



UNIVERSIDAD NACIONAL AUTÓNOMA DE MEXICO
PROGRAMA DE DOCTORADO EN GEOGRAFÍA

**EVALUACIÓN DEL USO DE IMÁGENES RADAR Y MULTIESPECTRALES CON
TÉCNICAS DE APRENDIZAJE PROFUNDO PARA MONITOREAR LA
DEFORESTACIÓN Y LA DEGRADACIÓN FORESTAL EN BOSQUES TROPICALES**

TESIS

**QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN GEOGRAFÍA**

PRESENTA:

JONATHAN VIDAL SOLÓRZANO VILLEGAS

TUTOR PRINCIPAL:

JEAN FRANÇOIS MAS

CENTRO DE INVESTIGACIONES EN GEOGRAFÍA AMBIENTAL

MIEMBROS DEL COMITÉ TUTOR

YAN GAO

CENTRO DE INVESTIGACIONES EN GEOGRAFÍA AMBIENTAL

JOSÉ ALBERTO GALLARDO CRUZ

CENTRO TRANSDISCIPLINAR UNIVERSITARIO PARA LA SUSTENTABILIDAD

MORELIA, MICHOACÁN, MÉXICO, OCTUBRE 2023



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradecimientos

Al Posgrado en Geografía y al Centro de Investigaciones en Geografía Ambiental (CIGA) de la Universidad Nacional Autónoma de México (UNAM), por darme la oportunidad de realizar un doctorado y fortalecer mi formación académica.

Al Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCyT) por otorgarme una beca de estudiante de tiempo completo para realizar mis estudios de posgrado.

Al Dr. Jean François Mas por integrarme al Laboratorio de Análisis Espacial y ayudarme a moldear este trabajo, y a los demás miembros de mi comité tutor, Dra. Yan Gao y Dr. J. Alberto Gallardo Cruz, por su atenta guía durante el desarrollo de esta tesis.

A los sinodales del examen de candidatura, Dr. Miguel Ángel Castillo Santiago y Dr. Adrián Ghilardi por la revisión de los avances de la tesis y por sus valiosas sugerencias.

A los demás sinodales del examen de grado, Dra. Hind Taud, Dr. Stéphane Couturier, Dr. José Luis Hernández Stefanoni por la revisión final de este trabajo, así como por sus comentarios que ayudaron a redondear mejor este documento y esbozar líneas de trabajo en el futuro.

Al proyecto de Ciencia Básica SEP-285349 “Análisis del patrón espacial de la degradación en selvas y bosques de México con percepción remota en múltiples escalas en el tiempo y espacio” del Consejo Nacional de Ciencia y Tecnología (CONACyT) por financiar parte de los datos registrados en la Cuenca del Río Ayuquila, Jalisco.

A la beca 14 DINV de la Universidad Iberoamericana por el financiamiento para cubrir el APC del artículo del tercer capítulo de la tesis.

Al proyecto PAPIME PE117519 “Herramientas para la enseñanza de la Geomática con programas de código abierto” de la Dirección General de Asuntos del Personal Académico de la Universidad Nacional Autónoma de México por los recursos para cubrir el APC del artículo del tercer capítulo de la tesis.

Al Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) de la Universidad Nacional Autónoma de México, proyecto PAPIIT IT103223 “Plataforma geoespacial para el análisis de los impactos ambientales por uso de leña y carbón vegetal a escala global” por el apoyo adicional para obtener el grado.

Al Programa de Apoyo a los Estudios de Posgrado (PAEP) por el apoyo brindado para asistir al XX Simposio Internacional SELPER en 2022 y presentar parte de los avances de esta tesis.

A René Tobar Díaz, Jaime Hernández y la Junta Intermunicipal del Río Ayuquila (JIRA) por su orientación y ayuda para realizar el trabajo de campo.

A mis ex compañeros del Centro del Cambio Global y la Sustentabilidad (CCGS), Alberto, Candelario, Ana, Matías, Rubén, Derio, Nikolay, Miguelina y Gabriela, por ayudarme directamente o indirectamente a generar parte de los datos que utilicé en esta tesis.

A Diana Jiménez Rodríguez por su interés y colaboración en torno al trabajo realizado en la Cuenca del Río Ayuquila.

Al Laboratorio Nacional de Análisis y Síntesis Ecológica (LANASE) y a los administradores de los servidores, Marco Villalba Quezada y Diego de Jesús Isla López, por la atención y las facilidades para realizar diversos análisis.

A mis papás Verónica y Vidal, y mis hermanos, Edgar y Jessy, por su cariño y constante apoyo.

A Valeria por embarcarnos en la aventura de realizar un posgrado juntos y su compañía en este proceso.

A Malta por forzarnos a levantarnos temprano y tomarnos descansos del trabajo.

A las personas con las que coincidí en Morelia y que han llenado de buenos ratos estos cuatro años, Fanny, Fer, Katia, Pau, Batún, Jaz, Jairo, Ileri, Alex, Diana, Haru, Ana, Geli, Piña, Karen A, Karla, Karen, Pacheco y Piña.

Índice

PRESENTACIÓN DEL TRABAJO	1
CAPÍTULO 1. INTRODUCCIÓN	4
CAMBIO GLOBAL: DEFORESTACIÓN Y DEGRADACIÓN FORESTAL	4
DEFINICIONES DE BOSQUE, DEFORESTACIÓN Y DEGRADACIÓN FORESTAL	7
USO DE LA PERCEPCIÓN REMOTA PARA EL ESTUDIO DE LA SUPERFICIE TERRESTRE	11
<i>Imágenes multiespectrales</i>	12
<i>Imágenes de radar de apertura sintética</i>	13
<i>Diferencias entre imágenes multiespectrales y SAR</i>	15
<i>Estudio de la deforestación y degradación mediante percepción remota</i>	16
<i>Programa Copernicus</i>	19
ALGORITMOS DE APRENDIZAJE AUTOMATIZADO	21
<i>Redes neuronales artificiales</i>	22
<i>Redes neuronales convolucionales</i>	25
<i>U-Net</i>	32
JUSTIFICACIÓN DEL TRABAJO	35
OBJETIVO GENERAL	36
OBJETIVOS PARTICULARES	36
CAPÍTULO 2. PATRONES ESPACIOTEMPORALES DE LAS OBSERVACIONES DE SENTINEL-2 A NIVEL DE IMAGEN Y PÍXEL SOBRE EL TERRITORIO MEXICANO ENTRE 2015 Y 2019.	37
RESUMEN	37
INTRODUCCIÓN	38
MATERIALES Y MÉTODOS	39
<i>Sitio de estudio</i>	39
<i>Imágenes Sentinel-2</i>	40
RESULTADOS	43
<i>Información a nivel de imagen</i>	43
<i>Información a nivel de píxel</i>	44
<i>El patrón interanual en las ecorregiones</i>	45
<i>El patrón intra-anual en las ecorregiones</i>	46
<i>Errores detectados en la máscara QA60</i>	48
DISCUSIÓN	50

<i>Consideraciones sobre la colección Sentinel-2 nivel 1C</i>	52
CONCLUSIONES	52
CAPÍTULO 3. CLASIFICACIÓN DE TIPOS DE CUBIERTA Y USOS DEL SUELO CON UNA U-NET: VENTAJAS DE COMBINAR IMÁGENES SENTINEL-1 Y SENTINEL-2	54
RESUMEN	54
MÉTODO	55
<i>Procesamiento de la información multiespectral y radar</i>	58
<i>Creación de datos de entrenamiento y validación</i>	59
<i>Entrenamiento de algoritmos</i>	60
<i>Validación de clasificación del área completa de estudio</i>	62
RESULTADOS	64
<i>U-Net</i>	64
<i>Comparación entre las U-Net MS, SAR y MS + SAR</i>	65
<i>Comparación entre las U-Net MS + SAR y RF MS + SAR</i>	67
<i>Clasificación del área completa de estudio</i>	67
DISCUSIÓN	69
<i>Implicaciones para el monitoreo de la región</i>	69
<i>Selección del algoritmo</i>	70
<i>Patrones por clase: selección de bandas</i>	72
<i>Exploración de hiperparámetros</i>	74
<i>Análisis de errores</i>	74
<i>Comparaciones con estudios similares</i>	76
<i>Aspectos metodológicos</i>	78
CONCLUSIONES	81
MATERIAL SUPLEMENTARIO	81
CAPÍTULO 4. DETECCIÓN DE LA DEFORESTACIÓN UTILIZANDO UN ENFOQUE DE APRENDIZAJE PROFUNDO ESPACIO TEMPORAL CON IMÁGENES DE RADAR DE APERTURA SINTÉTICA Y MULTIESPECTRALES.	86
RESUMEN	86
MÉTODOS	87
<i>Procesamiento de imágenes</i>	89
<i>Generación de datos de entrenamiento</i>	91
<i>Entrenamiento de algoritmos</i>	92

<i>Métricas de evaluación</i>	95
<i>Verificación de la clasificación del área completa de estudio</i>	96
<i>Paquetes de R utilizados</i>	98
RESULTADOS	98
<i>Patrones espectrales y temporales</i>	98
<i>Comparación entradas U-Net 3D</i>	100
<i>SVM</i>	101
<i>Comparación U-Net 3D y SVM MS + SAR</i>	101
<i>Análisis de errores</i>	102
<i>Evaluación de la exactitud del mapa y estimados de área</i>	103
DISCUSIÓN	105
<i>Implicaciones para el estudio de pérdida de vegetación arbórea</i>	106
<i>Comparación de algoritmos: U-Net 3D vs SVM</i>	106
<i>Selección de la entrada U-Net 3D</i>	109
<i>Análisis de errores</i>	110
<i>Diferencia de las métricas de datos de validación y verificación de mapa</i>	112
CONCLUSIONES	113
MATERIAL SUPLEMENTARIO	115
CAPÍTULO 5. DETECCIÓN DE LA DEGRADACIÓN FORESTAL EN LA CUENCA DEL RÍO AYUQUILA.	119
RESUMEN	119
MÉTODOS	120
<i>Esquema general</i>	120
<i>Área de estudio</i>	120
<i>Generación de datos de entrenamiento</i>	122
<i>Datos de percepción remota</i>	123
<i>Datos de entrenamiento, verificación y prueba</i>	124
<i>Algoritmos</i>	125
<i>Mapa de biomasa</i>	128
<i>Pruebas aleatorias</i>	129
RESULTADOS	130
<i>Biomasa del bosque</i>	130
<i>Patrones espectrales y temporales de los datos</i>	131
<i>Evaluación de los algoritmos</i>	133

<i>Mapas de AGB</i>	136
<i>Análisis de errores</i>	140
<i>Evaluaciones aleatorias</i>	141
DISCUSIÓN	144
<i>Estudio de la AGB en la zona de estudio</i>	144
<i>Comparación de algoritmos U-Net 3D LF y RF</i>	145
<i>Importancia de las bandas</i>	150
<i>Limitantes sobre el método de muestreo y el procesamiento de imágenes</i>	152
CONCLUSIONES	155
CAPÍTULO 6. DISCUSIÓN GENERAL	156
PRINCIPALES HALLAZGOS	156
IMPLICACIONES PARA EL MONITOREO DE LAS CUBIERTAS ARBÓREAS	159
EVOLUCIÓN DE LOS CAPÍTULOS DE LA TESIS	162
COMPARACIÓN MS + SAR, MS Y SAR	164
COMPARACIONES CON ESTUDIOS SIMILARES	168
ASPECTOS METODOLÓGICOS	169
LECCIONES APRENDIDAS Y RECOMENDACIONES PARA ESTUDIOS FUTUROS	169
CONCLUSIONES	172
REFERENCIAS	173

Índice de cuadros

- Cuadro 1.** Proporción de cada ecorregión con por lo menos una observación válida al año. Con un asterisco (*) se indican los años donde áreas menores a 0.005 % de la superficie de la ecorregión no mostraron observaciones válidas. 46
- Cuadro 2.** F1-score y su diferencia respecto al valor más alto ($\Delta F1$ -score) para cada clase con las tres U-Net MS + SAR, MS, SAR, así como RF MS + SAR. Abreviaturas: MS, bandas multiespectrales, SAR, bandas de apertura sintética, RF, Random forests. 65
- Cuadro 3.** Área, proporción del área de estudio y F1-score para de la clasificación para el área completa de estudio. Asimismo, se muestra el F1-score obtenido en los datos de validación, seguido de los estimados de área corregida y su intervalo del 95 % de confianza por cada clase. Abreviaturas: CI, intervalo de confianza. 67
- Cuadro S 1.** Matriz de confusión para los datos de validación, utilizando la U-Net MS + SAR. Las filas muestras las clases predichas, mientras que las columnas, las clases manualmente clasificadas. Las abreviaciones son: A: agua, AH: asentamientos humanos, BM: bosque maduro, BS: bosque secundario, C: caminos, P/A: pastizales / agricultura, PJ: plantaciones jóvenes, PM: plantaciones maduras, S: suelo, VA: vegetación acuática; UA, precisión del usuario, PA: precisión del productor. 81
- Cuadro S 2.** Matriz de confusión para los datos de validación, utilizando la U-Net MS. Las filas muestras las clases predichas, mientras que las columnas, las clases manualmente clasificadas. Las abreviaciones son: A: agua, AH: asentamientos humanos, BM: bosque maduro, BS: bosque secundario, C: caminos, P/A: pastizales / agricultura, PJ: plantaciones jóvenes, PM: plantaciones maduras, S: suelo, VA: vegetación acuática; UA, precisión del usuario, PA: precisión del productor. 82
- Cuadro S 3.** Matriz de confusión para los datos de validación, utilizando la U-Net SAR. Las filas muestras las clases predichas, mientras que las columnas, las clases manualmente clasificadas. Las abreviaciones son: A: agua, AH: asentamientos humanos, BM: bosque maduro, BS: bosque secundario, C: caminos, P/A: pastizales / agricultura, PJ: plantaciones jóvenes, PM: plantaciones maduras, S: suelo, VA: vegetación acuática; UA, precisión del usuario, PA: precisión del productor. 83
- Cuadro S 4.** Matriz de confusión para los datos de validación, utilizando RF MS + SAR. Las filas muestras las clases predichas, mientras que las columnas, las clases manualmente clasificadas. Las abreviaciones son: A: agua, AH: asentamientos humanos, BM: bosque maduro, BS: bosque secundario, C: caminos, P/A: pastizales / agricultura, PJ: plantaciones jóvenes, PM: plantaciones maduras, S: suelo, VA: vegetación acuática; UA, precisión del usuario, PA: precisión del productor. 84
- Cuadro S 5.** Matriz de confusión para los datos de verificación del mapa para el área completa de estudio, utilizando la U-Net MS + SAR. Las filas muestras las clases predichas, mientras que las columnas, las clases

<i>manualmente clasificadas. Las abreviaciones son: A: agua, AH: asentamientos humanos, BM: bosque maduro, BS: bosque secundario, C: caminos, P/A: pastizales / agricultura, PJ: plantaciones jóvenes, PM: plantaciones maduras, S: suelo, VA: vegetación acuática; UA, precisión del usuario, PA: precisión del productor.</i>	85
Cuadro 4. <i>Exactitud total y avgF1-score para cada clase con las tres U-Net 3D MS, SAR y MS + SAR, así como SVM MS + SAR y MS. Abreviaturas: MS, bandas multiespectrales, SAR, bandas de apertura sintética, SVM, Support Vector Machines. Evaluación sobre los datos de validación sin aumento artificial de los datos. La arquitectura que obtuvo el mayor avgF1-score se indica con un *.</i>	100
Cuadro 5. <i>Área, proporción del área de estudio y F1-score para de la clasificación para el área completa de estudio. Asimismo, se muestra el F1-score obtenido en los datos de validación, seguido de los estimados de área corregida y su intervalo del 95 % de confianza por cada clase. Abreviaturas: CI, intervalo de confianza.</i>	104
Cuadro S 6. <i>Matriz de confusión para los datos de prueba, utilizando la U-Net MS. Las filas muestras las clases predichas, mientras que las columnas, las clases manualmente clasificadas.</i>	115
Cuadro S 7. <i>Matriz de confusión para los datos de validación, utilizando la U-Net SAR. Las filas muestras las clases predichas, mientras que las columnas, las clases manualmente clasificadas.</i>	116
Cuadro S 8. <i>Matriz de confusión para los datos de validación, utilizando la U-Net MS + SAR. Las filas muestras las clases predichas, mientras que las columnas, las clases manualmente clasificadas.</i>	116
Cuadro S 9. <i>Matriz de confusión para los datos de validación, utilizando SVM MS. Las filas muestras las clases predichas, mientras que las columnas, las clases manualmente clasificadas.</i>	117
Cuadro S 10. <i>Matriz de confusión para los datos de validación, utilizando SVM MS + SAR. Las filas muestras las clases predichas, mientras que las columnas, las clases manualmente clasificadas.</i>	117
Cuadro S 11. <i>Matriz de confusión para los datos de evaluación de la calidad del mapa para el área completa de estudio, utilizando la U-Net MS. Las filas muestras las clases predichas, mientras que las columnas, las clases manualmente clasificadas. Todos los errores de comisión y omisión en la clase de no deforestación se detectaron en el estrato buffer.</i>	118
Cuadro 6. <i>Algoritmos que alcanzaron la menor raíz cuadrada del error cuadrático medio (RMSE) sobre los datos de prueba. Además, se muestra el RMSE sobre el conjunto de datos de entrenamiento, validación y prueba sin aumento artificial de los datos.</i>	134
Cuadro 7. <i>Biomasa aérea estimada (AGB) para el bosque tropical caducifolio (BTC) del área de estudio en 2019 y 2022, así como el área promedio por pixel. El área total del BTC es de 81353.24 ha.</i>	136
Cuadro 8. <i>Resultados obtenidos para el modelo RF aleatorizando la división de datos: 1) tradicional, respetando la división por parcela (9 píxeles por parcela) o 2) totalmente aleatorizado. Se presentan los</i>	

resultados con tres métodos 10-fold, Bootstrap y leave-one-out cross validation (LOO). Abreviaciones: EE, error estándar.

142

Índice de figuras

Figura 1. Esquema de una neurona.	23
Figura 2. Ejemplo de una red neuronal artificial. Las flechas rojas representan la propagación hacia adelante, mientras que las flechas azules representan la propagación hacia atrás.	24
Figura 3. Ejemplo de una capa convolucional de una red neuronal convolucional (CNN).	27
Figura 4. Ejemplo de una capa de reescalamiento de la información de una red neuronal convolucional (CNN).	28
Figura 5. Cuatro aplicaciones distintas de algoritmos de CNN para el procesamiento de imágenes (Fuente: Hoenser and Kuenzer, 2020).	30
Figura 6. Representación gráfica de una arquitectura U-Net con cuatro capas ocultas. La imagen en el extremo izquierdo representa la imagen utilizada como entrada, mientras que la del extremo derecho, la salida (una clasificación). Figura realizada mediante el código de https://github.com/Harislqbal88/PlotNeuralNet .	33
Figura 7. Representación gráfica de la arquitectura U-Net 3D. La imagen en el extremo izquierdo representa un par de imágenes utilizada como entrada, mientras que la del extremo derecho, la salida (una clasificación). Figura realizada mediante el código de https://github.com/Harislqbal88/PlotNeuralNet .	34
Figura 8. Ubicación geográfica de las ecorregiones nivel I de México de acuerdo con la clasificación de (INEGI et al., 2008). La escala de colores muestra un gradiente general de mayor a menor humedad (verde oscuro a naranja claro). Además, se muestran las escenas de Sentinel-2 que cubren la superficie terrestre de México (Fuente: ESA, 2020c). En los ejes superior e izquierdo se muestran las zonas UTM donde se encuentra el país, de acuerdo con el Sistema de Referencia de Cuadrícula Militar, mientras que en los ejes inferior y derecho, se muestran las coordenadas geográficas que delimitan cada una de estas zonas.	41
Figura 9. Número de imágenes Sentinel-2 1C disponibles para México entre 2015 y 2019 A) divididas por el satélite que las registró y B) por su porcentaje de cobertura de nubes.	44
Figura 10. Distribución espacial del número de observaciones válidas mensuales por píxel en 2019. Las franjas de mayor número de observaciones válidas corresponden a las áreas donde hubo un traslape entre órbitas adyacentes de la constelación Sentinel-2.	45
Figura 11. Promedio mensual de observaciones válidas por píxel de cada ecorregión del país durante el periodo 2015 – 2019.	47
Figura 12. Proporción de la superficie de cada ecorregión con por lo menos una observación válida de acuerdo con los mosaicos contruidos con imágenes de 1 mes (mensuales), 2 meses (bimestrales) y 3 meses (trimestrales) en el periodo de 2018 – 2019. Con un asterisco (*) se marcan los meses donde áreas menores a 0.005 % de la superficie de cada ecorregión no presentaron observaciones válidas.	48

Figura 13. Ejemplos de errores en la banda de evaluación de calidad (QA60) de las imágenes Sentinel-2 1C. En la esquina superior izquierda se muestra un área donde se identificó erróneamente una zona con un sustrato muy claro como nubes. Se muestra la imagen (compuesto RGB) y debajo, el número de observaciones válidas que presentó dicha área en 2019. Por su parte, en la esquina inferior derecha, se muestra un ejemplo de una imagen (compuesto RGB) con su respectiva banda QA60 donde no se detectó con precisión la presencia de nubes.	49
Figura 14. Localización del sitio de estudio, así como de las áreas clasificadas manualmente para entrenar y verificar el modelo U-Net.	57
Figura 15. Diagrama de flujo del procedimiento seguido en este trabajo.	58
Figura 16. Evaluación de las predicciones de la U-Net en tres distintas muestras de los datos de validación. (A) Compuesto RGB de Sentinel-2, (B) Áreas manualmente clasificadas, (C) Predicción con la U-Net utilizando las bandas multispectrales (MS) + radar de apertura sintética (SAR), (D) predicciones con la U-Net utilizando sólo las bandas MS, (E) predicciones de la U-Net utilizando las bandas SAR.	66
Figura 17. (A) Clasificación de cubierta y usos del suelo del área de estudio completa utilizando la U-Net MS + SAR, así como el compuesto RGB y los mapas de probabilidad de pertenencia a la clase correspondiente. (B) Ejemplo de un sitio dominado por bosque maduro con altas probabilidades de pertenencia. (C) Ejemplo de un área recientemente quemada con una alta presencia de nubes y sombras, donde se observan valores bajos de pertenencia.	68
Figura 18. Promedio de la señal espectral (reflectancia y coeficiente de retrodispersión, γ_0) de las clases incluidas en el sistema de clasificación de tipos de vegetación y usos del suelo.	71
Figura 19. Localización del área de estudio y de los polígonos de pérdida de bosque maduro y bosque secundario o plantaciones, delimitados mediante interpretación visual.	88
Figura 20. Diagrama de flujo del procedimiento para entrenar y validar las arquitecturas U-Net 3D y Support Vector Machines.	89
Figura 21. Esquema de la U-Net 3D utilizada para detectar la deforestación. La imagen del extremo izquierdo representa la serie de cuatro imágenes utilizada como entrada, mientras que la del extremo derecho, la salida que corresponde a una clasificación de sitios deforestados y no. Figura realizada mediante el código de https://github.com/Harislqbal88/PlotNeuralNet .	92
Figura 22. Diagrama de cajas que muestra el perfil espectral y temporal de las tres clases de interés en las seis bandas y cuatro periodos utilizados para entrenar los algoritmos U-Net 3D y SVM. Las fechas en el eje de las x están expresadas como año/mes. Estos datos fueron extraídos del conjunto de entrenamiento utilizado para SVM MS + SAR.	99

Figura 23. Ejemplos de áreas de validación donde se muestran los compuestos RGB iniciales de Feb – Abr 2019 (t1) y finales Feb – Abr 2020 (t4), así como las predicciones realizadas por las U-Net 3D MS, SAR, MS + SAR, así como SVM MS + SAR.	101
Figura 24. Ejemplos de áreas de validación donde se muestran los compuestos RGB de Feb – Abr 2019 y Feb – Abr 2020, así como las predicciones realizadas por las U-Net 3D MS + SAR, MS y SAR, así como SVM MS + SAR.	102
Figura 25. (A) Clasificación de las áreas deforestadas de 2019 a 2020 en el área de estudio utilizando la U-Net 3D MS y su mapa de probabilidad de pertenencia a la clase correspondiente. (B) Ejemplo de un sitio dominado por deforestación de bosque maduro. (C) Ejemplo de un área dominada por deforestación de bosque secundario o plantaciones.	105
Figura 26. Localización de la cuenca del río Ayuquila y de sitios de muestreo del bosque tropical caducifolio (BTC).	121
Figura 27. Esquema de la arquitectura U-Net 3D LF.	126
Figura 28. Histograma de biomasa aérea (AGB) de las parcelas muestreadas en este trabajo, divididas por año de registro.	131
Figura 29. Promedio de la reflectancia de los nueve píxeles por sitio de muestreo en las tres imágenes MS utilizadas para entrenar la U-Net 3D LF (t1: Ene-Abr, t2: May-Ago, t3: Sep-Dic). Además, el símbolo muestra el año del registro de los datos. El gráfico se encuentra dividido por los tres conjuntos de datos utilizados: entrenamiento, verificación y prueba.	132
Figura 30. Promedio del coeficiente de retrodispersión de los nueve píxeles por sitio de muestreo en las doce imágenes mensuales SAR utilizadas para entrenar la U-Net 3D LF. Además, el símbolo muestra el año del registro de los datos. Los datos están divididos por los tres conjuntos de datos utilizados: entrenamiento, verificación y prueba.	133
Figura 31. Comparación entre los valores observados y predichos por A) U-Net 3D LF y B) RF en los datos de entrenamiento, prueba y verificación. La línea recta muestra un ajuste perfecto entre los valores observados y predichos.	134
Figura 32. Importancia relativa de las variables predictivas para predecir AGB utilizando random forest.	135
Figura 33. Predicción de biomasa en el BTC realizada por el modelo U-Net 3D LF para 2019 y 2022.	137
Figura 34. Predicción de biomasa en el BTC realizada por el modelo RF para 2019 y 2022.	138
Figura 35. Cambios de biomasa en el BTC entre 2019 y 2022 de acuerdo con las predicciones realizadas por los modelos RF y U-Net 3D LF.	139
Figura 36. Ejemplos de aciertos o errores detectados en la diferencia del AGB entre los mapas predichos con random forest para 2019 y 2022. A) una pérdida de AGB, B) una falsa detección de pérdida de AGB, C) una falsa detección de ganancia de AGB, D) una posible ganancia de AGB. El cambio de AGB se calculó como al	

AGB de 2022 menos el AGB de 2019, de tal manera que los valores negativos implican una pérdida y los valores positivos una ganancia. 141

Figura 37. *Coefficiente de variación por píxel de acuerdo a las aleatorizaciones realizadas con el algoritmo random forest y usando un 10-fold cross validation para el año 2019 y 2022.* 143

Resumen

La deforestación y la degradación forestal son dos procesos que contribuyen al cambio global mediante la emisión de gases de efecto invernadero, la pérdida de biodiversidad y la reducción en la calidad de algunos servicios ecosistémicos. Gracias a su capacidad de estudiar grandes superficies y contar con un registro histórico, la percepción remota ha demostrado ser una herramienta esencial para cuantificar estos procesos. Además, el desarrollo de nuevos métodos de análisis como los algoritmos de aprendizaje profundo e imágenes de mayor resolución espectral, espacial y temporal gratuitas abren la posibilidad de desarrollar métodos que permitan mejorar las capacidades actuales de monitoreo. En este contexto, el presente trabajo evaluó el desempeño de los algoritmos de aprendizaje profundo con imágenes multiespectrales y de radar de apertura sintética para clasificar distintos tipos de coberturas del suelo, detectar a deforestación y la degradación forestal. Los resultados muestran que los algoritmos de aprendizaje profundo, en general, permiten obtener mejores resultados que su contraparte de aprendizaje automatizado, debido a la incorporación de patrones en las dimensiones espaciales y temporales. Sin embargo, dicho desempeño está condicionado al tamaño de muestra y la fuerza de la relación entre la señal remota y la clase o característica a evaluar. Por otro lado, la combinación de la información multiespectral y radar de apertura sintética en general fue beneficiosa, aunque en algunos casos el agregar la información radar brinda resultados similares que solo con la multiespectral. Este último resultado se debe a la falta de contribución de nueva información útil por parte de las imágenes de radar, en comparación con la que ya brindan las imágenes multiespectrales. Los resultados apuntan a que este tipo de técnicas permiten obtener resultados más precisos para identificar estos procesos, por lo cual, representan una herramienta atractiva para el desarrollo de futuras herramientas encaminadas al monitoreo de estos procesos.

Abstract

Deforestation and forest degradation are two global change drivers that contribute to greenhouse gases emissions, biodiversity loss and reduction in the quality of ecosystemic services. Due to their ability to study large extents and count with a historical record, remote sensing has demonstrated being an essential tool to monitor these processes. Besides, the development of new methods of analysis, such as deep learning and images with higher spatial, spectral and temporal resolution open the possibility to develop methods that enhance the current monitoring capacities. In this context, the current study evaluated the performance of deep learning algorithms with multispectral and synthetic aperture radar images to classify different land use land covers, identify deforestation and forest degradation. The results show that, in general, deep learning algorithms obtain more accurate results than its machine learning counterparts, due to the incorporation of spatial and temporal dimensions. Nonetheless, this performance is conditioned by sample size and the force of the relation between the remote signal and the class or attribute being evaluated. On the other hand, the combination of multispectral and synthetic aperture radar was, in general, beneficial, although in certain cases, adding the latter information resulted in similar capabilities than using only multispectral ones. This result is due to the lack of new useful information provided by the radar images, in comparison with the one already contributed by the multispectral ones. The results indicate that these techniques are capable of obtaining more accurate evaluations to identify deforestation and forest degradation; thus, they represent a very attractive alternative to develop future tools to monitor these processes.

Presentación del trabajo

La deforestación y la degradación forestal son dos procesos que se han asociado con algunos de los conductores del cambio global como la pérdida de biodiversidad, emisiones de gases de efecto invernadero, degradación de la provisión de los servicios ambientales, entre otros (Giam, 2017; Kindermann et al., 2008; Pearson et al., 2017; Pimentel, 1997; Putz and Redford, 2010). Por ello, de manera internacional se han planteado realizar acciones para reducir las tasas de deforestación y degradación forestal y promover un uso sostenible de los recursos forestales, como el Acuerdo de París, los Objetivos de Desarrollo Sostenible, las Metas de Aichi y el programa REDD+ (Deschamps Ramírez and Larson, 2017; UN, 2015, 2016; UNEP and CBD, 2010).

El primer paso para llevar a cabo acciones en este sentido es cuantificar la magnitud de estos fenómenos, y posteriormente, identificar su comportamiento a través del tiempo para saber si se está avanzando en la reducción de las tasas de deforestación y degradación forestal. Esto ha promovido que desde hace varios años se desarrollen métodos que permitan identificar y cuantificar estos procesos, hacerlo con mayor precisión, obtener estimaciones con menor error y con mayor resolución espacial (*e.g.*, Hansen et al., 2013; Isaienkov et al., 2020; Simoes et al., 2021; Zhu and Woodcock, 2014). A partir del desarrollo de estas técnicas, se podrá realizar un monitoreo para cuantificar el éxito o fracaso de acciones dirigidas en reducir las tasas de deforestación y degradación forestal. Por otro lado, el desarrollo de nuevos métodos de monitoreo de la superficie terrestre también ha buscado brindar información más detallada sobre la cubierta y el uso del suelo que permita mejorar las capacidades de monitoreo actuales. En conjunto, la generación de información detallada con estimaciones de superficie precisas permitirá mejorar las capacidades del monitoreo de la superficie terrestre.

Esta tesis pretende evaluar las capacidades del uso de algoritmos de aprendizaje profundo con imágenes multiespectrales y de radar para obtener información más detallada de la superficie terrestre, sobre todo para tratar de distinguir entre cubiertas arbóreas con distintas características y usos, *i.e.*, bosque maduro, bosque secundario y plantaciones. Además, evaluará las capacidades de dicho procedimiento para detectar la pérdida de este

tipo de cubiertas e identificar la degradación forestal, mediante la modelación de cambios en la biomasa aérea.

La tesis está organizada en cinco capítulos. El primero funge como una introducción a los conceptos y definiciones claves para el trabajo, así como una breve presentación del estudio de la superficie terrestre mediante sensores remotos y el uso de los algoritmos de aprendizaje profundo. En el capítulo dos se presenta una revisión del número de observaciones sin nubes para la superficie terrestre mexicana usando imágenes Sentinel-2, que sirvió como punto de partida para determinar los periodos en el que las imágenes presentaran la menor cobertura de nubes para estudiar las dos regiones de estudio. En el tercer capítulo se presenta la primera evaluación de un algoritmo de aprendizaje profundo, la U-Net, para realizar una clasificación de tipos de vegetación y usos del suelo, utilizando imágenes multiespectrales (Sentinel-2) y radar de apertura sintética (Sentinel-1). Este capítulo permitió probar si el método empleado era capaz de distinguir entre distintos tipos de cubierta arbórea. El capítulo cuatro, evalúa el uso de la versión 3D de este algoritmo, la U-Net 3D, la cual incorpora la dimensión temporal para detectar la deforestación y tratar de distinguir entre la pérdida de bosque maduro y bosque secundario o plantaciones. La definición de estas dos clases estuvo determinada por los resultados obtenidos del capítulo anterior. Por último, el quinto capítulo aborda el potencial de un algoritmo de aprendizaje profundo para detectar el proceso de degradación forestal, interpretado como una pérdida de biomasa aérea. Los trabajos del tercer y cuarto capítulo se realizaron en un área donde el bosque natural dominante es el bosque tropical húmedo, mientras que el quinto, en un área donde domina el bosque tropical caducifolio.

Los capítulos tres al cinco representan una evolución tanto en la complejidad del proceso estudiado, como de los algoritmos utilizados. Por ejemplo, se transita de una clasificación de distintas cubiertas y usos del suelo en un tiempo único, a identificar cambios de cubiertas entre dos tiempos, y, por último, identificar cambios en una característica del bosque (biomasa) entre dos tiempos. Por otro lado, los algoritmos utilizados progresan del uso de información en la dimensión espacial, a utilizar la temporal y la temporal con distintas dimensiones entre las bandas multiespectrales (MS) y de radar de apertura sintética (SAR). Además, las imágenes empleadas como insumos transitan

desde una sola imagen, hasta compuestos multitemporales. Por ello, cada capítulo se puede interpretar como un aumento en la complejidad de la tarea realizada.

Capítulo 1. Introducción

Cambio global: deforestación y degradación forestal

El cambio global hace referencia a los cambios y transformaciones a escala planetaria, producto de las actividades humanas. Dichas alteraciones derivan del cambio climático, el cambio de cubierta y uso de suelo, la contaminación de ecosistemas, la pérdida de biodiversidad y la modificación de los ciclos biogeoquímicos, entre otras (Rockstrom et al., 2009; Steffen et al., 2003). Estos procesos a su vez conllevan alteraciones con impactos en la salud y bienestar de la sociedad, así como impactos en los medios de vida (Díaz et al., 2015; IPCC, 2014). Por ejemplo, la alteración de los regímenes hidrológicos puede poner en riesgo la seguridad alimentaria de algunas regiones o aumentar su exposición a ciertos fenómenos meteorológicos extremos. Por ello, el cambio global se propone como el principal reto a enfrentar en el siglo XXI, así como la generación de estrategias para adaptarse o mitigar sus impactos (IPCC, 2014; Steffen et al., 2018).

Aunque algunos de los componentes del cambio global han ocurrido de manera natural en la historia del planeta, a partir de los años sesenta, se dio una aceleración en el desarrollo de las actividades humanas y sus impactos (IPCC, 2014; Rockstrom et al., 2009; Steffen et al., 2003). Este evento derivó de un aumento de la población global, un incremento del consumismo, la tecnificación de los métodos de producción y una mayor globalización, conocidas como la Gran Aceleración (Steffen et al., 2015; Steffen et al., 2018). Dicha aceleración estuvo y continúa estando íntimamente ligada al modelo económico capitalista dominante, lo cual ha provocado que dichos impactos se hayan agudizado o resulten en nuevas formas de interacciones (Baer, 2011; IPCC, 2014; Koch, 2012; Liu et al., 2019; Park, 2015).

Los bosques son un elemento clave en el ciclo del carbono, ya que capturan CO₂ de la atmósfera y lo almacenan en sus tejidos, siendo así de los mayores sumideros de carbono en los ecosistemas terrestres. Por ello, parte del interés de conservar las cubiertas boscosas radica en su papel como mitigadores del aumento del CO₂ atmosférico y el cambio climático. Sin embargo, a nivel mundial, cada año se deforestan grandes áreas para realizar diferentes actividades humanas (*e.g.*, agricultura, ganadería, construcción de

infraestructura, etc.). Este proceso conlleva una emisión de gases de efecto invernadero, producto de la quema o descomposición de los árboles (Houghton, 1990; Houghton, 1991). Además, la deforestación provoca una disminución del área ocupada por los sumideros de gases de efecto invernadero (sobre todo de carbono) y, por lo tanto, la cantidad que puede ser secuestrada en el futuro en estos sistemas (Gatti et al., 2021). Por ello, este proceso está relacionado con el cambio climático y la emisión de gases invernadero.

A pesar de que la degradación forestal no implica un cambio de uso de cubierta o uso de suelo, sí implica una disminución en alguna de sus características. Entre las más frecuentemente mencionadas, es la cantidad de carbono contenido o biomasa (IPCC, 2003). Por ello, no es común que se mencione como un causante del cambio global. Sin embargo, varios trabajos han sugerido la magnitud de las emisiones de gases de efecto invernadero por este proceso, puede ser equiparable o mayor al de la deforestación en ciertas regiones (Kruid et al., 2021; Pearson et al., 2017; Qin et al., 2021). Esto se debe a que comúnmente la degradación forestal implica una remoción parcial de la cobertura arbórea, pero debido a que puede abarcar áreas mucho más grandes que las deforestadas, implica una emisión absoluta mayor.

Por otro lado, tanto la deforestación como la degradación forestal provocan la pérdida de biodiversidad, debido a que dichas transformaciones están relacionadas con la pérdida de hábitat (Hanski; Rogan and Lacher, 2018). Aunque la degradación forestal no tiene un efecto tan dramático sobre la pérdida de biodiversidad, la tala selectiva de especies o la disminución de alguna característica clave también puede tener efectos sobre la pérdida de biodiversidad. Ambos procesos entonces contribuyen a la disminución de la biodiversidad de los ecosistemas, lo cual puede alterar el funcionamiento de los ecosistemas.

De igual manera, ambos procesos pueden alterar la calidad o la provisión de ciertos servicios ecosistémicos. Por ejemplo, al remover una cubierta arbórea se disminuye la posibilidad del suelo para infiltrar agua, así como para reducir la erosión. Además, los bosques pueden brindar algunos servicios más inmediatos a la población como provisión de alimento, medicinas y leña (Díaz et al., 2015; McFarland, 2018). Por su parte, aunque la degradación forestal no implica una remoción completa de la cubierta, sí puede facilitar la invasión de especies exóticas o la disminución en la provisión de alguno de los servicios

antes mencionado. Por ello, estos dos procesos están relacionadas con el cambio global en distintas dimensiones.

