation, 764 for forest, and the Bare Soil Index (BSI) were conducted. These data were represented on a raster map, and sampling points were superimposed on each category to determine the number of points coinciding both in the field and on the raster map. The results revealed a 93.5% match in the vegetation category, 95% in bare soil, and 91% in the forest category.

Key words: band, land cover, raster, satellite images, similarity of categories

Introducción

La clasificación digital de imágenes satelitales se ha posicionado como una herramienta indispensable en la caracterización y evaluación de la cobertura del suelo, permitiendo la identificación y clasificación automatizada de diferentes tipos de cobertura a una escala espacial significativa. En este sentido, la aplicación de tecnologías avanzadas, como la captura de datos de satélites Landsat 8, proporciona una oportunidad única para analizar la similitud entre las categorías de cobertura del suelo.

La investigación se adentra en la metodología de clasificación no supervisada de imágenes satelitales, empleando la información espectral capturada por Landsat 8 para discernir las variadas categorías de cobertura del suelo, como suelo desnudo, vegetación y bosque. Este enfoque, apoyado por el uso de software especializado en el procesamiento y análisis de datos geoespaciales, busca explorar la coherencia y la precisión de la clasificación digital en comparación con la realidad del terreno.

Además, se enfoca en la crucial fase de verificación en campo, destinada a validar y perfeccionar la clasificación obtenida mediante la recolección de datos in situ. La utilización de técnicas de muestreo y puntos muestrales en el terreno brinda una valiosa oportunidad para contrastar y calibrar la información derivada de las imágenes satelitales, asegurando así la exactitud y la confiabilidad de los resultados obtenidos.

En este contexto, la investigación se presenta como un aporte sustancial al campo de la teledetección y la

cartografía, proporcionando un análisis detallado y académicamente fundamentado sobre la similitud de cobertura del suelo entre datos satelitales y observaciones en campo, con implicaciones significativas para la gestión eficiente de los recursos naturales y la toma de decisiones medioambientales.

Materiales y métodos

Área de estudio

La investigación se llevó a cabo en un área de 2,10 km² (210 ha) con un perímetro de 6,55 km, situada al norte de la ciudad de Celendín, región Cajamarca. Los límites georreferenciados en WGS84 mediante coordenadas UTM se presentan en la figura adjunta (Tabla 1 y Figura 1).

Tabla 1. Coordenadas del área de interés

Punto	Este	Norte
A	814707	9241870
B	817162	9241731
C	817141	9240873
D	814735	9240965



Figura 1. Área de interés

Población y muestra de estudio

La población de estudio comprende la superficie delimitada por los puntos indicados, garantizando que los resultados obtenidos sean aplicables exclusivamente a esta zona. Se determinó el tamaño de la muestra con un nivel de confianza del 90% y un error del 10%, requiriendo un mínimo de 35 puntos muestrales para las categorías evaluadas. Kongalton R. et al. (2019) sugiere tomar aproximadamente el 1% de la superficie cartografiada, empleando la fórmula:

$$n = \frac{p q Z^2}{E^2}$$

Donde Z es la abscisa de la curva normal para un nivel dado de probabilidad, p es el porcentaje de aciertos, q es el porcentaje de errores (q=1-p), y E es el nivel permitido de error.

Jean François Mas et al. (2003) propone una cantidad mínima de unidades de muestreo, que en este caso son puntos, relacionándolas con su ubicación en mapas. El método de muestreo utilizado fue aleatorio estratificado, tomando al menos 30 puntos para cada categoría en el terreno y ubicándolos en el mapa generado mediante clasificación no supervisada (Tabla 2).

Tabla 2. Tamaño de la muestra por clases en función de p y d

*	p				
	90%	80%	70%	60%	50%
2,5%	553	983	1291	1475	1537
5,0%	138	246	323	369	384
10,0%	35	61	81	92	96

p: confiabilidad estimada de la clase

Trabajo de campo

En la zona de estudio, se recolectaron más de treinta puntos para cada categoría evaluada utilizando receptores GPS. Se empleó un software de manejo de datos geográficos para llevar a cabo la clasificación no supervisada, generando un mapa temático con las correspondientes categorías.

