UNA UNIVERSIDAD NACIONAL COSTA RICA

TRANSFORMA-INNOVA

INISEFOR

PROYECTO REDD+ COSTA RICA Pagos Basados en Resultados

GREEN CLIMATE FUND

PNUD

RESEÑA TÉCNICA:

CLASIFICACIÓN Y AUTOMATIZACIÓN DE CULTIVO DE CAFÉ Y COBERTURA ARBÓREA EN ZONA DE LOS SANTOS DE COSTA RICA

2020-2023

CONTENIDO

Agradecimientos	**05**
Introducción	**06**
Esquema metodológico	**08**
Datos utilizados	**08**
Implementación de algoritmo de clasificación de imágenes	**09**
Resultados	**11**
Conclusiones	**15**
Referencias	**16**

CRÉDITOS

Comité Directivo del proyecto
Sandra Sosa Cárcamo, Representante Residente, Programa de Naciones Unidas para Desarrollo (PNUD).
Franz Tattenbach Capra, ministro de Ambiente y Energía de Costa Rica.
Pascal Girot Pignot, director de la Escuela de Geología de la Universidad de Costa Rica.

Coordinación
Elena Vargas Fonseca, oficial del Programa Naturaleza, Clima y Energía, PNUD Costa Rica.
Maureen Ballestero Vargas, coordinadora del Proyecto REDD+ Pagos Basados en Resultados (PBR), PNUD Costa Rica.
Jairo Serna Bonilla Coordinación componente PNUD Programa TRANSFORMA INNOVA

Equipo de trabajo
Cristian Aguilar Barboza, consultor, Universidad Nacional.
Mauricio Vega Araya, Universidad Nacional.
Karina Sevilla Segura, Universidad Nacional.
Nicole Obando Morán, Universidad Nacional.
Marco Martínez Martínez, Programa de las Naciones Unidas para el Desarrollo (PNUD).
Luis Diego Monge Chinchilla, Instituto del Café de Costa Rica.
Yerlin Vargas Solano, Instituto del Café de Costa Rica.

Edición del documento
Maureen Godínez Chacón, asesora de Gestión de Conocimiento.

Comité Editorial de PNUD
José Daniel Estrada Sánchez, especialista de Monitoreo y Evaluación.
Rafaella Sánchez Mora, especialista de Género.
Charleene Cortez Sosa, especialista de Gestión de Conocimiento.
Glomara Iglesias Álvarez, especialista de Comunicación.

Fotografía
Proyecto REDD+ Pagos Basados en Resultados, PNUD.

Diseño y diagramación
Maricela Venegas Salas, especialista de Diseño Gráfico y Comunicación Visual.
Primera edición: (Febrero, 2025)

Programa de las Naciones Unidas para el Desarrollo [PNUD]. (2025) Reseña Técnica: clasificación y automatización de cultivo de café y cobertura arbórea en la Zona de Los Santos en Costa Rica. Costa Rica.
Está autorizada la reproducción total o parcial de esta publicación con propósitos educativos y sin fines de lucro, siempre que se utilice la referencia respectiva. Para el uso no se requiere ningún permiso especial del titular de los derechos. Este material se encuentra disponible en **https://pnud-conocimiento.cr**

Este documento se encuentra protegido por la Licencia Creative Commons CC BY-SA 4.0 HYPERLINK "https://creativecommons.org/licenses/by-sa/4.0/deed.es"Atribución/Reconocimiento-CompartirIgual-no comercial 4.0 Internacional a la siguiente dirección o correo:

Programa de las Naciones Unidas para el Desarrollo, Costa Rica.
Edificio Sigma, Edificio B, Piso 3, San Pedro de Montes de Oca, Costa Rica
Teléfono: (506) 2296-1544
Web: https://www.undp.org/costa-rica
Email: registry.cr@undp.org / comunicaciones.cr@undp.org

Derechos de propiedad intelectual: Programa de las Naciones Unidas para el Desarrollo (PNUD-Costa Rica) © 2025

AGRADECIMIENTOS

Las personas autoras desean agradecer el financiamiento del Proyecto TRANSFORMA/INNOVA y al Proyecto REDD+ Pagos Basados en Resultados 2014-2015, al Programa de las Naciones Unidas para el Desarrollo (PNUD), al Instituto del Café de Costa Rica (ICAFE), y a CoopeTarrazú R.L. por el apoyo técnico y de información en el proceso del estudio presentado.