Las causas de la deforestación y la degradación forestal incluyen factores relacionados con el mercado nacional e internacional (oferta, demanda, precios), aquellos relacionados con la gobernanza, como la presencia de incentivos gubernamentales e instrumentos legales que favorecen o desfavorecen ciertas actividades, y la tenencia de la tierra, así como motivaciones culturales, condiciones socioeconómicas y de accesibilidad, entre otras (Bullock et al., 2020; Burivalova et al., 2019; CICC, 2017; Díaz et al., 2015; Ellis et al., 2021; Gallardo-Cruz et al., 2021; Hodgdon et al., 2015; Hojas-Gascon et al., 2015; McFarland, 2018; Morales-Barquero et al., 2015; Runyan and D'Odorico, 2016; Sandler, 1993). Además, ambos procesos suelen estar presentes en una misma región, aunque también puede ser que uno de ellos sea el más frecuente. Por ejemplo, en algunas de las zonas tropicales del sureste del país, la expansión de las actividades pecuarias se mencionan como las principales causas del cambio de uso de suelo, aunque en ciertas regiones también la expansión de plantaciones y actividades agrícolas son importantes (CONAFOR, 2022; Fernández-Montes de Oca et al., 2015; Mwampamba et al., 2015; Romero and Albuquerque, 2018; Vaca et al., 2019); mientras que en otras regiones la misma presencia de actividades agropecuarias y el aprovechamiento del bosque como fuente de leña y postes se mencionan entre las principales causas de la degradación forestal (Borrego and Skutsch, 2019; Morales-Barquero et al., 2015; Salinas-Melgoza et al., 2018).

Para tratar de combatir o aminorar los efectos del cambio global, se han propuesto varias de las estrategias y metas a lograr para mitigar o adaptarse a ciertos componentes del cambio global se materializan en los objetivos de desarrollo sostenible, así como las metas nacionalmente determinadas (Gobierno de México, 2020; UN, 2015) o programas como REDD+ (Deschamps Ramírez and Larson, 2017). En particular, estos contienen metas específicas para enfrentar tanto la deforestación y la degradación forestal. Por ejemplo, la meta 15 de los objetivos de desarrollo sostenible considera los siguientes puntos: frenar la deforestación, reducir la degradación de hábitats naturales, restaurar bosques degradados, y disminuir la pérdida de biodiversidad para el 2020 (UN, 2015). Por su parte, entre las metas nacionalmente determinadas se encuentra alcanzar la tasa cero de deforestación para el

2030, conservar y restaurar ecosistémicas, incremento y permanencia de reservorios de carbono (Gobierno de México, 2020). Dichas metas y objetivos entonces requieren ser medidos y cuantificados para evaluar su progreso. Además, esta información puede ser utilizada para evaluar el desempeño de acciones enfocadas en reducir la deforestación y la degradación forestal.

Definiciones de Bosque, deforestación y degradación forestal

En la literatura se pueden encontrar definiciones muy distintas de degradación forestal y deforestación; sin embargo, estos conceptos suelen ir acompañados de una definición de bosque, la cual representa la condición inicial para que ocurran estos procesos y suele corresponder a una cubierta dominada por árboles (FAO, 2009, 2011, 2018; Lund, 2009; Sasaki and Putz, 2009). A partir de este concepto de bosque, la deforestación se puede considerar como una transición entre la categoría de bosque a no bosque (*i.e.*, una cubierta no dominada por árboles), mientras que la degradación, como la disminución de alguna característica del bosque (*e.g.*, diversidad de especies, biomasa, cobertura arbórea o resiliencia), pero que continúa siendo un bosque (Ryan et al., 2012; Sasaki and Putz, 2009; Thompson et al., 2013; Vásquez-Grandón et al., 2018).

Debido a que usualmente se desea trabajar con una definición medible u operacional, las definiciones de bosque, deforestación y degradación forestal se han puesto en términos de variables cuantificables en campo o mediante sensores remotos (*e.g.*, FAO, 2018; IPCC, 2003; Lund, 2009; Morales-Barquero et al., 2014; Vásquez-Grandón et al., 2018). Por ello, muchas de las definiciones se han puesto en términos de valores umbrales de algunas características estructurales de la vegetación. Aunque para la deforestación existen algunos umbrales más o menos aceptados a nivel internacional (ver siguiente párrafo), para la degradación forestal no existen definiciones tan ampliamente aceptadas, como se verá más adelante (FAO, 2009; Simula, 2009; Thompson et al., 2013). Sin embargo, cabe mencionar que no siempre existe una correspondencia entre las variables y valores umbrales para definir bosque y el método utilizado para identificar dicha clase (Fernández-Montes de Oca et al., 2021).

La definición de bosque más frecuentemente utilizada es la propuesta por la (FAO, 2018), la cual define a este como un área de mínimo 0.5 ha con un dosel de mínimo 5 m de altura y una cobertura arbórea mayor al 10 %. Adicionalmente, la misma definición presenta algunos requisitos adicionales como que el bosque debe contar con un mínimo de 20 m de ancho para corredores y excluye áreas con sistemas de producción agrícola y áreas urbanas (*e.g.*, plantaciones de palma de aceite o frutales y plantaciones agroforestales). Por último, dicha definición menciona que las áreas que se espera que tengan las propiedades del bosque en 5 años también se les considera bosque. Existen otros organismos nacionales e internacionales que coinciden con el umbral mínimo de 10 % de cobertura arbórea y 5 m de altura para definir un bosque, (*e.g.*, CBD, 2021; CONAFOR and SEMARNAT, 2019; UNFCCC, 2001), mientras que algunas incorporan las mismas variables, pero con intervalos más amplios (*e.g.*, 10 – 30 % de cobertura arbórea (IPCC, 2003), cobertura arbórea de 10 - 30 %, 0.05 - 1 ha y 2-5 m de altura (UNFCCC, 2011)). Cabe mencionar que, aunque la definición original de la FAO excluye a las plantaciones fuera de la definición de bosque, es común encontrar que estudios u otras definiciones ignoren dicho componente y se concentren únicamente en considerar las características estructurales de la vegetación para definir un bosque (Hansen et al., 2013; Milodowski et al., 2017). De acuerdo a esta breve revisión de definiciones, se pueden identificar al menos tres aspectos clave para definir un bosque (Fernández-Montes de Oca et al., 2021; Hojas-Gascon et al., 2015): 1) la unidad mínima de superficie, 2) la cobertura arbórea mínima y 3) la altura mínima de los árboles. En algunos casos, además se hace referencia al origen de dicho bosque (natural o antropogénico) y si una cobertura presentará las características de un bosque en menos de 5 años (FAO, 2018; Fernández-Montes de Oca et al., 2015). Ya que existe una gran variación de las características estructurales de los bosques, algunos autores han propuesto crear definiciones que se adapten a las condiciones locales del área de estudio, enfatizar sobre la exclusión de plantaciones como bosques o incluir aspectos como la diversidad o la provisión de recursos ecosistémicos (Hojas-Gascon et al., 2015; Milodowski et al., 2017; Putz and Redford, 2010).

En el caso de la deforestación, esta puede definirse como la pérdida de un bosque. De acuerdo a la (FAO, 2018), la deforestación se define como la transformación permanente de

un bosque a otro tipo de cubierta no arbórea o a la reducción permanente de la cobertura arbórea del dosel. Por ello, se considera que ocurre cuando se reduce la cobertura del dosel por debajo del 10 %, se reduce el área de la cubierta boscosa por debajo de 0.5 ha o se reduce la altura del dosel a una menor a 5 m de altura. Por otro lado, la deforestación también puede interpretarse como un cambio de uso de suelo de bosque a no bosque. Además, cabe mencionar que la mayoría de las definiciones de deforestación mencionan que dicha transformación debe tener una causa antrópica (FAO, 2018; IPCC, 2003; UNFCCC, 2001).

Por su parte, la degradación forestal ha sido definida a partir de diversas perspectivas y se pueden encontrar más de 50 definiciones de acuerdo a distintos organismos internacionales, con distintos criterios de medición (*e.g.*, aspectos sobre el suelo, agua, nutrientes, paisaje, biodiversidad, productividad de genes, diversidad genética) y de intervención humana (FAO, 2011; IPCC, 2003; Lund, 2009; Simula, 2009). Una de las definiciones de degradación forestal más citadas es: “la disminución de la capacidad del bosque para suministrar bienes y servicios” (FAO, 2018). Sin embargo, dicha definición abarca una variedad de conceptos muy amplia y formas de medirlo, por ejemplo, biomasa, estructura del bosque, diversidad de especies, entre otras (Thompson et al., 2013). Por ello, se han realizado diversos esfuerzos por crear criterio en común para definirla y medirla. Por ejemplo, la (CONAFOR and SEMARNAT, 2019), pone especial énfasis en la disminución de la capacidad de proveer servicios ambientales, de su productividad o de sus características estructurales, así como en las condiciones del suelo. Por su parte, la Ley General de Cambio Climático define a la degradación forestal como una disminución del carbono en la vegetación, ecosistema o suelo, en relación con un valor de referencia de un bosque que no ha sido perturbado (DOF, 2018; Leyva-Ovalle et al., 2017; Morales-Barquero et al., 2014). A partir de esta breve revisión se pueden identificar varios puntos clave para definir este concepto (Lund, 2009; Putz and Redford, 2010; Thompson et al., 2013; Vásquez-Grandón et al., 2018): 1) almacén de carbono y productividad, 2) diversidad biológica, 3) la afectación de su capacidad de recuperación, y 4) funciones protectoras. Cabe mencionar que todas estas definiciones también se pueden agrupar en dos corrientes (Chazdon, 2008; FAO, 2011; Morales-Barquero et al., 2014; Thompson et al., 2013): 1) las

que definen a la degradación forestal como un proceso cuantificado mediante la disminución de una característica a través del tiempo y 2) las que la definen como una comparación respecto a las características de un bosque conservado, donde no necesariamente se comparan las características de un bosque en dos tiempos distintos.

Dentro de las pocas definiciones operativas de degradación forestal, se encuentra la propuesta por la (Simula, 2009) en la que define un bosque degradado como aquel que se encuentra con un porcentaje de cobertura arbórea entre el 10 y 80 %, donde el 80 – 100 % corresponde al bosque conservado. En consecuencia, el proceso de degradación forestal ocurriría cuando un bosque pasa de un umbral del 80 - 100% de cobertura a uno entre 10 - 80 %. Aunque esta definición representa un primer esfuerzo por tratar de llegar a un consenso, la realidad es que no ha sido tan ampliamente reconocida ni usada como su contraparte de deforestación. Además, presenta el problema de que hay tipos de bosques que naturalmente pueden tener valores de cobertura arbórea entre 10 – 80 % sin estar en una condición degradada.

A pesar de que el problema de definir los conceptos de bosque, degradación forestal y deforestación, podría limitarse a un problema semiótico; en el contexto actual, tiene consecuencias importantes para el reporte de la superficie de bosque, deforestada y degradada en metas nacionalmente determinadas, de la CBD, UNFCCC, REDD+, entre otros, así como para la selección de áreas candidatas a entrar a programas como el pago por servicios ambientales o REDD+ (Putz and Redford, 2010). En muchos casos las propias definiciones no son específicas o no existe una forma de cuantificarlas de acuerdo a los métodos utilizados, de forma que no resultan 100 % operacionales (Fernández-Montes de Oca et al., 2021; Thompson et al., 2013). Por ejemplo, aunque un gran número de definiciones consideran que estos procesos deben ser resultado de acciones humanas, con los métodos que son cuantificados (*e.g.*, percepción remota) no se puede asegurar el origen de estos procesos, o en otros casos, algunas características de su uso (*e.g.*, distinción entre plantaciones y bosque) no pueden ser distinguidas, así como algunos criterios de área mínima.

A pesar de que el problema de la operacionalización y definición respecto a valores umbrales de la deforestación y degradación forestal es un aspecto que este trabajo no

pretende resolver, esta revisión tiene como objetivo mostrar las dificultades asociadas a cuantificar y medir procesos como la deforestación y degradación forestal. Debido a la naturaleza del presente estudio, se optará por definir la deforestación y la degradación con un enfoque en el carbono (IPCC, 2003) y se cuantificarán como procesos de cambio a través del tiempo. De tal manera, la deforestación se definirá como la pérdida total de cobertura arbórea, mientras que la degradación forestal, como la disminución de cobertura arbórea de un bosque que no implica la remoción total del bosque. También se tratará de incluir características sobre el origen del bosque o de su uso, sobre todo para distinguir entre tres clases de cobertura arbórea: bosque maduro, bosque secundario y plantaciones, que pueden relacionarse no sólo a usos distintos, sino, también a cubiertas arbóreas con características estructurales y de diversidad diferentes. En este caso, el término de bosque secundario se usará para referirse a un bosque joven que se desarrolló posterior a un evento de deforestación y que suele presentar árboles de menor talla y una composición distinta a la de uno maduro. La principal motivación para llevar a cabo esta distinción es para distinguir entre conductores de la deforestación y bajo un uso más intensivo, como el establecimiento de plantaciones, de áreas que pueden estar bajo dinámicas de un uso a largo plazo, *i.e.*, bosque secundario, de áreas bajo uso de muy poca intensidad o nula, *i.e.*, bosques maduros (Latam, 2020; Romero and Albuquerque, 2018). Estudios previos han utilizado definiciones similares para el estudio de la deforestación y distinguir aquella que proviene de bosque maduro, de la que proviene de otro tipo de cubierta arbórea (Kalamandeen et al., 2018). Además, consideramos que el uso de imágenes de radar permitirá brindar información sobre el arreglo geométrico de la cubierta arbórea, así como de algunas de sus características estructurales que ayudará a distinguir entre esas clases de cubierta arbórea. Por ello, se evaluará si el uso de algoritmos de aprendizaje profundo junto con imágenes multiespectrales y radar permitirá obtener clasificaciones con ese grado de detalle.

Uso de la percepción remota para el estudio de la superficie terrestre

La percepción remota o teledetección se define como el estudio de cualquier objeto mediante su señal electromagnética (Chuvieco, 2008). El fundamento teórico detrás de la

percepción radica en que la señal contiene información acerca de las características del objeto que se está midiendo. Este principio llevado al contexto del estudio de la superficie terrestre implica que la señal electromagnética brinda información sobre el tipo de cubierta y algunas de sus características (Aplin, 2004; Giri, 2020; Treitz and Rogan, 2004). Hoy en día, el uso de la percepción remota representa uno de los enfoques más utilizados para estudiar la dinámica de la superficie terrestre, debido a su capacidad para estudiar grandes extensiones y contar con un archivo histórico (Boyd and Danson, 2005; Melesse et al., 2007b; Pettorelli et al., 2018).

Existe una diversidad muy amplia de sensores que permiten obtener información de la superficie terrestre, los cuales incluyen sensores hiperspectrales, multispectrales, radar de apertura sintética (SAR), LiDAR, entre otros (Zhu et al., 2017a). Cada uno de ellos es sensible a un espectro electromagnético distinto y brinda información de distinta naturaleza. Por ejemplo, las imágenes multispectrales pueden contener información sobre el color de la superficie, mientras que las SAR pueden brindar información sobre el volumen de cubiertas arbóreas, diferenciar entre distintas distribuciones espaciales de (*e.g.*, distribución aleatoria o uniforme) o humedad de la superficie (Clerici et al., 2017; Joshi et al., 2016; Monsalve-Tellez et al., 2022). Por ello, se ha sugerido que el uso de más de un tipo de imagen producida mediante un sensor remoto ayudaría a obtener más información de la superficie y obtener mejores resultados para identificar eventos como la deforestación o la degradación forestal (Chen et al., 2018; Forkuor et al., 2020; Joshi et al., 2016; Sinha et al., 2019; Zhang et al., 2019c).

Imágenes multispectrales

Las imágenes multispectrales forman parte de la información producida por sensores pasivos. Esto quiere decir que son sensores que únicamente reciben la señal electromagnética producida por otra fuente (la cual suele ser el sol) que refleja la superficie terrestre (Awange and Kyalo Kiema, 2013; Zhu et al., 2017a). A partir de los distintos patrones de reflexión de la superficie en distintas longitudes de ondas, llamadas bandas, se puede inferir ciertas características de la superficie (Kerr and Ostrovsky, 2003; Lillesand and Kiefer, 2000). El intervalo de longitud de onda al que son sensibles estas imágenes suele abarcar el espectro visible de la luz (aprox. 450 – 700 nm) y del espectro infrarrojo

cercano (*ca.* 850 nm), medio (también llamado de onda corta, aprox. 1500 – 2300 nm) y térmico (también llamado de onda larga; aprox. 10 000 – 12 500 nm; Chuvieco, 2008).

La proporción de la señal que es reflejada por la superficie hacia el sensor se le conoce como reflectancia, la cual va a estar determinada por la longitud de onda de la señal y las características de la composición de la superficie (Aplin, 2004; Lillesand and Kiefer, 2000). De esta manera, la información proveniente de diferentes bandas brinda información sobre diferentes tipos de cubiertas. Por ejemplo, las superficies cubiertas por vegetación suelen tener baja reflectancia en el espectro del rojo y azul, debido a la absorción de estas longitudes por la clorofila, y alta en el espectro del verde y muy alta el infrarrojo cercano y medio, gracias a la reflexión de estas longitudes por las paredes celulares (Tucker and Sellers, 1986). Este patrón de reflectancia ayuda a distinguir las áreas cubiertas por vegetación de cualquier otro tipo de superficie. Además, se han desarrollado índices de vegetación que resaltan este patrón, los cuales a su vez se han relacionado con características de la vegetación como el índice de área foliar (LAI) o la productividad primaria de la vegetación (Colombo et al., 2003; Steltzer and Welker, 2006). Además, estos índices presentan la ventaja de resumir patrones en más de una banda, lo cual suele reducir el efecto de la contaminación atmosférica o de sombras causadas por nubes o la topografía; sin embargo, también suelen no reflejar variaciones en valores muy altos de biomasa (*i.e.*, saturación; Huete et al., 1997).

Imágenes de radar de apertura sintética

Las imágenes de radar de apertura sintética (SAR) son producidas por sensores activos, ya que el mismo sensor emite y recibe la señal electromagnética con la que obtiene información de la superficie terrestre (Trebits, 1987; Zhu et al., 2017a). En el caso de los sensores SAR, usualmente emiten una señal en la longitud de onda de las microondas (aprox. 1 – 130 cm), la cual es reflejada por la superficie terrestre y medida por el sensor (Moreira et al., 2013; Oliver and Quegan, 2004). Generalmente, la información contenida en las imágenes SAR responde a características geométricas (*i.e.*, volumen de un bosque o distribución espacial de los troncos) y de humedad de la superficie. A pesar de que los sensores radar suelen emitir una señal con una sola longitud de onda o frecuencia, comúnmente tienen más de una polarización para obtener mayor información de la

superficie terrestre. La polarización de la onda se refiere al plano lineal de la oscilación de la parte eléctrica de la onda electromagnética (Brown, 1967; Fitch, 1988). De tal manera, se puede emitir una onda electromagnética con un plano de polarización vertical u horizontal y de igual manera se puede recibir la señal en las mismas polarizaciones (Flores-Anderson et al., 2019). Por lo tanto, las bandas de los sensores SAR normalmente hacen referencia a la polarización en la que se emite la señal, seguida de la polarización en la que se recibe la señal (*e.g.*, VH, vertical - horizontal).

La proporción de la señal emitida que es reflejada por la superficie hacia el sensor se le conoce como coeficiente de retrodispersión (Moreira et al., 2013). La retrodispersión de una superficie va a estar determinada por (Sinha et al., 2015): 1) la longitud de onda de la señal, 2) la humedad de la superficie y 3) las características geométricas de la superficie. En el primer caso, la longitud de la banda del SAR (λ) va a determinar qué superficies se verán como rugosas y la capacidad de penetración de la señal en cubiertas forestales (Flores-Anderson et al., 2019; Oliver and Quegan, 2004). En general, las superficies cuya rugosidad (h) tenga una relación $h < \lambda / 32$ se verán como planas (baja retrodispersión), mientras que las que guarden una relación $h > \lambda / 2$, como rugosas (alta retrodispersión). Por su parte, las bandas de mayor longitud de onda (menor frecuencia) permitirán mayor penetración en las cubiertas forestales y son considerados mejores fuentes de información para el monitoreo de bosques (*e.g.*, bandas P y L; Moreira et al., 2013). En el segundo caso, mientras más húmeda esté una superficie, mayor será su retrodispersión (Flores-Anderson et al., 2019). Por último, de acuerdo a la polarización de la señal, se pueden distinguir tres tipos generales de respuesta (Fitch, 1988; Oliver and Quegan, 2004). El primero corresponde a superficies rugosas, donde la señal co-polarizada VV suele tener los valores más altos de retrodispersión. El segundo corresponde a áreas donde existe una dominancia de objetos verticales, donde se da un doble rebote de la señal, el cual provoca que la retrodispersión sea mayor en la señal co-polarizada, sobre todo en la HH. Por último, los cuerpos con volumen presentan valores altos de retrodispersión en la señal cruzada (HV o VH; Flores-Anderson et al., 2019; Sinha et al., 2015).

Utilizando las imágenes SAR se puede utilizar únicamente la información de la amplitud, fase o ambas (Flores-Anderson et al., 2019). La amplitud contiene la información

de la retrodispersión, mientras que la fase, contiene la parte imaginaria de la información SAR. Por último, en polarimetría se utiliza tanto la señal de retrodispersión, como la información de fase entre canales para monitorear sobre todo deformaciones de la superficie terrestre con precisión centimétrica.

Diferencias entre imágenes multiespectrales y SAR

Existen varias diferencias importantes entre una imagen multiespectral y una SAR. La primera es que las imágenes SAR no suelen ser sensibles a distintas longitudes de onda, sino que suelen ser sensibles a una única longitud de onda (Flores-Anderson et al., 2019; Zhu et al., 2017a). La segunda es que las imágenes SAR, debido a su longitud de onda, no son sensibles a las nubes (Oliver and Quegan, 2004; Sinha et al., 2015). Por lo tanto, permiten obtener información de la superficie terrestre en condiciones de alta nubosidad. Además, como las imágenes SAR son producidas por un sensor activo, se pueden obtener imágenes tanto en el día como en la noche. La tercera diferencia radica en la geometría con la que se registran ambos tipos de imágenes. Generalmente, las imágenes multiespectrales se registran en un ángulo de observación cercano al nadir, mientras que las imágenes SAR se registran en un ángulo oblicuo. Usualmente el sensor SAR se mueve en una dirección, mientras que la señal se emite y recibe en la dirección perpendicular a la del movimiento del sensor. Este modo de captura permite que la resolución de las imágenes sea mayor a la dada únicamente por el tamaño de la antena (ya que el sensor recibe información de la superficie en diferentes ubicaciones espaciales). Sin embargo, el ángulo de captura oblicuo también provoca que la superficie pueda interactuar de manera distinta con la señal de microondas dependiendo de la posición del sensor respecto a la superficie. Por ello, imágenes registradas con diferentes geometrías (*e.g.*, órbita ascendente y descendente) pueden contener información distinta de la misma superficie terrestre, sobre todo en áreas con un relieve accidentado. Por último, las imágenes multiespectrales suelen registrar características únicamente del dosel, mientras que las imágenes SAR suelen penetrar el dosel y obtener información de la estructura del bosque (dependiendo de la longitud de onda).

Estudio de la deforestación y degradación mediante percepción remota

Numerosos esfuerzos se han realizado para cuantificar la deforestación y la degradación forestal de los bosques tropicales e identificar sus patrones espaciales mediante sensores remotos (*e.g.*, Gao et al., 2020; Hirschmugl et al., 2017; Margono et al., 2012; Murray et al., 2015; Pearson et al., 2017; Xiao et al., 2019). Generalmente el uso de la información generada por sensores remotos se utiliza para desarrollar métodos *wall-to-wall*, es decir, que permiten hacer una comparación de cada píxel de una imagen en un tiempo dado contra el mismo píxel en otro tiempo.

Entre las distintas fuentes de información remota que se han utilizado para detectar la deforestación y degradación forestal se encuentran las imágenes multiespectrales (*e.g.*, Sentinel-2, sensibles al espectro visible e infrarrojo) y radar (*e.g.*, Sentinel-1, sensibles a las microondas), así como nubes de puntos LiDAR (Hirschmugl et al., 2017; Zhu et al., 2017a). Las más utilizadas han sido las imágenes multiespectrales debido a varias razones: fueron las primeras imágenes en desarrollarse y por lo tanto constan de los acervos históricos de imágenes más longevos, cuentan con una gran variedad de plataformas y resoluciones, y en la actualidad, muchas de éstas no tienen costo alguno para el usuario final (Banskota et al., 2014; Lu and Weng, 2007; Xie et al., 2008). Sin embargo, el constante desarrollo de nuevos sensores gratuitos continúa impulsando el uso de una mayor variedad de información proveniente de sensores remotos (Melesse et al., 2007b).

Para detectar la degradación y deforestación mediante sensores remotos, existen básicamente dos enfoques: el cualitativo y el cuantitativo (Mitchell et al., 2017). El primero radica en tratar de identificar estos procesos como transiciones entre categorías (*e.g.*, bosque a no bosque o bosque a bosque degradado; (Gebhardt et al., 2014; Rosete-Vergés et al., 2014); mientras que el segundo trata de identificar estos procesos con base en el seguimiento a través del tiempo de alguna variable biofísica (*e.g.*, la biomasa o la cobertura arbórea) para identificar las áreas que presentaron cambios de la variable cuantificada (Chen et al., 2018; Forkuor et al., 2020; Mitchell et al., 2017; Singh et al., 2018). Debido a la naturaleza de los procesos, el primer enfoque suele ser únicamente útil para la detección de la deforestación (aunque existen ejemplos para la degradación forestal también; *e.g.*, Badjana et al., 2017; Souza et al., 2013), mientras que el segundo ha sido más utilizado

para la detección de la degradación forestal (Bullock et al., 2018; Ryan et al., 2012; Souza et al., 2013). A pesar de ello, el último enfoque suele ser más difícil de utilizar, ya que: 1) la modelación de una variable continua suele ir acompañado de mayor error (*e.g.* Ryan et al., 2012), 2) existe una serie de errores en el proceso para relacionar las variables en campo y las imágenes (*e.g.*, errores de medición en campo, error de geoposicionamiento, entre otros; Réjou-Méchain et al., 2019), 3) requiere de datos de campo u otra fuente que permitan cuantificar la disminución de biomasa en el periodo evaluado, lo que puede ser difícil de obtener (aunque ver estudios realizados con LiDAR o con imágenes multiespectrales; Hirschmugl et al., 2017; Meyer et al., 2019; Singh et al., 2018), y 4) el potencial para detectar cambios en biomasa depende del error de los modelos, ya que cambios de una variable menores al error esperado del modelo no se pueden atribuir necesariamente a cambios en el bosque, sino a posibles errores del modelo (Hirschmugl et al., 2020).

Adicionalmente para estudiar la degradación forestal y deforestación mediante sensores remotos existen diversas dificultades (Chauvenet et al., 2015; Xie et al., 2008). Quizás la más evidente radica en que, frecuentemente, cambios en la señal remota no siempre corresponden a alteraciones de alguna característica biofísica de la vegetación (*e.g.*, biomasa aérea o porcentaje de cobertura arbórea; (Mitchell et al., 2017; Ryan et al., 2012). Por ejemplo, un cambio en la señal remota puede equivaler a una variación estacional (*e.g.*, caducifoliedad, presencia de flores, humedad en los tejidos vegetales, entre otros), o deberse a la presencia artefactos causados por la presencia de nubes o sombras (sobre todo en el caso de las imágenes multiespectrales; (Hirschmugl et al., 2017; Joshi et al., 2016; Singh et al., 2018). Por ello, se han propuesto dos estrategias para superar las dificultades asociadas al a detección de los procesos de degradación forestal y deforestación mediante sensores remotos, las cuales se pueden utilizar por separado o en conjunto: 1) contar con mediciones continuas; y 2) utilizar información remota complementaria. Ambos casos se ven beneficiados por utilizar dos tipos de imágenes (*e.g.*, imágenes multiespectrales y radar).

En el primer caso, el contar con un registro continuo permite caracterizar las variaciones estacionales, lo cual facilita la discriminación entre estos patrones y cambios reales provocados por una degradación forestal o deforestación (Banskota et al., 2014;

Reiche et al., 2018). En muchos casos, el uso exclusivo de imágenes multiespectrales no permite contar con un registro continuo, debido a las condiciones de nubosidad (Gómez et al., 2018; Hirschmugl et al., 2017; Mitchell et al., 2017). Por ello, existen estudios donde se han utilizado imágenes radar, las cuales son poco sensibles a las nubes para contar con medidas más frecuentes (Scarpa et al., 2018) y facilitar la identificación de los procesos de degradación forestal y deforestación (Godinho Cassol et al., 2019; Reiche et al., 2015; Zhang et al., 2019c).

En el segundo caso, el uso combinado de dos regiones del espectro electromagnético puede ser aprovechado para obtener información más detallada de la superficie terrestre (Baek et al., 2011; Joshi et al., 2016). Por ejemplo, al utilizar información radar y multiespectral se puede contar con información sobre la actividad fotosintética de la superficie, así como algunas características geométricas y de humedad (Bai et al., 2008; Joshi et al., 2016; Rignot et al., 1997). Por lo tanto, se ha reportado que el uso combinado de información multiespectral y radar ha permitido mejorar los resultados para detectar la degradación y la deforestación en algunos bosques tropicales (Hirschmugl et al., 2020; Hirschmugl et al., 2017; Singh et al., 2018) o modelar su biomasa (Crowley et al., 2019; Forkuor et al., 2020; Ghasemi et al., 2011; Sinha et al., 2015; Sinha et al., 2019). Por último, cabe resaltar que en los trabajos donde se evalúa las capacidades para monitorear deforestación y degradación forestal mediante imágenes multiespectrales o SAR, usualmente se menciona que la información multiespectral sola suele obtener mejores resultados que utilizando sólo las SAR (Heckel et al., 2020; Zhao et al., 2016).

En la literatura se pueden encontrar ejemplos sobre la cuantificación de la deforestación mediante el uso de imágenes ópticas a nivel global utilizando enfoques de clasificación con machine learning *e.g.*, Hansen et al., 2013; Song et al., 2018). Sin embargo, este no es el caso para el estudio de la degradación, donde no se encontró algún estudio a nivel global. Algunos estudios sugieren utilizar diferentes métodos para detectar la degradación forestal provocada por diferentes actividades (*e.g.*, extracción de leña, tala selectiva, fuegos), lo cual implica que se requeriría de una variedad de métodos para detectar la degradación forestal con información proveniente de sensores remotos (Gao et al., 2020). A pesar de estos esfuerzos, diversos trabajos han criticado este tipo de

evaluaciones globales, sobre todo para trabajar en contextos regionales (Milodowski et al., 2017; Sannier et al., 2016). Por ello, hasta ahora el estudiar los procesos a escalas regionales o locales parecen tener mejores resultados, *e.g.*, Hansen et al., 2016; Reiche et al., 2015; Singh et al., 2018; Whittle et al., 2012).

En el pasado, en México se había optado por utilizar un enfoque basado en las series de tipos de vegetación y uso de suelo del INEGI escala 1: 250 000 para detectar los procesos de deforestación y degradación (Sánchez Colón et al., 2009; SEMARNAT, 2016), o en conjunto con información del inventario forestal nacional (Leyva-Ovalle et al., 2017; Mas et al., 2009; Rosete-Vergés et al., 2014). A pesar de ello, muchos de estos productos no tienen la resolución a la que suelen ocurrir los procesos de disminución de biomasa vegetal o permiten detectarlos una vez que abarcan una superficie muy grande (Gebhardt et al., 2014; Ochoa-Gaona et al., 2007).

El primer esfuerzo en México para realizar este tipo de evaluaciones a mayor resolución es el programa MAD-Mex, donde parte del proceso utiliza las series de tipos de vegetación y uso de suelo del INEGI e información óptica con técnicas de machine learning (Gebhardt et al., 2014; Nielsen, 2007). A pesar de este esfuerzo, este método presenta algunas complicaciones para detectar la degradación y deforestación con percepción remota, derivadas del gran tamaño del país, la enorme heterogeneidad de las condiciones ambientales y de la vegetación, entre otras (Mas et al., 2016; Schmidt et al., 2016). Además, la escala de detección de este método es de 1: 100 000, la cual puede resultar aún muy gruesa para la detección de estos procesos a escala local o regional (Gebhardt et al., 2014).

Programa Copernicus

El programa Copernicus es el programa de observación de la Tierra de la agencia espacial europea que tiene como misión generar información para monitorear diferentes aspectos de la Tierra (ESA, 2019; Malenovský et al., 2012). Por ello, entre los sensores que se encuentran en este programa se encuentran el Sentinel-1 (SAR), Sentinel-2 (multiespectral), Sentinel-3 (SAR y multiespectral) y Sentinel-5 (multiespectral). De estos sensores, los primeros dos están enfocados al monitoreo de la superficie terrestre, mientras que el tercero y el cuarto, a la superficie acuática y atmosférica, respectivamente. Uno de los aspectos

más atractivos para utilizar la información generada por los sensores Sentinel-1 y -2 para estudiar la superficie terrestre radica en que todos son de distribución libre y presentan una alta resolución tanto temporal, como espectral y espacial, es decir, que registran información de la superficie en intervalos de 2-6 días, presentan entre 2 y 13 bandas y tienen un tamaño de píxel de 10 – 60 m.

Sentinel-1 es una constelación SAR de banda C (longitud de onda aprox. 5.5 cm) que cuenta con dos satélites en órbita (ESA, 2020a). Esta constelación puede registrar imágenes en cuatro modos distintos: Stripmap, Interferometric Wide Swath (IWS), Extra Wide Swath y Wave y tiene la capacidad de registrar la señal de microondas en cuatro polarizaciones: VV, HH, VH y HV (ESA, 2020a). Aunque no se describirán a detalle los cuatro modos de registro de información, cada modo tiene una resolución espacial distinta; sin embargo, el IWS es uno de los modos más utilizados ya que permite obtener un buen balance entre superficie cubierta y una alta resolución de píxel (10 m) y es el modo en el que la mayoría de la información es registrada (Heckel et al., 2020). Aunque en teoría este modo puede contar con bandas VV, HH, VH y HV, en la práctica, la mayoría de las imágenes para la superficie terrestre mexicana cuentan únicamente con las bandas VV y VH. La información que contienen estas bandas corresponde al coeficiente de retrodispersión (σ^0) en dB (ESA, 2020a).

Por su parte, Sentinel-2 es una constelación multispectral que cuenta con dos satélites en órbita que tiene un único modo de registro de la información. Las imágenes Sentinel-2 cuentan con 13 bandas espectrales. Por orden de longitud de onda, las bandas corresponden a coastal aerosol (CA, 443 nm), Blue (B, 490 nm), Green (G, 560 nm), Red (R, 665 nm), Red edge-1 (RE1, 705 nm), Red edge-2 (RE2, 740 nm), Red edge-3 (RE3, 783 nm), Near Infrared (NIR, 842 nm), Near Infrared narrow (NIRn, 865 nm), Water Vapor (WV, 945 nm), Cirrus (1380 nm), Short Wave Infrared-1 (SWIR1, 1910 nm), Short Wave Infrared-2 (SWIR2, 2190 nm; ESA, 2020b). Cuatro bandas tienen una resolución espacial de 10 m (B, G, R, NIR), seis con una resolución de 20 m (RE1, RE2, RE3, NIRn, SWIR1, SWIR2) y tres bandas con una resolución de 60 m (CA, WV y Cirrus; (ESA, 2020b). Además, las imágenes Sentinel-2 cuentan con una banda de la calidad de los píxeles para discriminar nubes y sombras de píxeles cuya reflectancia corresponde a la de la superficie terrestre.

Algoritmos de aprendizaje automatizado

Los algoritmos de aprendizaje profundo o *deep learning* son un conjunto de métodos que forman parte del campo del aprendizaje automatizado o *machine learning*, cuyo nombre hace referencia a que poseen arquitecturas profundas que les permiten incorporar relaciones no lineales y reconocer patrones complejos (Ghatak, 2019; Trask, 2019). Gracias a ello, en algunas aplicaciones se considera que los algoritmos de aprendizaje profundo pueden alcanzar resultados comparables con los de un humano (Buduma, 2017; LeCun et al., 2015). En la década de los 2010, el uso de algoritmos de aprendizaje profundo comenzó a permear en áreas interesadas en realizar análisis mediante imágenes, como en la caracterización de la superficie terrestre y su dinámica mediante percepción remota con resultados muy prometedores (Kemker et al., 2018; Li et al., 2018; Zhang et al., 2016).

Antes de la aparición de los algoritmos de aprendizaje profundo, gran parte de la investigación enfocada en generar mapas temáticos mediante percepción remota se basaba en buscar nuevas transformaciones de los datos que facilitaran la separación de las clases de interés (*e.g.*, índices de vegetación; Huete et al., 1997). A medida que se empezaron a desarrollar imágenes con mayor resolución espacial, espectral y temporal, se comenzaron a utilizar rasgos de las imágenes que no sólo tomaran en cuenta la información espectral (*e.g.*, reflectancia de las bandas), sino también las relaciones espaciales entre píxeles (*e.g.*, textura; (Franklin et al., 2001a; Franklin et al., 2001b; Puissant et al., 2005; Zhu and Yang, 1998). El desarrollo de estos rasgos fue acompañado de un aumento general en el desempeño de estos métodos para describir y monitorear la superficie terrestre.

Frecuentemente, estos trabajos utilizaron algoritmos de aprendizaje automatizado para determinar los valores umbrales de dichos rasgos que permitieran diferenciar entre las clases de interés con el menor error posible (Ali et al., 2015; Lary et al., 2016; Li et al., 2013). Por ejemplo, algunos de los algoritmos más populares incluyen: Decision Trees, Random Forests, Support Vector Machines o Redes neuronales artificiales (Banskota et al., 2014; Gómez et al., 2018; Hamdi et al., 2019; Otukei and Blaschke, 2010). En este tipo de trabajos “tradicionales” se requería de un paso que consistía en la ingeniería de rasgos (*feature engineering*), es decir, calcular una transformación de los datos que facilitaran la discriminación entre clases de interés (Alom et al., 2019; Kattenborn et al., 2021). Sin

embargo, *a priori* resultaba difícil saber cuál transformación será la óptima para resolver un problema de clasificación, sobre todo al cambiar de sitio de estudio o de esquema clasificatorio.

En este contexto, la aparición de los algoritmos de aprendizaje profundo supuso un cambio de paradigma, ya que no requieren de este paso de ingeniería de rasgos, sino que el mismo algoritmo es capaz de probar distintas transformaciones de los datos, e inclusive generar nuevas transformaciones que permitan maximizar la precisión de las clasificaciones (Chollet and Allaire, 2018; LeCun et al., 2015). Además, estos algoritmos presentan otras ventajas como el reconocer patrones a distintas escalas espaciales o temporales, lo cual les permite obtener mayor información de la superficie terrestre (Isaienkov et al., 2020; Pelletier et al., 2019; Zhang et al., 2016; Zhang et al., 2019c). Por otro lado, este tipo de algoritmos se ven beneficiados por un número muy grande de observaciones, ya que el error tiende a disminuir con el aumento en el número de muestras (Alom et al., 2019; Ball et al., 2017; Chollet and Allaire, 2018). Además, la diferencia en la precisión para alguna tarea utilizando un algoritmo de aprendizaje automatizado y uno profundo suele ser más evidente en estos casos.