Se combinaron las 7 bandas y luego se fusionaron con la banda 8 para mejorar la resolución de 30 m a 15 m, se convirtieron los valores de ND a radiancia, se combinaron las bandas B5, B4, B3 para identificar la vegetación. Para el suelo desnudo, se calculó el Índice BSI utilizando las bandas correspondientes. Para la categoría Bosque, se utilizó la combinación de las bandas B7, B6, B4. Después de las combinaciones, se llevó a cabo la clasificación no supervisada. Se utilizó una imagen de la zona descargada de Google Earth Pro y se realizó una clasificación no supervisada utilizando Corine Land Cover para la presentación de las coberturas.

Análisis de los datos

Después de fusionar las siete bandas, se procedió a combinarlas con la banda 8 con el objetivo de mejorar la resolución de las imágenes de 30 metros a 15 metros. Acto seguido, se llevó a cabo la conversión de los valores de ND a radiancia, seguida de la transformación a niveles de reflectancia. Con todas las imágenes ajustadas a niveles de reflectancia, se realizó una nueva combinación, utilizando las bandas B5, B4 y B3 con el propósito

de identificar las áreas de vegetación. Para determinar la categoría de suelo desnudo, se calculó el Índice BSI mediante la calculadora ráster y las bandas correspondientes.

En cuanto a la categoría de bosque, se empleó la combinación de las bandas B7, B6 y B4. Tras realizar estas combinaciones específicas, se procedió a llevar a cabo la clasificación no supervisada, utilizando un enfoque que permitió discernir automáticamente las diferentes categorías de cobertura del suelo.

Adicionalmente, se incorporó una imagen de la zona descargada de Google Earth Pro en el análisis. Se efectuó una clasificación no supervisada utilizando Corine Land Cover con el propósito de presentar y validar las diversas coberturas identificadas en el área de estudio. Este enfoque multifacético garantizó una evaluación exhaustiva y precisa de la cobertura del suelo en la región de interés.

Resultados y discusión

En la Clasificación no Supervisada, categoría Vegetación, de los 46 puntos registrados en campo al superponer sobre el resultado ráster solo se encontró tres desaciertos (Figura 2).

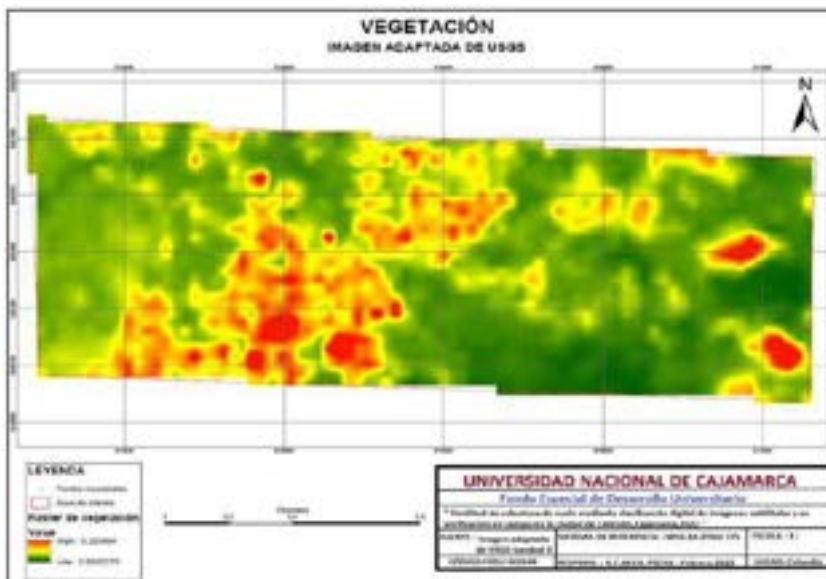


Figura 2. Ráster de la categoría Vegetación

En la Clasificación no Supervisada, categoría Suelo Desnudo, de los 40 puntos registrados en campo al superponer sobre el resultado ráster solo se encontró dos desaciertos (Figura 3).