INTRODUCCIÓN

El Sistema de monitoreo de cambios en el uso de la tierra en paisaje productivos (MOCUPP) es un esquema de trabajo conjunto en el cual interactúan sectores productivos, estado, academia y cooperación internacional, que genera información, periódica, independiente, y técnicamente sólida, sobre cultivos estratégicos (Piña, pastos, palma, musáceas y café) y áreas urbanas.

El sistema emplea técnicas de teledetección [1]desempeñan un papel crucial en la clasificación de las plantaciones de café y la evaluación de sus características en zonas tropicales. Un método común involucra el uso de imágenes satelitales para clasificar los tipos de cobertura terrestre, que incluye las plantaciones de café. Mediante el análisis espectral y temporal es posible diferenciar entre diversas clases de cobertura como bosques, cultivos y plantaciones de café. Estas últimas son particularmente desafiantes debido a que, mayormente, están organizadas en configuraciones con sombra y en ambientes de altas pendientes. Estudios como el de Asner et al. (2006) han demostrado la efectividad de la teledetección para delinear con precisión los límites de las plantaciones de café y monitorear su extensión espacial a lo largo del tiempo.

Además, la teledetección permite evaluar parámetros biofísicos de las plantas de café, como las texturas, que proporcionan información valiosa sobre su condición fisiológica y facilitan la implementación de patrones de manejo, como las podas. La integración de información espaciotemporal a partir de imágenes satelitales o aéreas permite desarrollar modelos para estimar variables biofísicas clave, lo que mejora la clasificación y el monitoreo de las plantaciones de café en regiones tropicales (Paz-Kagan et al., 2018).

En particular, el uso de teledetección de acceso abierto (tanto en datos como en herramientas) se alínea con las tendencias internacionales en este campo. Los procesos sostenibles de monitoreo, reporte y verificación se han transformado significativamente, mejorando la toma de decisiones basada en datos geoespaciales. Esta tecnología de acceso abierto permite que las partes interesadas en los procesos productivos utilicen los mismos datos para llevar a cabo sus análisis, por lo que promueve así una transparencia que es esencial en la actualidad. Dicha accesibilidad estimula la colaboración, la innovación y la transparencia al aprovechar los datos geoespaciales para tomar decisiones informadas (Shao et al., 2018).

Las últimas tendencias que complementan el uso de datos abiertos incluyen la integración de la inteligencia artificial (IA) y el aprendizaje automático, para mejorar la precisión y eficiencia de la recopilación y el análisis de datos geoespaciales. Estas innovaciones abordan los desafíos en la clasificación de coberturas terrestres complejas, como el café. Le et al. (2022) destacan el potencial de los modelos de aprendizaje profundo (AP) para la clasificación del café, al utilizar datos de teledetección. Este enfoque podría reemplazar la interpretación manual de imágenes y mejorar las técnicas de clasificación supervisadas y no supervisadas. En particular, este estudio representa una de las primeras aplicaciones de técnicas basadas en AP para la clasificación de cultivos de café en Centroamérica.

1 Teledetección: Es la técnica que obtiene datos observados de La Tierra a través de satélites o sensores remotos, por medio de la captura de las variaciones espectrales de las ondas radiomagnéticas (Chuvieco, E. 2002).

ESQUEMA METODOLÓGICO

El área que se determinó como área de estudio, corresponder a 29.03,7 hectáreas. Se basó en 117 puntos y 11 polígonos de fincas cafetaleras en la zona conocida como Los Santos donde se encuentra el cantón de Tarrazú, en San José, Costa Rica.