El uso de los algoritmos de aprendizaje automatizado o aprendizaje profundo usualmente se implementan en un esquema supervisado, lo cual se refiere a que al algoritmo se le proveen muestras de las entradas (*e.g.*, imágenes multiespectrales) y las salidas objetivo (*e.g.*, una clasificación; (Cireşan et al., 2003; Gómez et al., 2018; Krizhevsky et al., 2017; Lary et al., 2016; Li et al., 2013; Otakei and Blaschke, 2010; Rogan et al., 2008). De esta manera, el algoritmo ajusta de manera iterativa sus parámetros para minimizar el error entre las predicciones realizadas y las salidas objetivo. Cuando este procedimiento trata de reconocer variables cualitativas (*e.g.*, bosque, no bosque) se le conoce como clasificación, por el contrario, cuando se trata de variables cuantitativas (*e.g.*, biomasa) se le llama regresión (Lary et al., 2016).

Redes neuronales artificiales

El modelo más sencillo utilizado en aprendizaje profundo es el de una red neuronal artificial (ANN), la cual consta de varias capas de neuronas conectadas entre sí. Cada neurona se puede interpretar como una unidad que convierte sus entradas en una salida

mediante tres pasos: 1) una transformación lineal de sus entradas, 2) la suma de todas las transformaciones de sus entradas y 3) una función de activación (Figura 1; (Buduma, 2017). La transformación lineal de cada entrada depende de dos parámetros: peso (w) y sesgo (b) y es transformada mediante la fórmula $wx + b$. El número de pesos de una neurona será igual al número de atributos (x) que reciba como entradas más un único valor de sesgo (Buduma, 2017; Ghatak, 2019). Una vez que se transformaron linealmente todas las entradas de acuerdo a sus pesos, se suman estas transformaciones y después se le aplica la función de activación. Esta función permite incorporar transformaciones no lineales entre la entrada y la salida, lo cual posibilita la captura de patrones más complejos. Algunas de las funciones de activación más comunes incluyen ReLU (Rectified Linear Unit), tan (tangente), sigmoide, entre otras (Chollet and Allaire, 2018).

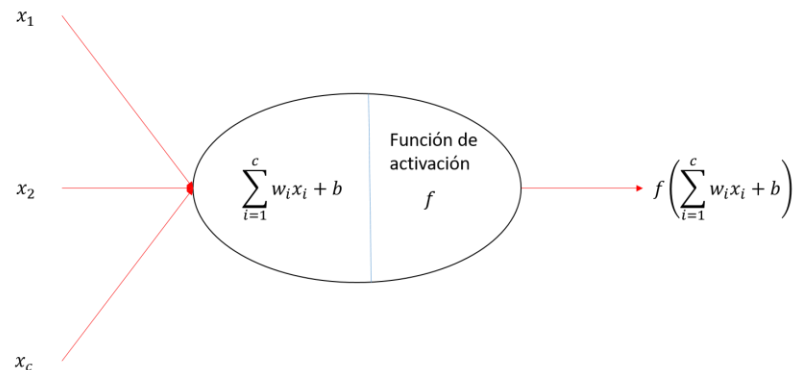


Figura 1. Esquema de una neurona.

Al interconectar varias neuronas entre sí, se conforma una ANN, en la cual el conjunto conectado de entradas y salidas de las neuronas van a obtener una salida (Figura 2). En una ANN, al conjunto de salidas de una capa neuronal se le conoce como activación y se indica como $a^{[l]}$, donde l corresponde al nivel de capa (Ghatak, 2019). Debido a que las muestras de entrenamiento no se consideran una capa, éstas se indican como $a^{[0]}$, es decir la activación en la capa 0. Sin embargo, en las capas subsecuentes las activaciones representan el resultado de la función llevada a cabo por cada neurona y sus parámetros. De manera conjunta, a todas las activaciones de una capa se le denota $a^{[l]}$, mientras que a todos

los pesos de una capa $W^{[l]}$ y los sesgos $b^{[l]}$. Finalmente, el resultado de la última capa ($a^{[L]}$) contiene el resultado predicho (\hat{y}) para una única muestra, a partir de todos sus atributos (x_{1-3} ; Ghatak, 2019). En la Figura 2, los atributos (x_{1-3}) podrían corresponder a la información espectral de una imagen en las bandas roja, verde y azul, mientras que la salida podría corresponder a una predicción de clase (\hat{y}), por ejemplo, si la imagen muestra un bosque o no.

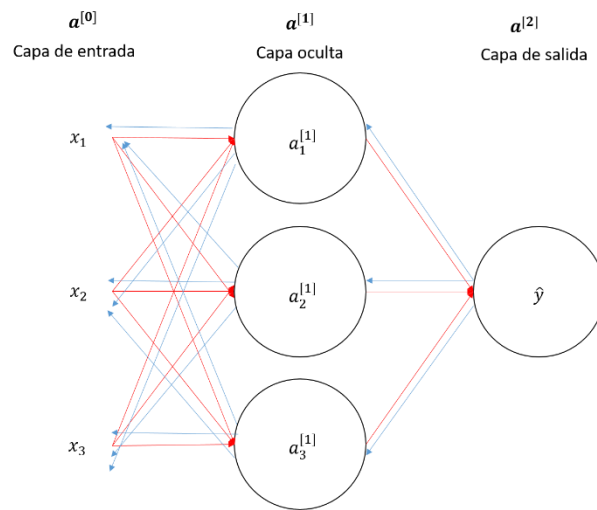


Figura 2. Ejemplo de una red neuronal artificial. Las flechas rojas representan la propagación hacia adelante, mientras que las flechas azules representan la propagación hacia atrás.

Existen dos fases principales en el entrenamiento de una ANN (Chollet and Allaire, 2018): 1) la propagación hacia delante (*forward propagation*) y 2) la propagación hacia atrás (*back propagation*). En la primera fase, se pasan todas las muestras de entrenamiento por la red y se obtiene una predicción para cada muestra. A continuación, se calcula una función de pérdida (*loss function*) que representa una medida de la sumatoria del error entre la clase predicha (\hat{y}) y la clase real (y) de todos los ejemplos de entrenamiento (Ghatak, 2019). Cabe mencionar que el resultado de la predicción, depende enteramente de los parámetros de pesos (w) y sesgos (b) de las neuronas, así como de las funciones de activación utilizadas (f). Una vez que se calculó la función de pérdida, se puede calcular la derivada del valor de la pérdida en función de cada uno de los parámetros anteriores utilizando la regla de la cadena (Chollet and Allaire, 2018). De tal manera, el algoritmo

puede identificar la magnitud del cambio esperado sobre el resultado predicho en función de alteraciones en cada uno de los parámetros de las neuronas. Mediante este mecanismo iterativo de pases hacia adelante y hacia atrás, la red es capaz de fijar los parámetros de las neuronas (w y b) que permitan minimizar la función de pérdida, este procedimiento representa el “aprendizaje” de una ANN (Buduma, 2017).

Por último, las ANN requieren de una función que les indique en qué dirección modificar sus parámetros en función del valor de pérdida de cada iteración, a las cuales se les conoce como funciones de optimización (Chollet and Allaire, 2018). Esto permite que el algoritmo mueva sus parámetros en la dirección de un mínimo global para alcanzar una solución óptima. Dentro de estas funciones se encuentran SGD (*stochastic gradient descent*), RMSprop, Adam, entre otros (Chollet and Allaire, 2018).

Redes neuronales convolucionales

Dentro del aprendizaje profundo se puede encontrar una cantidad enorme de tipos de modelos y arquitecturas de redes neuronales, utilizadas en distintas áreas y con distintos propósitos. Por ejemplo, se pueden encontrar redes neuronales recurrentes, redes adversarias generativas, redes neuronales convolucionales, entre otras (Alom et al., 2019). A pesar de ello, para aplicaciones enfocadas en trabajar con imágenes, el modelo más común son las redes neuronales convolucionales (CNN; (Alom et al., 2019; Ball et al., 2017; Mountrakis et al., 2018; Yuan et al., 2020; Zhang et al., 2016; Zhu et al., 2017b).

La estructura de una CNN es muy parecida a una ANN; sin embargo, la principal diferencia radica en que en lugar de aprender los pesos de muchas neuronas, en una CNN el aprendizaje se da en los pesos de los filtros. Los filtros se pueden interpretar como una ventana de tamaño fijo con ciertos pesos, los cuales permiten reconocer de patrones espectrales-espaciales entre los píxeles (w ; Figura 3; Hoeser and Kuenzer, 2020). Entre las principales ventajas de las CNN respecto a las ANN al trabajar con imágenes radica en que pueden aprender jerarquías espaciales de patrones (Chollet and Allaire, 2018; Zhang et al., 2016). Esto quiere decir que una vez que reconocen un patrón lo pueden identificar en cualquier posición de una imagen y pueden aprender patrones espaciales a distintas escalas. Por otro lado, otra ventaja de las CNN respecto a las ANN radica en que requieren de un número mucho menor de parámetros, ya que los pesos de los filtros se fijan respecto a áreas

de las imágenes y no respecto a todas las neuronas (Alom et al., 2019). Esto provoca que el número de parámetros de una CNN dependa del número de filtros, de capas y el tamaño de los filtros; mas no, del tamaño de las imágenes de entrada, como sí ocurriría con una ANN tradicional. A pesar de las diferencias entre las ANN y las CNN, estas últimas trabajan con el mismo principio iterativo de las ANN para encontrar el valor de los parámetros que permitan minimizar el valor de la función de pérdida mediante pasos hacia delante y hacia atrás, y utilizan las mismas funciones de activación y de optimización.

Los filtros de una CNN, además de los pesos y el sesgo, tienen un par de parámetros adicionales (*padding* y *stride*) que indican cuántos píxeles se agregan en los bordes exteriores de cada imagen de entrada para evitar una reducción del tamaño de la imagen de salida (*padding*) y la distancia en píxeles en las que se van a evaluar las relaciones espectrales-espaciales (*stride*). Estos parámetros, en conjunto con el tamaño de los filtros determinan las dimensiones de la salida. Cada filtro tiene el mismo número de bandas que la imagen de entrada para hacer una multiplicación por banda; sin embargo, al final todos los valores del mismo píxel se suman para dar lugar a una imagen de un solo canal por filtro (Figura 3).

La arquitectura típica de una CNN consiste en una entrada, una serie de capas ocultas y una salida (LeCun et al., 2015). La entrada consiste de las muestras de entrenamiento que se utilizarán como base para que el algoritmo encuentre los atributos que permitan estimar de mejor manera la variable de interés (la salida). Por su parte, en las capas ocultas se pueden diferenciar dos tipos de subcapas: las capas de convolución (*convolutional layer*), y las capas de reescalamiento de la información (*pool layer*; Song et al., 2019; Zhang et al., 2016). Al igual que en una neurona de una ANN, la capa de convolución de una CNN está compuesta por dos partes (Figura 3): 1) los filtros convolucionales y sus parámetros y 2) la función de activación. Como se mencionó anteriormente, los filtros permitirán identificar patrones espectrales-espaciales, mientras que la función de activación permite introducir relaciones no lineares dentro de la red, lo cual facilita la identificación de patrones complejos. Las capas de reescalamiento de la información corresponden a capas en las cuales se reduce la cantidad de información tratando de resumir patrones importantes (Figura 4; (Zhang et al., 2016). Por ejemplo, una capa de reescalamiento muy utilizada es la

de tomar el valor máximo de una imagen de entrada en secciones de ventanas determinadas para reducir a la mitad las dimensiones espaciales de la imagen, pero manteniendo la información más relevante (Figura 4). Las principales diferencias entre una capa convolucional y una de reescalamiento son: 1) los filtros de la capa convolucional contienen parámetros a ser entrenados (pesos y sesgo), mientras que los de una de reescalamiento no, y 2) la capa de reescalamiento modifica las dimensiones espaciales de la imagen (alto y ancho), mientras que la convolucional modifica la dimensión espectral de la imagen (número de bandas). Existen otros tipos de capas como las capas completamente conectadas o las capas *softmax*, utilizadas sobre todo en la parte final de las redes para clasificar el resultado en alguna de las clases de interés (Alom et al., 2019; Song et al., 2019). Sin embargo, éstas pueden estar presentes o no en distintas arquitecturas de redes.

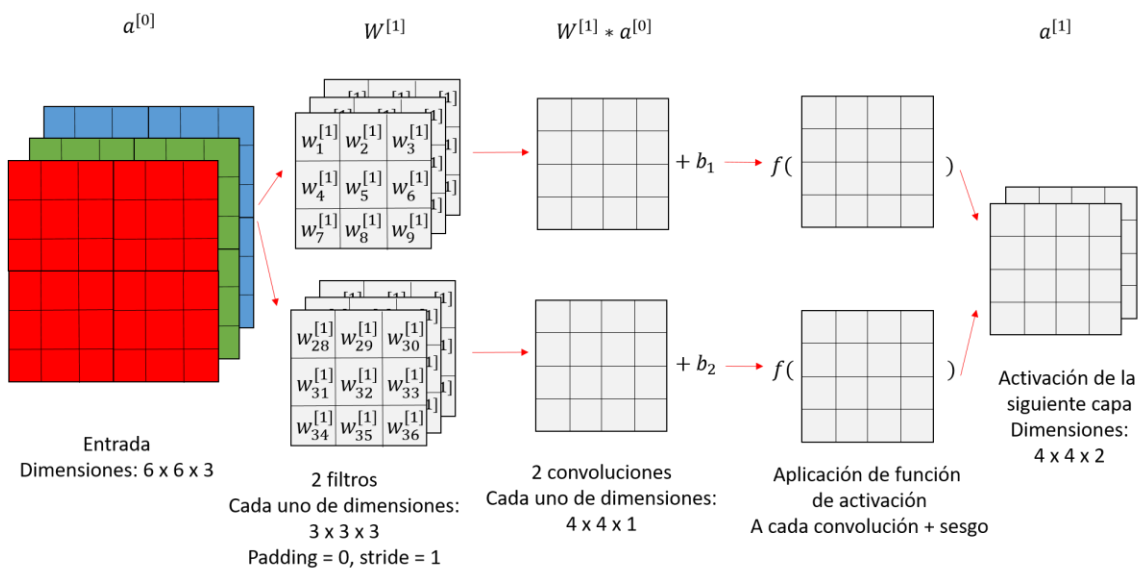


Figura 3. Ejemplo de una capa convolucional de una red neuronal convolucional (CNN).

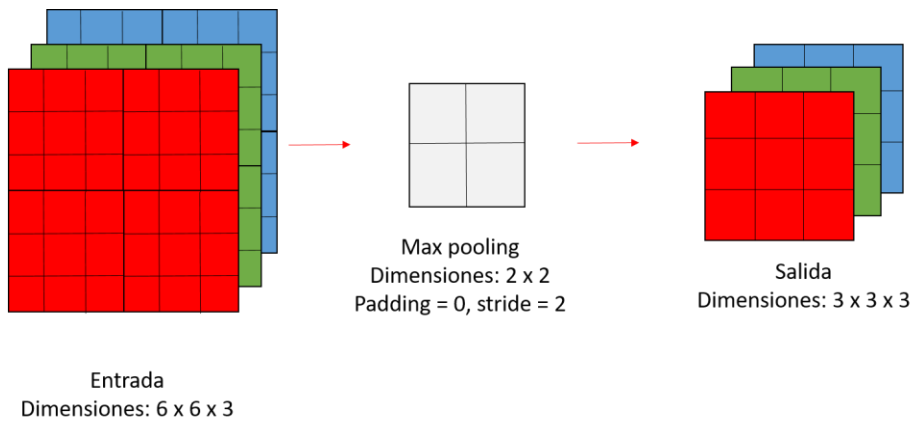


Figura 4. Ejemplo de una capa de reescalamiento de la información de una red neuronal convolucional (CNN).

Los filtros convolucionales operan sobre arreglos de 3D llamados tensores, es decir, matrices con tres ejes de información: alto, ancho y profundidad. Las primeras dos se pueden interpretar como la posición de un píxel en una imagen, mientras que la profundidad se puede interpretar como el valor de ese en distintas bandas de una imagen. Aunque propiamente la dimensión espectral constituye otra dimensión de los datos, en la práctica, no se toma como tal, sino que sólo se toma como diferentes atributos de una misma observación. Por ello, los datos que cuentan con dos dimensiones espaciales (alto y ancho) y una espectral, usualmente se procesan mediante convoluciones de 2D, las cuales sintetizan los patrones espectrales – espaciales en todas las bandas de las observaciones. Por su parte, en arquitecturas CNN más complejas se pueden realizar convoluciones 3D sobre tensores que incluyan la dimensión temporal. En este caso, las convoluciones se realizan sobre las dos dimensiones espaciales y una temporal.

Hiperparámetros

Los hiperparámetros se refieren a los parámetros que controlan el aprendizaje de cualquier algoritmo de aprendizaje profundo. Algunos de los más importantes incluyen la función de pérdida, el algoritmo de optimización, la tasa de aprendizaje, el número de iteraciones (*epochs*), el número de capas ocultas y de filtros por capa oculta, las funciones de activación y algunos otros parámetros como el tamaño de lote (*batch*) o el uso de alguna

técnica de regularización (Chollet and Allaire, 2018; Goodfellow et al., 2016; Mueller and Massaron, 2019; Trask, 2019).

La función de pérdida es la función que el algoritmo tratará de minimizar ya que es una medida del error entre los datos de entrada y las predicciones. El algoritmo de optimización determinará la función mediante la cual el algoritmo se moverá para encontrar la solución óptima. Por su parte, la tasa de aprendizaje va a determinar la rapidez con la que el algoritmo irá moviéndose hacia la solución óptima, basado en el algoritmo de optimización. Una tasa muy alta de aprendizaje puede hacer que la solución diverja en lugar de converger, mientras que un valor muy pequeño puede demorar mucho tiempo en obtener una solución o llegar a un mínimo de la función de pérdida local en lugar de uno global (Alom et al., 2019). Por su parte, el número de iteraciones determinará cuántos pases completos por los datos se realizarán antes de que el algoritmo se detenga. Un valor muy pequeño puede provocar que no se llegue a una solución donde se alcance un óptimo global. En el caso del número de capas y filtros, éstos determinarán el poder de abstracción de la red.

Normalmente se asume que una red más profunda y con más unidades ocultas será capaz de identificar atributos más complejos en las imágenes; sin embargo, también tiene mayor probabilidad de sobreajustar sus parámetros a los datos de entrenamiento, sobre todo si se usa un conjunto de datos pequeño (Cheng et al., 2017; Wagner et al., 2019). Por su parte, las funciones de activación van a determinar las relaciones no lineales que se van a incorporar en la CNN, así como la naturaleza de la salida (Chollet and Allaire, 2018). El tamaño de lote determinará el número de observaciones que se va a utilizar para ir actualizando los parámetros de la red sin necesidad de realizar un pase completo por los datos (una época) para hacer una actualización (LeCun et al., 2015). Por último, los parámetros de regularización evitarán en mayor o menor medida un sobreajuste del modelo a los datos de entrenamiento, lo cual, permite obtener resultados con mayor posibilidad de generalización (Goodfellow et al., 2016).

Tipos de CNN

De acuerdo con (Hoeser and Kuenzer, 2020) las CNN se pueden clasificar en cuatro tipos de aplicaciones (Figura 5): 1) reconocimiento de imágenes, 2) segmentación de imágenes, 3) detección de objetos y 4) segmentación de casos. Las primeras dos aplicaciones

corresponden a tareas enfocadas en resolver un problema de clasificación, ya que buscan asignar una clase a una imagen completa (reconocimiento de imágenes) o a cada píxel de una imagen (segmentación de imágenes; Yuan et al., 2020). Por el contrario, las últimas dos aplicaciones están interesadas en resolver un problema de regresión y clasificación, ya que buscan delimitar la posición espacial de los objetos de interés (detección de objetos), y adicionalmente, asociarle una clase al conjunto de píxeles que representa el objeto (segmentación de caso; Alom et al., 2019; Song et al., 2019).

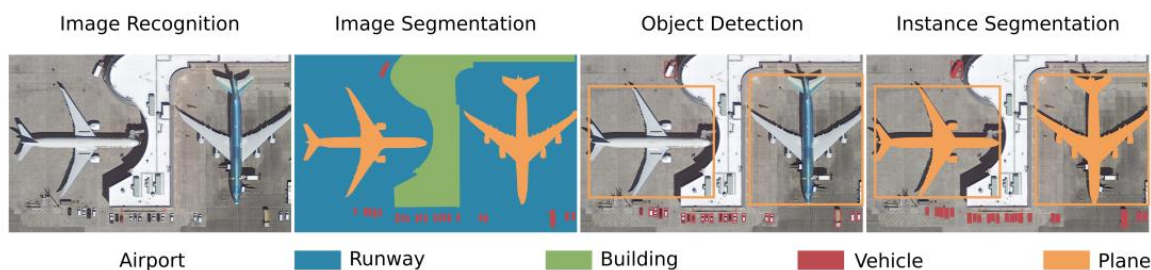


Figura 5. Cuatro aplicaciones distintas de algoritmos de CNN para el procesamiento de imágenes (Fuente: Hoeser and Kuenzer, 2020).

Debido al interés del estudio nos enfocaremos a describir las redes destinadas a realizar tareas de segmentación, ya que permiten asociarle una etiqueta a cada píxel de una escena utilizando su contexto, ya sea espacial o temporal (Alom et al., 2019; Song et al., 2019). Sin embargo, si se desea conocer el panorama más general del uso de los algoritmos de aprendizaje profundo para diferentes aplicaciones se puede consultar (Alom et al., 2019; Ball et al., 2017; Fan et al., 2019; Hoeser et al., 2020; Hoeser and Kuenzer, 2020; Li et al., 2018; Ma et al., 2019; Mountrakis et al., 2018; Yu et al., 2017; Yuan et al., 2020; Yuan et al., 2021; Zhang et al., 2019a; Zhang et al., 2016; Zhao et al., 2014; Zhu et al., 2017b). Cabe resaltar que aunque el término segmentación de imágenes es comúnmente utilizado en el contexto de algoritmos de aprendizaje profundo para generar clasificaciones, dicho término también ha sido utilizado previamente para referirse a la segmentación espaciotemporal para crear objetos en el contexto de la clasificación orientada a objetos. En este sentido, adoptaremos el término de segmentación de imágenes para referirnos a la clasificación de imágenes utilizando algoritmos de aprendizaje profundo.

Arquitecturas de segmentación de imágenes

Dentro de los modelos de segmentación de imágenes, existen varios tipos de arquitecturas; sin embargo, las principales arquitecturas se pueden clasificar de acuerdo al número de dimensiones sobre las que se realizan las convoluciones; una sola dimensión (1-D CNN; temporal), dos dimensiones (2-D CNN; espacial) y tres (3-D; espacial y temporal; Pelletier et al., 2019; Zhu et al., 2017b). Dentro de las CNN utilizadas para tareas de segmentación semántica existen esencialmente dos tipos de modelos: decodificadores simples (*naive decoders*) y codificadores – decodificadores (*encoder – decoder*; Hoeser and Kuenzer, 2020). En los primeros se suele reconstruir la imagen utilizando algún algoritmo de reescalamiento hacia arriba como el vecino más cercano o bilinear. Por el contrario, los modelos codificadores – decodificadores reconstruyen la imagen utilizando el contexto espacial de la misma imagen para obtener una imagen de mayor resolución, mediante conexiones saltadas (*skip connections*), y son los que permiten obtener mejores resultados. La salida de un codificador - decodificador suele constar de una imagen que contiene la probabilidad de pertenencia de cada píxel a cada una de las clases de interés (Zhang et al., 2016). A partir de esta salida, se puede obtener una clasificación al asignar a cada píxel la clase con mayor probabilidad de pertenencia.

A pesar de que muchos de estos modelos de redes convolucionales se desarrollaron para aplicaciones médicas, su uso en aplicaciones de observación de la Tierra ha tenido resultados prometedores, sobre todo para clasificar imágenes en distintos tipos de cubierta y usos del suelo, así como para detectar disturbios en cubiertas forestales (Huang et al., 2018; Neves et al., 2020; Wagner et al., 2020b; Wagner et al., 2019). Una revisión de las principales arquitecturas utilizadas para estudios enfocados en la segmentación de imágenes mostró que existe una tendencia por utilizar arquitecturas más sencillas que las utilizadas para el reconocimiento de imágenes (*e.g.*, ResNet o VGG), probablemente porque en estudio de observación de la Tierra se suelen tener pocos datos de entrenamiento (Hoeser and Kuenzer, 2020). Además, en estos casos, el elegir modelos más sencillos evita sobreajustar los parámetros de la CNN a los datos de entrenamiento (Cheng et al., 2017). La arquitectura más utilizada para tareas de segmentación de imágenes son las arquitecturas codificantes - descodificantes, y en particular el modelo U-Net o derivados de esta

(Diakogiannis et al., 2020; Hoeser et al., 2020; Ronneberger et al., 2015). Este modelo fue desarrollado para aplicaciones médicas, sin embargo, ya ha probado su utilidad para obtener clasificaciones o detectar cambios en la superficie terrestre (Bragagnolo et al., 2021b; Clark and McKechnie, 2020; Du et al., 2020; Flood et al., 2019; Giang et al., 2020; Hamdi et al., 2019; Huang et al., 2018; Isaienkov et al., 2020; Neves et al., 2020; Robinson et al., 2019; Stoian et al., 2019; Ulmas and Liiv, 2020; Wagner et al., 2020a; Wagner et al., 2020b; Wagner et al., 2019; Yi et al., 2019; Zhang et al., 2018). Otras arquitecturas codificadoras – decodificadoras frecuentemente utilizadas son DeconvNet, SegNet, RefineNet, U-Net ++, DeepLabV3+, W-net y algunas versiones completamente convolucionales de redes preexistentes (Alom et al., 2019; Gargiulo et al., 2020; Song et al., 2019).

U-Net

La idea principal de la U-Net, al ser un algoritmo basado en CNN, es que, a diferencia de enfoques de *machine learning* como *random forests* o *support vector machines*, donde cada píxel recibe una categorización de acuerdo únicamente a su información espectral, los píxeles serán clasificados de acuerdo a sus características espectrales y las relaciones espaciales que guarden con sus vecinos a distintas escalas. Esto supone que la U-Net y las CNN en general, son capaces de identificar patrones espaciales como la detección de bordes, la forma o la textura. Dichos patrones quedan registrados en los valores de los pesos de los filtros.

Como lo indica su nombre, la arquitectura U-Net como cualquier arquitectura codificante – descodificante basada en CNN consta de dos partes (Figura 6; Ronneberger et al., 2015). En la primera sección, cada capa oculta consta de un par de subcapas convolucionales, seguidas de una de reescalamiento. Las capas convolucionales se encargan de codificar la información espectral y espacial, mientras que las de reescalamiento se encargan de reducir la dimensión espacial para mantener únicamente los patrones más relevantes (*i.e.*, el valor máximo). Cabe recordar que en los pesos de los filtros convolucionales es donde se “aprenderán” los patrones espectrales y espaciales que permitan reconocer las clases de interés en la imagen de entrada. En la parte final de esta sección de la arquitectura, se obtiene una imagen con un gran número de bandas, en las

cuales se resumen los patrones codificados por los filtros convolucionales a diferentes escalas. Posteriormente, esta imagen se debe de reescalar en el sentido inverso para obtener una imagen en la resolución original de las probabilidades de pertenencia por clase por píxel. Este procedimiento se realiza en la parte descodificante de la U-Net, la cual corresponde a una versión en espejo de la parte codificante. En esta sección de la U-Net, cada capa oculta está conformada por una subcapa de reescalamiento hacia arriba seguida de un par de subcapas convolucionales. Para proveer de información de mayor resolución espacial a las operaciones de reescalamiento hacia arriba, se utilizan las conexiones saltadas para concatenar la imagen de mayor resolución proveniente de la parte codificante con la imagen de menor resolución de la parte descodificante (Figura 6). El último paso de la U-Net consiste en aplicar una función de activación, la cual calcula la probabilidad de pertenencia por clase de cada uno de los píxeles de la imagen en función de la salida de la parte descodificante. En la mayoría de las arquitecturas basadas la U-Net se utiliza la función Softmax para calcular la probabilidad de pertenencia de cada píxel a cada una de las clases de interés.

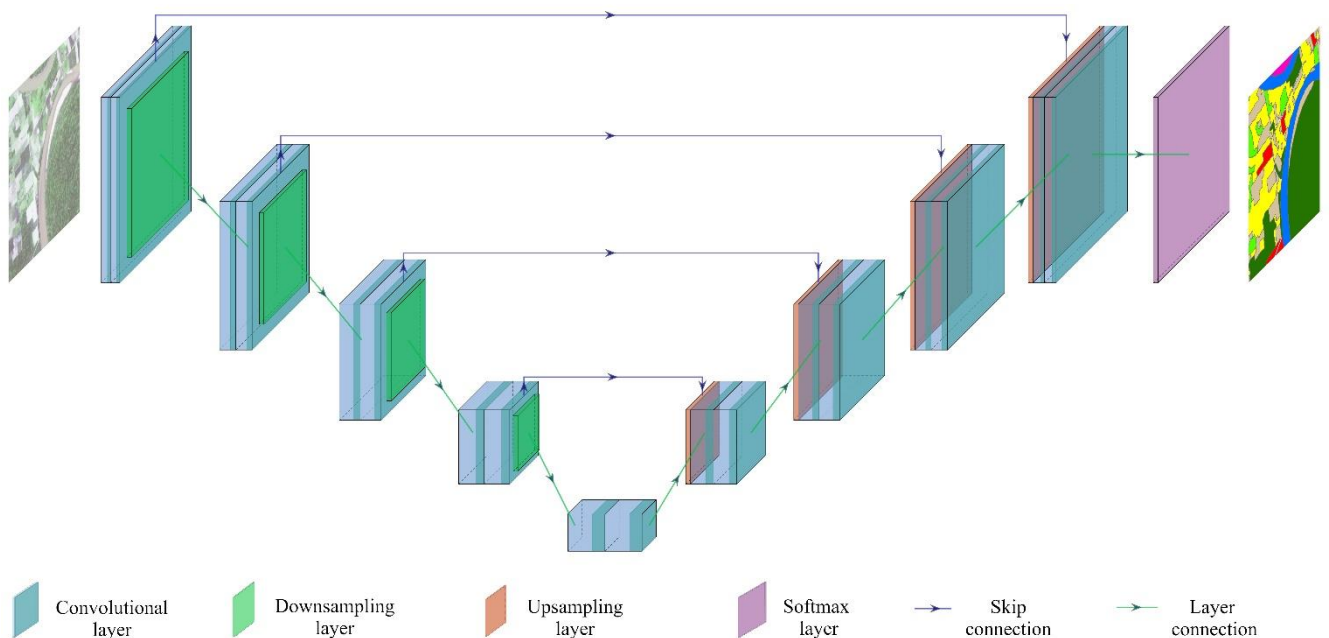


Figura 6. Representación gráfica de una arquitectura U-Net con cuatro capas ocultas. La imagen en el extremo izquierdo representa la imagen utilizada como entrada, mientras que la del extremo derecho, la salida (una clasificación). Figura realizada mediante el código de <https://github.com/HarisIqbal88/PlotNeuralNet>.

Una vez descrita la arquitectura U-Net tradicional, la cual realiza convoluciones en la dimensión espacial para identificar patrones espectrales y espaciales en una imagen, podemos pensar en que se puede crear una U-Net de tres dimensiones, en la que aparte de realizar las convoluciones en la dimensión espacial, las haga en la dimensión temporal. En este sentido, a la U-Net tradicional la llamaremos U-Net simple o U-Net 2D, mientras que a la versión que incluya convoluciones en la dimensión temporal, una U-Net 3D (Isaienkov et al., 2020). De igual manera que una U-Net 2D trata de resumir relaciones espectrales entre píxeles contiguos en el espacio, la U-Net 3D tratará de resumir relaciones espectrales entre píxeles contiguos en el espacio y en el tiempo (Figura 7). Al ver la estructura de una U-Net 3D, se puede observar que la arquitectura es exactamente la misma, lo único que cambia es que las convoluciones se realizan en tres dimensiones en lugar de dos. De igual manera, los filtros convolucionales y de reescalamiento en la U-Net 3D tendrán tres dimensiones.

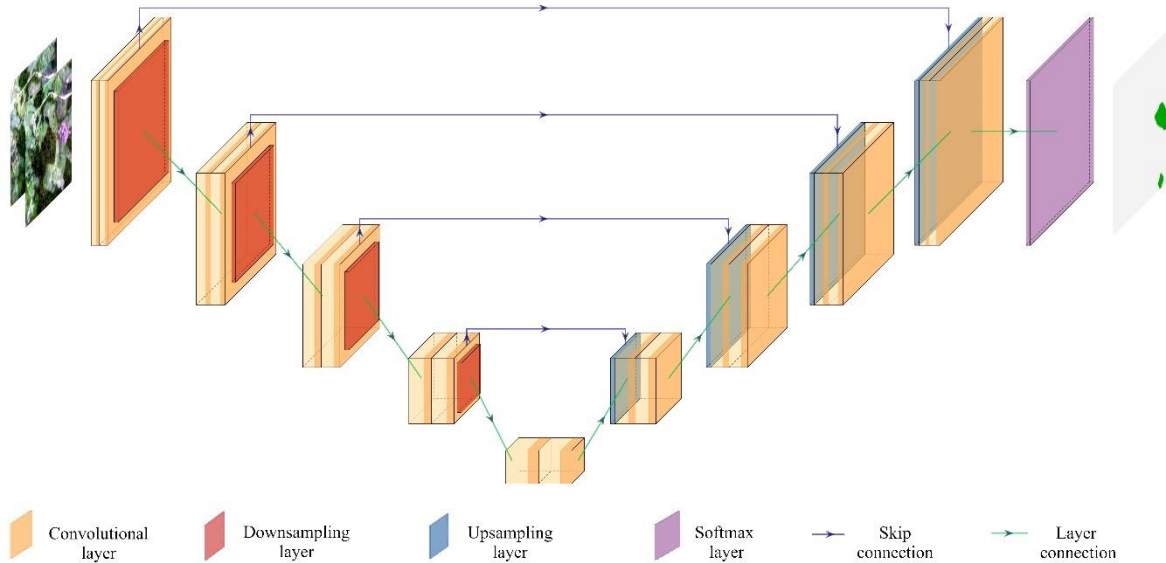


Figura 7. Representación gráfica de la arquitectura U-Net 3D. La imagen en el extremo izquierdo representa un par de imágenes utilizada como entrada, mientras que la del extremo derecho, la salida (una clasificación).

Figura realizada mediante el código de <https://github.com/HarisIqbal88/PlotNeuralNet>.

Justificación del trabajo

Actualmente, existe una necesidad de generar herramientas que permitan mejorar la precisión para caracterizar la cubierta arbórea y monitorear su dinámica, para conocer su estado actual, cuantificar la magnitud de afectación por parte de procesos como la deforestación o la degradación forestal y detectar su tendencia temporal. Adicionalmente, este interés es mayor en zonas tropicales donde se concentra la mayor parte de la deforestación y degradación forestal a nivel mundial, así como los bosques con mayor biodiversidad y carbono secuestrado (FAO, 2020).

Hasta el día de hoy, en México no se conoce algún trabajo que pretenda evaluar el uso conjunto de imágenes radar y ópticas; así como tampoco un esfuerzo por evaluar el uso de algoritmos de aprendizaje profundo en conjunto con estos insumos. A pesar de ello, existen diversos esfuerzos regionales y nacionales basados en información óptica o en una combinación de imágenes e información de campo (CONAFOR, 2022; Gebhardt et al., 2014; Mendoza and Dirzo, 1999; Ochoa-Gaona and González-Espinosa, 2000; Rosete-Vergés et al., 2014; SEMARNAT, 2016; Vidal et al., 2014).

Por ello, este estudio pretende evaluar el potencial de nuevos enfoques como el uso de algoritmos de aprendizaje profundo y el empleo de imágenes multiespectrales y radar para brindar información más detallada respecto a algunas características o usos de la cubierta arbórea, así como evaluar si permite mejorar la precisión para describir y cuantificar los procesos de deforestación y degradación forestal de la cubierta arbórea en zonas tropicales de México. De acuerdo a estudios previos, la aplicación de este enfoque metodológico permitirá obtener mejores resultados que acercamientos previos debido a que cuenta con las siguientes dos ventajas: 1) incorpora información de reflectancia, geometría y humedad de la superficie terrestre, 2) además de la información espectral de las imágenes, incluye información contextual, ya sea en la dimensión espacial o también en la dimensión temporal. Derivado de estos antecedentes, entonces surgen las preguntas de este estudio: ¿el enfoque metodológico permitirá mejorar las capacidades para describir la cubierta arbórea de una región? ¿permitirá mejorar la precisión total para discriminar áreas deforestadas y áreas forestales degradadas que métodos previamente disponibles? El trabajo se

desarrollará en dos sitios de estudio, uno donde domine el proceso de deforestación sobre la degradación forestal y otro donde ocurra lo contrario.

Por último, cabe mencionar que el enfoque utilizado en este trabajo representará un método con alta posibilidad de ser replicado e implementado para el monitoreo de la superficie terrestre ya que se utilizará información remota gratuita (imágenes Sentinel-1 y Sentinel-2) y se utilizaron únicamente programas de distribución libre (R, Python, Google Earth Engine).

Objetivo general

1. Evaluar el uso de imágenes radar y multiespectrales con técnicas de aprendizaje profundo para monitorear la deforestación y degradación forestal en dos tipos de bosques tropicales.

Objetivos particulares

1. Evaluar el número de observaciones sin nubes disponibles en el país en el acervo de imágenes Sentinel-2.
2. Evaluar el uso de un algoritmo de aprendizaje profundo e imágenes multiespectrales y radar para clasificar la cubierta en distintos tipos de vegetación y usos del suelo en un bosque tropical húmedo.
3. Evaluar el uso de un algoritmo de aprendizaje profundo en imágenes multiespectrales y radar para detectar áreas deforestadas en un bosque tropical húmedo.
4. Evaluar el uso de una técnica de aprendizaje profundo en imágenes multiespectrales y radar para detectar áreas degradadas en un bosque tropical caducifolio.

Capítulo 2. Patrones espaciotemporales de las observaciones de Sentinel-2 a nivel de imagen y píxel sobre el territorio mexicano entre 2015 y 2019.

Jonathan V. Solórzano; Jean-François Mas; Yan Gao; J. Alberto Gallardo-Cruz

Artículo publicado en: Revista de Teledetección (2020) 56 (Número especial), 103-115

Resumen

Actualmente, las imágenes Sentinel-2 son uno de los acervos multiespectrales y gratuitos de mayor resolución temporal, espectral y espacial para monitorear la superficie terrestre. Sin embargo, la posibilidad de utilizar este acervo para distintas aplicaciones está condicionada por el número de observaciones sin nubes disponibles para una ventana espaciotemporal determinada. Por ello, este artículo tuvo el objetivo de analizar el número de observaciones de Sentinel-2 disponibles para el territorio mexicano a nivel de imagen y de píxel. En el primer caso, se contabilizó el total de imágenes disponibles por año y su porcentaje de nubosidad; mientras que, en el segundo, se calculó el número de observaciones despejadas por píxel. Además, para tomar en cuenta la diversidad del territorio, se evaluó el promedio mensual de las observaciones por píxel de cada una de las siete ecorregiones del país, así como la proporción de su superficie con por lo menos una observación despejada en intervalos mensuales, bimestrales, trimestrales y anuales. Los resultados mostraron que el número de observaciones válidas por píxel variaron entre 0 y 121 observaciones al año y entre 0 y 6.58 al mes. Adicionalmente, se observó que en el periodo 2017 – 2019 se pueden obtener observaciones de todo el país en ventanas anuales, mientras que en el periodo 2018 – 2019, se pueden obtener observaciones en intervalos mensuales o trimestrales, dependiendo de la ecorregión. Finalmente, consideramos que los resultados de este trabajo servirán de guía para los usuarios interesados en utilizar estas imágenes para distintos estudios.

