



Clasificación supervisada de cobertura de suelos en la cuenca del Río Grande usando imágenes satelitales, Celendín, Cajamarca, 2023

Supervised classification of land cover in the Río Grande basin using satellite imagery, Celendín, Cajamarca, 2023

Agustín Medina Chávez¹, Leyla Guadaly Díaz Cadenillas¹, Rocío Elizabeth Mori Sánchez¹, José Luis Vera Díaz¹

¹Facultad de Ingeniería, Universidad Nacional de Cajamarca, Cajamarca, Perú

Resumen

Se hizo una categorización preliminar supervisada de cobertura de suelos en la cuenca de Río Grande, tomando en cuenta las categorías que mejor podían distinguirse en las imágenes Landsat 8 descargadas de USGS (United States Geological Survey) de la zona de interés. Se usó un tamaño muestral de puntos de verificación mayor a cuarenta lo que da un grado de confianza del noventa por ciento y un error de diez por ciento. Los puntos fueron tomados en campo con un navegador GPS para cada una de las categorías. Se recorrió la cuenca y tomó puntos, para cada categoría, en los lugares donde actualmente se presenta esas categorías. Con las imágenes ya corregidas se realizó la combinación de bandas 543 para vegetación, 652 para suelo agrícola y el índice (BSI) de suelo desnudo para la misma categoría. A la clasificación supervisada realizada sobre un mapa ráster se superpuso los puntos tomados en campo encontrándose para vegetación 99%, para suelo desnudo 51.67 % y para suelo agrícola 75% con 70, 60, 114 puntos respectivamente para cada una de las categorías.

Palabras clave

Ráster, banda, USGS

Abstract

A preliminary supervised categorization of land cover in the Rio Grande basin was made, taking into account the categories that could best be distinguished in the Landsat 8 images downloaded from USGS (United States Geological Survey) of the area of interest. A sample size of verification points greater than forty was used, which gives a confidence level of ninety percent and an error of ten percent. The points were taken in the field with a GPS navigator for each of the categories. The basin was toured and points were taken, for each category, in the places where those categories currently occur. With the already corrected images, the combination of bands 543 for vegetation, 652 for agricultural soil and the bare soil



index (BSI) for the same category was carried out. The points taken in the field were superimposed on the supervised classification carried out on a raster map, finding 99% for vegetation, 51.67% for bare soil and 75% for agricultural soil with 70,60,114 points respectively for each of the categories.

Keywords

Raster, band, USGS

Introducción

La ausencia de información sobre la cobertura de suelos, por diversas circunstancias, los gobiernos locales y los nacionales descuidan invertir en el conocimiento sobre el uso y cobertura de suelos en los territorios que administran a fin de tomar las mejores decisiones en beneficio de la población. Las tecnologías para recopilar información geográfica han tenido grandes avances en especial en geomática. El Perú tiene hasta un satélite para incrementar el conocimiento sobre diferentes aspectos territoriales, pero no se lo está aprovechando adecuadamente. Este trabajo es un primer intento y constituye, dentro de la clasificación supervisada, la fase de entrenamiento en donde se va comparando las categorías que se conocen de la zona, las cuales son determinadas en un software de manejo de datos geográficos, y a su vez se toma puntos muestrales en cada categoría para ver el porcentaje de coincidencias verificando de esta manera, mediante el software, el porcentaje de coincidencia. No hay trabajos de este tipo que hayan sido realizados en esta cuenca. El Gobierno Regional de Cajamarca ha hecho una clasificación a nivel macro, pero hay fuertes discrepancias cuando se realiza la verificación a nivel de cuenca o a escala local. Esto ha motivado a recorrer la cuenca en altitudes altas, medias y bajas con respecto al recorrido del Río Grande, que atraviesa casi toda la cuenca de sur a norte y tomándose puntos para cada categoría y comparando con las categorías determinadas a partir de la clasificación supervisada. Servirá de base para seguir haciendo más investigaciones en esta cuenca y superando las limitaciones sobre imágenes de alta calidad en resolución espacial y temporal.

Material y métodos

Ubicación del área de estudio

La zona de estudio será en un área de 9010 ha y un perímetro 47.51 Km, ubicada en Celendín, región Cajamarca. Los límites del área de estudio se indican en los mapas que se



indican. La ciudad de Celendín es la capital de la provincia de Celendín y se ubica en el distrito de Celendín, región Cajamarca.