Este proceso se apoya en herramientas y fuentes de datos de última generación de MOCUPP[1] Café, financiado por la iniciativa alemana para el clima (IKI) que son completamente accesibles, sin costo alguno, lo que permite su procesamiento de manera libre y en la nube. Además, Además, es un proceso totalmente automatizado que elimina cualquier intervención manual por parte de técnicos, y reduce los sesgos asociados. Incluye, además, un componente de validación estadística riguroso, que sigue estándares internacionales para la clasificación de usos y coberturas del suelo basados en teledetección.

DATOS UTILIZADOS

La principal fuente de datos empleada proviene de las imágenes del programa Copernicus, de la unidad de Observación de la Tierra de la Agencia Espacial Europea (AEE). Específicamente, se utilizan los datos satelitales de Sentinel-1 y Sentinel-2. Sentinel-1 es un satélite equipado con Radar de Apertura Sintética (SAR por sus siglas en inglés), un sensor activo capaz de capturar información de la superficie terrestre incluso bajo condiciones de nubosidad (Flores-Anderson et al., 2019). Por otro lado, Sentinel-2 es un satélite que opera principalmente en el espectro óptico e infrarrojo, recolectando datos a través de una amplia gama de bandas espectrales[2].

Adicionalmente, se emplearon datos obtenidos de modelos de elevación digital (MED) para modelar variables topográficas claves en la segmentación de cultivos, tales como la pendiente y la elevación. La eficacia de estas fuentes de datos para el mapeo de cultivos ha sido reconocida en diversos estudios (Meng et al., 2023; Phan et al., 2020).

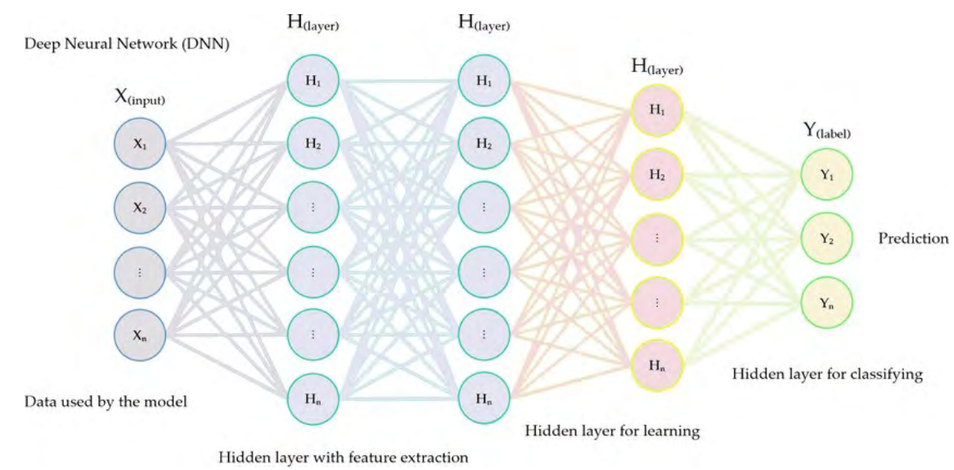
1 (Meng et al., 2023; Phan et al., 2020).
2 Para acceder a más información visitar los siguientes enlaces:
Sentinel-1: https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-1
Sentinel-2: https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-2
DEM: https://developers.google.com/earth-engine/datasets/catalog/NASA_NASADEM_HGT_ 001

IMPLEMENTACIÓN DE ALGORITMO DE CLASIFICACIÓN DE IMÁGENES

Como se mencionó en el presente estudio, se implementó un algoritmo de AP, específicamente una red neuronal densa (RND). Estas redes neuronales artificiales son altamente eficientes en aprender niveles complejos de abstracción y representación de datos, lo que facilita tareas de reconocimiento de patrones y clasificación con una eficiencia y precisión superiores a los enfoques tradicionales de aprendizaje automático. Recientemente, se han popularizado en el procesamiento de datos de teledetección, gracias al incremento de las capacidades computacionales en la nube. (Maskell et al., 2021).

En la Figura 1 se muestra un esquema de la arquitectura de una RND.

Figura 1. Arquitectura de una RND para clasificación.



Fuente: Prado et al., (2021).

Como todo proceso de clasificación de tipo supervisado, se requiere una serie de etapas principales, entre ellas: los datos de entrenamiento, la homogeneización de las covariables o datos de entrada (en este caso las imágenes satelitales y otros datos auxiliares), así como el proceso de validación.