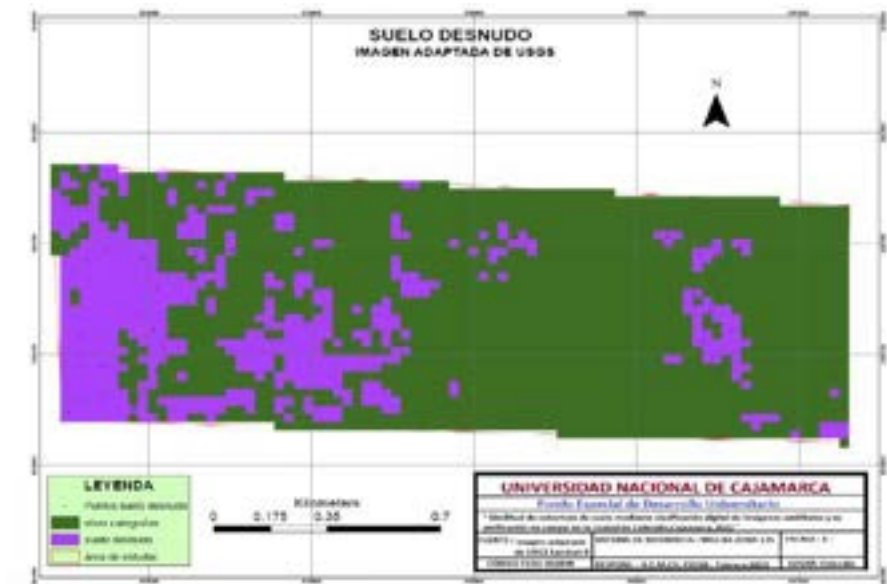


Figura 3. Ráster de la categoría Suelo Desnudo

En la Clasificación no Supervisada, categoría Bosque, de los 45 puntos registrados en campo al superponer sobre el resultado ráster solo se encontró cuatro desaciertos (Figura 4).

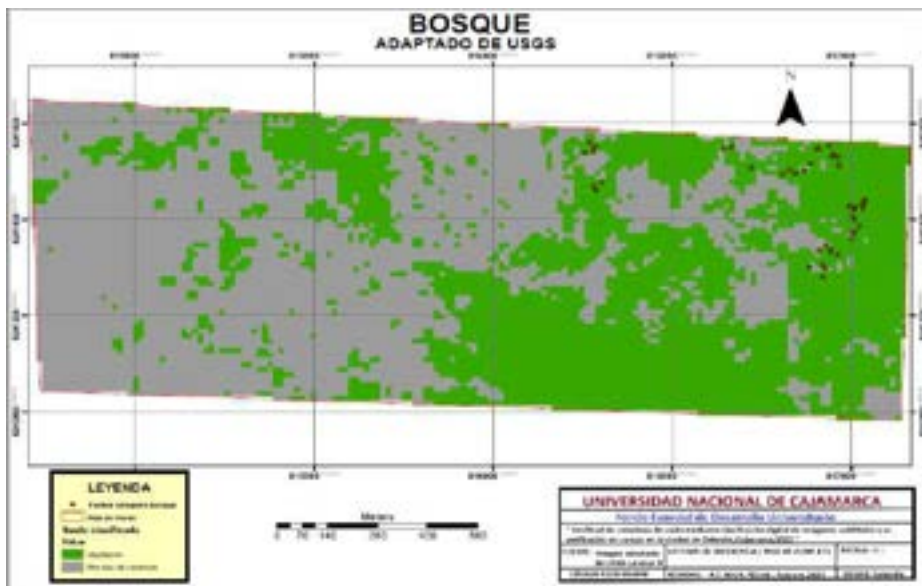


Figura 4. Ráster de la categoría Bosque

De la misma manera, ante los resultados presentados en las figuras, se pudo elaborar la siguiente tabla resumen:

Tabla 3. Resumen de los resultados obtenidos

Categoría	Número de puntos muestrales	Porcentaje similitud	Nº puntos con aciertos	Confiabilidad estimada en%	Porcentaje de error
Vegetación	46	93,5	43	90	10
Suelo Desnudo	40	95	38	90	10
Bosque	45	91	41	90	10

Jean François Mas (2003) señala discrepancias significativas entre la clasificación de suelos realizada a través de fotografías aéreas y la cobertura verificada en campo. Estas discrepancias superan el 20%, y un 23% de las áreas no se corresponden con la cobertura evaluada en el terreno, resaltando la complejidad y los desafíos asociados con la clasificación de suelos a partir de fuentes remotas.