Palabras clave: México, Ecorregiones, Observaciones sin nubes, Sentinel-2 1C, imágenes satelitales ópticas.

Introducción

El estudio de la superficie terrestre mediante imágenes satelitales se ha consolidado como un área de continua investigación y desarrollo durante el siglo XXI (Boyd and Danson, 2005; Melesse et al., 2007a). Por ello, hoy en día se puede encontrar una diversidad enorme de estudios que utilizan la teledetección en aplicaciones forestales, urbanas, biológicas, geológicas, arqueológicas, entre otras (Agapiou et al., 2014; Ramoelo et al., 2015; Van der Meer et al., 2014; Xiao et al., 2019).

Actualmente, uno de los proyectos de obtención de imágenes satelitales gratuitas más importantes es el programa Copernicus y su constelación de satélites Sentinel que incluyen sensores radar (Sentinel-1 y -3) y multiespectrales (Sentinel-2, -3, -4 y -5; (Berger et al., 2012; Malenovský et al., 2012). Entre ellos, la constelación Sentinel-2 está destinada a monitorear la superficie terrestre y de mares interiores mediante sensores multiespectrales (ESA, 2018). En comparación con misiones multiespectrales gratuitas anteriores, ésta se destaca por tres aspectos principales (Berger et al., 2012; Drusch et al., 2012): 1) su alta resolución temporal (tiempo de revisita de 3 – 5 días en su periodo de máxima capacidad de funcionamiento), 2) su alta resolución espectral (cuenta con 13 bandas en el espectro visible, borde del rojo e infrarrojo) y 3) su alta resolución espacial (bandas con tamaño de píxel de 10 m, 20 m y 60 m). Por ello, Sentinel-2 se perfila como uno de los programas más utilizados en un futuro.

Un factor clave para poder estudiar un fenómeno mediante teledetección radica en la compatibilidad entre las escalas espaciotemporales del fenómeno de interés y de la información remota (Comber and Wulder, 2019; Li and Roy, 2017). En el caso de las imágenes multiespectrales, uno de los principales fenómenos meteorológicos que dificultan el registro de información de la superficie terrestre es la presencia de nubes (Coluzzi et al., 2018; Main-Knorn et al., 2017; Wilson and Jetz, 2016). Por ello, el conocer el número de observaciones válidas disponibles (*i.e.*, observaciones en condiciones despejadas o sin nubes) de un acervo de imágenes determinado para un área y periodo de tiempo particular permitirá: 1) conocer su distribución espaciotemporal, 2) determinar el mejor periodo para obtener mediciones válidas, 3) determinar la resolución temporal en la que se podría

realizar algún tipo de monitoreo y 4) finalmente, evaluar la factibilidad de utilizar estas imágenes para un tipo de estudio en particular (Whitcraft et al., 2015; Wulder et al., 2016).

Previamente, se han realizado trabajos de revisión y descripción de los acervos de imágenes Landsat para determinados países (*e.g.*, Solórzano et al., 2020; White and Wulder, 2014). A pesar de la información valiosa que proveen estos esfuerzos, la mayoría se ha limitado a hacer caracterizaciones a nivel de imagen; por lo tanto, no permiten identificar las diferencias de condiciones de nubosidad entre regiones de un territorio (aunque ver Wilson and Jetz, 2016; Zhou et al., 2019). Además, la posibilidad de realizar descripciones a nivel de píxel había estado restringida por el enorme poder de cómputo requerido para realizarlas, pero el día de hoy, gracias al desarrollo de plataformas de procesamiento en la nube como Google Earth Engine (Gorelick et al., 2017), es posible realizarlas.

En este contexto, el objetivo de este trabajo fue caracterizar el acervo de observaciones Sentinel-2 disponibles para el territorio continental mexicano y sus principales ecorregiones, a nivel de escena y de píxel en el periodo 2015 – 2019. Esta caracterización permitirá identificar los principales patrones temporales y espaciales de las observaciones despejadas y servirán para guiar estudios futuros interesados en utilizar este acervo de imágenes.

Materiales y métodos

Sitio de estudio

México es un país que presenta una enorme heterogeneidad de condiciones climáticas y edafológicas, producto de su compleja fisiografía y ubicación intertropical (Espinosa et al., 2008). Esta heterogeneidad se ve reflejada en la diversidad de ecosistemas terrestres presentes en el país, la cual incluye desde desiertos hasta selvas húmedas (Espinosa et al., 2008). Cada uno de estos ecosistemas presenta un régimen de precipitación distinto, lo cual, a su vez, determina los principales patrones temporales de la presencia de nubes sobre su superficie (Sudmanns et al., 2019; Wilson and Jetz, 2016). Para caracterizar con mayor detalle el número de observaciones válidas por píxel en cada tipo de ecosistema, se utilizó la ecorregionalización en nivel I propuesta por (INEGI et al., 2008). De esta manera, se

dividió al país en siete ecorregiones: California Mediterránea, Desiertos de América del Norte, Elevaciones Semiáridas Meridionales, Grandes Planicies, Selvas Cálido Húmedas, Selvas Cálidos Secas y Sierras templadas (Figura 8).

Imágenes Sentinel-2

Las imágenes Sentinel-2 son distribuidas en escenas (“tiles” o “granules”) que abarcan un área de 110 km x 110 km (12,100 km²), las cuales se sobrelapan 10 km entre escenas contiguas (ESA, 2020c). Estas escenas siguen el Sistema de Referencia de Cuadrícula Militar (MGRS por sus siglas en inglés) con una proyección UTM / WGS84 (ESA, 2020c). En el resto de este trabajo, se usará el término escena para referirse a la clave única de las imágenes, mientras que el término imágenes se utilizará para referirse al registro de una escena en una fecha particular. En el caso de México, un total de 303 escenas cubren su superficie terrestre, las cuales están ordenadas, de acuerdo con las zonas UTM en las que se ubica el país (Figura 8).

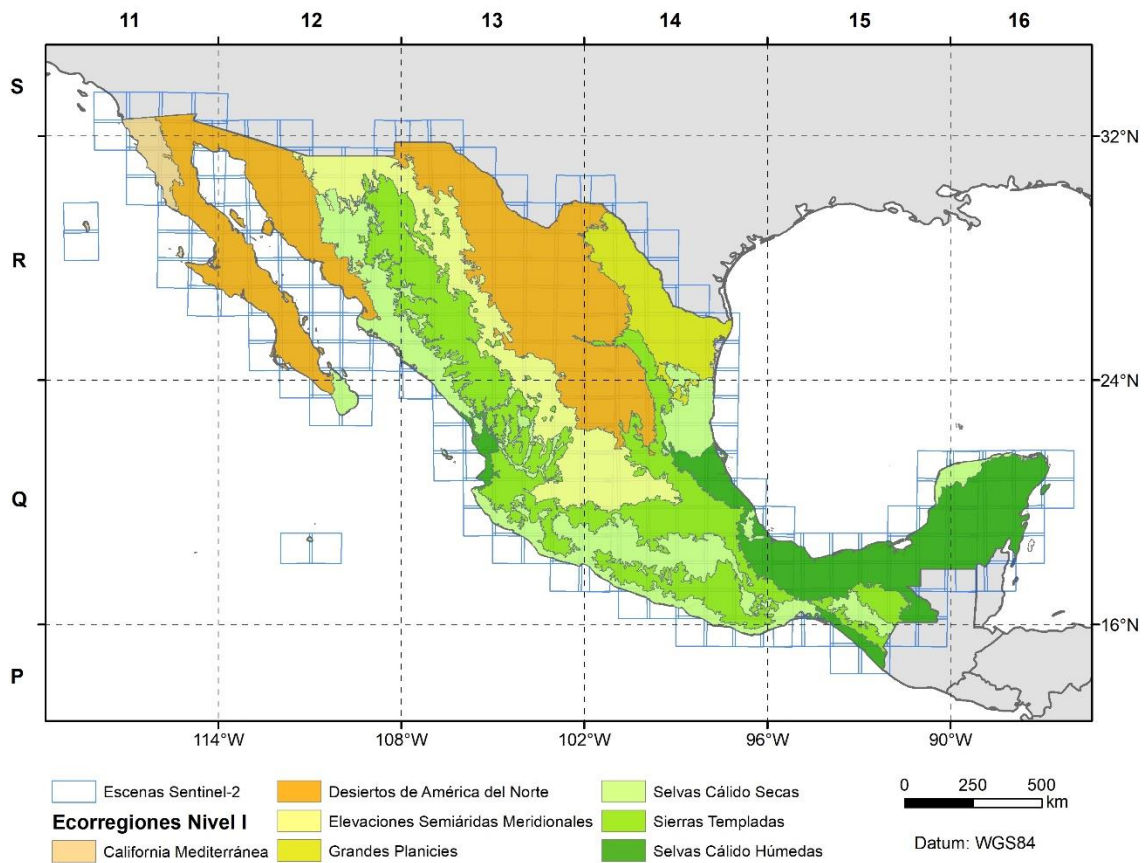


Figura 8. Ubicación geográfica de las ecorregiones nivel I de México de acuerdo con la clasificación de (INEGI et al., 2008). La escala de colores muestra un gradiente general de mayor a menor humedad (verde oscuro a naranja claro). Además, se muestran las escenas de Sentinel-2 que cubren la superficie terrestre de México (Fuente: ESA, 2020c). En los ejes superior e izquierdo se muestran las zonas UTM donde se encuentra el país, de acuerdo con el Sistema de Referencia de Cuadrícula Militar, mientras que en los ejes inferior y derecho, se muestran las coordenadas geográficas que delimitan cada una de estas zonas.

La información a nivel de imagen y de píxel de las imágenes Sentinel-2 se obtuvo en la API de JavaScript de Google Earth Engine (GEE; Gorelick et al., 2017). Actualmente, existen dos colecciones de Sentinel-2 disponibles: la 1C (reflectancia sobre la atmosfera) y 2A (reflectancia de la superficie). Las tres principales diferencias entre estas colecciones son: 1) el nivel de procesamiento, 2) el algoritmo utilizado para detectar nubes y sombras, y 3) el periodo para el cual se encuentran disponibles. A pesar de que la colección 2A es la que tiene un mayor nivel de procesamiento y calidad, las primeras imágenes disponibles del territorio mexicano datan de diciembre de 2018 (Google Earth Engine, 2019). Por el

contrario, las primeras imágenes de la colección 1C son de julio de 2015; por ello, se eligió esta última colección para realizar el análisis. En GEE, se consultaron todas las imágenes Sentinel-2 de la colección 1C que abarcaban la superficie terrestre de México, utilizando un polígono suavizado del contorno de la superficie del país entre 01/07/2015 (mes del lanzamiento de Sentinel-2A; ESA, 2016) y 31/12/2019.

En primer lugar, se construyó una rutina en GEE para obtener la información de las imágenes que cubrieron la superficie terrestre mexicana en el periodo de estudio. Estos datos se obtuvieron de los metadatos e incluyeron: el número total de imágenes disponibles, el sensor que registró cada imagen (*i.e.*, Sentinel-2A o 2B), el porcentaje de la imagen cubierta por nubes y su fecha de captura.

En segundo lugar, se construyó una rutina para obtener la información a nivel de píxel, es decir, el número de observaciones despejadas por píxel. Para ello se utilizó la información de la banda QA60 con una resolución de píxel de 60 m. Esta banda contiene la información de la evaluación de los píxeles, por lo tanto, permite diferenciar las observaciones despejadas, *i.e.*, sin nubes (Bit 1 = 0) de las áreas cubiertas por nubes densas y cirrus (Bit 10 ó Bit 11 = 1; GEE, 2020). En la colección 1C, esta banda utiliza la reflectancia del espectro del azul y dos infrarrojos de onda corta para detectar la presencia de nubes en cada imagen (Coluzzi et al., 2018; ESA, 2020b). El criterio para definir las observaciones válidas por píxel constó de dos requisitos: 1) tratarse de una observación sin nubes, de acuerdo con la banda QA60 y 2) eliminar observaciones repetidas en las áreas de solapamiento entre imágenes contiguas. Este número de observaciones válidas por píxel se calculó para cada mes y cada año del periodo de estudio (01/07/2015 – 31/12/2019). Esta información se obtuvo con la resolución de píxel nativa de la banda QA60, es decir, de 60 m. Posteriormente, se calculó el número de observaciones por píxel para cada ecorregión y se estimó la proporción de su superficie con por lo menos una observación válida anual. El código para obtener la información a nivel de imagen y de píxel puede ser consultado en la siguiente liga: <https://github.com/JonathanVSV/PerPixelObsSentinel-2/blob/master/Script.js>.

Por último, para identificar el mejor periodo intra-anual para obtener observaciones válidas de cada ecorregión se utilizaron sólo los datos de 2018 – 2019, ya que en el periodo

2015 – 2017 el número de imágenes disponibles no fue constante en todos los meses. Esto se debió al incremento paulatino en la capacidad de funcionamiento de la constelación (ESA, 2018). En este procedimiento, se construyeron mosaicos mensuales, bimestrales y trimestrales que comenzaron en cada mes de 2018 y 2019 (terminando el 28/02/2020 para el mosaico trimestral de diciembre de 2019). Para cada mosaico se promedió la proporción de la superficie con por lo menos una observación válida en los dos años evaluados. El análisis de los datos y la elaboración de los gráficos se realizaron en R 3.6.2 (R Core Team, 2020) utilizando el paquete tidyverse (Wickham et al., 2019).

Resultados

Información a nivel de imagen

En total, se encontraron 108,242 imágenes que cubrieron la superficie terrestre mexicana entre 2015 y 2019, de las cuales, la mayoría fueron registradas entre 2018 y 2019 (34,952 y 35,933 imágenes, respectivamente; Figura 9A). Por su parte, en 2015 – 2017 se encontraron un número menor de imágenes (16,779, 16,671 y 3,907 imágenes en 2017, 2016, 2015, respectivamente; Figura 9A).

Adicionalmente, se encontró que la mayor parte de las imágenes tuvieron una calidad intermedia – alta, con un porcentaje de cobertura de nubes menor al 70 % (entre un 61.5 y 65.5 % de las imágenes disponibles por año; Figura 9B). De éstas, la mayoría correspondieron a las de mayor calidad, *i.e.*, 0 – 10 % de cobertura por nubes (entre un 44.50 y 47.6 % del total de imágenes por año; Figura 9B).

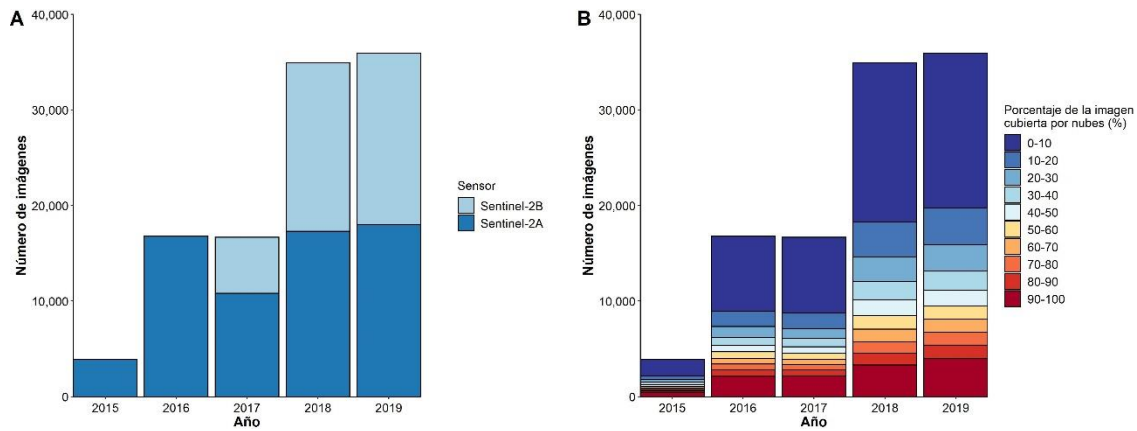


Figura 9. Número de imágenes Sentinel-2 1C disponibles para México entre 2015 y 2019 A) divididas por el satélite que las registró y B) por su porcentaje de cobertura de nubes.

Información a nivel de píxel

El número de observaciones válidas por píxel de la superficie terrestre mexicana fue de entre 0 y 121 al año; sin embargo, esta distribución fue heterogénea tanto en el tiempo como en el espacio. Al igual que en los resultados a nivel de imagen, el número de observaciones por píxel aumentó gracias al inicio de operaciones del Sentinel-2B en 2017, así como al incremento en la capacidad de funcionamiento de ambos sensores (Figura 9 y Figura 11).

Además, se observaron franjas, alineadas en un eje NE – SO, con un mayor número de observaciones válidas por píxel que corresponden al solapamiento entre órbitas adyacentes. Estas franjas presentaron casi el doble de observaciones válidas respecto a las demás áreas donde no se presenta dicho traslape (Figura 10).

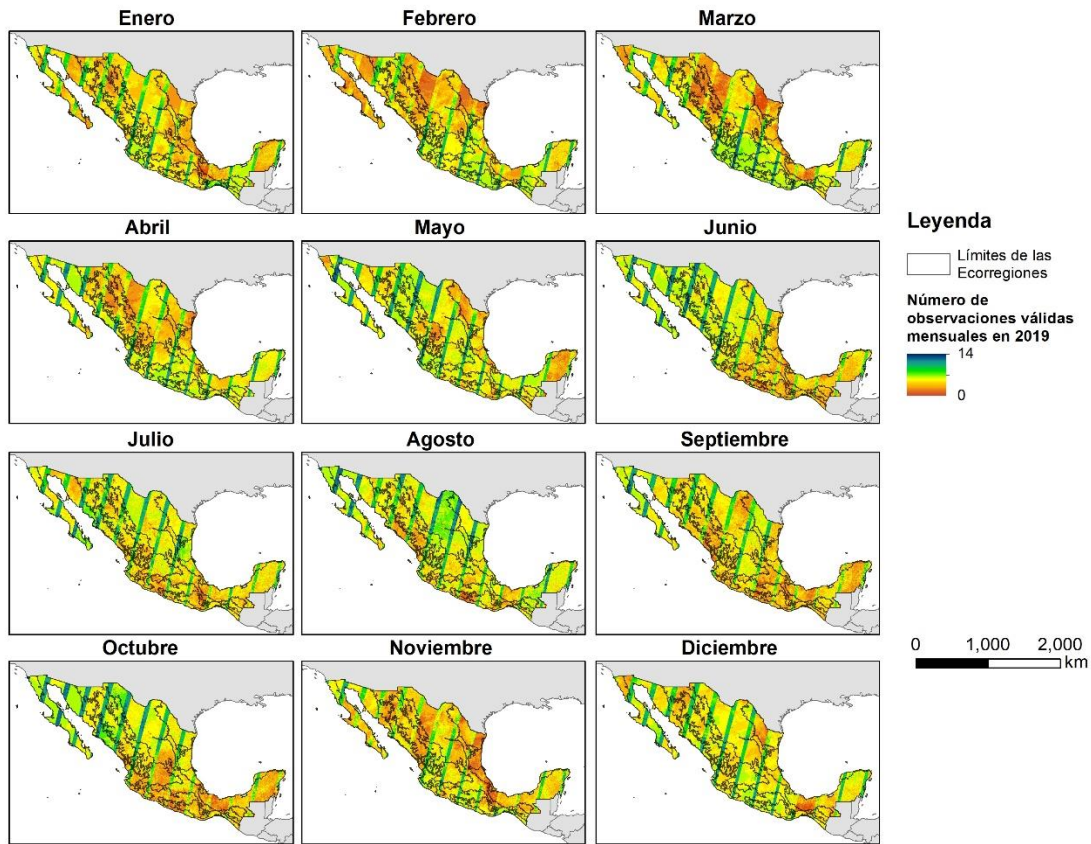


Figura 10. Distribución espacial del número de observaciones válidas mensuales por píxel en 2019. Las franjas de mayor número de observaciones válidas corresponden a las áreas donde hubo un traslape entre órbitas adyacentes de la constelación Sentinel-2.

El patrón interanual en las ecorregiones

Las ecorregiones de California Mediterránea, Elevaciones Semiáridas Meridionales, Grandes Planicies y Selvas Cálido Secas presentaron por lo menos una observación válida para toda su superficie en el periodo 2016 – 2019, mientras que las ecorregiones de Selvas Cálido Húmedas y Sierras Templadas, sólo en el periodo 2017 – 2019 (Cuadro 1). Un caso, excepcional fue Desiertos de América del Norte, donde en todos los años evaluados una porción muy pequeña de su superficie no mostró observación válida alguna.

Cuadro 1. Proporción de cada ecorregión con por lo menos una observación válida al año. Con un asterisco (*) se indican los años donde áreas menores a 0.005 % de la superficie de la ecorregión no mostraron observaciones válidas.

Ecorregión	2015	2016	2017	2018	2019
California Mediterránea	99.91 %	100 %	100 %	100 %	100 %
Desiertos de América del Norte	93.26 %	99.99 %	99.99* %	99.99* %	99.99* %
Elevaciones Semiáridas					
Meridionales	93.25 %	100 %	100 %	100 %	100 %
Grandes Planicies	77.47 %	100 %	100 %	100 %	100 %
Selvas Cálido Húmedas	68.89 %	99.99* %	100 %	100 %	100 %
Selvas Cálido Secas	97.12 %	100 %	100 %	100 %	100 %
Sierras Templadas	92.83 %	99.99* %	100 %	100 %	100 %

El patrón intra-anual en las ecorregiones

En general, las ecorregiones más áridas (California Mediterránea, Desiertos de América del Norte, Elevaciones Semiáridas Meridionales) presentaron un mayor número de observaciones válidas por mes que las ecorregiones más húmedas (Grandes Planicies, Selvas Cálido-Secas, Selvas Cálido-Húmedas y Sierras Templadas; Figura 11). El promedio mensual de las observaciones válidas por píxel se encontró entre 0 y 6.58 observaciones (Figura 11). A pesar de ello, el promedio de observaciones por píxel fue menor en 2015 – 2017 (0 – 4.56 observaciones válidas) que en 2018 – 2019 (2 – 6.58 observaciones válidas; Figura 11).

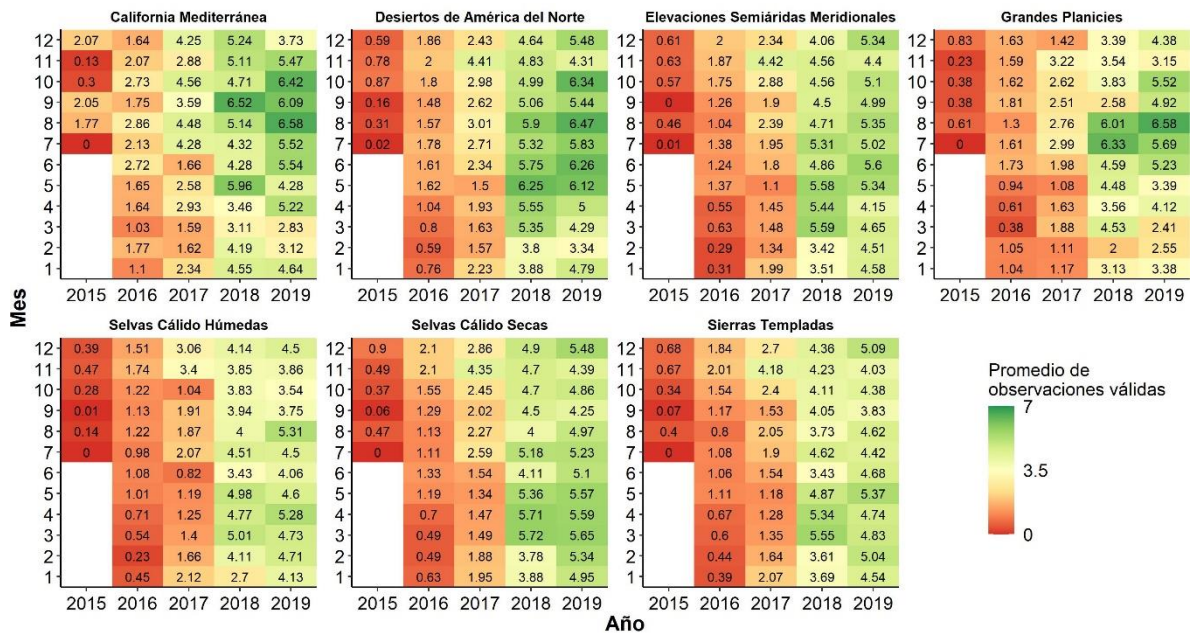


Figura 11. Promedio mensual de observaciones válidas por píxel de cada ecorregión del país durante el periodo 2015 – 2019.

Para las siete ecorregiones se requirió de un distinto número de meses para construir mosaicos multitemporales que presentaran por lo menos una observación válida para toda su superficie en el periodo 2018 – 2019. A pesar de ello, el 95 % de la superficie de todas las ecorregiones mostró por lo menos una observación válida en la mayoría de los meses, un criterio comúnmente utilizado para determinar una alta calidad de las imágenes (Figura 12; Whitcraft et al., 2015; Zhou et al., 2019). La única excepción a este umbral se observó en Grandes planicies en el mes de marzo (Figura 12). Por su parte, al realizar los mosaicos bimestrales, el porcentaje de la superficie de cada ecorregión con por lo menos una observación válida fue mayor al 99.65 %; mientras que los mosaicos trimestrales, fue mayor al 99.85 % (Figura 12).

California Mediterránea fue la única ecorregión donde el 100 % de su superficie presentó por lo menos una observación válida en los mosaicos mensuales de septiembre, octubre y noviembre (Figura 12). Las demás ecorregiones requirieron de realizar mosaicos bimestrales o trimestrales para obtener el total de su superficie con por lo menos una observación válida. Grandes Planicies obtuvo por lo menos una observación válida para la totalidad de su superficie en los mosaicos bimestrales de abril – mayo, mayo – junio, junio

– julio y julio – agosto, mientras que en Selvas Cálido Secas, en abril – mayo (Figura 12). Para Elevaciones Semiáridas Meridionales esto ocurrió en los compuestos trimestrales de septiembre – noviembre y octubre – diciembre, para Selvas Cálido Húmedas, en los compuestos trimestrales de enero – marzo, febrero – abril y diciembre – febrero, mientras que para Sierras Templadas fue en el mosaico trimestral de diciembre – febrero (Figura 12). En el caso de Desiertos de América del Norte, en todas las evaluaciones bimestrales y trimestrales, una pequeña proporción de su superficie no presentó observaciones válidas, lo cual fue consistente con lo observado en las evaluaciones anuales.

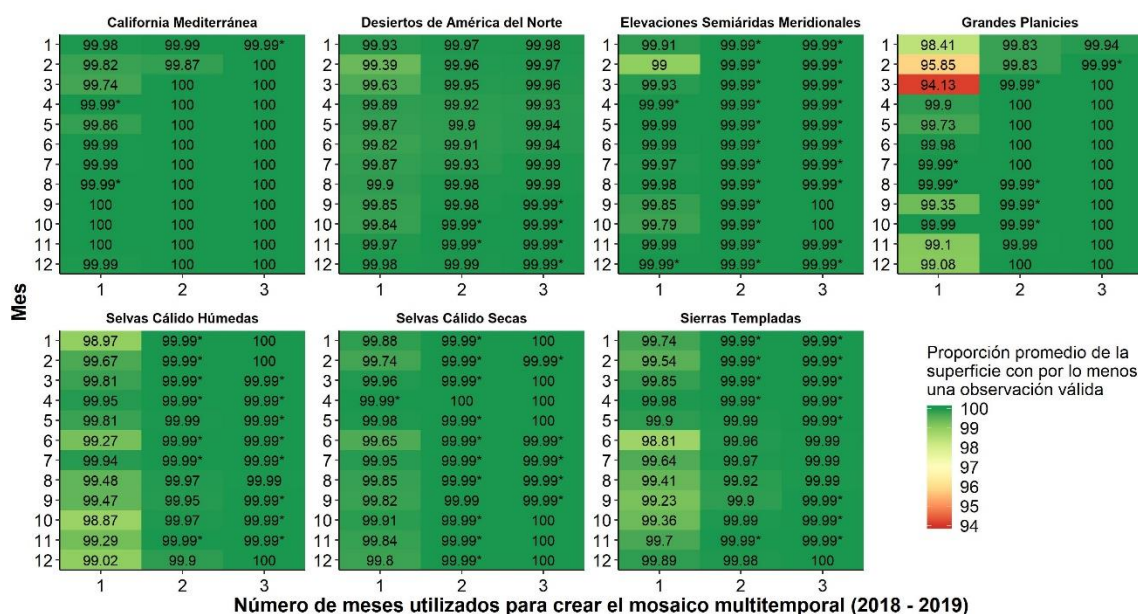


Figura 12. Proporción de la superficie de cada ecorregión con por lo menos una observación válida de acuerdo con los mosaicos construidos con imágenes de 1 mes (mensuales), 2 meses (bimestrales) y 3 meses (trimestrales) en el periodo de 2018 – 2019. Con un asterisco (*) se marcan los meses donde áreas menores a 0.005 % de la superficie de cada ecorregión no presentaron observaciones válidas.

Errores detectados en la máscara QA60

Por último, se realizó una inspección manual de algunas zonas donde, de acuerdo con reportes previos, se esperaba que hubiera errores en la detección de nubes por parte de la banda QA60 (Coluzzi et al., 2018). Esta inspección permitió detectar dos fallas de la banda de evaluación de calidad de los píxeles: 1) la falta de detección de algunas nubes pequeñas o los bordes de estas en la ecorregión de Selvas Cálido Húmedas (*i.e.*, error de omisión) y

2) la confusión de áreas prácticamente desprovistas de vegetación con sustratos de color muy claro con nubes en Desiertos de América del Norte (*i.e.*, error de comisión; Figura 6). Esta última observación permitió explicar por qué los Desiertos de América del Norte fue la única ecorregión para la cual áreas muy pequeñas no mostraron observaciones válidas en el periodo 2016 – 2019 (Figura 13).

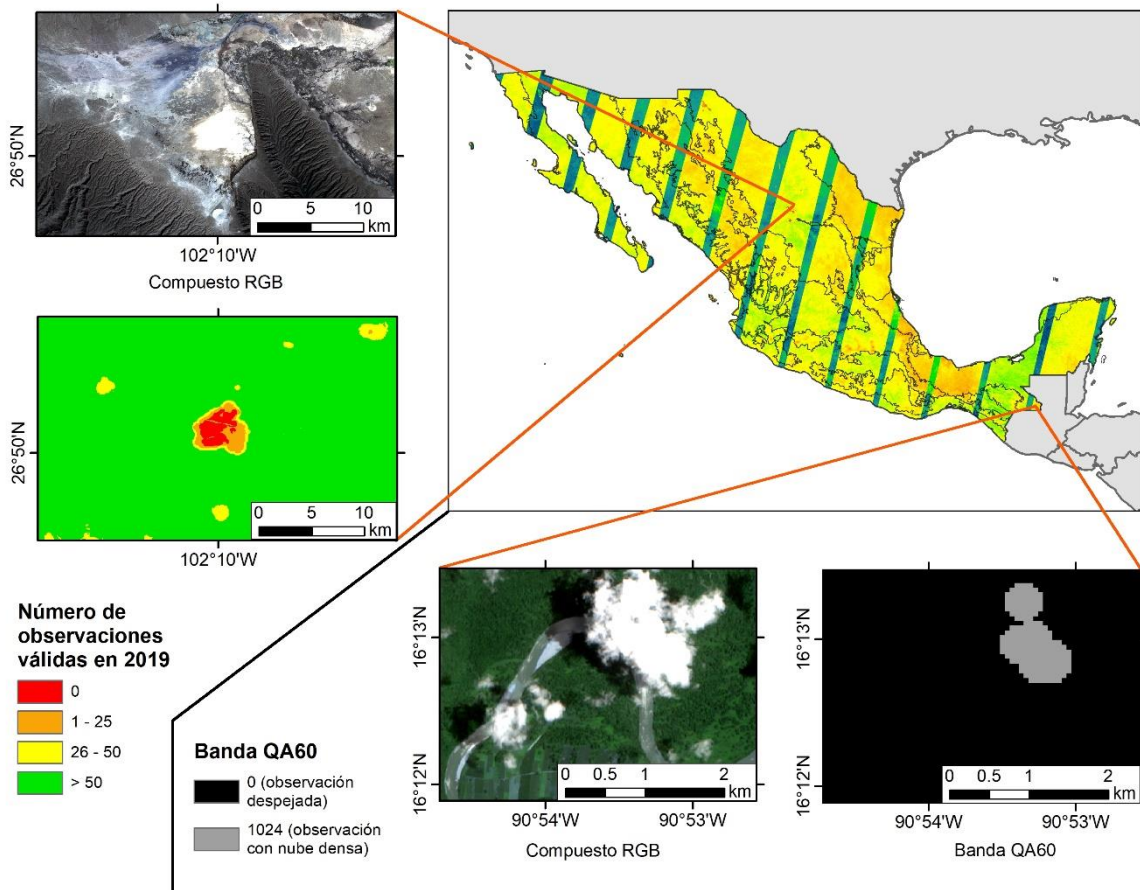


Figura 13. Ejemplos de errores en la banda de evaluación de calidad (QA60) de las imágenes Sentinel-2 1C.

En la esquina superior izquierda se muestra un área donde se identificó erróneamente una zona con un sustrato muy claro como nubes. Se muestra la imagen (compuesto RGB) y debajo, el número de observaciones válidas que presentó dicha área en 2019. Por su parte, en la esquina inferior derecha, se muestra un ejemplo de una imagen (compuesto RGB) con su respectiva banda QA60 donde no se detectó con precisión la presencia de nubes.

Discusión

El conocer los principales patrones espaciotemporales de un acervo de imágenes es una labor poco abordada; sin embargo, resulta esencial para determinar: 1) el mejor periodo para obtener observaciones sin nubes, 2) la frecuencia con la que se podría realizar un monitoreo en particular y finalmente y 3) la viabilidad de utilizar un acervo para estudiar un fenómeno de interés. En este estudio se identificaron los principales patrones espaciotemporales de las observaciones válidas de Sentinel-2 1C para la superficie terrestre mexicana.

Los resultados obtenidos en este estudio permitieron detectar dos patrones interanuales importantes: 1) el primer año con una cobertura completa del país con observaciones válidas fue 2017 (a excepción de áreas en Desiertos de América del Norte) y 2) en el periodo 2018 – 2019 se encontró un acervo de imágenes más grande y homogéneo. Debido a que estos patrones fueron provocados por el aumento en las capacidades de la misión, se esperaría observar un patrón similar a nivel global, aunque podrá variar en mayor o menor grado dependiendo de la latitud.

Por otro lado, se observó que el potencial de utilizar el acervo de imágenes Sentinel-2 para estudiar la superficie terrestre mexicana es heterogéneo en el espacio y respondió a dos factores: 1) a las distintas condiciones climáticas relacionadas con la presencia de nubes de cada ecorregión y 2) las características de la órbita de los sensores. Además, se observó que los periodos para obtener observaciones válidas en cada ecorregión fueron distintos.

Varias de las aplicaciones para las que se planeó la misión Sentinel-2 incluyen el monitoreo de cambios en la cobertura terrestre, de la vegetación, de cuerpos de hielo y de costas y mares interiores (Malenovský et al., 2012). De acuerdo con nuestros resultados, resultaría factible realizar cualquiera de estos monitoreos anualmente, a partir de 2017 en cualquier ecorregión del país. Sin embargo, las ventanas temporales para contar con imágenes de mayor calidad o construir mosaicos multitemporales serán distintas dependiendo de la ecorregión. En total, se distinguieron tres periodos para obtener las imágenes de mayor calidad o mosaicos multitemporales de mayor calidad: 1) en los primeros y últimos meses del año (Selvas Cálido Húmedas, Sierras Templadas), 2) en los meses de septiembre – noviembre (California Mediterránea, Elevaciones Semiáridas

Meridionales y Desiertos de América del Norte) y 3) entre los meses de abril y julio (Grandes Planicies y Selvas Cálido Secas). Estos periodos presentaron una alta relación con el periodo de secas de cada ecorregión. Adicionalmente, a partir de 2018, algunos monitoreos se podrán realizar en intervalos mensuales en algunas ecorregiones (p.ej., California Mediterránea), mientras que, en otras, en periodos trimestrales (e.g., Selvas Cálido Húmedas y Sierras Templadas).

Cabe mencionar que para ciertas aplicaciones el uso de Sentinel-2 será mejor que en otras debido a la coincidencia del periodo de menor nubosidad con el de mayor frecuencia de los eventos de interés. Por ejemplo, para la identificación de cicatrices de incendios, donde la época de mayor ocurrencia de eventos coincide con la época de mayor número de observaciones válidas por píxel (*i.e.*, época de secas; Filipponi, 2018; Roteta et al., 2019). Sin embargo, en el caso del monitoreo de la vegetación y variables biofísicas relacionadas con su productividad, por ejemplo, el índice de área foliar (LAI) o el índice de vegetación de diferencia normalizada (NDVI), los meses de mayor actividad fotosintética suele coincidir con el periodo de mayor presencia de nubes; por lo tanto, con menor número de observaciones válidas por píxel (Hansen et al., 2013). Esto podría resultar en un monitoreo subóptimo de las variables biofísicas relacionadas con la productividad de la vegetación, sobre todo para las ecorregiones más húmedas o muy estacionales (e.g., Selvas Cálido Húmedas o Selvas Cálido Secas). Para estos casos, se podría complementar la información con otros sensores del mismo programa Copernicus u otros (e.g., Sentinel-1; Heckel et al., 2020).

Adicionalmente, se encontraron franjas donde el número de observaciones válidas prácticamente se duplican, debido a la configuración de las órbitas de los satélites Sentinel-2A y 2B. Estas áreas se encontraron en todas las ecorregiones en las que se dividió el país; por lo cual, representan sitios ideales para utilizar las imágenes Sentinel-2 para monitorear su superficie o para probar métodos basados en este acervo. Debido a la órbita de la constelación Sentinel-2, estas franjas son más amplias en regiones polares y más angostas en regiones cercanas al Ecuador (ESA, 2020c; Li and Roy, 2017). En el caso de México, estas franjas tienen un ancho aproximado de entre 30 y 70 km.

Consideraciones sobre la colección Sentinel-2 nivel 1C

Trabajos previos han reportado que la precisión de detección de nubes y sombras en la colección Sentinel-2 nivel 1C puede variar entre 29.44 % - 97.44 % (Coluzzi et al., 2018). Sin embargo, los autores reportan que la precisión depende de tres factores: la humedad del ecosistema (menor precisión en ambientes húmedos), la agregación espacial de las nubes (mayor precisión en imágenes con nubes agregadas) y el color del sustrato (menor precisión en superficies desprovistas de vegetación y con sustrato de color claro). Estos resultados también se observaron en la superficie terrestre mexicana.

Dado este panorama, el uso de la colección Sentinel-2 nivel 1C debe realizarse con cautela, sobre todo al utilizarse para estudiar las ecorregiones más húmedas (*e.g.*, Selvas Cálido Húmedas, Sierras Templadas) y áridas (*e.g.*, California Mediterránea y Desiertos de América del Norte). En el primer caso, la banda QA60 presentará una mayor omisión de detección de nubes, mientras que en las más áridas, la misma banda presentará un mayor número de áreas identificadas erróneamente como nubes. Por otro lado, se esperaría que en ecorregiones como Elevaciones Semiáridas Meridionales, Grandes Planicies y Selvas Cálido Secas, la presencia de este tipo de errores en la banda QA60 fuera menor.