1. Datos de entrenamiento

Se construyó un proceso de fotointerpretación con las clases semánticas de interés a clasificar: coberturas de café, cobertura arbórea y otras tierras. Para lo que se utilizó la plataforma Collect Earth Online, que es una herramienta eficiente y permite el trabajo de equipos de forma sincronizada. Para los años de interés 2020 y 2023 se definió de forma aleatoria un total de 9.000 muestras. Dichas muestras fueron sometidas a un proceso de verificación para comprobar su utilidad al entrenar la RND.

2. Preparación de covariables

Este punto incluye todas aquellas labores de preprocesamiento de los datos de entrada del modelo, que son todos aquellos productos derivados de las imágenes Sentinel-1, Sentinel-2 y datos topográficos, como lo son el enmascaramiento de nubes, creación de mosaicos espaciotemporales, computación de índices espectrales y de radar, productos de reducción de dimensión de datos, computación de bandas texturales, estandarización numérica, entre otras. En total, se computaron más de 80 covariables o datos de entrada, los cuales fueron sometidos a procesos de prueba y error para definir los que mejor ayudaron a discriminar el café de otras coberturas mediante una RND.

3. Testeo de la RND

Todo algoritmo de aprendizaje automático implica el ajuste de una serie de parámetros o hiperparámetros que permitan optimizar su rendimiento en términos computacionales y de precisión. Para este caso, se sometieron diversos escenarios de clasificación de la RND y se seleccionó el producto con mejor ajuste en términos de pérdida de entrenamiento y precisión de la red neuronal.

4. Postproceso automático

Uno de los efectos inherentes a un proceso de clasificación a partir de imágenes es el conocido efecto sal y pimienta (speckle). Para reducir dicho efecto se aplicó un algoritmo basado en objetos, el cual permitió definir la unidad mínima de mapeo para las fincas de café de 0.5 ha, tal y como se dispone en la metodología del Monitoreo del Cambio de Uso y Cobertura de la Tierra en Paisajes Productivos (MOCUPP). De esta forma se obtuvo un producto consistente y comparable entre los años de interés.

5. Validación

El proceso automático se sometió a un procedimiento de validación riguroso basado en lo planteado en Olofsson et al., (2014), el cual se ajusta a criterios o estándares internacionales para este tipo de procesos. Es importante tener en cuenta que la clase de café se sometió a una revisión con un nivel de confianza del 95%. De igual manera, la validación se desarrolló teniendo en cuenta la unidad mínima de mapeo de 0,5 ha.

RESULTADOS

El Cuadro 1 y el Cuadro 2 ilustran las matrices de confusión para los años 2020 y 2023, y detallan, entre otras cosas, la estimación de área en hectáreas (ha) para cada clase semántica junto con su respectiva incertidumbre. Esto representa el rango de variación posible del valor verdadero en torno a la estimación. Por ejemplo, para la clase de café, esta incertidumbre se traduce en que el área real puede variar en un ± 2,23% respecto a la estimada, lo que ofrece un nivel de confianza del 95% en las mediciones. Esta información es clave para evaluar la precisión de la clasificación y la fiabilidad de las estimaciones de área.

En el Cuadro 1, se muestran los resultados para el proceso de clasificación del año 2020, en el que se resaltan los altos porcentajes del productor del mapa (PP) en la evaluación del producto realizada. De tal manera, el parque cafetalero para el año 2020 en el área de estudio es de 14.210,7 ha, con un porcentaje de incertidumbre del 2,23%.