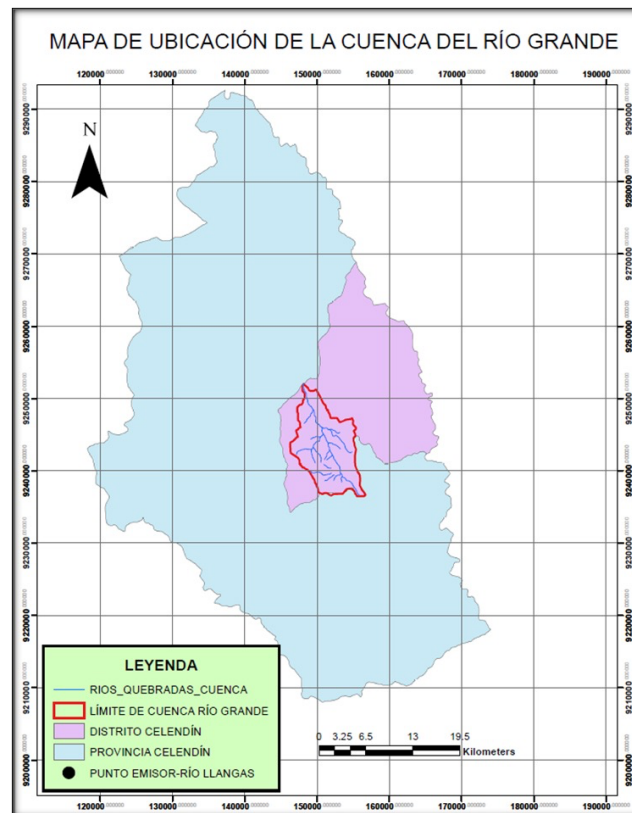


Figura 1. Mapa de ubicación de la cuenca del Río Grande.

Población

La población de estudio es la superficie encerrada dentro del perímetro de la cuenca del Río Grande.

Muestra

El tamaño de la muestra se tomó con una confianza del 90% y un error del 10% para lo cual se requirió tomar 35 puntos muestrales como mínimo para las categorías que se evaluaron.

En cuanto al tamaño de la muestra Kongalton R. et al. (2019), sugiere tomar una superficie aproximada al 1% de la superficie cartografiada y como se trata de medir una variable binomial (acierto-error), normalmente, se emplea.

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

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



Jean François Mas et al (2003) nos hace un resumen de la cantidad de cantidad de unidades de muestreo, que para nuestro caso son puntos que se debe toma en el terreno y se contrasta con su ubicación sobre los correspondientes mapas. El método de muestreo para el presente trabajo es aleatorio estratificado, pues se tomaron para cada una de las categorías 30 puntos como mínimo en el terreno y se ubicó en el mapa elaborado mediante clasificación no supervisada.

El mismo autor nos presente el siguiente cuadro que relacionan los elementos de la fórmula antes citada.

Tabla 1

Indicadores de crecimiento de los cuyes reproductores machos, evaluados luego del destete y por 10 semanas

	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

*medio intervalo de confianza

Nuestro trabajo se aproxima a $P=90\%$ y $E= 10\%$ unas aproximadamente 35 unidades muestrales.

Metodología

El estudio se desarrolló mediante tres fases principales: trabajo de campo, trabajo de gabinete y tratamiento de datos.

En la fase de trabajo de campo, se recolectaron más de treinta puntos georreferenciados por cada una de las categorías evaluadas dentro del área de estudio, utilizando receptores GPS de precisión. Estos puntos sirvieron como insumo para la validación y entrenamiento de la clasificación de imágenes satelitales.

Posteriormente, en la fase de gabinete, se empleó un software especializado en sistemas de información geográfica para realizar una clasificación supervisada de las coberturas. Como resultado, se generó un mapa temático que representó las distintas categorías analizadas.



Previamente, las imágenes satelitales fueron sometidas a un conjunto de preprocesamientos. En primer lugar, se realizó la combinación de las siete bandas multiespectrales con la banda pancromática (banda 8), con el objetivo de mejorar la resolución espacial de 30 m a 15 m mediante técnicas de fusión de imágenes. Seguidamente, los niveles digitales (ND) fueron convertidos a radiancia y posteriormente a reflectancia, con el fin de estandarizar los valores espectrales y permitir comparaciones consistentes.

Una vez obtenidas las imágenes en niveles de reflectancia, se procedió a la combinación de bandas B5, B4 y B3 para la identificación de coberturas vegetales. Para la determinación de áreas de suelo desnudo, se calculó el Índice de Suelo Desnudo (BSI) utilizando la calculadora ráster y las bandas espectrales correspondientes. Asimismo, la categoría de suelo agrícola fue identificada mediante la combinación de bandas B7, B6 y B4, lo cual permitió resaltar características específicas de este tipo de cobertura.