Las evaluaciones sobre el algoritmo de detección de nubes de la colección de Sentinel-2 nivel 2A en diferentes tipos de ambientes reportan una precisión mayor a la del nivel 1C (84 - 92 %; Main-Knorn et al., 2017). Debido a esta diferencia en la precisión de detección de nubes, ciertos usuarios podrían estar más interesados en utilizar la colección 2A en fechas previas a diciembre de 2018. Una alternativa para obtener las imágenes en nivel 2A previo al 2019 sería descargar las imágenes en nivel 1C y procesarlas de manera local utilizando el algoritmo “sen2cor” dentro del programa SNAP (ESA, 2020d).

Conclusiones

Sentinel-2 se perfila como uno de los acervos gratuitos de imágenes multiespectrales más utilizados debido a su alta resolución espacial, espectral y temporal. Por ello, este estudio permitió identificar los principales patrones espaciotemporales de la disponibilidad de observaciones a nivel de imagen y de píxel de la colección de imágenes Sentinel-2 nivel 1C sobre la superficie terrestre mexicana entre el 2015 y 2019. Se espera que este trabajo

promueva el uso de estas imágenes para realizar diferentes tipos de estudios en México, así como sugerir posibles capacidades de análisis con este conjunto de imágenes. A pesar de que esta evaluación fue nacional, es muy probable que patrones similares sean observados en otros sitios con ecorregiones similares, sobre todo aquellas que se encuentren en latitudes cercanas a los trópicos. Por último, para facilitar la implementación de este tipo de evaluaciones en otras partes del mundo, se puso a disponibilidad de cualquier usuario el código utilizado en este trabajo.

Capítulo 3. Clasificación de tipos de cubierta y usos del suelo con una U-Net: ventajas de combinar imágenes Sentinel-1 y Sentinel-2

Jonathan V. Solórzano; Jean-François Mas; Yan Gao; J. Alberto Gallardo-Cruz
Síntesis del artículo publicado en: *Remote Sensing* (2021) 13(18), 3600; y el capítulo del libro: *Integración global: el nuevo entorno de los territorios locales* (2023), (Borrego, Ramírez, Astier [eds.]), CIGA – UNAM, pp. 315 – 343.

Resumen

Actualmente, la U-Net es uno de los algoritmos de aprendizaje profundo más populares para realizar clasificaciones de tipos de cubierta y usos del suelo; sin embargo, pocas veces se ha utilizado con imágenes de radar de apertura sintética (SAR) y multiespectrales (MS). Por otro lado, la diferenciación entre plantaciones y bosques resulta crítica, especialmente en las zonas tropicales, debido a los diferentes servicios ecosistémicos que proveen, así como la biodiversidad que albergan. En este estudio, se entrenó una U-Net utilizando información proveniente de imágenes Sentinel-1 y Sentinel-2 por separado y combinadas, *i.e.*, MS, SAR y MS + SAR; mientras que un algoritmo random forest (RF) se entrenó con información MS + SAR para evaluar la diferencia de selección de algoritmo. El sistema de clasificación constó de diez clases, las cuales incluyeron bosque maduro y bosque secundario, así como plantaciones maduras y jóvenes. Los resultados más precisos fueron obtenidos con la U-Net MS + SAR, con una exactitud total de 0.76 y un *F1-score* promedio de 0.58. A pesar de que las U-Net MS + SAR y MS dieron resultados similares para todas las clases, para las plantaciones maduras y el bosque secundario, la adición de las bandas SAR permitieron aumentar el *F1-score* entre 0.08 y 0.11 (0.62 vs 0.54 y 0.45 vs 0.34, respectivamente). Consecuentemente, en comparación con MS + SAR RF, MS + SAR U-Net obtuvo un *F1-score* más alto para casi todas las clases. Nuestros resultados indican que el uso de la U-Net con imágenes SAR y MS permitieron obtener más altos valores de

exactitud total y *F1-score* para un mapa de tipos de cubierta y usos del suelo detallado, en comparación con los otros métodos evaluados.

Palabras clave: aprendizaje profundo, imágenes multiespectrales y radar de apertura sintética (SAR), redes neuronales convolucionales (CNN), mosaico de un paisaje tropical, mapeo de tipos de cubiertas y usos del suelo.

Método

La zona de estudio se encuentra en la región suroriental de la Selva Lacandona, Chiapas y cubre el municipio de Marqués de Comillas y Benemérito de las Américas, así como áreas aledañas en el municipio de Ocosingo (Figura 14). El área de estudio abarca aprox. 3,040 km² y forma parte de la región de la Selva Lacandona, la cual se caracteriza por presentar una temperatura media anual de 24 a 26 °C y una precipitación media anual de 2 266 mm concentrada en los meses de mayo a diciembre (Centro Nacional de Investigaciones Agrarias, 1982; SEMARNAP, 2000b). La altitud de la zona de estudio varía entre los 50 y 700 m s.n.m. y el tipo de vegetación predominante es Selva Alta Perennifolia o Bosque Tropical Húmedo (Carabias et al., 2015; SEMARNAP, 2000b). Además, la región de estudio presenta pobladores provenientes de distintos estados de la República, grupos indígenas y refugiados guatemaltecos, producto de su compleja historia de colonización (Berget et al., 2021; Cano Castellanos, 2018; Carabias et al., 2015; de Vos and Marion, 2015).

La región Selva Lacandona alberga algunos de los ecosistemas más importantes del sureste mexicano en términos de la biodiversidad que alberga y los servicios ecosistémicos que provee (*e.g.*, captura de carbono, recarga de agua, entre otros (Carabias et al., 2015). Sin embargo, estos dos municipios presentan una alta tasa de deforestación desde hace más de 20 años, lo cual los posiciona como una de las regiones del país con mayores tasas de deforestación (CONAFOR, 2020; Fernández-Montes de Oca et al., 2015; Mendoza and Dirzo, 1999; Ortiz-espejel and Toledo, 1998). Por lo tanto, es común encontrar parches de vegetación recién desmontada para sembrar pastos que sirvan como forraje para el ganado o establecer plantaciones de hule o palma de aceite (Castillo-Santiago et al., 2007; Fernández-Montes de Oca et al., 2021; Vaca et al., 2012). Muchos de estos cambios de uso

de suelo han sido promovidos por programas gubernamentales con un diseño e implementación pobre, sobre todo para incentivar el sector ganadero y aceitero (Cano Castellanos, 2014; Montes de Oca et al., 2015). Por otro lado, en la región también existen programas o instrumentos destinados a promover la conservación de grandes extensiones o parches de bosque natural, como la presencia de áreas protegidas o el pago por servicios ambientales (DOF, 2017; Izquierdo-Tort et al., 2021; SEMARNAP, 2000a; SEMARNAT, 2022). Por ello, esta región presenta varias condiciones interesantes para probar un sistema de monitoreo de la dinámica de la cubierta arbórea: por un lado, se pueden encontrar fuerzas que empujan la degradación forestal o deforestación (*e.g.*, obtención de madera, ocurrencia de fuegos, abrir tierras para expansión de la frontera agrícola, ganadera o de palma de aceite) y fuerzas que empujan hacia la permanencia del bosque (*e.g.*, pago por servicios ambientales o áreas naturales protegidas).

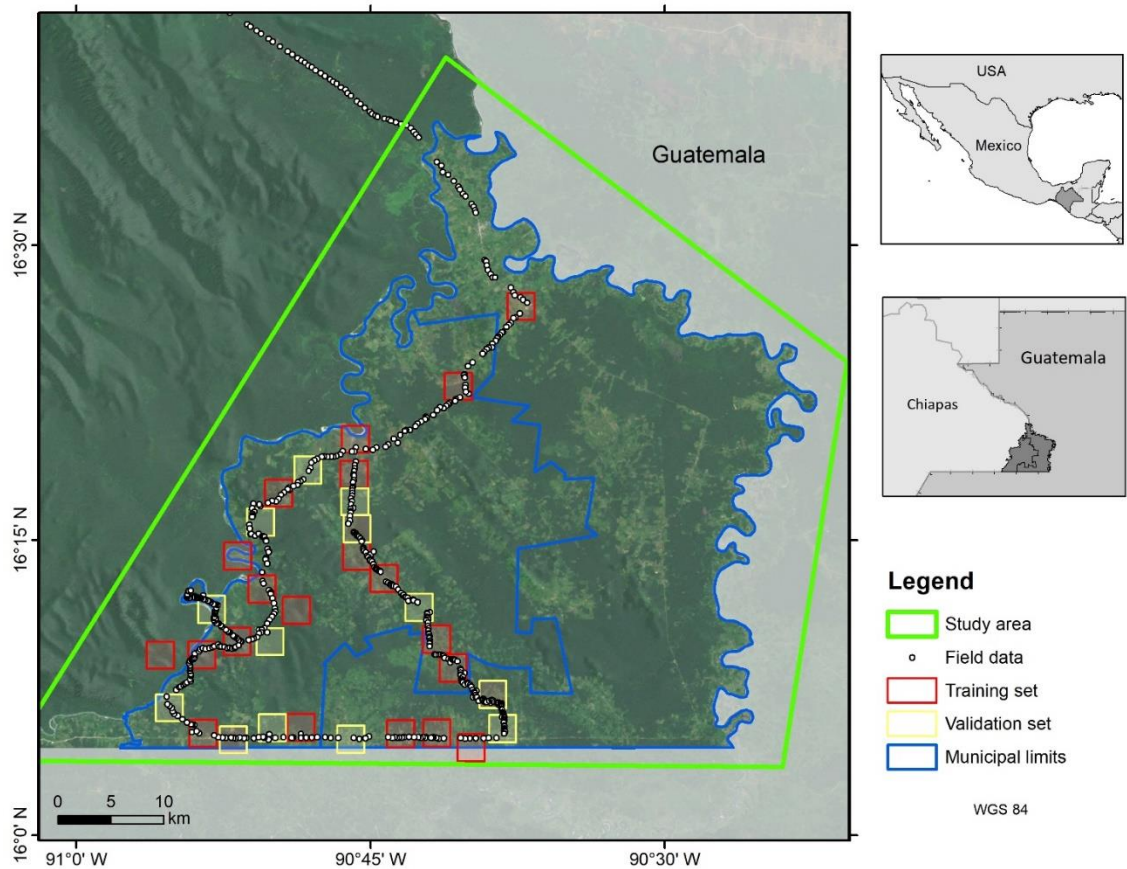


Figura 14. Localización del sitio de estudio, así como de las áreas clasificadas manualmente para entrenar y verificar el modelo U-Net.

El método siguió tres fases: 1) el preprocesamiento de las imágenes satelitales y la clasificación manual, 2) el entrenamiento y validación de un algoritmo de aprendizaje profundo (U-Net) y uno de aprendizaje automatizado (random forests) y 3) la selección de la mejor arquitectura, la clasificación del área completa de estudio y la validación del mapa para el área de estudio (Figura 15).

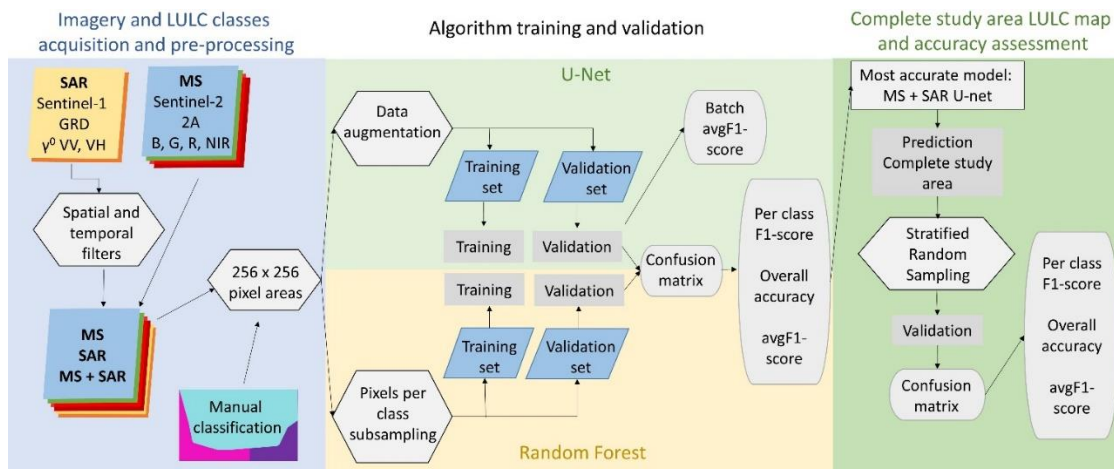


Figura 15. Diagrama de flujo del procedimiento seguido en este trabajo.

Procesamiento de la información multiespectral y radar

Como insumos para realizar las predicciones de las clases de cubierta terrestre con el modelo U-Net, se utilizaron imágenes de radar de apertura sintética (SAR) y multiespectrales (MS) registradas por los sensores Sentinel-1 y -2, respectivamente. Para las imágenes SAR, consultó la información de la colección de imágenes Sentinel-1 GRD (*Ground Range Detected*), la cual contiene la información de retrodispersión en dB (σ^0) en dos bandas VV (polarización vertical - vertical) y VH (vertical – horizontal). Posteriormente, estas imágenes fueron corregidas utilizando el ángulo de registro de la información (γ^0 ; Cassol et al., 2019; Small, 2011). En el caso de las imágenes multiespectrales, se consultó la colección Sentinel-2 2A, que corresponde a la reflectancia debajo de la atmosfera (*bottom of atmosphere*), y únicamente se utilizaron las bandas de mayor resolución espacial (10 m): azul (B), verde (G), rojo (R) e infrarrojo cercano (NIR).

Primero, se identificaron las imágenes Sentinel-2 que tuvieran la menor cobertura de nubes más cercanas a la fecha de captura de información en campo de los tipos de vegetación y cubierta de la zona (*i.e.*, marzo de 2019). Esta imagen correspondió a una registrada el 4 de julio de 2019. En el caso de las imágenes Sentinel-1, para reducir el efecto del *speckle*, se promedió el valor de la retrodispersión de las imágenes registradas entre el 4 de junio de 2019 y el 4 de julio de 2019, lo cual abarcó 9 imágenes. Después, se aplicó un filtro 3×3 píxeles para asignar el promedio de la ventana al píxel central.

Finalmente, se construyeron tres conjuntos de datos para evaluar el desempeño de la U-Net para realizar predicciones de tipos de vegetación y usos del suelo: MS (4 bandas), SAR (2 bandas) y MS + SAR (6 bandas). Todo este procedimiento se realizó en la API de Javascript de Google Earth Engine (Gorelick et al., 2017).

Creación de datos de entrenamiento y validación

Los datos de entrenamiento constaron de clasificaciones manuales de 33 áreas de 256×256 píxeles. Estas áreas se seleccionaron en sitios donde se tuvieran datos del tipo de vegetación o uso de suelo registrados en campo (300 puntos). Además, se consultaron otras imágenes satelitales para clasificar los píxeles de estas 33 áreas: 1) imágenes Sentinel-2 de 2016 – 2020 para contar con un registro multitemporal, 2) imágenes Planet de enero y marzo de 2019 (3.12 m de píxel; Planet Team, 2017) e 3) imágenes de muy alta resolución provistas por Google Earth, Yandex y Bing (resolución píxel ≤ 1 m). Estas 33 áreas se clasificaron manualmente en diez clases: bosque maduro, bosque secundario, vegetación inundable, plantaciones maduras, plantaciones jóvenes, pastizales / agricultura, suelo, caminos, agua, asentamientos humanos. En el caso del conjunto de datos MS y MS + SAR se agregaron otras dos clases al esquema de clasificación: *nubes* y *sombras*. Estas dos clases se delinearon manualmente mediante la interpretación visual de la imagen MS, aunque fueron excluidas de la evaluación final de los resultados. Estas clasificaciones se transformaron a formato ráster con una resolución de 10 m. Finalmente, los mapas de clasificación correspondieron a rústers binarios de 10 – 12 bandas, *i.e.*, una banda por clase.

A continuación, se recortó la imagen compuesta por las bandas MS y SAR para trabajar únicamente con los datos de los 33 cuadros de 256×256 píxeles (Figura 15). Además, para facilitar el aprendizaje de la U-Net, se estandarizaron estos datos mediante la resta de la media de cada banda y se dividió este valor entre la desviación estándar de cada banda (Mueller and Massaron, 2019). Después, se aumentó de manera artificial el conjunto de datos para evitar el sobreajuste del modelo (Isaienkov et al., 2020; Wagner et al., 2019). Este procedimiento se realizó mediante la división de cada cuadro de 256×256 píxeles en 9 cuadros de 128×128 píxeles utilizando un desplazamiento de 64 píxeles en la dirección tanto vertical como horizontal y se construyó una imagen en espejo en un sentido vertical como horizontal. De este número, las imágenes provenientes de 30 cuadros se utilizaron

como datos de entrenamiento (810 imágenes), mientras que las provenientes de los 3 cuadros restantes se utilizaron como datos de validación (81 imágenes; Fig. 1). Este procedimiento se realizó en R 4.0.3 (R Core Team, 2021) utilizando los paquetes raster (Hijmans, 2020), rray (Vaughan, 2020) y reticulate (Ushey et al., 2020).

Entrenamiento de algoritmos

U-Net

Para construir la arquitectura U-Net se utilizó el paquete unet (Falbel and Zak, 2020), mientras que el entrenamiento del modelo se realizó utilizando el software R 4.0.3 y el paquete keras (Allaire and Chollet, 2018) con tensorflow (Abadi et al., 2016) como *backend*. Por su parte, para realizar una búsqueda de la combinación de hiperparámetros óptimos se realizaron varias iteraciones utilizando el método de *early stopping* para evitar el sobreajuste, en el cual, el entrenamiento se detuvo cuando la métrica de precisión no aumentó en una centésima después de 10 épocas. La combinación de hiperparámetros explorados constó del tamaño de lote (16, 32, 48), el número de capas ocultas (2, 3, 4), el número de filtros en la primera capa (32, 64) y la probabilidad de *dropout* (0, 0.1, 0.2, 0.3, 0.4, 0.5; Srivastava et al., 2014). Otros parámetros del modelo incluyeron el uso de filtros convolucionales de 3 x 3 píxeles, la función de entropía cruzada categórica como función de pérdida, el optimizador Adam (Kingma and Ba, 2015), una normalización por lote y una inicialización normal de He para inicializar los pesos de los filtros (He et al., 2015). La exploración de hiperparámetros se realizó utilizando el paquete tfruns (Allaire, 2018) y se generaron en total 90 modelos.

Para evaluar los resultados obtenidos de la exploración anterior se calculó la precisión total (*OA*) de la clasificación predicha por el modelo U-Net como el número de píxeles correctamente clasificados entre el total de píxeles de acuerdo a la ecuación (1). Además, se calculó la métrica *F1-score* como el promedio armónico de la precisión (ecuación 2) y la sensibilidad (ecuación 3) de las clases predichas (ecuación 4; Goutte and Gaussier, 2005). Aunque los términos precisión y sensibilidad son más utilizados en la comunidad de aprendizaje automatizado, en la comunidad de percepción remota, estos términos equivalen a la precisión del usuario (precisión) y la precisión del productor (sensibilidad; Maxwell et al., 2021). Posteriormente, se calculó el promedio del *F1-score* (*avgF1-score*) obtenido

para cada clase para obtener una medida del desempeño del modelo sobre todas las clases (ecuación 5).

$$OA = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

$$p = \frac{TP}{TP + FP} \quad (2)$$

$$r = \frac{TP}{TP + FN} \quad (3)$$

$$F_1 \text{score} = 2 \frac{pr}{p+r} \quad (4)$$

$$avgF_1 \text{score} = \frac{1}{C} \sum_{c=1}^C F_1 \text{score}_c \quad (5)$$

donde TP corresponde al número de observaciones verdaderas positivas, TN , verdaderos negativos, FP , falsos positivos, FN , falsos negativos, C al número de clases y c a cada clase. Por su parte, OA corresponde a precisión total, p a precisión, r , sensibilidad, $F_1 \text{score}_c$ al $F_1 \text{score}$ de la clase c y $avgF_1 \text{score}$ al $F_1 \text{score}$ promedio. Estas dos métricas se monitorearon en el entrenamiento de la U-Net. La mejor configuración de hiperparámetros se seleccionó como la que obtuvo el $avgF_1 \text{score}$ más alto sobre los datos de validación.

La exploración de los hiperparámetros se realizó con el paquete `tfruns` (Allaire, 2018). En total se entrenaron 90 arquitecturas para cada conjunto de datos de entrada (MS, SAR, MS + SAR), lo que resultó en un total de 270 modelos explorados. El entrenamiento y la validación de estas arquitecturas se realizó en una tarjeta de video NVIDIA RTX 2060 con 6 GB de memoria.

Random forests

El algoritmo random forests es uno de los más frecuentemente utilizados en trabajos de aprendizaje automatizado, así como para realizar clasificaciones (Heckel et al., 2020; Lary

et al., 2016; Talukdar et al., 2020). Este algoritmo es un método de ensamble basado en árboles de decisión, en el cual se entrenan varios árboles a partir de un conjunto de observaciones y variables predictivas muestreadas al azar. Después el algoritmo asigna la clase de cada píxel como aquella que obtuvo más votos en los árboles individuales (Breiman, 2001).

La clasificación con random forests se realizó con la información que permitió la más alta *OA* en los datos de verificación (MS + SAR). En esta clasificación, se utilizaron los mismos datos de entrenamiento y validación que la U-Net, pero sin el esquema de aumento artificial de los datos. De esta manera, los datos de entrenamiento correspondieron a 23 áreas de 256 x 256 píxeles, mientras que los de validación a 10 áreas de las mismas dimensiones. Debido a algunas restricciones de memoria, no se pudo utilizar el conjunto completo de los datos de entrenamiento para entrenar el random forests, sino que se hizo un submuestreo aleatorio, en el cual se asignó un mismo número de observaciones a cada clase (Sawangarreerak and Thanathamthee, 2020). Después de realizar este proceso, los datos de entrenamiento constaron de 6057 puntos por clase, ya que este número correspondió al menor número de observaciones por clase en los datos de entrenamiento (*i.e.*, vegetación acuática). Este algoritmo se entrenó utilizando 500 árboles y dos variables predictoras seleccionadas al azar en cada árbol y utilizando el paquete randomForest (Liaw and Wiener, 2002). Posteriormente, este algoritmo entrenado se utilizó para realizar la predicción sobre los datos de validación y obtener tanto la matriz de confusión como las mismas métricas calculadas para la U-Net, *i.e.*, *OA* y *avgF1-score*.

Validación de clasificación del área completa de estudio

La arquitectura U-Net que obtuvo la mayor precisión fue MS + SAR. Por ello, esta arquitectura se utilizó para predecir la clasificación de tipos de vegetación y usos del suelo del área de estudio completa. Un problema comúnmente reportado en estudios utilizando la U-Net es que los bordes de los cuadros sobre las que se realizan las predicciones tienen menor calidad que la de los píxeles ubicados lejos de los bordes (Flood et al., 2019; Ronneberger et al., 2015). Esto es provocado porque los píxeles en los bordes tienen un menor número de píxeles vecinos que su contraparte más cercana al centro de los cuadros. Esto promueve que las convoluciones más profundas integren una menor cantidad de datos

en los bordes. Para reducir este efecto, se dividió al área de estudio en dos cuadrículas de 128 x 128 píxeles de manera que los bordes de los cuadros de una de ellas se superpusieran con el centro de los cuadros de la otra. Posteriormente, se construyó el mosaico final asignándole la clase con mayor probabilidad de pertenencia en alguna de las dos cuadrículas.

La precisión de la clasificación para el área de estudio se evaluó mediante un diseño aleatorio estratificado (Card, 1982; Cochran, 1977; Olofsson et al., 2014). En este diseño, el número de puntos por clase se distribuyeron de acuerdo a la proporción ocupada por cada una de ellas en el mapa, a excepción de las clases más raras, donde se asignó un número mayor al de su proporción de área. El total de número de unidades de muestreo para la verificación se calculó de acuerdo a la ecuación 6 (Cochran, 1977):

$$n = \frac{(\sum_i^q W_i S_i)^2}{(s(\hat{O}))^2 + (\frac{1}{N}) \sum_i^q W_i S_i^2} \approx \left(\frac{\sum_i^q W_i S_i}{s(\hat{O})} \right)^2 \quad (6)$$

donde $S(\hat{O})$ es el error estándar deseado de la OA. En este estudio se adoptó un valor ligeramente mayor que el recomendado por Oloffson et al. (2013), 0.015. W_i es la proporción de área ocupada por cada clase en el mapa, mientras que i corresponde a cada clase y q , al número de clases. S_i es la desviación estándar de la clase i , la cual se calcula como $S_i = \sqrt{U_i(1 - U_i)}$ donde u_i significa la precisión del usuario esperada a priori. En este caso, se utilizó el valor de la precisión de usuario obtenida en los datos de validación (Cuadro S 1 y Cuadro S 3).

Para las clases menos frecuentes, las cuales ocuparon menos del 5 % del mapa (todas las clases a excepción de bosque maduro, pastizal / agricultura y suelo), se asignaron 25 puntos de verificación. Para las clases más dominantes, se asignó un número de puntos proporcional a su área en el mapa. En total se verificaron 448 puntos. Este proceso se llevó a cabo mediante interpretación visual de la misma información utilizada para generar los datos de entrenamiento, *i.e.*, imágenes Sentinel-2 2016 – 2019, imágenes Planet, Google Earth, Bing y Yandex.

Finalmente, se utilizó la herramienta de evaluación de la precisión de OpenForis (FAO, 2017) para calcular los estimados de área sin sesgo, así como sus intervalos de confianza. Estos estimados se basaron en la matriz de confusión resultante de la verificación siguiendo el diseño estratificado aleatorio, de acuerdo a las ecuaciones 7 y 8. Adicionalmente, el *F1-score* por clase, *OA* y *avgF1-score* fueron calculados.

$$S(\hat{p}_{.k}) = \sqrt{\sum_i W_i^2 \frac{n_{ik} \left(1 - \frac{n_{ik}}{n_i}\right)}{n_{i.} - 1}} \quad (7)$$

$$95\% \text{ CI } \hat{A}_k = \left(A \sum_{i=1}^q W_i \frac{n_{ik}}{n_i} \right) \pm 1.96 (A S(\hat{p}_{.k})) \quad (8)$$

Donde $S(\hat{p}_{.k})$ corresponde al error estándar de la proporción de área estimada para la clase k , W_i es la proporción de área ocupada por la clase i , n_{ik} significa el número de observaciones en la celda i,k de la matriz de confusión, mientras que n_i es la suma de la fila de la clase i . Por último, $95\% \text{ CI } \hat{A}_k$, significa el intervalo de confianza del 95% del área estimada para la clase k , A es el área total del mapa, mientras que q representa el número de clases.

Resultados

U-Net

Exploración de hiperparámetros

La exploración de los hiperparámetros mostró que la arquitectura U-Net que alcanzó la *OA* más alta fue MS + SAR, con una *OA* de 0.76 y *avgF1-score* de 0.59 en los datos de verificación (Cuadro 2). Por su parte, la U-Net MS alcanzó una *OA* de 0.75, *avgF1-score* de 0.55, mientras que U-Net SAR, una *OA* de 0.65 y *F1-score* de 0.39. En el caso de la *OA* y *avgF1-score* obtenidos en los datos de entrenamiento se observó un patrón similar, ya que MS + SAR presentó una $OA = 0.91$ y $avgF1-score = 0.86$, en el caso de U-Net MS una $OA = 0.91$ y $avgF1-score = 0.86$, mientras que para U-Net SAR una $OA = 0.72$ y $avgF1-score = 0.60$. Dos de las mejores arquitecturas, MS + SAR y MS, constaron de tres capas ocultas, un tamaño de lote de 16 y 64 filtros en la primera capa. Sin embargo, MS + SAR, presentó

un *dropout* más alto (0.5), mientras que MS de 0.1. En el caso de la mejor arquitectura U-Net SAR, esta presentó dos capas ocultas, un tamaño de lote de 8, 64 filtros en la primera capa y un *dropout* intermedio (0.3). Finalmente, un error de 0.15 se detectó en la interpretación visuales realizadas sobre los datos de entrenamiento sin tener acceso a la información de campo.

Cuadro 2. *F1-score* y su diferencia respecto al valor más alto ($\Delta F1-score$) para cada clase con las tres U-Net MS + SAR, MS, SAR, así como RF MS + SAR. Abreviaturas: MS, bandas multiespectrales, SAR, bandas de apertura sintética, RF, Random forests.

Clase	U-Net						RF	
	MS + SAR		MS		SAR		MS + SAR	
	F1-Score	$\Delta F1-Score$	F1-Score	$\Delta F1-Score$	F1-Score	$\Delta F1-Score$	F1-Score	$\Delta F1-Score$
Agua	0.96	0.01	0.97	0	0.94	0.03	0.94	0.03
Asentamientos humanos	0.87	0	0.87	0	0.45	0.42	0.29	0.58
Bosque maduro	0.86	0	0.86	0	0.79	0.07	0.69	0.17
Bosque secundario	0.45	0	0.34	0.11	0.20	0.25	0.33	0.12
Caminos	0.35	0.08	0.43	0	0	0.43	0.23	0.2
Pastizales / Agricultura	0.78	0	0.77	0.01	0.69	0.09	0.63	0.15
Plantaciones jóvenes	0.11	0	0.03	0.08	0.01	0.1	0.11	0
Plantaciones maduras	0.62	0	0.54	0.08	0.47	0.15	0.43	0.19
Suelo	0.65	0	0.64	0.01	0.30	0.35	0.57	0.08
Vegetación acuática	0.15	0	0.10	0.05	0	0.15	0.04	0.11
Precisión total (OA)	0.76	0	0.76	0	0.65	0.11	0.53	0.23
avgF1-score total	0.58	0	0.55	0.03	0.39	0.19	0.43	0.15

Comparación entre las U-Net MS, SAR y MS + SAR

Al comparar las mejores arquitecturas de U-Net de acuerdo a sus entradas, la que obtuvo un *F1-score* más alto fue U-Net MS + SAR, sin embargo, para algunas clases, U-Net MS obtuvo *F1-scores* más altos (Cuadro 2). Por ejemplo, MS + SAR obtuvo un *F1-score* más alto para tres clases: plantaciones maduras, bosque maduro y plantaciones jóvenes ($\Delta F1-$

$score \geq 0.05$), pero más bajos en la clase de caminos ($\Delta FI-score = 0.08$). Adicionalmente, otras cinco clases obtuvieron un $FI-score$ similar en la U-Net MS + SAR y U-Net MS ($\Delta FI-score \leq 0.01$; Cuadro 2). En comparación con la U-Net SAR, ambos modelos, U-Net MS + SAR y MS, tuvieron un $FI-score$ más alto para todas las clases, sin embargo, para la clase de agua, el $FI-score$ fue bastante similar en cualquiera de las tres U-Net (Cuadro 2; $\Delta FI-score \leq 0.03$). La diferencia en poder de discriminación entre clases de las tres U-Net también fue evidente al comparar las predicciones realizadas por cada modelo y la clasificación manual (Figura 16). A pesar de que hubo diferencias particulares entre los tres modelos, en general, las clases con mayores valores de $FI-score$ fueron bosque maduro, agua y asentamientos humanos, mientras que las clases de plantaciones jóvenes, caminos y vegetación acuática presentaron los $FI-score$ más bajos (Cuadro 2).

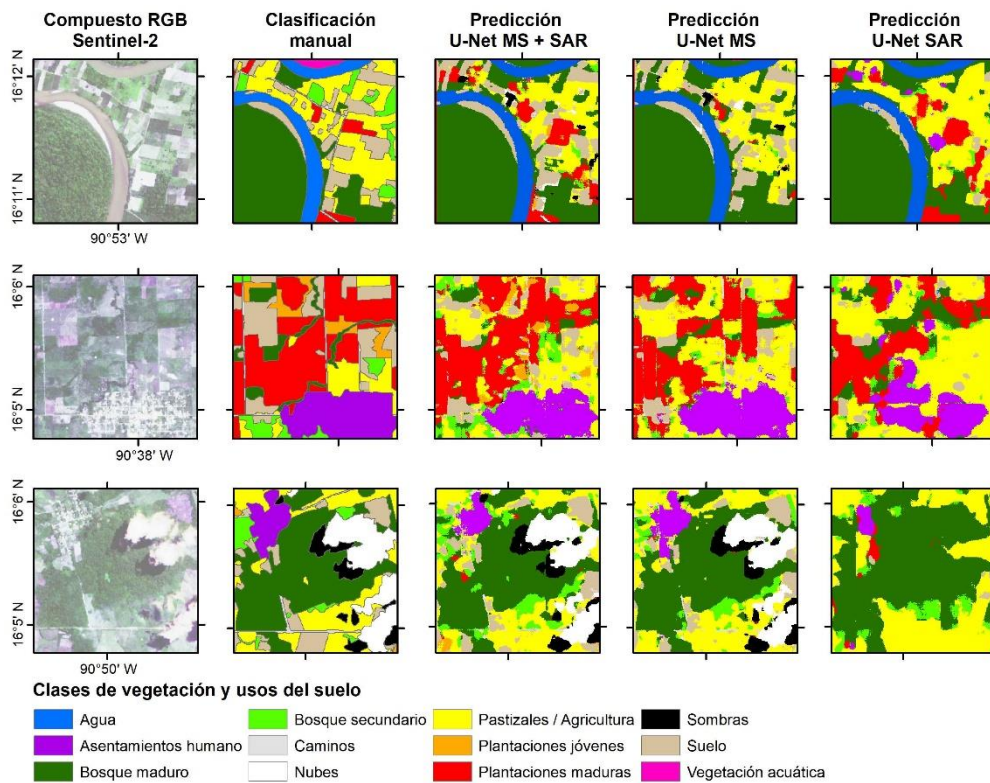


Figura 16. Evaluación de las predicciones de la U-Net en tres distintas muestras de los datos de validación.

(A) Compuesto RGB de Sentinel-2, (B) Áreas manualmente clasificadas, (C) Predicción con la U-Net utilizando las bandas multispectrales (MS) + radar de apertura sintética (SAR), (D) predicciones con la U-Net utilizando sólo las bandas MS, (E) predicciones de la U-Net utilizando las bandas SAR.

Comparación entre las U-Net MS + SAR y RF MS + SAR

Al comparar los resultados obtenidos de la U-Net MS + SAR con RF MS + SAR, fue evidente que la primera obtuvo un mayor *avgF1-score* y *OA* que su equivalente RF ($\Delta OA = 0.23$ and $\Delta F1-score = 0.15$; Cuadro 2). A pesar de que el mismo patrón fue evidente al comparar los *F1-score* por clases ($0.08 \leq \Delta F1-score \leq 0.58$), la clase de agua mostró un *F1-score* ligeramente más bajo en RF MS + SAR ($\Delta F1-score \leq 0.03$), mientras que la clase de plantaciones jóvenes mostró el mismo *F1-score* en U-Net MS + SAR y RF MS + SAR (Cuadro 2).

Clasificación del área completa de estudio

La evaluación de la precisión sobre el área completa de estudio mostró una *OA* de 0.77 y un *avgF1-score* de 0.68. En esta verificación, las clases que obtuvieron más alto *F1-score* fueron agua, asentamientos humanos, bosque maduro, pastizales / agricultura (Cuadro 3). Por el contrario, las clases con el *F1-score* más bajo fueron plantaciones jóvenes, bosque secundario, suelo y caminos (Cuadro 3).

Cuadro 3. Área, proporción del área de estudio y *F1-score* para de la clasificación para el área completa de estudio. Asimismo, se muestra el *F1-score* obtenido en los datos de validación, seguido de los estimados de área corregida y su intervalo del 95 % de confianza por cada clase. Abreviaturas: CI, intervalo de confianza.

Clase	Clasificación el área completa de estudio U-Net MS + SAR			Datos de validación U-Net MS + SAR	Estimados de área	
	Área (ha)	Proporción del área de estudio (%)	F1-Score	F1-Score	Área corregida	IC 95%
Agua	4617.35	2.08	0.96	0.96	4780.39	319.57
Asentamientos humanos	2494.09	1.12	0.94	0.87	3322.42	1082.37
Bosque maduro	93,341.09	41.97	0.87	0.86	90772.02	5449.89
Bosque secundario	6195.65	2.79	0.54	0.45	5824.51	2793.85
Caminos	806.18	0.36	0.60	0.35	6111.42	2790.73
Pastizales / Agricultura	84,572.15	38.03	0.79	0.78	76282.87	6516.50
Plantaciones jóvenes	4048.08	1.82	0.31	0.11	9434.56	3823.09

Plantaciones maduras	8076.54	3.63	0.56	0.62	7511.87	3233.03
Suelo	17,769.69	7.99	0.55	0.65	12,162.86	4080.16
Vegetación acuática	467.06	0.21	0.70	0.15	6184.96	3848.47

Al analizar de manera visual el mapa de probabilidad de la clasificación del área de estudio, resultó evidente que la mayoría del bosque maduro obtuvo probabilidades altas de corresponder a esa clase (Figura 17). Por el contrario, las áreas que se encontraron entre los límites de distintas clases, así como en áreas con alta nubosidad o recientemente quemadas, obtuvieron valores de probabilidad de pertenencia a su clase bajos (Figura 17).

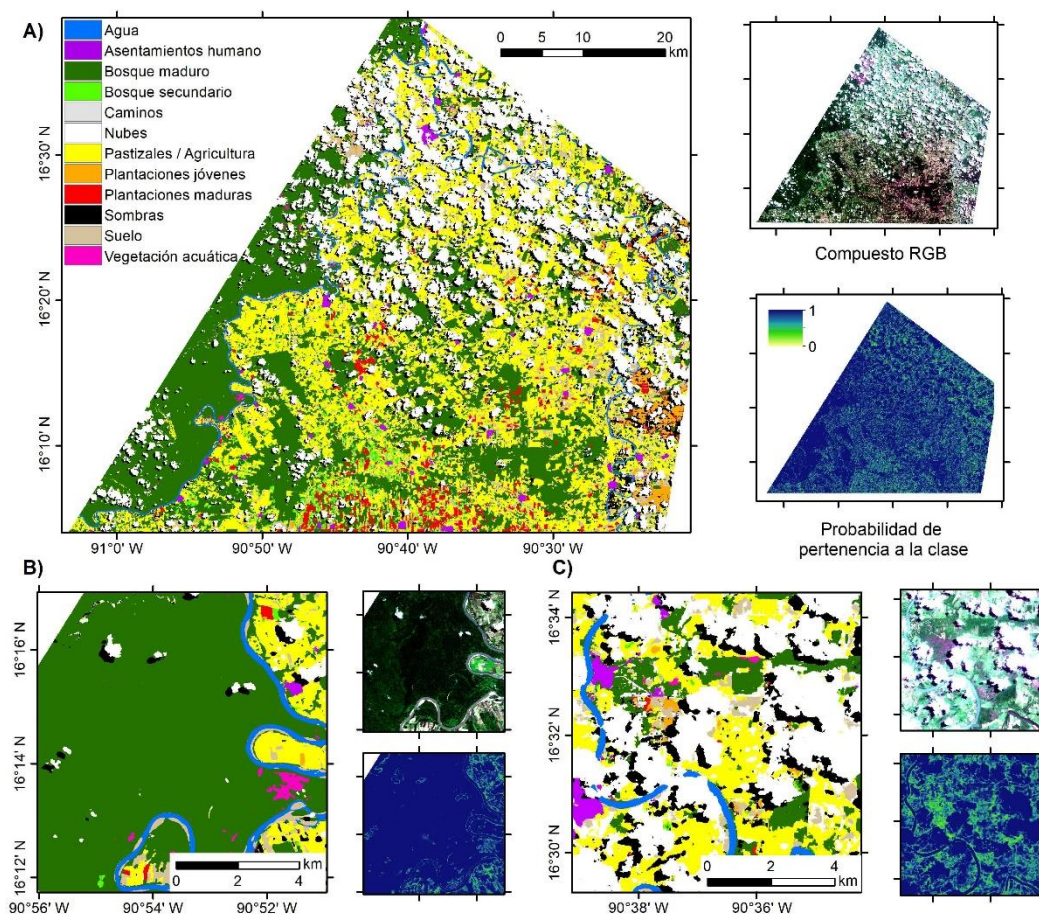


Figura 17. (A) Clasificación de cubierta y usos del suelo del área de estudio completa utilizando la U-Net MS + SAR, así como el compuesto RGB y los mapas de probabilidad de pertenencia a la clase correspondiente. (B) Ejemplo de un sitio dominado por bosque maduro con altas probabilidades de pertenencia. (C) Ejemplo de un área recientemente quemada con una alta presencia de nubes y sombras, donde se observan valores bajos de pertenencia.