Cuadro 1

2020

PROPORCIONES EN CADA CLASE DE ESTRATO

Estrato	Muestras	Café	Cobertura arbórea	Otras tierras	Area (ha)	Incertidumbre (%)	PP(%)
Café	387	0,9884	0,0065	0,0052	14.210,67	2,23%	95,90%
Cobertura arbórea	195	0,0282	0,9667	0,0051	10.005,28	2,74%	98,20%
Otras tierras	117	2,0641	0,0171	0,9188	4.867,87	5,96%	97,30%
						Precisión global (PG%)	**96,90%**

Fuente: Elaboración propia, 2024.

muestran la distribución espacial del parque cafetalero en el área de estudio. La alta precisión global (PG) obtenida superó el 96% en ambos años, 2020 y 2023, lo que indica un grado de exactitud muy elevado en la clasificación de las imágenes. Además, las bajas incertidumbres para la clase de café, del 2,23% en 2020 y del 1,80% en 2023, refuerzan la confiabilidad del método específicamente para la identificación de plantaciones de café. En este aspecto, se resalta la resolución espacial (con pixeles de 10*10m), radiométrica y espectral de los datos, así como las características propias del contexto de trabajo (pendientes, complejidad de cultivo, entre otras).

Para 2023, y según el Cuadro 2, se obtuvo una mejor evaluación del proceso de clasificación en general, al alcanzar un 97,3% comparado con el 2020 de un 96,90%. Particularmente, el parque cafetalero 2023 es de 14.559,4 ha, con una incertidumbre del 1,80%.

Cuadro 2. Matriz de confusión de datos de clasificación 2023

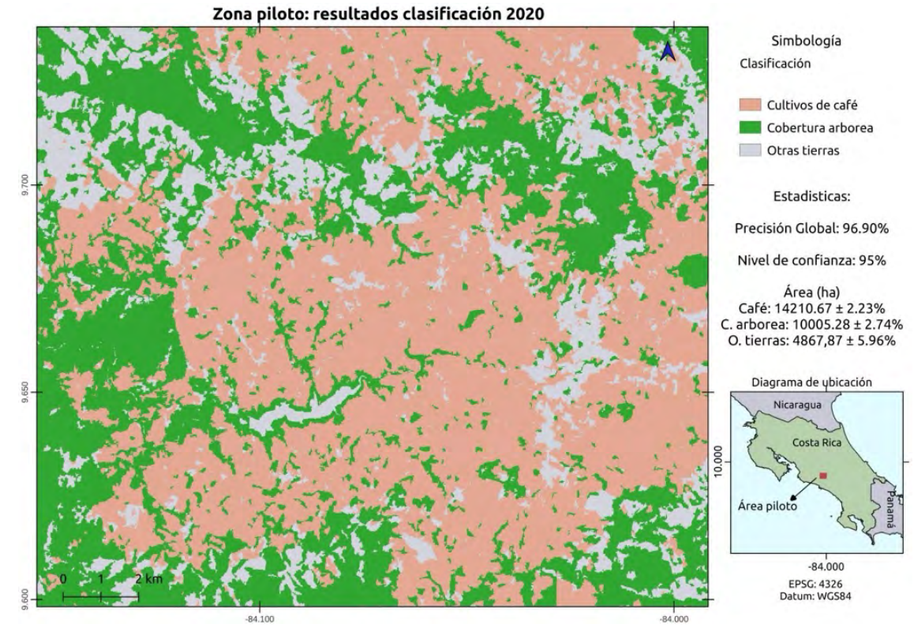
PROPORCIONES EN CADA CLASE DE ESTRATO

Estrato	Muestras	Café	Cobertura arbórea	Otras tierras	Area (ha)	Incertidumbre (%)	PP(%)
Café	392	0,9885	0,0026	0,0089	14.559,41	1,80%	97,70%
Cobertura arbórea	200	0,01	0,965	0,025	9.900,28	2,62%	99,20%
Otras tierras	114	0,0526	0,0088	0,9386	4.627,08	6,73%	92,00%
						Precisión global (PG%)	**97,30%**

Fuente: Elaboración propia, 2024.