Finalmente, en la fase de tratamiento de datos, los puntos muestrales obtenidos en campo fueron superpuestos sobre el mapa temático generado. Se evaluó la concordancia entre la clasificación y los puntos de referencia, contabilizando aquellos que coincidían o no con su categoría correspondiente. A partir de esta comparación, se determinó el porcentaje de coincidencia, lo cual permitió estimar la precisión del proceso de clasificación.

Resultados y discusión

Clasificación supervisada categoría Vegetación

De los 46 puntos registrados en campo al superponer sobre el resultado ráster solo se encontró tres desaciertos.

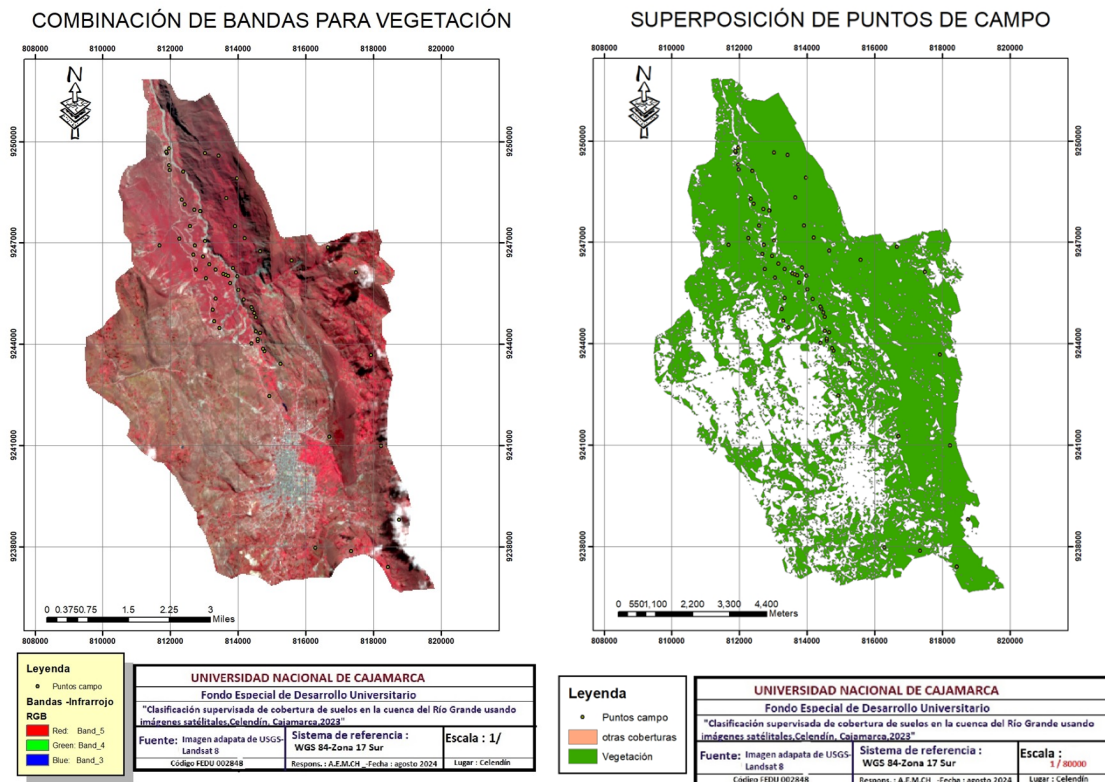


Figura 2. Ráster categoría Vegetación.

Clasificación supervisada categoría Suelo desnudo

De los 40 puntos registrados en campo al superponer sobre el resultado ráster solo se encontró dos desiertos.