Discusión

Este estudio mostró que el uso de imágenes MS y SAR con el algoritmo U-Net permite obtener una clasificación detallada de la cubierta y uso de suelo con una *OA* de 0.78. Este método cuenta con dos ventajas que le permiten obtener mejores resultados que algoritmos previamente disponibles para discriminar entre las clases de interés (Du et al., 2020; Giang et al., 2020): contar con información de la composición de la superficie, características geométricas y de humedad (Mercier et al., 2019; Poortinga et al., 2019), así como resumir patrones espectrales y espaciales de las imágenes.

En particular, los resultados de este estudio indican que el método permite clasificar las áreas de bosque y plantaciones maduras con una precisión de usuario de 0.89 y 0.76 y precisión del productor de 0.9 y 0.73, respectivamente (Cuadro S 1). Por su parte, la clase de bosque secundario obtuvo una precisión del usuario más baja, de 0.37 y del productor de 0.28. A pesar de que los *F1-score* obtenidos para el bosque secundario y plantaciones maduras no fueron tan altos como los obtenidos para el bosque maduro, la U-Net MS + SAR permitió un aumento de entre 0.08 y 0.25 en comparación con los otros métodos probados.

Implicaciones para el monitoreo de la región

El interés por discriminar entre las tres clases de cubierta arbórea radicó en poder diferenciar entre tipos de vegetación con características muy distintas en términos de biodiversidad, carbono almacenado, provisión de servicios ecosistémicos, así como de prácticas de aprovechamiento de los recursos naturales (Gibson et al., 2011; Wright, 2005). Los bosques maduros suelen corresponder a áreas de alta biodiversidad y con grandes reservas de carbono, mientras que las plantaciones y los bosques secundarios, usualmente albergan un menor contenido de carbono y biodiversidad (Fernández-Montes de Oca et al., 2021; Tropek et al., 2014). Además, los bosques maduros suelen encontrarse bajo un uso extensivo o nulo; por su parte, las plantaciones suelen corresponder a áreas bajo un uso intensivo, mientras que los bosques secundarios usualmente están asociados a ciclos de descanso de prácticas agrícolas (Isaac-Márquez et al., 2016; Ochoa-Gaona et al., 2007).

En la zona de estudio se ha reportado que la palma de aceite es una creciente causa de la pérdida de áreas cubiertas por bosque maduro, sobre todo en los últimos 10 años (Latam, 2020). En parte, esto ha sido fomentado por el crecimiento de la demanda de aceite de palma a nivel mundial en las industrias alimentaria, cosmética, farmacéutica y energética (Romero and Albuquerque, 2018; WWF (World Wildlife Fund), 2020). En muchas ocasiones, este cambio de uso de suelo ha ido acompañado de diversas afectaciones tanto sociales como ambientales que incluyen el despojo de tierras, la explotación laboral, la alteración del ciclo hidrológico, la degradación de los suelos, la pérdida de hábitat, entre otras (Isaac-Márquez et al., 2016; Meijaard et al., 2018; Romero and Albuquerque, 2018).

A pesar de que la diferenciación entre tipos de cubiertas arbóreas permite obtener mayor detalle de los procesos que ocurren en la superficie terrestre, esta discriminación ha sido muy difícil de lograr mediante sensores remotos y algoritmos previamente disponibles. Por ello, la aparición de nuevas herramientas, como los algoritmos de aprendizaje profundo y nuevas imágenes satelitales con mejores capacidades, ayudarán a obtener información de mayor calidad. Este tipo de métodos permitirán cuantificar la influencia de las presiones de mercado nacional e internacional sobre el cambio de uso de suelo de la región, *e.g.*, la expansión de plantaciones de palma de aceite, así como establecer si dicha expansión se da a costa de bosques maduros o secundarios. Además, esta información ayudará a estimaciones futuras sobre las emisiones de carbono asociadas a la deforestación y la degradación forestal, así como del cálculo de tasas de pérdida de hábitat natural o cambios en la calidad de servicios ecosistémicos (Chi et al., 2016). Por último, el contar con herramientas de esta naturaleza ayudará a distinguir áreas arbóreas bajo un uso intensivo, muchas veces asociados a una degradación de suelos (*e.g.*, plantaciones de palma), de otras zonas dominadas por árboles, pero bajo prácticas con un menor impacto ambiental (*e.g.*, áreas con ciclos de descanso agrícola).

Selección del algoritmo

Los resultados muestran claramente que el uso de la U-Net permite aumentar la precisión obtenida para generar mapas temáticos de un área de interés en comparación con algoritmos de aprendizaje automatizado como RF. En particular, la diferencia entre el *avgF1-score* y la *OA* de la U-Net MS + SAR y RF MS + SAR fue de 0.15 y 0.23,

respectivamente. Adicionalmente, nueve de las diez clases de interés obtuvieron un *FI-score* más alto en la U-Net que en su contraparte RF, gracias a la inclusión de patrones en la dimensión espacial. Estudios previos, enfocados en comparar RF con alguna técnica de aprendizaje profundo muestran resultados similares (de Bem et al., 2020; Du et al., 2020; Giang et al., 2020; Ienco et al., 2019; Kussul et al., 2017; Robinson et al., 2019; Stoian et al., 2019).

Al comparar los resultados de la U-Net MS + SAR y su contraparte RF, se detectaron tres tipos de patrones. El primero constó de las clases que obtuvieron un *FI-score*, de entre 0.08 y 0.19, mayor en la U-Net MS + SAR (Cuadro 3). La mayoría de las clases presentaron este patrón (*i.e.*, todas excepto asentamientos humanos, agua y plantaciones jóvenes). Este resultado significa que la gran mayoría de las clases se beneficiaron de incluir la dimensión espacial para discriminar entre las clases. Sin embargo, dada la diferencia relativamente baja, esto quiere decir que la gran mayoría de la información utilizada para discriminar entre clases proviene de la dimensión espectral (Figura 18).

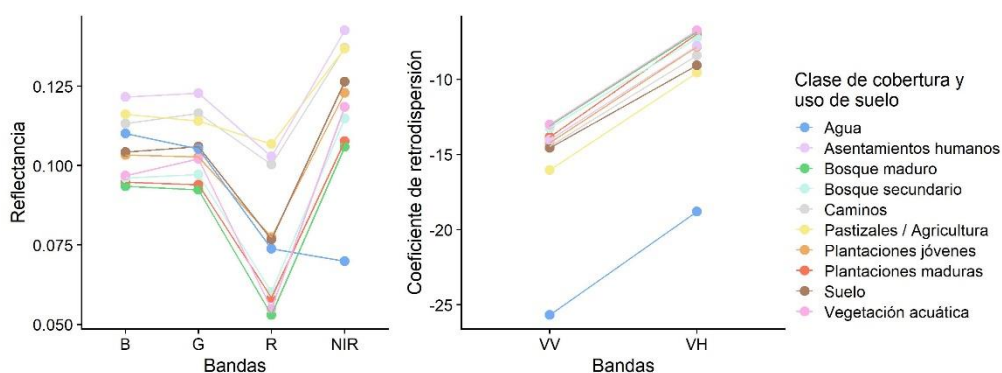


Figura 18. Promedio de la señal espectral (reflectancia y coeficiente de retrodispersión, γ^0) de las clases incluidas en el sistema de clasificación de tipos de vegetación y usos del suelo.

El segundo caso correspondió a la clase asentamiento humanos, donde el incorporar la dimensión espacial claramente tuvo un efecto dramático sobre el *FI-score* para discriminar dicha clase ($\Delta FI-score = 0.58$; Cuadro 3). Debido a que los píxeles que corresponden a esta clase pueden tener reflectancias o una señal de retrodispersión muy distintas (*e.g.*, edificios, casas, árboles, pastizales), su correcta detección depende fuertemente de la información espacial de la imagen. Por ejemplo, el promedio de la reflectancia o la retrodispersión de

esta clase es bastante similar a otras clases que se pueden encontrar en un asentamiento humano, como caminos, pastizales o agricultura, bosques secundarios o suelo. Por ello, la U-Net MS + SAR obtuvo una capacidad de identificar esta clase mucho mayor a RF, la cual utiliza únicamente de la información espectral.

El tercer caso correspondió a las clases de agua y plantaciones jóvenes, donde la diferencia de *F1-score* entre U-Net y RF fue mínima (Cuadro 2). Para la clase de agua, esto se puede deber a que se puede distinguir fácilmente de las demás clases utilizando únicamente la información espectral (Figura 18). Por lo tanto, el incorporar la dimensión espacial aumenta ligeramente el *F1-score* para dicha clase. En el caso de las plantaciones jóvenes, de hecho, la inclusión de la dimensión espacial no representó un aumento en las capacidades para identificar dicha clase. Este caso se discute con más detalle en la sección de análisis de errores.

Patrones por clase: selección de bandas

Aunque la U-Net MS + SAR fue la arquitectura que logró el mayor *avgF1-score*, la U-Net MS obtuvo valores muy cercanos. Esto quiere decir que la mayoría de las capacidades de clasificación de la U-Net MS + SAR provienen de la información MS. Este resultado es similar a estudios previos donde se comparan los resultados obtenidos para algoritmos entrenados con MS + SAR, MS o SAR para realizar clasificaciones de tipos de vegetación y usos del suelo (Heckel et al., 2020; Hirschmugl et al., 2020; Khan et al., 2020; Tavares et al., 2019; Van Tricht et al., 2018).

Al comparar los *F1-score* por clase en las tres U-Net (MS + SAR, MS y SAR) se encontraron cinco tipos de respuestas. La primera corresponde a un *F1-score* ligeramente menor en MS + SAR, en comparación con MS. La clase de camino fue la única que presentó este patrón. Debido a que esta clase consiste de líneas estrechas de píxeles, la clasificación se puede realizar mejor utilizando sólo la información MS, ya que es la que presenta la mayor resolución espacial. Aunque SAR en teoría tiene la misma resolución espacial, es probable que gracias a los filtros espaciales que se aplicaron sobre la imagen para reducir el *speckle*, dicha información no tenga la misma calidad de información a 10 m de píxel como los datos MS.

El segundo patrón constó de las clases que obtuvieron un *F1-score* intermedio a alto y que obtuvieron un *F1-score* mayor en la U-Net MS + SAR que en MS o SAR. Las dos clases que mostraron dicho patrón fueron plantaciones maduras y bosque secundario (Cuadro 2). Esto probablemente se debió a la habilidad de la información SAR de penetrar levemente en el dosel y adquirir información del arreglo geométrico y textura de las plantaciones maduras y bosques secundarios. Como estudios previos han mostrado, el uso de información SAR adicional a la MS ayuda a discriminar entre plantaciones y bosques maduros, sobre todo cuando la dimensión temporal es incluida (Mercier et al., 2019; Poortinga et al., 2019). A pesar de que la penetración de la banda C (como Sentinel-1) no es tan profunda como la de bandas L o P (Flores-Anderson et al., 2019), la polarización dual de esta banda (VV y VH) brinda información que ayudar en la correcta detección de estas dos clases.

El tercer grupo de respuestas corresponde a las clases donde la diferencia entre el *F1-score* obtenido en la U-Net MS + SAR y MS fue despreciable ($\Delta F1\text{-score} \leq 0.01$; Cuadro 2). La mayoría de las clases mostraron este patrón: pastizales / agricultura, asentamientos humanos, bosques maduros, suelos y agua. Para estas cinco clases fue evidente que la información MS contenía la mayoría de la información para identificarlas; por lo cual, el añadir las bandas SAR no aumentó las capacidades de la U-Net para identificarlas.

El cuarto patrón corresponde a la clase de agua, donde su *F1-score* fue muy similar en cualquiera de las tres U-Net (Cuadro 2). Esto parece estar relacionado con que la señal del agua es bastante diferente a las demás clases utilizando ya sea MS o SAR. Una conclusión similar se detectó al comparar U-Net MS + SAR y RF MS + SAR.

El último patrón lo mostraron las clases de vegetación acuática y plantaciones jóvenes. A pesar de que su *F1-score* fue mayor en la U-Net MS + SAR que la U-Net MS, su valor fue bastante bajo (0.11 – 0.15). Por ello, concluimos que la U-Net MS + SAR mostró una baja capacidad para identificar dichas clases. Este tema se ahonda con mayor profundidad en la sección de análisis de errores y en la evaluación de la calidad del área completa de estudio.

Exploración de hiperparámetros

La exploración de los hiperparámetros de la U-Net mostró que, aunque la combinación de hiperparámetros afectan la *OA* y el *avgF1-score*, dichos valores cambian en un intervalo de 0.05 – 0.11 entre los valores más altos y bajos. Cabe destacar que las arquitecturas U-Net que obtuvieron los mejores resultados correspondieron a arquitecturas relativamente simples, con dos o tres capas ocultas. Este resultado puede estar relacionado con el tamaño relativamente pequeño con el que se entrenaron dichos algoritmos. Estudios previos han reportado que las CNN con un número más grande de capas ocultas o filtros son capaces de detectar patrones más complejos con mejores precisiones (Buduma, 2017; LeCun et al., 2015; Li et al., 2018). Sin embargo, cuando se tiene un conjunto de datos de entrenamiento limitados, como el caso de este estudio, el principal problema con arquitecturas más complejas radica en que tienden a sobreajustarse a los datos de entrenamiento. Por ello, el escoger arquitecturas más simples, en estos casos, reduce las probabilidades de sobreajustar los pesos de la red a los datos de entrenamiento, aunque también limita el poder de abstracción de la CNN (Cheng et al., 2017; Hoerer et al., 2020; Wagner et al., 2019).

Análisis de errores

La incorrecta discriminación de ciertas clases utilizando la U-Net consistió de cinco posibilidades: 1) similitud en la señal información espacial y espectral, 2) similitud conceptual en la definición de las clases, 3) errores espaciales asociados a la arquitectura propia de la U-Net, 4) un reducido número de observaciones por clase y 5) errores asociados a la diferencia de fechas entre las imágenes Sentinel-1 y -2 utilizadas para entrenar la red y las imágenes VHR (Google, Bing y Yandex). Las primeras tres posibilidades estuvieron relacionadas con una limitación espacial de las imágenes Sentinel-1 y -2.

El primer caso fue el error más comúnmente detectado. Aunque los datos de campo y de las imágenes VHR ayudaron a determinar la clase de cada polígono, la resolución de las imágenes fue demasiado gruesa para distinguir entre las clases (Ulmas and Liiv, 2020). Por ejemplo, sin importar el tipo de imágenes utilizadas, a 10 m de resolución, las plantaciones jóvenes no se pueden distinguir de cubiertas herbáceas (pastizales / agricultura) ni las plantaciones en una fase de crecimiento intermedia de una plantación madura. Por otro

lado, la mayoría de las confusiones detectadas en las cubiertas arbóreas se relacionaron con este error, ya que las tres categorías arbóreas presentaron una señal espectral y espacial similar. En este caso, el uso de múltiples interpretes ayudaría a reducir el error de las interpretaciones visuales, así como de proveer de métricas del error esperado en los datos de entrenamiento (McRoberts et al., 2018).

El segundo tipo de errores estuvo relacionado con las decisiones arbitrarias para determinar los bordes o límites entre distintos tipos de cubiertas. A pesar de que los mismos criterios se siguieron para determinar las clases mediante interpretación visual, la delimitación conceptual y espacial de algunas clases no fue enteramente claro. Por ejemplo, los límites entre agua y suelo o entre clases arbóreas y caminos, donde los límites usualmente correspondieron a píxeles con señales mixtas de los dos tipos de cubiertas. Por ello, una cantidad mínima del error puede estar asociado a este tipo de confusiones.

El tercer caso está relacionado con el efecto de borde asociado a las predicciones realizadas por la U-Net. Por ejemplo, cuando se analizó la probabilidad de pertenencia a cada clase de los píxeles, se observó que los bordes solían presentar una más baja probabilidad de pertenencia que áreas ubicadas más al centro de los cuadros. Este efecto se puede deber a los filtros de reescalamiento de la información y pérdida de detalle en la dimensión espacial en la U-Net. A pesar de que la U-Net implementa las conexiones saltadas para reducir el efecto de los filtros de reescalamiento, puede que no sea suficiente para restaurar el detalle espacial de la imagen original, como estudios previos lo han reportado (He et al., 2019; Isaienkov et al., 2020; Ronneberger et al., 2015; Stoian et al., 2019). Por lo tanto, como se confirmó en la verificación de la clasificación del área de estudio completa, ciertos errores se asociaron con los límites de los polígonos, ya fuera por un aumento en el tamaño del polígono (*e.g.*, aumento de la superficie cubierta por nubes) o por un sesgo espacial leve (*e.g.*, límites entre caminos y plantaciones).

El cuarto tipo de error permitiría explicar por qué las clases más raras fueron las que obtuvieron los menores *FI-score*. A pesar de que las CNN son capaces de identificar clases raras, el número de muestras también limita las capacidades de las CNN para aprender a detectar una clase en particular. Por ello, es probable que el tamaño de muestra para las clases más raras como plantaciones jóvenes y vegetación acuática se haya debido a su

tamaño reducido de muestra (Ulmas and Liiv, 2020). Sin embargo, el caso de la vegetación acuática se discute en la parte final de la discusión.

Finalmente, una fuente de error mínima radica en la diferencia en las fechas de adquisición de datos de las imágenes VHR y Sentinel-1 y -2 utilizadas. En la mayoría de los casos dicha diferencia fue despreciable; sin embargo, para tres puntos clasificados como plantaciones jóvenes en la evaluación del mapa de la clasificación obtenida por la U-Net MS + SAR, no se pudo constatar si se trataba de dicha clase o no en las imágenes VHR. En estos tres casos, las áreas se clasificaron como pastizales / agricultura o suelo, dependiendo únicamente de la información presente en las imágenes Planet o Sentinel-2. A pesar de que este fue una fuente de error mínima, un acercamiento de anotaciones esparcidas podría reducir el error al permitir que no todas las áreas de entrenamiento requieran una etiqueta de clase (Stoian et al., 2019).

Comparaciones con estudios similares

Al comparar el *FI-score* y *OA* de la clasificación obtenida en este trabajo con otras clasificaciones realizadas utilizando la U-Net, se observó que este trabajo se encontró entre los valores más bajos (Du et al., 2020; Flood et al., 2019; Giang et al., 2020; Hamdi et al., 2019; Huang et al., 2018; Isaienkov et al., 2020; Neves et al., 2020; Robinson et al., 2019; Stoian et al., 2019; Ulmas and Liiv, 2020; Wagner et al., 2020a; Wagner et al., 2020b; Wagner et al., 2019; Yi et al., 2019; Zhang et al., 2018). Sin embargo, esto puede ser explicado por la combinación de tres factores: 1) un sistema de clasificación más detallado, 2) el uso de imágenes de menor resolución espacial y 3) un conjunto de datos de entrenamiento más pequeño.

En el primer caso, otros estudios han utilizados esquemas de clasificación más sencillos para los cuales la discriminación podría ser más fácil de obtener (*e.g.*, bosque / no bosque; Flood et al., 2019; Neves et al., 2020; Wagner et al., 2019). A pesar de que estos estudios tienen distintas motivaciones, la discriminación espectral y espacial de estas clases también afectará el desempeño de los algoritmos de aprendizaje profundo para resolver ciertas tareas. Otros estudios han utilizado sistemas jerárquicos de clasificación para poder calcular el desempeño de este tipo de algoritmos a distintos niveles (Ulmas and Liiv, 2020). Un esquema similar se podría haber adoptado en este trabajo para obtener un *FI-score* más

alto al trabajar con clases más gruesas, por ejemplo, al juntar las clases de plantaciones o de bosque en una sola.

Respecto a la segunda situación, estudios previos han utilizado imágenes de muy alta resolución (*e.g.*, Worldview o imágenes aéreas) para obtener altas precisiones en sus tareas de clasificación (Clark and McKechnie, 2020; Du et al., 2020; Giang et al., 2020; Hamdi et al., 2019; Huang et al., 2018; Neves et al., 2020; Wagner et al., 2020a; Wagner et al., 2020b; Wagner et al., 2019; Zhang et al., 2018). Este tipo de imágenes permiten una mejor discriminación de los distintos tipos de cubiertas, en comparación con imágenes de resolución intermedia o alta, ya que proveen de información de mayor detalle. Esto se traduce en que el desempeño de la U-Net depende fuertemente de la resolución de la imagen de entrada al modelo (Clark and McKechnie, 2020; Datcu et al., 2020; Hoeser et al., 2020). Por ejemplo, si en este estudio se hubiera utilizado una imagen de muy alta resolución, es probable que la precisión para detectar clases como plantaciones jóvenes o caminos hubiera mejorado. Adicionalmente, otros estudios han mejorado el desempeño de la U-Net al incorporar la dimensión temporal en las convoluciones (Isaienkov et al., 2020; Stoian et al., 2019), fusionado varias entradas con distinta resolución espacial cada una (Robinson et al., 2019) o utilizando versiones personalizadas de la U-Net (Gargiulo et al., 2020). Estas alternativas podrían ser interesantes de explorar en un futuro para mejorar el desempeño de la U-Net para desarrollar clasificaciones de tipo de vegetación y usos del suelo.

La tercera circunstancia está relacionada al tamaño de los datos de entrenamiento. Generalmente se menciona que el mayor potencial de los algoritmos de aprendizaje profundo es observable cuando se cuenta con un conjunto de datos de entrenamiento muy grande (Chollet and Allaire, 2018; LeCun et al., 2015). Sin embargo, en aplicaciones de la observación de la Tierra, los datos de entrenamiento suelen ser escasos, ya que requieren de una gran cantidad de tiempo y de recursos (Hoeser and Kuenzer, 2020; Yuan et al., 2021). Por ello, no es sorprendente que frecuentemente se utilicen técnicas para incrementar el tamaño de los datos de manera artificial (*data augmentation*). Otros estudios han utilizado técnicas más agresivas que las utilizadas en este trabajo para aumentar el volumen de datos, por ejemplo, utilizando rotación de las imágenes en 90°, 180°, 270°, así como modificar los

valores de brillo, saturación y matiz (Clark and McKechnie, 2020; Du et al., 2020; Wagner et al., 2020b). A pesar de que estas técnicas ayudan a que el algoritmo encuentra patrones más generales, suelen ser computacionalmente demandantes por lo que frecuentemente uno se tiene que ajustar a los recursos disponibles. En este caso, se optó por un esquema que permitiera un entrenamiento veloz, dados los recursos computacionales disponibles, para acortar el tiempo demorado para realizar la exploración de hiperparámetros.

Estudios previos han utilizado la transferencia de aprendizaje (*transfer learning*) para compensar la disponibilidad de un conjunto de datos de entrenamiento pequeños y alcanzar altas precisiones en tareas de clasificación con algoritmos de aprendizaje profundo (Alom et al., 2019; Ball et al., 2017; Zhao et al., 2019). En este trabajo no se utilizó esta técnica ya que la mayoría de las CNN se han entrenado utilizando imágenes RGB, mientras que las entradas utilizadas en este trabajo contaron con las bandas NIR y dos bandas SAR. Adicionalmente, muchas de las CNN pre-entrenadas disponibles han sido entrenadas con imágenes con una geometría, resolución y sistemas de clasificación muy distintos a los utilizados en aplicaciones de la Observación de la Tierra (Hoeser and Kuenzer, 2020; Li et al., 2018). Por ello, se optó por entrenar la U-Net sólo con los datos de entrenamiento disponibles.

A pesar de que es claro que estos tres factores juegan un papel para determinar el desempeño de la U-Net para obtener clasificaciones, la forma en la que interactúan no resulta tan claro. Estudios futuros deberían tratar de describir estas interacciones para entender los efectos de estos factores sobre el desempeño para obtener clasificaciones.

Aspectos metodológicos

Aunque es relativamente común que se implemente el protocolo desarrollado por (Cochran, 1977; Olofsson et al., 2014) para evaluar la precisión de un mapa, en estudios que utilizan algoritmos de aprendizaje profundo, raramente son utilizados (pero ver Clark and McKechnie, 2020; Du et al., 2020). Además, estamos conscientes de que si un conjunto de datos de verificación es lo suficientemente grande, con una distribución espacial aleatoria y no ha sido expuesto a la U-Net durante la fase de entrenamiento, debería de obtener resultados similares a los obtenidos con el protocolo antes mencionado. Sin embargo, al comparar el valor de *F1-score* por clases obtenidos en los datos de verificación y en la

verificación estratificada del área de estudio completa, se puede observar que para tres clases la diferencia fue mayor a 0.2 (*i.e.*, vegetación acuática, caminos, plantaciones jóvenes). Esta diferencia puede estar relacionada a dos situaciones en las cuales se infló de manera artificial el error para estas tres clases: 1) contaron con un tamaño de muestra muy pequeño, 2) y su ubicación espacial en los datos de verificación.

En el primer caso, el efecto de la inflación del error resulta más evidente en las clases poco frecuentes, debido probablemente al número limitado de observaciones. Gracias a que los datos de verificación vienen de un procedimiento de aumento artificial de los datos, estos no son completamente independientes, ya que muchas imágenes sólo son una imagen girada o espejeada de otra. Por lo tanto, si un píxel fue clasificado de manera errónea en una observación, es probable que también se clasificara de manera errónea en una versión en espejo o girada. Esto permitiría explicar por qué el *F1-score* obtenido en los datos de validación fue mucho menor que el obtenido en la muestra estratificada ($0.20 \leq \Delta F1\text{-score} \leq 0.25$).

En el segundo caso, al comparar el *F1-score* obtenido para vegetación acuática en los datos de verificación y del muestreo estratificado se observó una diferencia significativa ($\Delta F1\text{-score} = 0.55$), por lo cual, se buscó otro factor que fuera responsable de tal diferencia. Al analizar visualmente la ubicación espacial de esta clase en los datos de verificación se encontró que normalmente se encontraba en los bordes de los cuadros de las imágenes de entrenamiento, lo cual fue un aspecto completamente ignorado en el registro de los datos de entrenamiento y validación. Por ello, debido a su ubicación en los bordes, es probable que las predicciones realizadas para esta clase contuvieran mayor error. Esto a su vez, provocó que el error calculado para esta clase en los datos de verificación fuera mucho mayor que el detectado en el muestreo estratificado del mapa completo.

Consideramos que el utilizar el protocolo de verificación de Olofsson et al., (2014) permite obtener estimados más precisos del potencial de la U-Net para generar mapas de clasificación y reducir el error asociado a predicciones de mala calidad realizadas en los bordes de los datos. Adicionalmente, al utilizar este protocolo se puede obtener un estimado de área por clase con un intervalo de confianza. Por ello, sugerimos que futuros estudio implementen dicho protocolo para evaluar mapas.

Cabe destacar que se optó por trabajar con una imagen de un solo tiempo para maximizar la concordancia entre la información de campo y la imagen MS, en lugar de maximizar la cobertura de la superficie. A pesar de que pocos estudios han evaluado el efecto de la composición multitemporal sobre los resultados de una clasificación (Chuvieco et al., 2005; Praticò et al., 2021), prácticamente ninguno lo ha evaluado utilizando algoritmos de aprendizaje profundo. Estos estudios reportan que usualmente los mejores resultados se obtienen al obtener mosaicos multitemporales de una ventana temporal corta. Sin embargo, en áreas con alta nubosidad durante todo el año, como el sitio estudiado en este trabajo, aun utilizando mosaicos multitemporales no se puede asegurar que se obtenga un compuesto de alta calidad o de una cobertura del área de estudio completa. Estudios futuros deberían abordar este dilema.

Por otro lado, en la literatura se ha utilizado frecuentemente al índice kappa para resumir los resultados de una matriz de confusión, ya que permite tomar en cuenta los valores marginales de la matriz de confusión. Sin embargo, varios trabajos han recomendado que no se utilice dicho índice debido a que resulta uno difícilmente interpretable o no aporta información adicional a la precisión total (Pontius and Millones, 2011; Stehman, 1997). Por ello, nosotros decidimos restringirnos al uso de la *OA* y el *F1-score* como métricas para elegir la mejor combinación de hiperparámetros y la precisión del usuario y del productor como métricas para describir la matriz de confusión.

Adicionalmente, aunque el algoritmo U-Net en teoría es capaz de incorporar patrones espaciales que incluyan aspectos como la forma o la textura de los objetos para realizar la clasificación, el aprendizaje de la U-Net queda registrado en los pesos de los filtros, los cuales resultan difíciles de interpretar en términos de formas o texturas. Por ello, algunos estudios han buscado métodos que permitan entender lo que este tipo de algoritmos “aprenden” mediante pruebas post-aprendizaje sobre imágenes con diferentes patrones espectrales y espaciales (Chollet and Allaire, 2018). Este tipo de métodos definitivamente son una opción atractiva para tratar de entender en qué radica el “aprendizaje” de estos algoritmos y tratar de deducir si la importancia de distintas características de la imagen como el color, la textura, los bordes o la forma.

Conclusiones

El uso de CNN para aplicaciones de observación de la Tierra ha permitido mejorar las capacidades para obtener clasificaciones detalladas de la superficie terrestre. Sin embargo, resulta esencial evaluar el papel de distintos tipos de imágenes e información para obtener mapas temáticos con especial énfasis en discriminar entre distintos tipos de cubierta arbórea. En este estudio se encontró que la U-Net con un conjunto de datos pequeño obtuvo mejores resultados que su contraparte RF. El mapa que obtuvo la mayor precisión total fue U-Net MS + SAR, seguido de MS y por último, SAR. Adicionalmente, U-Net MS + SAR obtuvo *F1-score* más altos para las clases de bosque maduro y plantaciones. Concluimos que el mejor desempeño de la U-Net en comparación con RF se debe sobre todo a la incorporación de rasgos en la dimensión espacial, además de la espectral, para generar la clasificación. Este estudio demostró las capacidades de una CNN para obtener un mapa de tipos de vegetación y usos del suelo a partir de imágenes de mediana resolución espacial, con énfasis en discriminar entre bosque maduro y secundario, así como plantaciones jóvenes y maduras. Finalmente, el desarrollo de este tipo de métodos ayudará a monitorear la superficie terrestre con mayor detalle, lo cual brindará información para profundizar en el entendimiento de las relaciones socioambientales y brindar información que permita una mejor gestión de los recursos naturales.

Material suplementario

Cuadro S 1. Matriz de confusión para los datos de validación, utilizando la U-Net MS + SAR. Las filas muestran las clases predichas, mientras que las columnas, las clases manualmente clasificadas. Las abreviaciones son: A: agua, AH: asentamientos humanos, BM: bosque maduro, BS: bosque secundario, C: caminos, P/A: pastizales / agricultura, PJ: plantaciones jóvenes, PM: plantaciones maduras, S: suelo, VA: vegetación acuática; UA, precisión del usuario, PA: precisión del productor.

	A	AH	BM	BS	C	P/A	PJ	PM	S	VA	UA
A	59215	1	775	9	0	855	0	44	272	170	0.97
AH	0	159892	627	2010	2810	6917	13	1278	1926	0	0.91
BM	1595	5081	1372026	68833	3245	55782	3938	116552	3980	6001	0.82
BS	0	8288	67390	143531	2463	67145	5531	49581	1287	0	0.41
C	0	3488	895	986	21636	8587	653	2364	10066	0	0.44
P/A	322	11583	35162	64813	23834	902138	22190	25472	88484	0	0.76

PJ	0	197	164	1104	927	8338	5159	12528	2668	0	0.17		
PM	128	238	26673	15303	3790	22157	26695	246789	1667	0	0.72		
S	1041	1745	3046	1363	12314	48149	2280	1413	181030	21	0.69		
VA	0	0	240	0	0	115	0	0	0	594	0.63		
PA	0.95	0.84	0.91	0.48	0.30	0.80	0.08	0.54	0.61	0.09			
											Precisión total (OA)	0.76	
												avgF1-score por lote	0.72
												avgF1-score total	0.58

Cuadro S 2. Matriz de confusión para los datos de validación, utilizando la U-Net MS. Las filas muestran las clases predichas, mientras que las columnas, las clases manualmente clasificadas. Las abreviaciones son: A: agua, AH: asentamientos humanos, BM: bosque maduro, BS: bosque secundario, C: caminos, P/A: pastizales / agricultura, PJ: plantaciones jóvenes, PM: plantaciones maduras, S: suelo, VA: vegetación acuática; UA, precisión del usuario, PA: precisión del productor.

	A	AH	BM	BS	C	P/A	PJ	PM	S	VA	UA		
A	61225	0	916	38	0	1104	0	82	409	451	0.95		
AH	0	171136	1782	4903	5038	8727	130	4255	5070	0	0.85		
BM	324	2401	1386121	71614	3906	57945	14346	136575	3827	5497	0.81		
BS	1	1888	43919	89761	870	43758	5930	45045	338	317	0.39		
C	0	1256	1393	1382	26717	7045	561	1861	11403	0	0.52		
P/A	45	12870	38984	105371	24740	945076	33760	58029	103450	109	0.71		
PJ	0	1	171	368	129	2316	1001	3178	1180	0	0.12		
PM	0	381	34138	23195	2072	23140	10017	206132	1857	1	0.68		
S	536	771	1631	768	8216	32971	697	899	162686	13	0.76		
VA	0	0	25	62	0	327	0	8	0	397	0.48		
PA	0.98	0.9	0.91	0.3	0.37	0.84	0.02	0.45	0.55	0.06			
												Precisión total (OA)	0.76
												avgF1-score por lote	0.72
												avgF1-score total	0.55

Cuadro S 3. Matriz de confusión para los datos de validación, utilizando la U-Net SAR. Las filas muestran las clases predichas, mientras que las columnas, las clases manualmente clasificadas. Las abreviaciones son: A: agua, AH: asentamientos humanos, BM: bosque maduro, BS: bosque secundario, C: caminos, P/A: pastizales / agricultura, PJ: plantaciones jóvenes, PM: plantaciones maduras, S: suelo, VA: vegetación acuática; UA, precisión del usuario, PA: precisión del productor.

	A	AH	BM	BS	C	P/A	PJ	PM	S	VA	UA	
A	59834	0	1291	4	0	1014	0	25	3088	222	0.91	
AH	0	79076	19238	7312	2662	14585	3997	23998	5583	0	0.51	
BM	1727	70460	1512101	138688	13216	150845	6467	183508	24125	6561	0.72	
BS	0	2610	50157	48959	2383	41897	2351	18848	2739	0	0.29	
C	0	16	1	1	7	103	6	32	30	0	0.04	
P/A	124	26365	115271	103875	48227	944129	33311	50653	199658	1	0.62	
PJ	0	150	77	201	202	2382	203	2014	463	0	0.04	
PM	64	12542	36230	14267	6766	25098	17395	182438	12689	1	0.59	
S	552	514	2610	1030	1690	16572	2876	2725	60535	1	0.68	
VA	0	0	0	0	0	0	0	0	0	0	0	
PA	0.96	0.41	0.87	0.16	0	0.79	0	0.39	0.20	0		
											Precisión total (OA)	0.65
											<i>avgF1-score</i> por lote	0.57
											<i>avgF1-score</i> total	0.39

Cuadro S 4. Matriz de confusión para los datos de validación, utilizando RF MS + SAR. Las filas muestran las clases predichas, mientras que las columnas, las clases manualmente clasificadas. Las abreviaciones son: A: agua, AH: asentamientos humanos, BM: bosque maduro, BS: bosque secundario, C: caminos, P/A: pastizales / agricultura, PJ: plantaciones jóvenes, PM: plantaciones maduras, S: suelo, VA: vegetación acuática; UA, precisión del usuario, PA: precisión del productor.

	A	AH	BM	BS	C	P/A	PJ	PM	S	VA	UA	
A	8542	1	215	67	0	439	4	16	127	73	0.90	
AH	0	9104	1760	1185	1512	10147	266	539	2672	2	0.28	
BM	0	322121083		4389	29	1180	411	12033	10	296	0.85	
BS	0	2804	37231	24686	470	20566	2146	13487	288	56	0.24	
C	3	6247	525	627	5367	12179	229	620	7774	0	0.15	
P/A	12	2459	3373	5325	1385	93773	1725	1220	6354	16	0.80	
PJ	0	4343	3747	3381	933	19031	3211	6912	3575	1	0.07	
PM	1	1451	17256	4370	129	2552	2065	23980	126	56	0.46	
S	8	2015	322	162	1110	12449	542	244	26977	6	0.59	
VA	0	133	19409	3474	22	3068	174	615	6	551	0.02	
PA	0.99	0.30	0.58	0.51	0.48	0.53	0.29	0.40	0.55	0.49		
											Precisión total (OA)	0.53
											<i>avgF1-score</i> por lote	-
											<i>avgF1-score</i> total	0.43

Cuadro S 5. Matriz de confusión para los datos de verificación del mapa para el área completa de estudio, utilizando la U-Net MS + SAR. Las filas muestran las clases predichas, mientras que las columnas, las clases manualmente clasificadas. Las abreviaciones son: A: agua, AH: asentamientos humanos, BM: bosque maduro, BS: bosque secundario, C: caminos, P/A: pastizales / agricultura, PJ: plantaciones jóvenes, PM: plantaciones maduras, S: suelo, VA: vegetación acuática; UA, precisión del usuario, PA: precisión del productor.

	A	AH	BM	BS	C	P/A	PJ	PM	S	VA	UA
A	23	0	0	0	0	0	0	0	0	0	1.00
AH	0	22	0	0	0	0	0	0	0	0	1.00
BM	0	0	113	2	0	2	0	2	0	4	0.88
BS	1	0	5	18	1	9	1	1	0	2	0.44
C	0	1	0	0	12	0	0	1	0	0	0.8
P/A	0	0	4	1	1	100	5	2	5	3	0.75
PJ	0	1	0	0	0	3	6	1	0	1	0.46
PM	0	0	4	3	1	1	9	18	0	0	0.46
S	0	1	1	0	10	5	4	0	20	0	0.42
VA	0	0	3	0	0	0	0	0	0	15	0.83
PA	0.92	0.88	0.86	0.72	0.48	0.83	0.24	0.72	0.80	0.60	
											Precisión total (OA) 0.77
											avgF1-score por lote -
											avgF1-score total 0.68

Capítulo 4. Detección de la deforestación utilizando un enfoque de aprendizaje profundo espacio temporal con imágenes de radar de apertura sintética y multiespectrales.