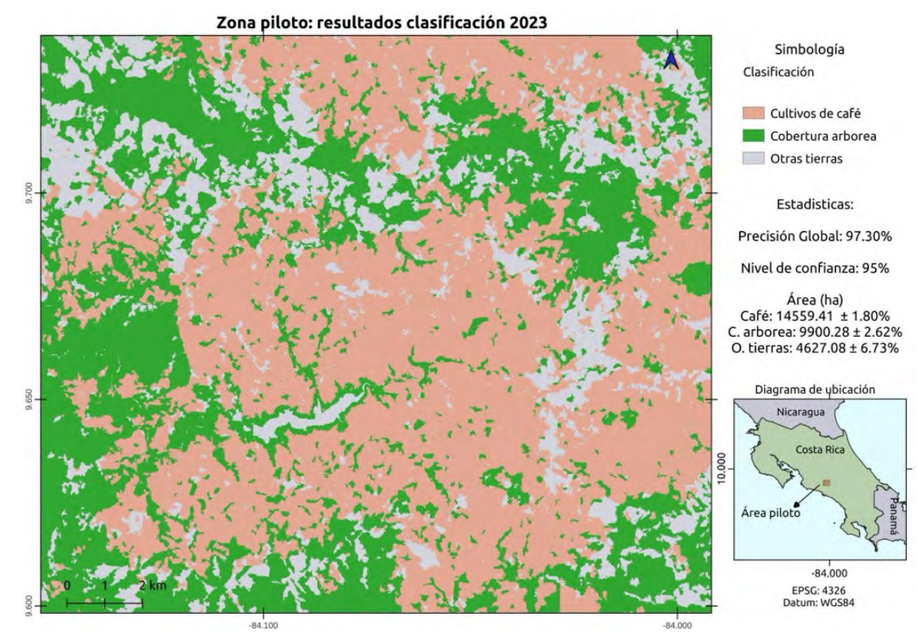
En las Figuras 2 y 3, se detalla la distribución espacial de los resultados de clasificación en forma de mapa pared a pared para el área de estudio del pilotaje.

Figura 2: Resultado de clasificación para 2020



Fuente: Elaboración propia, 2024.

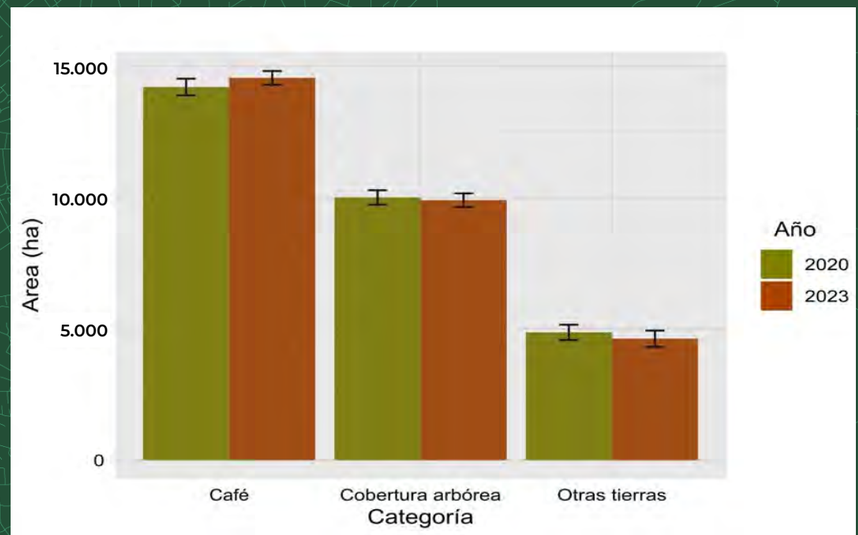
Figura 3: Resultado de clasificación para 2023.



Fuente: Elaboración propia, 2024.

En la Figura 4, se presenta la comparación de las hectáreas por categoría con sus respectivas barras de error, las cuales provienen de los Cuadros 1 y 2. Como se puede notar en esta figura, hay una leve tendencia a un aumento, para el año 2023 en la cobertura de la clase café y una disminución, para este mismo año, en la clase de Otras tierras. Sin embargo, esto se debe analizar, siempre, tomando en cuenta las barras de error respectivas. La cobertura arbórea, por su parte, muestra una tendencia a mantenerse estable, sin cambios significativos, para los años de análisis.

Figura 4. Comparación de áreas con intervalos de error por categoría en 2020 y 2023.



Fuente: Elaboración propia, 2024.