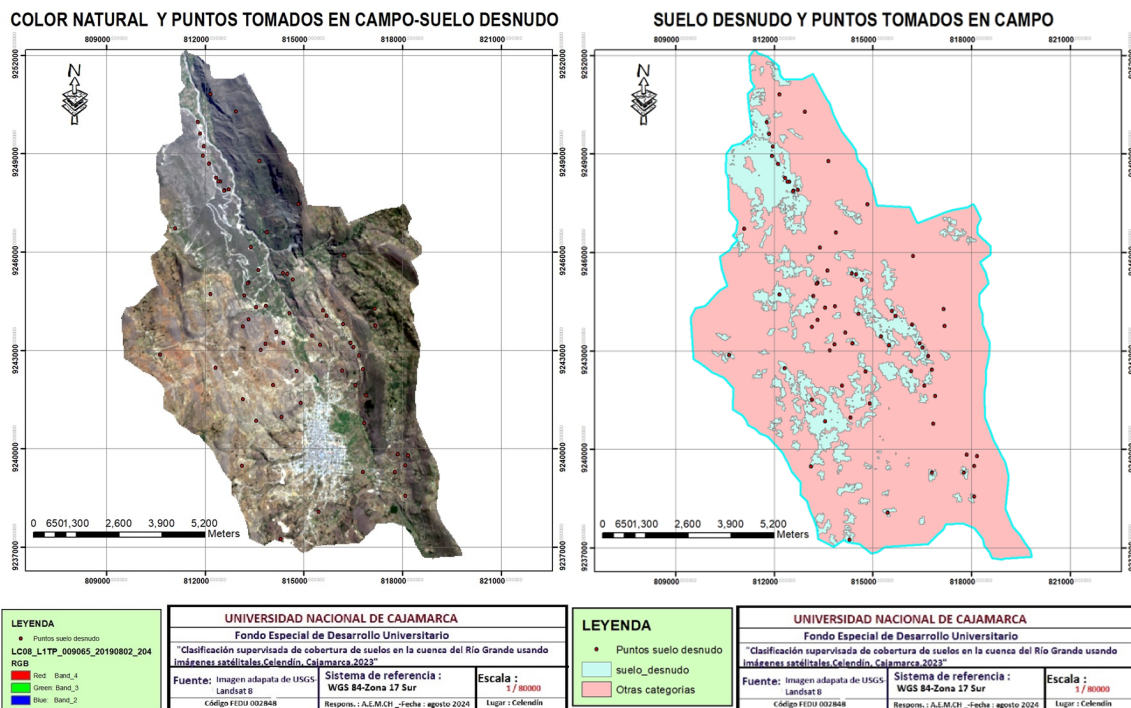


Figura 3. Ráster categoría Suelo desnudo.

Clasificación supervisada categoría Suelo agrícola

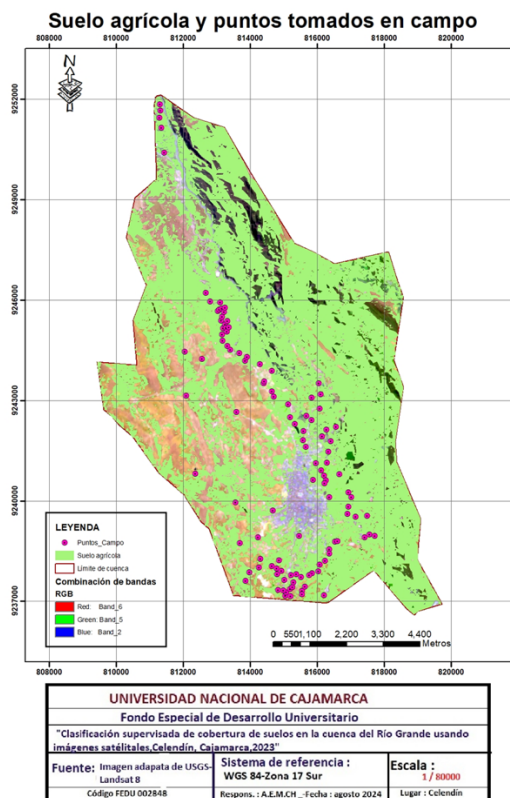


Figura 3. Ráster categoría Suelo agrícola.

**Tabla 2**

Resumen de los resultados obtenidos

Categoría	Número de puntos muestrales	Porcentaje similitud	Nº puntos con aciertos	Confiabilidad estimada en %	Porcentaje de error
Vegetación	70	99	69	90	10
Suelo desnudo	60	51.67	31	90	10
Suelo agrícola	114	75	86	90	10

Mollan y Bazán (2015) nos mencionan las desventajas de la clasificación supervisada que pueden conducir a errores en la clasificación supervisada los cuáles dependen de la calidad de las imágenes teniendo en cuenta la resolución espectral y espacial. En este mismo aspecto, Jean François Mas et al (2003) nos menciona sobre otros aspectos que tienen influencia en la clasificación como son la fecha de las imágenes utilizadas y la fecha de verificación en campo que ocasiona otro problema que se debe tener en cuenta la variabilidad del uso del suelo.

Teniendo en cuenta estos aspectos que son transversales a los problemas de verificación de imágenes han ocurrido o pueden ocurrir en esta etapa preliminar de entrenamiento y verificación en cualquier trabajado de categorización de suelos usando imágenes. En base a esto podemos arribar a las siguientes conclusiones para el presente trabajo.