Jonathan V. Solórzano; Jean-François Mas; J. Alberto Gallardo-Cruz; Yan Gao; Ana Fernández-Montes de Oca

Artículo publicado en: ISPRS Journal of Photogrammetry and Remote Sensing (2023)
199, 87-101

Resumen

La deforestación es un conductor del cambio global que contribuye a las emisiones de carbono a la atmósfera, a la pérdida de biodiversidad y a la degradación de servicios ecosistémicos. Usualmente, este proceso se ha cuantificado y monitoreado utilizando la percepción remota. El desarrollo de algoritmos de aprendizaje profundo, así como imágenes de mayor resolución espacial y temporal ha mejorado las capacidades de monitoreo de este proceso. El objetivo de este estudio fue evaluar el potencial de un algoritmo de aprendizaje profundo espacio temporal, la U-Net 3D, con imágenes multiespectrales y radar de apertura sintética para detectar la deforestación en un bosque tropical húmedo en el sureste mexicano entre 2019 y 2020. Las entradas de la U-Net 3D constaron de 147 áreas de 256 por lado que contenían, al menos un área deforestada. Estas entradas constaron de: 1) cuatro compuestos sin nubes (Feb – Abr 2019, May – Sep 2019, Oct 2019 – Ene 2020 y Feb – Abr 2020) con seis bandas cada uno (cuatro bandas Sentinel-2 y dos Sentinel-1), y 2) polígonos de áreas deforestadas delineadas mediante interpretación visual. El sistema de clasificación constó de tres clases: no deforestación, pérdida de bosque maduro y pérdida de bosque secundario / plantaciones. El mapa de clasificación obtuvo una precisión total de 0.97 y un *F1-score* promedio de 0.94. Los estimados de área corregidos fueron $3\ 195.26 \pm 1\ 132.88$ ha (0.48 ± 0.17 %) para la pérdida de bosque maduro y $4\ 234.85 \pm 1\ 912.01$ (0.64 ± 0.29 %) para la pérdida de bosque secundario o plantaciones. La mayoría de los errores de la clasificación se encontraron en los bordes

entre clases contiguas, causados por la confusión de pérdida de bosque secundario / plantaciones con pérdida de cobertura herbácea o asociada con artefactos en las imágenes.

Palabras clave: bosque tropical húmedo, pérdida de bosque, cambio de cubierta y uso del suelo (LULCC), redes neuronales convolucionales (CNN), U-Net 3D, Sentinel-1 y Sentinel-2.

Métodos

El área de estudio fue la misma que en el capítulo anterior, es decir, la región suroriental de la Selva Lacandona, Chiapas que cubre el municipio de Marqués de Comillas y Benemérito de las Américas, así como áreas aledañas en el municipio de Ocosingo (Figura 19). Sin embargo, en este caso, se amplió el área de estudio para contar con un mayor número de observaciones de áreas deforestadas, lo cual abarcó aprox. 5,900 km².

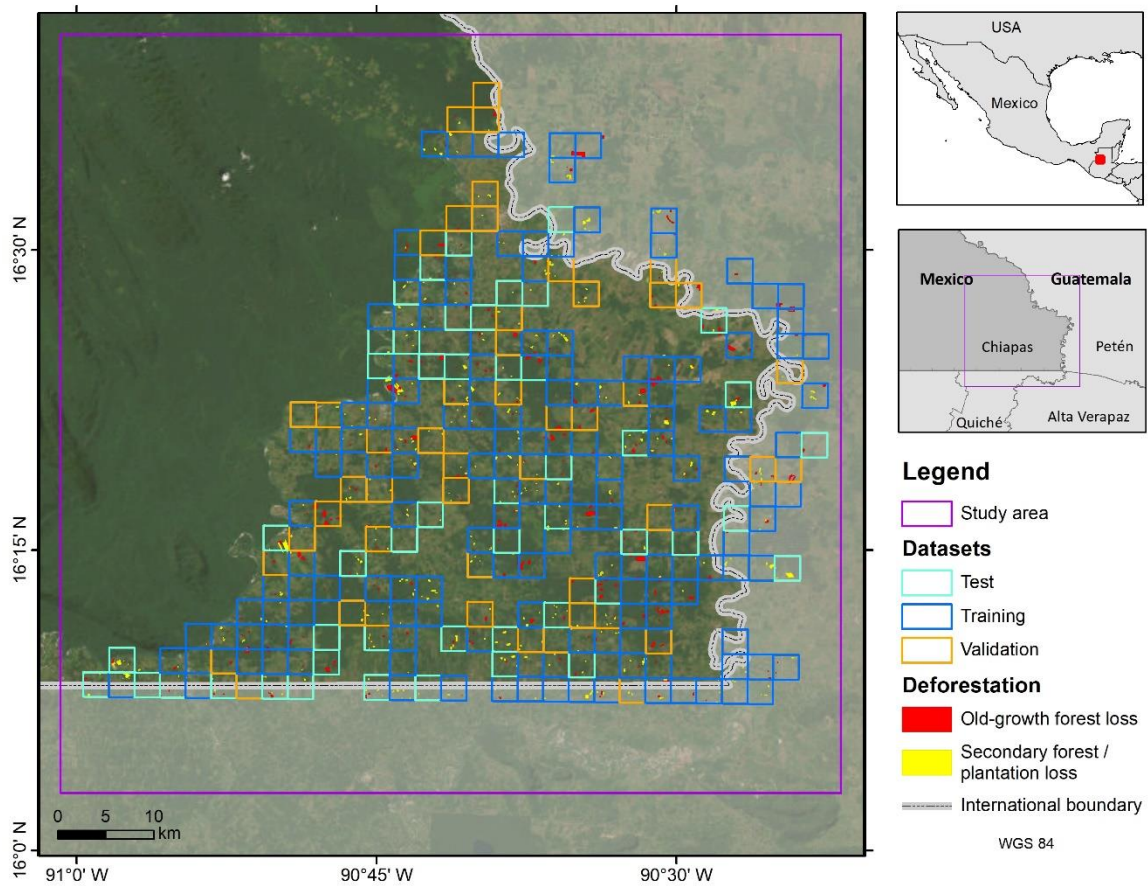


Figura 19. Localización del área de estudio y de los polígonos de pérdida de bosque maduro y bosque secundario o plantaciones, delimitados mediante interpretación visual.

El método siguió tres fases: 1) el preprocesamiento de las imágenes satelitales y la clasificación manual, 2) el entrenamiento y validación de un algoritmo de aprendizaje profundo (DL), *i.e.*, U-Net 3D y uno de aprendizaje automatizado (ML), *i.e.*, support vector

machines (SVM) y 3) la selección de la mejor arquitectura, la clasificación del área completa de estudio y la verificación del mapa para el área de estudio (Figura 20).

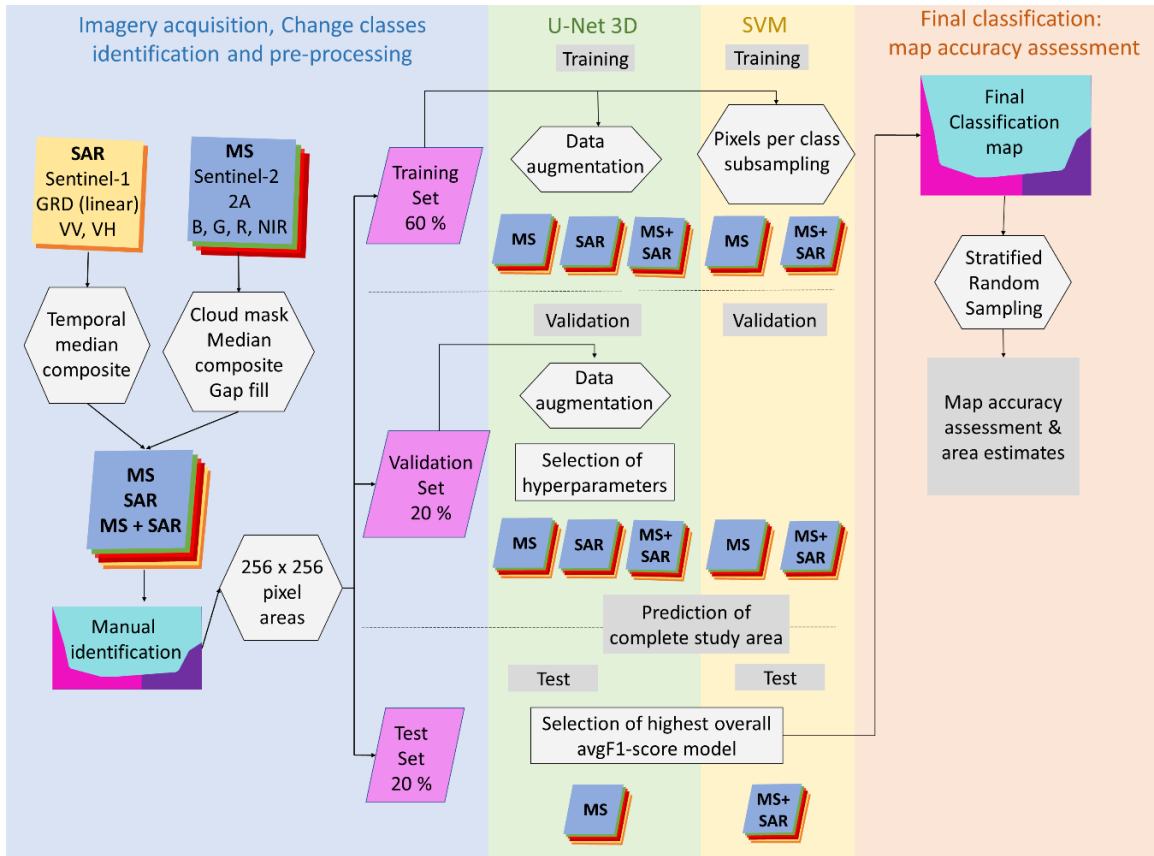


Figura 20. Diagrama de flujo del procedimiento para entrenar y validar las arquitecturas U-Net 3D y Support Vector Machines.

Procesamiento de imágenes

Como insumos para realizar las predicciones de los cambios en las clases de cubierta terrestre con el modelo U-Net 3D, se utilizaron imágenes radar de apertura sintética (SAR) y multiespectrales (MS) registradas por los sensores Sentinel-1 y -2, respectivamente. El primer paso consistió en definir el periodo más corto en el que se podía obtener un mosaico multitemporal con la mínima presencia de nubes para las imágenes MS, el cual correspondió a Feb – Abr 2019 y 2020. Posteriormente se definieron dos periodos intermedios para contar con una mayor resolución temporal en la evaluación: May – Sep

2019 y Oct 2019 – Ene 2020. En cada periodo el compuesto multitemporal se calculó como la mediana de la reflectancia para su correspondiente periodo, tanto para las imágenes MS como SAR.

En el caso de las imágenes Sentinel-2 se utilizó la información proveniente de la colección 2A que corresponde a valores de reflectancia por debajo de la atmosfera. Para enmascarar las nubes y sombras de estas imágenes se utilizó el método *s2cloudless*, utilizando las bandas SCL, QA60 e NIR (Skakun et al., 2022). Tras realizar este procedimiento, se aplicó un filtro para calcular el promedio de reflectancia en ventanas de 5 x 35 píxeles y rellenar los pocos píxeles sin datos remanentes. La imagen final contó con cuatro bandas, rojo (R), verde (G), azul (B) e infrarrojo cercano (NIR), que corresponden a las bandas de mayor resolución, *i.e.*, 10 m. Aunque las imágenes Sentinel-2 cuentan con más bandas, estas no se utilizaron por su menor resolución (*e.g.*, SWIR o RE), ya que se han reportado mejores resultados con imágenes de mayor resolución en conjunto con algoritmos basados en CNN (Clark and McKechnie, 2020; Hoeser et al., 2020; Matosak et al., 2022; Torres et al., 2021).

En el caso de los datos Sentinel-1 se utilizaron las imágenes de la colección *Ground Range Detected* (GRD), registradas en el modo *Interferometric Wide Swath* (IWS). Estas imágenes contienen la información de retrodispersión en escala lineal en dos bandas, VV (polarización emisión y recepción vertical - vertical) y VH (vertical – horizontal). Debido a que estudios previos han reportado que el uso de filtros espaciales para reducir el *speckle* en imágenes SAR puede reducir las capacidades de detección de los algoritmos basados en CNN, no se aplicó ningún filtro (Ortega Adarme et al., 2022).

Después de este procedimiento se obtuvieron los mosaicos MS y SAR para los cuatro periodos mencionados: Feb – Abr 2019 (t1) May – Sep 2019 (t2), Oct 2019 – Jan 2020 (t3) and Feb – Abr 2020 (t4). Para comparar entre los tipos de entradas para entrenar los algoritmos, se crearon tres tipos de imágenes, MS + SAR con 6 bandas (R, G, B, NIR, VV y VH), MS (R, G, N, NIR) y SAR (VV y VH). Las imágenes se estandarizaron mediante la resta del promedio de cada banda y la división entre su desviación estándar. Todo este preprocesamiento de las imágenes se realizó en la API de Javascript de Google Earth Engine (Gorelick et al., 2017).

Generación de datos de entrenamiento

Para generar los datos de entrenamiento se consultaron los datos del Global Forest Change (GFC) para 2019 y 2020, realizada a partir de imágenes Landsat (Hansen et al., 2013), una evaluación de pérdida de vegetación arbórea para 2019 y 2020 realizada a partir de mosaicos anuales de Landsat del índice normalizado de sequía multibanda (NMDI; (Solórzano et al., 2018) y la clasificación de tipos de vegetación y uso de suelo generada para 2019 en el capítulo anterior (Solórzano et al., 2021). Los polígonos de deforestación identificados tenían que ser observables entre la primera imagen de los mosaicos MS y la última (feb – abr 2019 y 2020), ya que se supuso que esos serían los polígonos que podrían reconocer los algoritmos utilizados.

Un total de 1084 polígonos de pérdida de vegetación arbórea se delinearon, los cuales se clasificaron en dos clases: pérdida de bosque maduro ($n = 380$) y pérdida de bosque secundario o plantaciones ($n = 704$). Adicionalmente, para todas las áreas que no correspondieran a alguno de esos dos cambios, se le asignó la clase de no cambio. De esta manera el sistema de clasificación contó con tres clases: no cambio, pérdida de bosque maduro y pérdida de bosque secundario o plantaciones. Estos datos se rasterizaron a la misma resolución que las imágenes MS y SAR.

A pesar que las entradas para el algoritmo SVM corresponden a datos por píxel, la U-Net 3D requiere de pequeñas imágenes sobre las que detectará patrones espaciales relevantes. Por ello, para usar la misma información para entrenar, verificar y probar los dos algoritmos, el sitio de estudio se dividió en una cuadrícula de celdas de 256 píxeles por lado. Después, se eliminaron las áreas que no se superpusieron con algún polígono de deforestación, lo cual resultó en un total de 245 cuadrados. Estas áreas se utilizaron para cortar las imágenes de entrada, *i.e.*, MS, SAR y MS + SAR, así como sus correspondientes etiquetas (*i.e.*, esquema de clasificación). De estas áreas 147 se usaron como datos de entrenamiento (60 %), 49 como datos de validación (20 %) y 49 como datos de prueba (20 %; Figura 19). Para entrenar el SVM, estos datos se ingresaron por píxel, mientras que para la U-Net 3D, por cuadrado. Los datos sufrieron un distinto procesamiento para fungir como entradas para SVM y U-Net 3D, por lo que estos detalles se mencionan en la sección de cada algoritmo.

Entrenamiento de algoritmos

U-Net 3D

Para construir el modelo U-Net 3D se agregó la dimensión temporal a la entrada de la U-Net, así como a los filtros convolucionales y de reescalamiento. Esta arquitectura de U-Net 3D es similar a la presentada en trabajos anteriores (Çiçek et al., 2016; Isaienkov et al., 2020; Ji et al., 2018). Debido a que la entrada de los datos de entrenamiento sólo contaba con cuatro imágenes en la dimensión temporal, los filtros de reescalamiento de las primeras dos capas ocultas fueron de $2 \times 2 \times 2$ píxeles para reducir las tres dimensiones a la mitad en cada pase (*i.e.*, tiempo, alto, ancho). Tras realizar esas dos reducciones, los filtros de reescalamiento tuvieron dimensiones de $1 \times 2 \times 2$ píxeles, para sólo continuar reduciendo la dimensión espacial (Figura 21). Por el contrario, los filtros convolucionales siempre tuvieron una dimensión de $3 \times 3 \times 3$ píxeles. La U-Net 3D fue creada utilizando una versión modificada del paquete unet para usar convoluciones de tres dimensiones en lugar de 2 (Falbel and Zak, 2020).

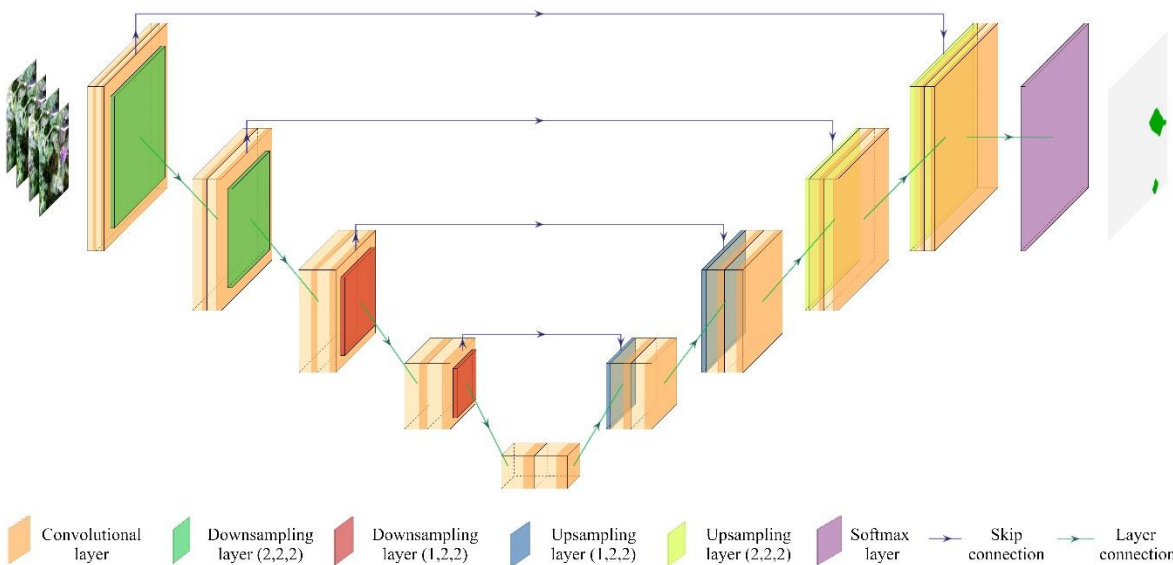


Figura 21. Esquema de la U-Net 3D utilizada para detectar la deforestación. La imagen del extremo izquierdo representa la serie de cuatro imágenes utilizada como entrada, mientras que la del extremo derecho, la salida que corresponde a una clasificación de sitios deforestados y no. Figura realizada mediante el código de

<https://github.com/HarisIqbal88/PlotNeuralNet>.

Por su parte, para realizar una búsqueda de la combinación de hiperparámetros óptimos se realizaron varias iteraciones entrenando el algoritmo por 9 épocas y utilizando un tamaño de lote de 16 muestras. La combinación de hiperparámetros explorados constaron de: entre 4 y 5 capas ocultas y tres valores de probabilidad de *dropout* (0.1, 0.3, 0.5); (Srivastava et al., 2014). Otros parámetros del modelo incluyeron el uso de 64 filtros en la primera capa, el optimizador Adam, una normalización por lote y una inicialización normal de He para inicializar los pesos de los filtros (He et al., 2015). Por su parte, se utilizó una función de pérdida compuesta por la suma ponderada de la entropía cruzada categórica y el *avgF1-score* (ecuación 9; similar a la propuesta por (Isaienkov et al., 2020):

$$Loss = 0.2 * CCE + 0.8 * (1 - avgF1-score) \quad (9)$$

donde CCE corresponde a la entropía cruzada categórica (*categorical crossentropy*), mientras que *avgF1-score* corresponde al promedio por clases del *F1-score* (ecuación 5).

Para contar con un número suficientemente grande de observaciones en los datos de entrenamiento y de verificación, se aplicó un proceso de aumento artificial de los datos. En este proceso, cada cuadrado de 256 píxeles por lado se dividió en nueve cuadrados de 128 píxeles por lado, utilizando un desfase vertical y horizontal de 64 píxeles, lo cual resultó en 1323 imágenes para los datos de entrenamiento y 441 para los de validación (Figura 19). Se generaron en total seis modelos por arquitectura U-Net 3D (*i.e.*, MS + SAR, MS o SAR). El modelo que obtuvo el mayor *avgF1-score* en los datos de validación se eligió como el mejor modelo por tipo de entrada, resultando en tres mejores arquitecturas: U-Net 3D MS, U-Net 3D SAR y U-Net 3D MS + SAR. El *avgF1-score* se calculó como el promedio de los *F1-score* por clases, que a su vez correspondió al promedio armónico de la precisión (precisión del usuario) y la sensibilidad (precisión del productor) de las clases predichas (ecuación 4; Zou et al., 2004).

Posteriormente, para obtener la clasificación del área completa de estudio, se dividieron las imágenes en un tamaño que pudiera ser utilizado por la U-Net, *i.e.*, cuadrados de 128 píxeles por lado. Después, se obtuvo una clasificación para cada cuadrado de 128

píxeles por lado y se hizo un mosaico a partir de ellos. Para evitar obtener predicciones de menor calidad en las esquinas de los cuadrados de 128 píxeles, debido a que tienen un vecindario de píxeles más pequeño (Flood et al., 2019; Ronneberger et al., 2015), se realizaron predicciones sobre los cuadrados provenientes de dos cuadrículas. Estas dos cuadrículas presentaron un desfase vertical y horizontal de 64 píxeles para que el centro de los cuadrados de una se superpusiera con las esquinas de los cuadros de la otra. A continuación, se construyeron mosaicos del área de estudio completa con los datos de cada cuadrícula por separado para obtener dos clasificaciones y dos mapas de probabilidad. Posteriormente, la clase de cada píxel se definió como la que presentó la mayor probabilidad en cualquiera de las dos imágenes. Finalmente, se evaluó esta clasificación utilizando los datos de prueba sin ningún aumento artificial, *i.e.*, 49 cuadrados de 256 píxeles por lado.

Support Vector Machines

El algoritmo support vector machines (SVM) es uno frecuentemente para realizar clasificaciones con imágenes satelitales, sobre todo con la función de base radial (Boser, 1992; Dabija et al., 2021; Sheykhmousa et al., 2020). Este algoritmo no paramétrico supone que los datos son separables linealmente en un espacio de atributos, por lo cual se define un hiperplano, a partir de los datos que caen en la frontera de la decisión, para realizar la clasificación (Mountrakis et al., 2011). A las observaciones utilizadas para definir el hiperplano se le llaman vectores de apoyo (*support vectors*), de ahí el nombre del algoritmo. Debido a que comúnmente los datos no se pueden separar con un simple plano, se utiliza una función de *kernel* para transformar el espacio de los atributos en uno que ayude a separar las clases con planos lineales. En este caso, se empleó el kernel de función de base radial (*radial based*). En este algoritmo se definen normalmente dos parámetros: costo y sigma. El primero consiste de un valor de penalización de predecir una observación en el lado errónea del hiperplano, mientras que sigma es un parámetro de la precisión. En este caso se utilizó un costo de 1 y una sigma calculado de manera automática a partir de un procedimiento heurístico (Kuhn and Wickham, 2020).

A diferencia de la U-Net 3D, SVM sólo utiliza la información espectral para realizar la clasificación, por lo que las imágenes de distintos tiempos se concatenaron en la dimensión

espectral para formar imágenes de 8 bandas (SAR), 16 (MS) y 24 (MS + SAR). Los datos de entrenamiento de la SVM correspondieron a una submuestra aleatoria de 63 070 píxeles por clase de los datos de entrenamiento originales. Este número correspondió al número de observaciones de la clase más rara (*i.e.*, pérdida de bosque maduro) en los datos de entrenamiento y tuvo el propósito de balancear el número de observaciones por clase. En el caso del SVM, debido a que es un método por píxel, las predicciones se realizaron directamente sobre los datos de validación y prueba sin ningún aumento artificial de los datos. En el caso de SVM, se ajustaron los modelos con MS y MS + SAR únicamente, debido a que fueron los que obtuvieron el mayor *avgF1-score* utilizando la U-Net 3D.

Métricas de evaluación

Se utilizaron tres métricas para evaluar el desempeño de los dos algoritmos: la exactitud total (*OA*), *F1-score* por clase y el *F1-score* promedio (*avgF1-score*; Zou et al., 2004). Estas métricas fueron exactamente las mismas a las utilizadas en el capítulo anterior.

$$OA = \frac{TP + TN}{TP + FP + TN + FN} \quad (10)$$

$$UA = \frac{TP}{TP + FP} \quad (11)$$

$$PA = \frac{TP}{TP + FN} \quad (12)$$

$$F1 = \frac{2 UA PA}{UA + PA} \quad (13)$$

$$avgF1 = \frac{1}{C} \sum_{c=1}^C F1_c \quad (14)$$

Donde *OA* corresponde a la exactitud total, TP a los verdaderos positivos, TN a los verdaderos negativos, FP a los falsos positivos, FN a los falsos negativos. UA corresponde a la precisión del usuario, PA, a la precisión del productor, F1 al *F1-score*, C al total de clases, c a cada clase y avgF1 al promedio de *F1-scores*. Además, el intervalo de confianza

del 95 % se calculó para la *OA* y *avgFI-score* utilizando *10-fold cross validation* para detectar diferencias significativas entre los algoritmos. Finalmente, la mejor arquitectura se eligió como la que obtuvo el *avgFI-score* más alto en los datos de prueba y se utilizó para hacer el mapa de clasificación final.

Verificación de la clasificación del área completa de estudio

El paso final fue evaluar la precisión del mapa producido utilizando el algoritmo que obtuvo el mayor *avgFI-score* en los datos de prueba y utilizando un muestreo aleatorio estratificado (Card, 1982; Cochran, 1977; Olofsson et al., 2014). Estudios previos han sugerido utilizar un estrato de buffer para verificar un mapa con pesos de área muy dispares, como fue el caso del mapa final de este estudio. El uso de un estrato buffer permite reducir el tamaño de los intervalos de confianza potencialmente grandes para los estimados de área de las clases más raras (Arévalo et al., 2020; Olofsson et al., 2020). Esta situación usualmente es provocada por la presencia de omisiones en las clases raras (*i.e.*, deforestación) en la clase más común (*i.e.*, no deforestación). Por lo tanto, el estrato de buffer se crea alrededor de las áreas de las clases raras para potencialmente contener estas omisiones y reducir el área por la que estos errores modifican los intervalos de confianza de estas clases. En este estudio se utilizó un buffer de 12 píxeles alrededor de las clases de cambio (pérdida de bosque maduro y pérdida de bosque secundario / plantaciones), siguiendo las recomendaciones de (Olofsson et al., 2020).

En este diseño, el número de puntos por clase se distribuyeron de acuerdo a la proporción ocupada por cada una de ellas en el mapa, a excepción de las clases más raras, donde se asignó un número mayor al de su proporción de área. El número total de puntos para la verificación se calculó de acuerdo a la ecuación 5, utilizando un valor de $S(\hat{\theta})$ de 0.0035 (Cochran, 1977).

$$n = \frac{(\sum_i^q w_i s_i)^2}{(s(\hat{\theta}))^2 + (\frac{1}{N}) \sum_i^q w_i s_i^2} \approx \left(\frac{\sum_i^q w_i s_i}{s(\hat{\theta})} \right)^2 \quad (15)$$

donde $S(\hat{\theta})$ es el error estándar deseado de la *OA*. En este estudio se adoptó un valor ligeramente mayor que el recomendado por Olofsson et al. (2013), 0.015. W_i es la proporción de área ocupada por cada clase en el mapa, mientras que i corresponde a cada

clase y q , al número de clases. S_i es la desviación estándar de la clase i , la cual se calcula como $S_i = \sqrt{U_i(1 - U_i)}$ donde u_i significa la precisión del usuario esperada a priori. En este caso, se utilizó el valor de la precisión de usuario obtenida en los datos de validación (Cuadro S 6).

El conjunto de datos de verificación constó de 400 puntos, de los cuales 225 correspondieron a la clase de no deforestación, 75 al estrato de buffer de no deforestación, 50 en la clase de pérdida de bosque maduro y 50 en la clase de pérdida de bosque secundario o plantación. La verificación se llevó a cabo mediante interpretación visual de imágenes Sentinel-2 2016 – 2021, compuestos mensuales Sentinel-2 2019 – 2021, imágenes Planet (Planet Team, 2017), Google Earth, Bing y Yandex. Finalmente, se utilizó la herramienta de evaluación de la precisión de OpenForis (FAO, 2017) para calcular los estimados de área sin sesgo, así como sus intervalos de confianza de acuerdo a las ecuaciones 7 y 8. Estos estimados se basaron en la matriz de confusión resultante de la verificación siguiendo el diseño estratificado aleatorio. Además, se calculó el *FI-score* por clase, *OA* y *avgFI-score*.

$$S(\hat{p}_{.k}) = \sqrt{\sum_i W_i^2 \frac{\frac{n_{ik}(1 - \frac{n_{ik}}{n_i})}{n_i - 1}}{n_{i.} - 1}} \quad (16)$$

$$95\% \text{ CI } \hat{A}_k = \left(A \sum_{i=1}^q W_i \frac{n_{ik}}{n_i} \right) \pm 1.96 (A S(\hat{p}_{.k})) \quad (17)$$

Donde $S(\hat{p}_{.k})$ corresponde al error estándar de la proporción de área estimada para la clase k , W_i es la proporción de área ocupada por la clase i , n_{ik} significa el número de observaciones en la celda i,k de la matriz de confusión, mientras que n_i es la suma de la fila de la clase i . Por último, $95\% \text{ CI } \hat{A}_k$, significa el intervalo de confianza del 95% del área estimada para la clase k , A es el área total del mapa, mientras que q representa el número de clases.

Paquetes de R utilizados

Todo el método a excepción de la creación de los mosaicos multitemporales se realizó en un entorno de R 4.1.1 (R Core Team, 2021). En el procesamiento de las imágenes se utilizaron los paquetes raster (Hijmans, 2020), rray (Vaughan, 2020) y reticulate (Ushey et al., 2020). Para manejar y rasterizar la información manualmente clasificada se utilizaron los paquetes sf (Pebesma, 2018) y fasterize (Ross, 2020), mientras que para entrenar la U-Net 3D se utilizaron los paquetes unet (Falbel and Zak, 2020), tfruns (Allaire, 2018) y keras (Allaire and Chollet, 2020) con tensorflow como *backend* (Abadi et al., 2016). Para entrenar el SVM se utilizó el paquete tidymodels (Kuhn and Wickham, 2020) y kernlab (Karatzoglou et al., 2004), mientras que todas las métricas de evaluación se calcularon utilizando yardstick (Kuhn and Vaughan, 2021).

Resultados

Patrones espectrales y temporales

Los patrones espectrales y temporales de las tres clases en los datos de entrenamiento mostraron lo esperado. Por ejemplo, las áreas de pérdida de bosque maduro o bosque secundario / plantaciones mostraron una disminución en la reflectancia del NIR, así como en los coeficientes de retrodispersión VV y VH en la última imagen (t4). Al contrario, en estas dos clases hubo un incremento en las bandas R, G y B en la última imagen (Figura 22). Por su parte, las áreas no deforestadas no mostraron un cambio dramático en la reflectancia o retrodispersión, sino que solamente un decremento pequeño en las bandas R, G y B en la última imagen. Estos patrones probablemente estuvieron relacionados al hecho de que la mayoría de las pérdidas ocurren entre la tercera y cuarta imagen (durante la época de secas) y que la cubierta resultante suele corresponder a suelo desnudo, el cual provoca el aumento en la reflectancia en las bandas del espectro visible.

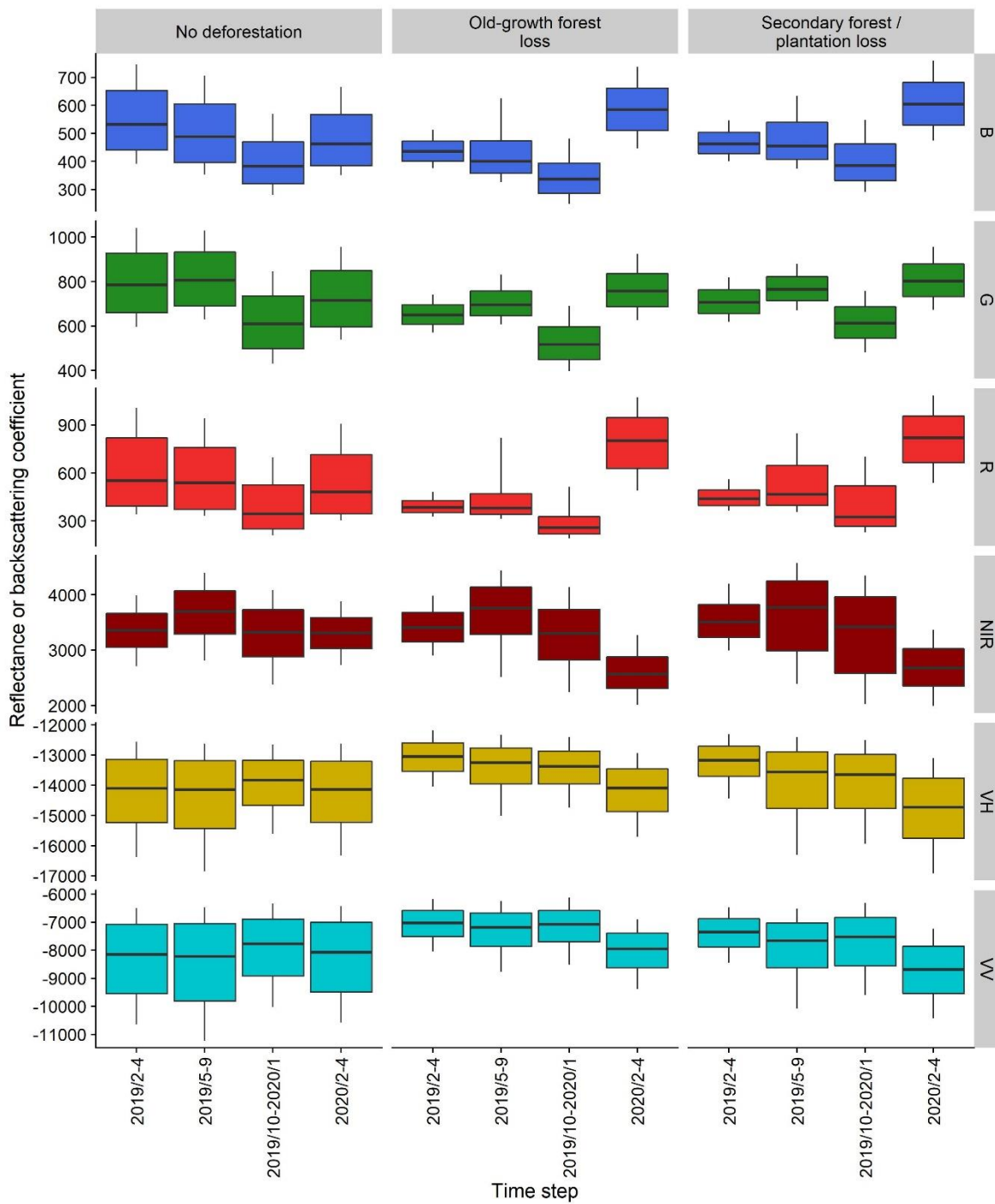


Figura 22. Diagrama de cajas que muestra el perfil espectral y temporal de las tres clases de interés en las seis bandas y cuatro periodos utilizados para entrenar los algoritmos U-Net 3D y SVM. Las fechas en el eje de las x están expresadas como año/mes. Estos datos fueron extraídos del conjunto de entrenamiento utilizado para SVM MS + SAR.

Comparación entradas U-Net 3D

Los resultados mostraron que la arquitectura que alcanzó el mayor *avgF1-score* fue el modelo U-Net 3D MS construido con un *dropout* de 0.1 y cuatro capas ocultas. Esta arquitectura logró una precisión de 0.99 y un *avgF1-score* de 0.73, sobre los datos de validación (Cuadro S 6). Por su parte, el modelo U-Net 3D MS + SAR obtuvo una precisión y *avgF1-score* prácticamente igual al anterior, con una diferencia menor a 0.03 en su *avgF1-score* sobre los datos de validación (Cuadro S 7). La arquitectura U-Net 3D SAR alcanzó la menor precisión y *avgF1-score* de las U-Nets (Cuadro 4). Todas las arquitecturas U-Net presentaron cuatro capas ocultas y un valor *dropout* de 0.1. De acuerdo a los intervalos de confianza del 95 %, tanto la *OA* como el *avgF1-score* de la U-Net 3D MS, SAR y MS + SAR fueron significativamente distintos ($SE < 0.001$).

Al analizar las clases de manera individual, la U-Net 3D MS obtuvo los valores más altos de *F1-score* para todas las clases. Por su parte, la U-Net 3D MS + SAR mostró capacidades similares de clasificación a la U-Net 3D MS, mientras que U-Net 3D SAR tuvo un mayor error de omisión que las demás arquitecturas (Figura 23).

Cuadro 4. Exactitud total y *avgF1-score* para cada clase con las tres U-Net 3D MS, SAR y MS + SAR, así como SVM MS + SAR y MS. Abreviaturas: MS, bandas multiespectrales, SAR, bandas de apertura sintética, SVM, Support Vector Machines. Evaluación sobre los datos de validación sin aumento artificial de los datos. La arquitectura que obtuvo el mayor *avgF1-score* se indica con un *.

Algoritmo	Imágenes	<i>OA</i>	<i>avgF1-score</i>	No deforestación	Per class <i>F1-score</i>	
					Pérdida de bosque maduro	Pérdida de bosque secundario o plantación
U-Net 3D	MS*	0.99	0.73	1	0.73	0.45
	SAR	0.98	0.62	0.99	0.51	0.24
	MS + SAR	0.99	0.70	1	0.68	0.44
SVM	MS + SAR*	0.95	0.52	0.98	0.35	0.24
	MS	0.95	0.51	0.97	0.36	0.18

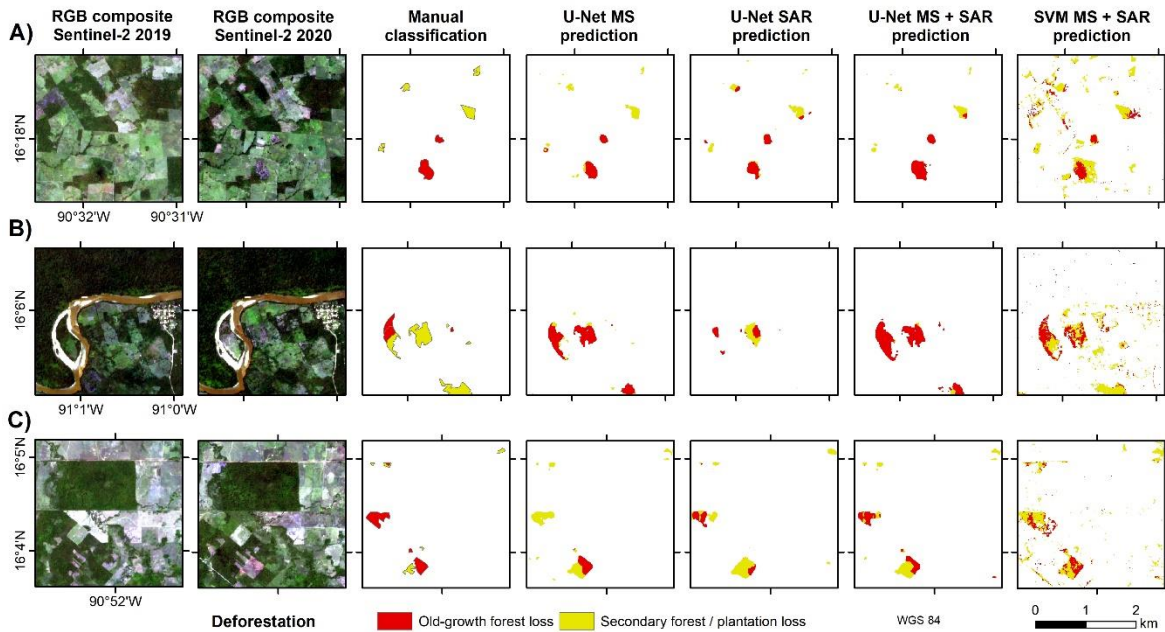


Figura 23. Ejemplos de áreas de validación donde se muestran los compuestos RGB iniciales de Feb – Abr 2019 (t1) y finales Feb – Abr 2020 (t4), así como las predicciones realizadas por las U-Net 3D MS, SAR, MS + SAR, así como SVM MS + SAR.