CONCLUSIONES

Este estudio muestra la efectividad de las técnicas de detección remota, particularmente el análisis de imágenes satelitales con consideraciones espectrales y temporales, para clasificar plantaciones de café en ambientes tropicales y condiciones de cultivo muy desafiantes. El uso de datos de Copernicus, específicamente de Sentinel (tanto 1 como 2) de acceso abierto, promueve la transparencia y la colaboración para el seguimiento del parque cafetalero.

La integración de técnicas de AP es prometedora para mejorar la precisión y eficiencia de las tareas de clasificación de usos y coberturas de la tierra.

La rigurosa metodología empleada en este estudio, con énfasis en herramientas avanzadas, automatización y validación estadística, arrojó resultados altamente precisos y confiables (más del 96% de precisión global) para la clasificación de plantaciones de café para los años de 2020 y 2023.

Basados en los datos obtenidos, se destaca una tendencia sutil observada en 2023, que indica un ligero aumento en la cobertura de café y una disminución correspondiente en la clase de otras tierras. Sin embargo, es crucial interpretar estos hallazgos con cautela, al considerar los errores asociados al proceso de clasificación y a la evaluación empleada. Esto subraya la importancia de un análisis exhaustivo para casos particulares, así como la respectiva verificación en campo de ciertos casos. Esto para reconocer la variabilidad de las posibles fuentes de error y con esto extraer conclusiones significativas en las fincas.

Este estudio destaca el potencial de la teledetección para procesos de clasificación automatizada para actividades de monitoreo, reporte y verificación. Con esto se sembrarán las bases para la detección de cambios en el uso y cobertura de la tierra en paisajes productivos.

REFERENCIAS

Asner, Gregory P., Eben N. Broadbent, Paulo J. C. Oliveira, and José N. M. Silva. 2006. Condition and fate of logged forests in the Brazilian Amazon. Edited by William C. Clark, Harvard University, Cambridge, MA. 103 (34) 12947-12950. En: https://doi.org/10.1073/pnas.0604093103

Baeta, R., Nogueira, K., Menotti, D., & Dos Santos, J. A. (2017). Chuvieco, E. (2002) Teledetección Ambiental. Editorial Ariel. Learning Deep Features on Multiple Scales for Coffee Crop Recognition. 2017 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 262–268. En: https://doi.org/10.1109/SIBGRAPI.2017.41