Por otro lado, Illán-Fernández et al. (2022) emplea la clasificación Corine Land Cover y reporta un prometedor 89% de similitud utilizando la matriz de confusión. Sin embargo, identifican limitaciones al utilizar la combinación de bandas mediante clasificación no supervisada, específicamente en la incapacidad para ofrecer una separación detallada de los diferentes tipos de vegetación. Similarmente, la clasificación de bosques carece de una combinación precisa que permita una distinción clara entre áreas boscosas y aquellas que no lo son. En respuesta a estas limitaciones, optaron por utilizar un índice específico para la identificación de suelos desnudos, permitiendo así una separación más efectiva de esta categoría.

De manera alternativa, en la misma área de estudio, llevaron a cabo una clasificación no supervisada utilizando una imagen de mayor resolución espacial descargada de Google Earth Pro. A pesar de tener solo tres bandas, los resultados mejoraron considerablemente, y la clasificación se realizó utilizando la simbología de Corine Land Cover. Este hallazgo subraya la importancia de la resolución espacial, indicando que a medida que esta aumenta, la calidad de la clasificación digital de las coberturas del suelo se ve notablemente mejorada.

En resumen, tanto las observaciones de Mas como las conclusiones de Illán-Fernández et al. subrayan la relevancia de considerar cuidadosamente la fuente de datos, la resolución espacial y las metodologías empleadas al clasificar la cobertura del suelo, destacando la necesidad de enfoques más refinados y precisos para una caracterización más exacta y fiable de las coberturas en el entorno estudiado.

Conclusiones

En el porcentaje de similitud de la cobertura de suelo encontrado en la categoría vegetación fue de 93,5%, en

suelo desnudo 95% y en bosque 91% encontrándose en promedio 93% de similitud comparado con la clasificación no supervisada usando imágenes Landsat 8 haciendo composición de bandas.

Referencias

Chuvieco Salinero, E. (2006). Teledetección Ambiental: La observación de la Tierra desde el Espacio.

Chuvieco Salinero, E., & Sancho Comíns, J. (1996). Tratamiento visual y digital de las imágenes espaciales: Aplicaciones docentes. *Didáctica geográfica*, 14, 17-28.

Chuvieco, E. (1996). *Fundamentos de teledetección espacial*, Tercera Edición Revisada. Ediciones RIALP, S.A. Madrid.

Congalton, R. (1988). A Comparison of Sampling Schemes Used in Generating Error Matrices for Assessing the Accuracy of Maps Generated from Remotely Sensed Data. *Photogrammetric Engineering and Remote Sensing*.

Flores Cesáreo, J. C., Bustamante González, A., Vargas López, S., Cajuste, L., Escobedo, F. J., & Valadez Ramírez, M. (2020). Cartografía del uso del suelo en la subcuenca Huaquechula, Puebla, México, con un índice combinado de imágenes de satélite. *Investigaciones geográficas*, 101.

García Rodríguez, M. del P., Sanz Donaire, J. J., Pérez González, M. E., & Navarro Madrid, Á. (2012). *Guía práctica de teledetección y fotointerpretación*. Universidad Complutense de Madrid.

Illán-Fernández, E., Pérez-Morales, A., & Romero-Díaz, A. (2022). Fiabilidad en la detección de las superficies selladas empleando datos del programa Copernicus | *Boletín de la Asociación de Geógrafos Españoles*.

Jean Francois Mas. (2003). Monitoring land-cover changes: A comparison of change detection techniques. *International Journal of Remote Sensing*, 20(1), 139-152.

Kogut. (2022). Los Landsat orbitan alrededor de la Tierra en órbita circular.

Kongalton R., Russell G., & Kass Green. (2019). *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, Third Edition. Routledge & CRC Press.

Robinove, C. J. (1981). The logic of multispectral classification and mapping of land. *Remote Sensing of Environment*, 11, 231-244.