Conclusiones

- En base a las imágenes Landsat descargadas en forma libre de USGS con una resolución espacial de 30m x 30m se hizo la clasificación solo en tres categorías: Vegetación, suelo desnudo y suelo de uso agrícola y por la calidad de las imágenes no hizo otras.
- Se delimitó la cuenca hidrográfica del Río Grande.
- Se usó imágenes del Landsat 8 y las combinaciones de bandas para obtener cada uno de los mapas temáticos para cada categoría.
- Se validó la clasificación supervisada con un número de puntos mayor a 40 encontrándose los por porcentajes de coincidencia : Para cobertura vegetal 99%, para suelo desnudo 51.67% y para suelo agrícola 75%.



Referencias

- Bacaro, A. y Pérez, M. (2020). Clasificación de coberturas de suelo de la Cuenca del Río Tunuyán Superior. Ediciones. INTA.
https://www.researchgate.net/publication/341608942_Clasificacion_de_coberturas_de_suelo_de_la_Cuenca_del_Rio_Tunuyan_Superior.
- Chuvieco, E. (1996). Fundamentos de teledetección espacial (3era edición). Editorial Rialp.
- Diccionario SIG de ESRI Support. <https://support.esri.com/es-es/gis-dictionary/band>.
- Ferrelli, F., Brendel, A., Perillo, G. M. E., & Piccolo, M. C. (2020). Validación de productos satelitales a partir de mediciones in situ para el monitoreo de coberturas del suelo en el sur de la Región Pampeana (Argentina).
- García Floriano, A. (2007). Modelo de clasificación supervisada. Obtenido de <https://www.scielo.org.mx/pdf/cys/v23n4/1405-5546-cys-23-04-1619.pdf>
- Gimenez Lozano, Y. (2010). Clasificación no supervisada: El método de k-medias.
- Gobierno Regional de Cajamarca (2010). Zonificación Ecológica y Económica-Base para el ordenamiento territorial de Cajamarca.
https://zeeot.regioncajamarca.gob.pe/sites/default/files/Libro_Oficial_ZEE_lite.pdf.
- Kongalton R., Russell G., & Kass Green. (2019). Assessing the Accuracy of Remotely Sensed Data: Principles and Practices, Third Edition. Routledge & CRC Press.
<https://www.routledge.com/Assessing-the-Accuracy-of-Remotely-Sensed-Data-Principles-and-Practices/Congalton-Green/p/book/9780367656676>.
- Lazo Pillaga, L. R. (2014). Evaluación de métodos de clasificación basados en píxeles y en objetos para la obtención de mapas de uso y cobertura del suelo a partir de imágenes satelitales Rapideye (Master's thesis, Universidad del Azuay).
- Mollan Velásquez, L., & Bazan Velásquez, H. (2015). Clasificación no supervisada de coberturas vegetales. Tesis. ¿Recuperado el 01 de 03 de 2024, de <https://repositorio.unal.edu.co/bitstream/handle/unal/34075/21276-72316-1-PB.pdf?sequence=1&isAllowed=y>.
- Muguerza Rivero, J. (2011). Aportaciones a la clasificación no supervisada. Tesis. ¿Recuperado el 01 de 03 de 2024, de <https://addi.ehu.es/bitstream/handle/10810/13910/2010%20Gurrutxaga%20I.pdf?sequence=1>.
- NASA. (2013). LANDSAT 8. Recuperado el 01 de 03 de 2024, de <https://landsat.gsfc.nasa.gov/satellites/landsat-8/>



- Rojas Briones, K., & Velásquez Mollan, L. (2000). clasificación no supervisada de coberturas vegetales con indicadores sensores remotos: Landsat 8. Tesis. Recuperado el 01 de 03 de 2024, de http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0304-28472005000100004.
- Silva Abanto, F. (2019). Clasificaciones no supervisada y supervisada aplicada a una imagen satelital Landsat-8 de la zona del PNN Sierra Nevada de Santa Mart. Tesis. Recuperado el 01 de 03 de 2024, de https://rstudio-pubs-static.s3.amazonaws.com/538955_4d16a27af0414388ab9f202af8ab80d1.html Tesis. Recuperado el 01 de 03 de 2024, de 68 https://cms.dm.uba.ar/academico/carreras/licenciatura/tesis/2010/Gimenez_Yani_na.pdf
- Williamson, C. E., Saros, J. E., Vincent, W. F., & Smol, J. P. Lakes and reservoirs as sentinels, integrators, and regulators of climate change. *Limnology and Oceanography*, vol. 54(6part2), p. 2273- 2282, 2009. https://doi.org/10.4319/lo.2009.54.6_part_2.2273. <https://studylib.es/doc/8967224/combinaci%C3%B3n-rgb-de-bandas-para-landsat-8>