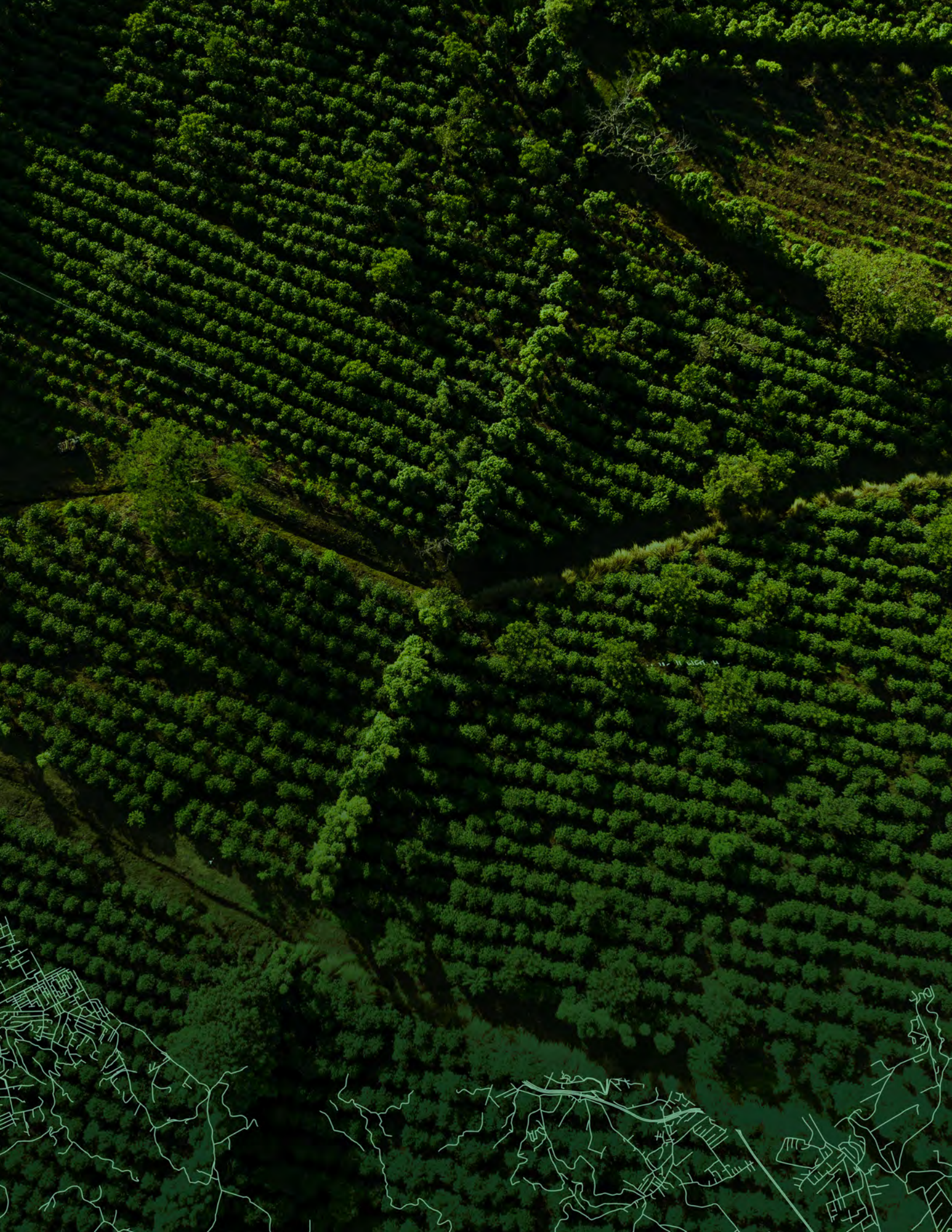
Flores, A., Herndon, K., Thapa, R., & Cherrington, E. (2019). Synthetic Aperture Radar (SAR) Handbook: Comprehensive Methodologies for Forest Monitoring and Biomass Estimation. En: https://doi.org/10.25966/NR2C-S697

Le, Q. T., Dang, K. B., Giang, T. L., Tong, T. H. A., Nguyen, V. G., Nguyen, T. D. L., & Yasir, M. (2022). Deep Learning Model Development for Detecting Coffee Tree Changes Based on Sentinel-2 Imagery in Vietnam. IEEE Access, 10, 109097–109107. En: https://doi.org/10.1109/ACCESS.2022.3203405

Korkovelos, A.; Khavari, B.; Sahlberg, A.; Howells, M.; Arderne, C. The Role of Open Access Data in Geospatial Electrification Planning and the Achievement of SDG7. An OnSSET-Based Case Study for Malawi. Energies 2019, 12, 1395. En: https://doi.org/10.3390/en12071395

Meng, X., Xie, S., Sun, L., Liu, L., & Han, Y. (2023). Evaluation of temporal compositing algorithms for annual land cover classification using Landsat time series data. International Journal of Digital Earth, 16(1), 2574–2598. En: https://doi.org/10.1080/17538947.2023.2230958

Phan, T. N., Kuch, V., & Lehnert, L. W. (2020). Land Cover Classification using Google Earth Engine and Random Forest Classifier—The Role of Image Composition. Remote Sensing, 12(15), 2411. En: https://doi.org/10.3390/rs12152411

Shao, W.; Hu, Y.; Yang, J.; Nunziata, F.; Sun, J.; Li, H.; Zuo, J. 2018. An Empirical Algorithm to Retrieve Significant Wave Height from Sentinel-1 Synthetic Aperture Radar Imagery Collected under Cyclonic Conditions. Remote Sens., 10, 1367. En: https://doi.org/10.3390/rs10091367

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., & Wulder, M. A. (2014). Good practices for estimating area and assessing accuracy of land change. Remote Sensing of Environment, 148, 42–57. En: https://doi.org/10.1016/j.rse.2014.02.015