SVM

El SVM MS + SAR obtuvo los valores más altos de *avgF1-score*, en comparación con SVM MS y su diferencia fue estadísticamente significativa ($SE < 0.001$). Los *F1-scores* por clase fueron también más altos en SVM MS + SAR, a excepción de la clase de pérdida de bosque maduro (Cuadro S 9 y 10). En general, las predicciones de SVM mostraron un alto error de comisión en las clases de pérdida de bosque maduro y bosque secundario / plantaciones. Por el contrario, el error de omisión fue considerablemente más pequeño (Cuadro S 9 y 10).

Comparación U-Net 3D y SVM MS + SAR

Al comparar entre los dos tipos de algoritmos entrenados, se observó que la U-Net 3D MS fue superior a SVM MS + SAR, tanto en la *OA*, como en el *avgF1-score* y los *F1-scores* de cada clase (Cuadro S 6 y Cuadro S 10). La clase que mostró la menor diferencia en *F1-score* entre U-Net 3D MS y SVM MS + SAR fue la clase de no deforestación ($\Delta F1-score = 0.02$), mientras que la diferencia para las clases de deforestación fue mayor ($0.38 \geq \Delta F1-score \geq 0.21$). Además, las predicciones de SVM presentaron un mayor error de comisión

(Cuadro S 6 y Cuadro S 10). De nuevo, las diferencias en su *avgF1-score* fueron estadísticamente significativas ($SE < 0.001$).

Análisis de errores

Tres tipos de errores se detectaron al comparar visualmente los resultados obtenidos por la U-Net MS y los polígonos delineados manualmente: 1) asociados a bordes o caminos (Figura 24A), 2) relacionados con inconsistencias temporales entre las bandas MS y SAR, así como residuos de artefactos en las imágenes (Figura 24A y B), y 3) discriminación entre clases (Figura 24C y D). Estos errores son discutidos a profundidad en la Discusión.

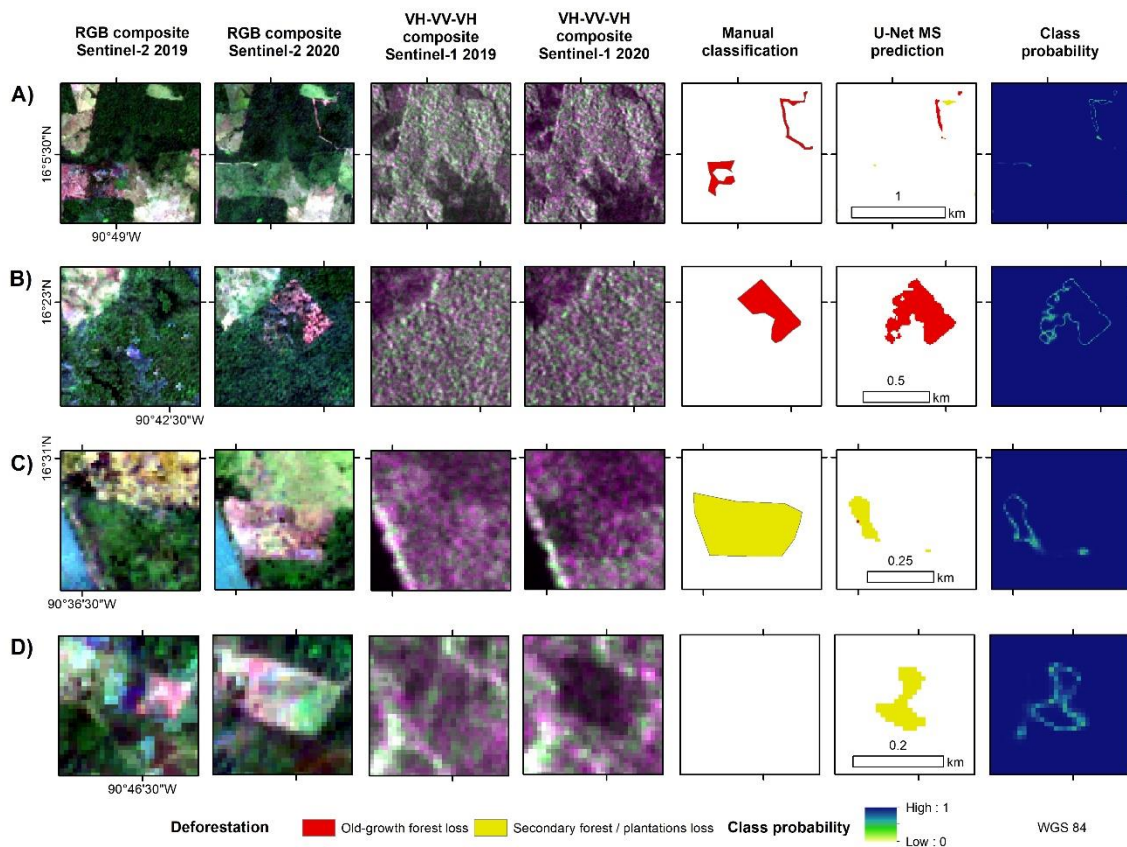


Figura 24. Ejemplos de áreas de validación donde se muestran los compuestos RGB de Feb – Abr 2019 y Feb – Abr 2020, así como las predicciones realizadas por las U-Net 3D MS + SAR, MS y SAR, así como SVM MS + SAR.

Evaluación de la exactitud del mapa y estimados de área

La evaluación de la precisión sobre el área completa de estudio mostró una *OA* de 0.97 y un *avgF1-score* de 0.94 (Cuadro 5; Cuadro S 11). En esta verificación, las clases que obtuvieron más alto *F1-score* fueron no deforestación, seguida de deforestación de bosque maduro y deforestación de bosque secundario o plantaciones (Cuadro 5). Además, el estimado de área corregida para la clase de pérdida de bosque maduro presentó el menor intervalo de confianza, seguido de pérdida de bosque maduro / plantación y no deforestación (Cuadro 5). El estrato de buffer ayudó a reducir la amplitud de los intervalos de confianza, en particular, el de la clase de pérdida de bosque secundario / plantación, debido a tres observaciones de errores de omisión en la clase de no deforestación dentro del estrato buffer. Adicionalmente, no se observaron errores de omisión en la clase de no deforestación ignorando el buffer, por lo que los estimados corregidos de áreas y sus intervalos de confianza se sumaron con el estrato buffer para obtener los estimados para la clase de no deforestación (Cuadro S 11). Finalmente, los estimados de área corregidos para cada clase fueron: no deforestación ($654\,327.22 \pm 2\,172.42$ ha), pérdida de bosque maduro ($3\,195.26 \pm 1\,132.88$ ha) y pérdida de bosque secundario / plantaciones ($4\,234.84 \pm 1\,912.01$ ha).

La mayoría de la clasificación presentó una alta probabilidad de corresponder a su clase asignada, debido a que la mayoría del mapa se clasificó como no deforestación. Por el contrario, los bordes de los polígonos de alguna clase de deforestación mostraron una probabilidad intermedia de corresponder a su clase (Figura 25).

Cuadro 5. Área, proporción del área de estudio y *F1-score* para de la clasificación para el área completa de estudio. Asimismo, se muestra el *F1-score* obtenido en los datos de validación, seguido de los estimados de área corregida y su intervalo del 95 % de confianza por cada clase. Abreviaturas: CI, intervalo de confianza.

U-Net 3D MS						
Clase	Clasificación el área completa de estudio			Datos de validación	Estimados de área	
	Área (ha)	Proporción del área de estudio (%)		F1-Score	Área corregida	IC 95%
		F1-Score	F1-Score	F1-Score		
No deforestación	656440.65	99.19	0.99	1.00	654327.22	2172.42
Deforestación bosque maduro	2923.34	0.44	0.94	0.73	3195.26	1132.88
Deforestación bosque secundario o plantaciones	2393.35	0.36	0.89	0.45	4234.85	1912.01

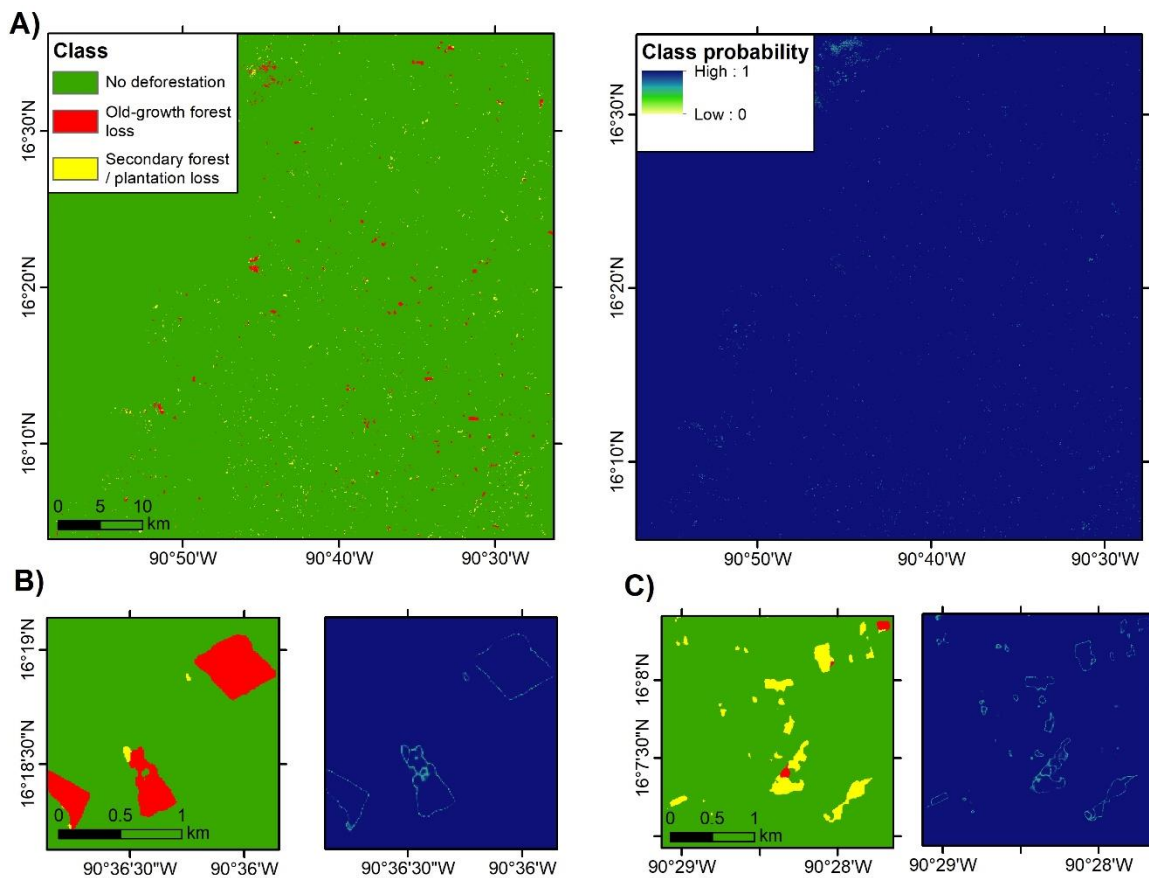


Figura 25. (A) Clasificación de las áreas deforestadas de 2019 a 2020 en el área de estudio utilizando la U-Net 3D MS y su mapa de probabilidad de pertenencia a la clase correspondiente. (B) Ejemplo de un sitio dominado por deforestación de bosque maduro. (C) Ejemplo de un área dominada por deforestación de bosque secundario o plantaciones.

Discusión

Los resultados de este estudio demostraron que el uso de un algoritmo basado en CNN permite obtener mejores resultados para detectar áreas deforestadas en comparación con un algoritmo de ML. Por otro lado, se mostró que el uso de información MS + SAR con una U-Net 3D, permite obtener resultados con alta *OA* y *avgF1-score* para detectar áreas deforestadas y diferenciar entre bosque maduro y bosque secundario o plantaciones ($OA = 0.99$ y $avgF1-score = 0.70$). De hecho, la verificación del mapa del área de estudio completa muestran que se pueden obtener clasificaciones muy precisas con este tipo de algoritmos ($OA = 0.96$ y $avgF1-score = 0.94$).

Implicaciones para el estudio de pérdida de vegetación arbórea

El contar con un método que permita discriminar entre una deforestación de un bosque maduro y un bosque secundario o una plantación abre la posibilidad de conocer si la mayoría de la deforestación se da a partir de bosque maduros o no. Debido a que los bosques maduros suelen presentar mayor biodiversidad, biomasa y provisión de algunos servicios ambientales, la diferenciación de la deforestación de este tipo de bosque de otro tipo de cubierta arbórea permite hacer evaluaciones que consideren estas diferencias y realizar un monitoreo más preciso (Fernández-Montes de Oca et al., 2021; Gutiérrez-Vélez et al., 2011; Kalamandeen et al., 2018; Mercier et al., 2019; Putz and Redford, 2010; Tropek et al., 2014). Por otro lado, la capacidad de distinguir entre la pérdida de distintos tipos de cubierta arbórea permite obtener información sobre el uso del suelo e identificar, por ejemplo, si una pérdida de cubierta arbórea corresponde a una extensión de la frontera agrícola o un área simplemente bajo un ciclo de descanso y uso (Cano Castellanos, 2014; Gallardo-Cruz et al., 2021; Ochoa-Gaona et al., 2007). Además, estos resultados, se pueden utilizar como línea base para estudio futuros para identificar impulsores del cambio de tipos de cubiertas y usos del suelo (Irvin et al., 2020; Masolele et al., 2021). Finalmente, la alta resolución de este tipo de métodos permite detectar áreas posiblemente no detectadas en evaluaciones previas (Hansen et al., 2013; Matosak et al., 2022; Torres et al., 2021), lo cual podría representar un incremento importante en los estimados de área perdida para la región (Reiche et al., 2021).

Comparación de algoritmos: U-Net 3D vs SVM

Al comparar los resultados de U-Net 3D MS y SVM MS + SAR se concluyó que la primera logró una mejor clasificación tanto en todo el sistema de clasificación como por clase. Los resultados mostraron que la arquitectura que alcanzó el mayor *avgF1-score* fue el modelo U-Net 3D MS construido con un *dropout* de 0.1 y cuatro capas ocultas. Esta arquitectura logró una precisión de 0.99 y un *avgF1-score* de 0.73, sobre los datos de validación (Cuadro S 6). Por su parte, el modelo U-Net 3D MS + SAR obtuvo una precisión y *avgF1score* prácticamente igual al anterior, con una diferencia menor a 0.03 en su *avgF1-score* sobre los datos de validación (Cuadro S 8). La arquitectura U-Net 3D SAR alcanzó la menor precisión y *avgF1-score* de las U-Nets (Cuadro S 7). Todas las arquitecturas U-Net

presentaron cuatro capas ocultas y un valor *dropout* de 0.1. De acuerdo a los intervalos de confianza del 95 %, tanto la *OA* como el *avgF1-score* de la U-Net 3D MS, SAR y MS + SAR fueron significativamente distintos ($SE < 0.001$).

Al analizar las clases de manera individual, la U-Net 3D MS obtuvo los valores más altos de *F1-score* para todas las clases. Por su parte, la U-Net 3D MS + SAR mostró capacidades similares de clasificación a la U-Net 3D MS, mientras que U-Net 3D SAR tuvo un mayor error de omisión que las demás arquitecturas (Cuadro 4). Estudios previos han encontrado resultados similares al comparar un método basado en CNN y uno ML para la detección de áreas deforestadas o realizar clasificaciones (Adarme et al., 2020; de Bem et al., 2020; Ji et al., 2018; Matosak et al., 2022). Por otro lado, la UA para todas las clases fue mayor en la U-Net 3D MS (Cuadro S 6), en comparación con los SVM (Cuadro S 9), mientras que lo opuesto ocurrió solamente para la PA de la clase de pérdida de bosque secundario / plantaciones. Esto implica que SVM tiene un error más grande de comisión, pero menor error de omisión en una sola clase, lo cual también puede ser apreciado en la Figura 23. Estos resultados son similares a estudios previos, donde se reporta que los algoritmos basados en CNN permiten obtener clasificaciones con menor ruido y, por lo tanto, suelen requerir un menor post procesamiento, en comparación con un algoritmo ML (Adarme et al., 2020; de Bem et al., 2020).

El mejor desempeño de U-Net 3D MS, en comparación con SVM MS + SAR, probablemente se debió al uso de la información en el dominio espacial, así como a la separación de las dimensiones temporales (lo cual no ocurre en SVM, donde las dos dimensiones se encuentran concatenadas en la dimensión espectral). Debido a que la mayoría de las áreas deforestadas se perdieron entre el t3 y t4, la estructura de la U-Net 3D fue capaz de retener los patrones temporales más importantes para detectar los dos tipos de pérdida (en particular por el uso de los filtros de reescalamiento temporal). Por el contrario, debido a que SVM no tiene un mecanismo de filtro temporal, la presencia de tres imágenes de “cubierta de bosque” pudo haber dificultado la detección de la pérdida entre la información del t3 y t4.

Adicionalmente, la falta de uso de filtros espaciales en las imágenes SAR para reducir el *speckle* pudo haber impactado negativamente las capacidades de detección de la SVM,

ya que no considera la dimensión espacial. Por el contrario, la U-Net 3D fue capaz de sobreasar el ruido causado por el *speckle*, ya que utiliza información en la dimensión espacial (Ortega et al., 2019). El uso de la información espacial probablemente fue crítico para lograr la discriminación entre la pérdida de bosque maduro y pérdida de bosque secundario / plantaciones (Solórzano et al., 2021). Además, la U-Net 3D SAR, que fue la que mostró el peor desempeño, mostró un mayor *avgF1-score* que SVM MS + SAR, lo cual confirma las mejores capacidades de discriminación de la U-Net 3D.

Aunque este estudio no comparó el algoritmo U-Net 3D contra otras versiones de la U-Net con menos dimensiones, trabajos similares han reportado que el utilizar la dimensión temporal por separado ayuda a aumentar, aunque sea ligeramente, la precisión obtenida en distintas tareas de clasificación o detección de cambios (Giannopoulos et al., 2022; Isaienkov et al., 2020; Ji et al., 2018; Masolele et al., 2021). Admitimos que esto es un aspecto que debería ser abordado por estudios futuros; sin embargo, nuestro principal objetivo era mostrar las posibles ventajas de utilizar algoritmos basados en CNN, en comparación con algoritmos ML, así como de las posibles ventajas de utilizar imágenes MS y SAR.

Existen otros algoritmos para detectar la deforestación, además de los basados en CNN, sin embargo, cada uno presenta sus propias ventajas y desventajas. Por ejemplo, los métodos de series de tiempo trabajan sobre todo en las dimensiones temporales y espectrales con distintos métodos que incluyen algoritmos ML y DL (Simoes et al., 2021; Verbesselt et al., 2012; Zhu and Woodcock, 2014). La principal ventaja de estos métodos es que son capaces de trabajar en una colección de imágenes enmascaradas (para evitar nubes y sombras), por lo cual, tienen una mayor resolución temporal para detectar eventos de deforestación que algoritmos como la U-Net 3D (Hostert et al., 2015). Sin embargo, la mayor ventaja de la U-Net 3D es que permite incorporar la información espacial, lo cual podría ser particularmente relevante para diferenciar entre tipos de pérdidas de cubierta arbórea e incrementar la precisión para detectar áreas recientemente deforestadas (Adarme et al., 2020; de Bem et al., 2020; Ji et al., 2018; Matosak et al., 2022; Solórzano et al., 2021). Sin embargo, los artefactos remanentes en los compuestos multitemporales podrían tener un mayor impacto al usar algoritmos como la U-Net 3D, en comparación con un

método basado en series de tiempo. Finalmente, la decisión para utilizar algún método en particular podría no ser clara en primera instancia, ya que cada método tendrá sus propias ventajas y desventajas, aunque algunos de ellos podrían ensamblarse para obtener resultados más precisos.

Selección de la entrada U-Net 3D

En este estudio se seleccionó el modelo U-Net 3D MS como el que alcanzó la máxima precisión y *avgF1-score*. Sin embargo, la diferencia con el segundo mejor modelo (MS + SAR) fue bastante pequeña. Esta selección favoreció al modelo con mejores capacidades para reconocer la deforestación de bosque maduro y bosque secundario o plantaciones. Para esta clase, la U-Net 3D MS presentó menores errores de omisión y comisión, en comparación con la U-Net 3D MS + SAR (Cuadro S 6 y Cuadro S 7).

A diferencia de evaluaciones anteriores, las bandas SAR no ayudan mucho en el esquema de clasificación para diferenciar tipos de cubierta arbórea (Joshi et al., 2016; Reiche et al., 2018; Solórzano et al., 2021). No obstante, al usar SVM sí presentó un mejor desempeño al utilizar MS + SAR que sólo MS; por lo tanto, la ventaja de utilizar MS + SAR parece depender del algoritmo o método utilizado. Basados en nuestros resultados y de otros estudios, la ventaja de combinar MS y SAR parece depender de la aplicación, la zona de estudio, el sistema de clasificación, el método utilizado para delinear las clases de interés, imágenes utilizadas y el método utilizado para analizar la información (Adarme et al., 2020; Joshi et al., 2016; Reiche et al., 2018). Por lo tanto, consideramos que esta ventaja depende del balance entre la información adicional que proveen las bandas SAR, en combinación con las MS, así como la concordancia espacial y temporal entre las bandas MS y SAR. Conjeturamos que ante la presencia de mosaicos MS de baja calidad (*i.e.*, alta presencia de artefactos de nubes y sombras), las bandas SAR pueden ser cruciales para obtener mejores resultados. Además, suponemos que cuando la concordancia temporal y espacial es optimizada, la ventaja de utilizar ambas bandas aumenta (este aspecto se discute en la sección de análisis de errores). Debido a que la mayoría de la deforestación en nuestro sitio de estudio ocurre en áreas planas, suponemos que el desfase espacial entre las bandas MS y SAR fue mínimo (Ye et al., 2021). Esto además es apoyado por el hecho de que SVM obtuvo los mejores resultados utilizando MS + SAR, en lugar de solo MS.

Debido a que Sentinel-1 es un sensor de banda C, sus capacidades para distinguir entre diferentes tipos de cubiertas y usos del suelo son limitadas (Delgado-Aguilar et al., 2017; Reiche et al., 2018), especialmente en comparación con otros sensores SAR de baja frecuencia (*i.e.*, bandas L y P; Deus, 2016; Joshi et al., 2016; Pompeu Pavanelli et al., 2018). Sin embargo, el uso de información SAR banda C ha sido útil para detectar ciertos tipos de cubiertas y usos del suelo, como la palma de aceite, la cual se encuentra en nuestro sitio de estudio (Danylo et al., 2021; Monsalve-Tellez et al., 2022; Solórzano et al., 2021). En este estudio, la falta de o mínima ventaja de utilizar las bandas SAR probablemente estuvo relacionada a que la mayoría de las áreas de pérdida de bosque secundario / plantaciones correspondieron a pérdidas de bosque secundario. A pesar de que este estudio no mostró una clara ventaja de utilizar MS + SAR, asumimos que este patrón probablemente no sea general para diferentes aplicaciones y sitios de estudios. Estudios futuros deberían abordar esta pregunta en distintos contextos y utilizando diferentes métodos.

El *avgF1-score* de este estudio es comparable con el reportado en estudios previos para detectar la deforestación basados en arquitecturas CNN (Adarme et al., 2020; Bragagnolo et al., 2021a, b; Maretto et al., 2021; Matosak et al., 2022; Torres et al., 2021). En todos estos estudios la *OA* fue mayor a 0.95 y el *avgF1-score* mayor a 0.63, lo cual confirma las capacidades de arquitecturas similares a la U-Net para detectar la deforestación.

Análisis de errores

El primer caso de errores se observa en la falta de detección de deforestaciones asociadas a caminos y la presencia de probabilidades bajas de pertenencia a la clase correspondiente en los bordes de las áreas deforestadas detectadas (Figura 25). Este tipo de error se explica por la combinación de tres factores: las dificultades de la U-Net 3D para delinear objetos con pocos píxeles de ancho (Giannopoulos et al., 2022; He et al., 2019; Isaienkov et al., 2020; Stoian et al., 2019) y el efecto de borde. El primero probablemente es causado por el uso de las capas de reescalamiento espacial en la U-Net, mientras que el último está asociado a que las fronteras entre clases suelen contener información mezclada de distintos tipos de cubiertas que puede ser un reto de definir, inclusive por un intérprete visual (Adarme et al., 2020; Bragagnolo et al., 2021a, b; Isaienkov et al., 2020; Torres et al., 2021). Aunque la

resolución de las imágenes utilizadas sea de 10 m, es probable que el área mínima de deforestación detectable con este método sea ligeramente mayor a esta resolución.

El segundo tipo de errores constó de confusiones relacionadas con artefactos en las imágenes, de los cuales se distinguieron dos circunstancias (Figura 25). La primera correspondió a áreas de deforestación no detectadas por la U-Net 3D, donde en el primer tiempo de la evaluación ya se puede observar el inicio de una deforestación; por lo tanto, el cambio “típico” de una deforestación no se apreció en el periodo evaluado. La segunda situación de errores asociados con artefactos constó de una falta de concordancia entre las bandas MS y SAR, donde la deforestación únicamente era visible en el compuesto MS y no en las bandas SAR. Esta diferencia probablemente fue causada por el hecho de que la presencia de nubes en las bandas MS provocaron una distribución de valores de píxel sesgada al final del periodo Feb – Abr, mientras que la de SAR sí estuvo centrada en dicha ventana temporal. A pesar de que intentamos evitar estas inconsistencias al utilizar ventanas temporales cortas para crear los compuestos multitemporales, no fue suficiente para evitarlas.

Aunque raros, la mayoría de los artefactos en las imágenes estuvieron causados por información remanente de nubes o sombras, o por diferencias en la distribución de los datos sin nubes o sombras de píxeles contiguos. Quizás por esta razón no resulta sorprendente que varios estudios hayan utilizado imágenes de una sola fecha en aplicaciones con algoritmos basados en CNN (Adarme et al., 2020; Bragagnolo et al., 2021a; de Bem et al., 2020; Isaienkov et al., 2020; Ji et al., 2018; Pelletier et al., 2019; Solórzano et al., 2021; Wagner et al., 2019). Sin embargo, el utilizar imágenes de una sola fecha en zonas de alta nubosidad como la de este estudio, impide hacer un monitoreo para el área de estudio completa. Por ello, otros estudios también han utilizado compuestos de medianas para obtener evaluaciones sobre un área completa, como fue en este trabajo (Giannopoulos et al., 2022; Maretto et al., 2021; Masolele et al., 2021; Matosak et al., 2022). Dependiendo de las condiciones de nubosidad del área de estudio y de las imágenes utilizadas, puede existir una disyuntiva entre utilizar una sola imagen con menos artefactos o usar un compuesto multitemporal que cubra la totalidad de un área de interés. Esta sería una línea de investigación interesante de abordar en un futuro, sobre todo en el uso de algoritmos

basados en CNN que incorporan la dimensión espacial. Además, sería atractivo abordar la disyuntiva entre la profundidad de la dimensión temporal utilizada en la U-Net 3D y la calidad de los compuestos multitemporales. Por ejemplo, podríamos haber utilizado una ventana temporal más amplia para construir los mosaicos, lo cual habría resultado en un mosaico MS con menos artefactos, pero probablemente con una mayor discordancia temporal con las bandas SAR. Por el contrario, al utilizar una ventana más reducida, podríamos haber obtenido un mosaico MS con más artefactos, pero menor discordancia temporal con las bandas SAR.

El último tipo de confusión está relacionado con la similitud en la señal remota entre las clases del sistema de clasificación, el cual fue el tipo de error más común (Figura 25C y D; Cuadro S 6). Este tipo de errores se pueden apreciar sobre todo en áreas que el intérprete clasificó como deforestación de bosque secundario o plantaciones y que la U-Net 3D MS +SAR clasificó como no deforestación (*i.e.*, error de comisión; Figura 25C) o viceversa (error de comisión; Figura 25D). En el primer caso, probablemente correspondió a bosque secundario o plantaciones perdidas que resultaron en una cubierta herbácea en la última imagen. Esto provocó que estas áreas no fueran identificadas correctamente, ya que no mostró el patrón temporal típico de la mayoría de las pérdidas, es decir, una disminución dramática en las bandas NIR, VV y VH en el t4, debido a su cubierta herbácea. Esto podría ser un indicador de que estos casos raros no fueron lo suficientemente comunes en los datos de entrenamiento como para que la U-Net aprendiera a identificarlos. En el segundo caso, probablemente correspondió a áreas de cubierta herbácea que resultó en una disminución importante de las bandas NIR, VV y VH en el t4. Por lo tanto, es probable que la U-Net confundió esta disminución con el patrón de la pérdida de bosque secundario / plantaciones. Finalmente, estas confusiones seguramente tuvieron su expresión en los patrones captados por las convoluciones temporales y espaciales, pero resulta más difícil de identificar y discutir.

Diferencia de las métricas de datos de validación y verificación de mapa

Al comparar las métricas de evaluación obtenidas en los datos de validación y evaluación de la calidad del mapa se observó un mayor *FI-score* para las dos clases de deforestación y un menor *FI-score* para la clase de no deforestación en esta última (Cuadro S 6 y Cuadro S

11). A primera instancia este resultado parece contradictorio, sin embargo, cada métrica está evaluando un aspecto distinto. La evaluación sobre los datos de prueba permitió examinar el desempeño del algoritmo sobre datos no expuestos al algoritmo, pero que compartían ciertas características con los datos de entrenamiento (Kattenborn et al., 2022) y que contenían una mayor proporción de las clases raras (*i.e.*, pérdida de bosque maduro y pérdida de bosque secundario / plantaciones). Por el contrario, los datos del muestreo estratificado aleatorio tuvieron el propósito de evaluar la calidad del mapa. De hecho, un grupo de los puntos de entrenamiento se encontraron en los datos de evaluación de la calidad del mapa (deforestación de bosque maduro: 29 de 50, deforestación de bosque secundario o plantaciones: 20 de 50, no deforestación 29 de 201 puntos). Esto podría haber inflado la *OA* y *avgF1-score* del mapa; sin embargo, también muestra que una buena proporción de las áreas deforestadas fueron delineadas manualmente y, por lo tanto, se pudo obtener un mapa altamente preciso. Por ende, la evaluación sobre los datos de prueba puede ser considerado un estimado más realista del desempeño de la U-Net 3D sobre datos nuevos; mientras que la evaluación de la calidad del mapa, como una valoración de la calidad del mapa y los estimados de área por clase. Esto quiere decir que en la evaluación del mapa, se encontró una proporción menor de píxeles de la clase con menor *avgF1-score* (*i.e.*, deforestación), en comparación con los datos de prueba.

Finalmente, el mapa de probabilidades de la clasificación podría utilizarse para mejorar este mapa, ya que se puede utilizar como un indicador de áreas que potencialmente podrían contener errores de omisión de áreas deforestadas. Estudios previos con métodos basados en series de tiempo, han sugerido que la suma acumulada de residuos se puede utilizar para informar sobre áreas de posibles errores de omisión. Por lo tanto, se puede utilizar en lugar del estrato buffer para crear un estrato que potencialmente contenga los errores de omisión (Olofsson et al., 2020). Píxeles con una probabilidad intermedia de corresponder a su clase (*e.g.*, 0.5 – 0.7) deberían de contener la mayoría de los errores de omisión.

Conclusiones

Los algoritmos de aprendizaje profundo basados en CNN han permitido mejorar las capacidades para detectar y cuantificar la deforestación en la superficie terrestre. Por un

lado, permiten mejorar la precisión para detectar estos eventos, y por otro, mejorar el detalle en los sistemas de clasificación para diferenciar entre deforestaciones de bosque maduro y otro tipo de cubierta arbórea. Este estudio demostró que la U-Net 3D permite obtener mejores resultados que SVM, mientras que el uso de MS permitió la mejor precisión, seguida cercanamente de MS + SAR. Por ello, este estudio fue incapaz de encontrar una ventaja de utilizar MS + SAR con la U-Net 3D (ΔOA y $\Delta avgF1-score \leq 0.002$), aunque sí se detectó dicha ventaja con SVM. El uso de un estrato de buffer ayudó en la reducción de los intervalos de confianza de los estimados de área, lo cual es crítico para evitar un traslape con el cero, en especial para las dos clases de deforestación. Estudios futuros debería abordar el rol de los compuestos multitemporales sobre el desempeño de algoritmos basados en CNN y evaluar la disyuntiva entre optimizar la calidad y la concordancia temporal entre las bandas MS y SAR. Finalmente, el desarrollo de este tipo de trabajos brindará herramientas para establecer métodos que permitan incrementar las capacidades para detectar la deforestación de bosques tropicales y proveer de estimados más precisos del área perdida para informar decisiones encaminadas a reducir este proceso.

Material suplementario

Cuadro S 6. Matriz de confusión para los datos de prueba, utilizando la U-Net MS. Las filas muestran las clases predichas, mientras que las columnas, las clases manualmente clasificadas.

	Clases			Precisión del usuario	<i>F1-score</i>
	No deforestación	Deforestación bosque maduro	Deforestación bosque secundario o plantaciones		
No deforestación	3156693	2961	6449	1.00	1
Deforestación bosque maduro	3751	16280	1772	0.75	0.73
Deforestación bosque secundario o plantaciones	10858	3283	9217	0.39	0.45
Precisión del productor	0.99	0.72	0.53		
			Precisión total (<i>OA</i>)		0.99
			<i>avgF1-score</i>		0.73

Cuadro S 7. Matriz de confusión para los datos de validación, utilizando la U-Net SAR. Las filas muestran las clases predichas, mientras que las columnas, las clases manualmente clasificadas.

	Clases			Precisión del usuario	<i>F1-score</i>
	No deforestación	Deforestación bosque maduro	Deforestación bosque secundario o plantaciones		
No deforestación	3142166	5869	8129	0.99	0.99
Deforestación bosque maduro	10653	13882	2387	0.52	0.56
Deforestación bosque secundario o plantaciones	18483	2773	6922	0.25	0.30
Precisión del productor	0.99	0.62	0.40		
			Precisión total (OA)		0.98
			<i>avgF1-score</i>		0.62

Cuadro S 8. Matriz de confusión para los datos de validación, utilizando la U-Net MS + SAR. Las filas muestran las clases predichas, mientras que las columnas, las clases manualmente clasificadas.

	Clases			Precisión del usuario	<i>F1-score</i>
	No deforestación	Deforestación bosque maduro	Deforestación bosque secundario o plantaciones		
No deforestación	3158082	3197	7319	1.00	1
Deforestación bosque maduro	7373	16824	2875	0.62	0.68
Deforestación bosque secundario o plantaciones	5847	2503	7244	0.46	0.44
Precisión del productor	1.00	0.75	0.41		
			Precisión total (OA)		0.99
			<i>avgF1-score</i>		0.70

Cuadro S 9. Matriz de confusión para los datos de validación, utilizando SVM MS. Las filas muestran las clases predichas, mientras que las columnas, las clases manualmente clasificadas.

	Clases			Precisión del usuario	F1-score
	No deforestación	Deforestación bosque maduro	Deforestación bosque secundario o plantaciones		
No deforestación	3011964	716	1143	1.00	0.97
Deforestación bosque maduro	48326	16307	2939	0.24	0.36
Deforestación bosque secundario o plantaciones	111012	5501	13356	0.10	0.18
Precisión del productor	0.95	0.72	0.77		
			Precisión total (OA)		0.95
			avgF1-score		0.51

Cuadro S 10. Matriz de confusión para los datos de validación, utilizando SVM MS + SAR. Las filas muestran las clases predichas, mientras que las columnas, las clases manualmente clasificadas.

	Clases			Precisión del usuario	F1-score
	No deforestación	Deforestación bosque maduro	Deforestación bosque secundario o plantaciones		
No deforestación	3022186	907	1793	1.00	0.98
Deforestación bosque maduro	43272	15442	7801	0.23	0.35
Deforestación bosque secundario o plantaciones	97048	4946	17869	0.15	0.24
Precisión del productor	0.96	0.72	0.65		
			Precisión total (OA)		0.98
			avgF1-score		0.52

Cuadro S 11. Matriz de confusión para los datos de evaluación de la calidad del mapa para el área completa de estudio, utilizando la U-Net MS. Las filas muestran las clases predichas, mientras que las columnas, las clases manualmente clasificadas. Todos los errores de comisión y omisión en la clase de no deforestación se detectaron en el estrato buffer.

	Clases			Precisión del usuario	<i>F1-score</i>
	No deforestación	Deforestación bosque maduro	Deforestación bosque secundario o plantaciones		
No deforestación	300	1	3	0.99	0.99
Deforestación bosque maduro	0	45	5	0.9	0.94
Deforestación bosque secundario o plantaciones	3	0	47	0.94	0.89
Precisión del productor	0.99	0.98	0.85		
			Precisión total (<i>OA</i>)		0.97
			<i>avgF1-score</i>		0.94

Capítulo 5. Detección de la degradación forestal en la cuenca del río Ayuquila.

Resumen

La degradación forestal es un proceso que contribuye a las emisiones de carbono a la atmósfera, así como a la disminución de la calidad de algunos servicios ecosistémicos o cambios en las características de un bosque. Aunque este proceso se ha estudiado mediante percepción remota, ha sido más difícil de monitorear en comparación con la deforestación. Con la aparición de nuevas formas de analizar la información proveniente de sensores remotos, así como de imágenes de mayor resolución espacial y temporal, se abren nuevas posibilidades de analizar y monitorear este proceso. El objetivo de este estudio fue evaluar el potencial de dos algoritmos, uno de aprendizaje profundo espacio temporal (U-Net 3D con fusión tardía; U-Net 3D LF) y uno de aprendizaje automatizado (random forest), con imágenes multiespectrales y radar de apertura sintética para detectar la degradación forestal en un bosque tropical seco en el occidente de México entre 2019 y 2022. Los datos para entrenar estos algoritmos constaron de: 1) 63 parcelas de 500 m² donde se obtuvo su biomasa aérea (AGB) mediante un muestreo en campo, 2) cuatro compuestos de Sentinel-2 (Ene-Mar, Abr-Jun, Jul-Sep, Oct-Dic) para los dos años de muestreo (2019 y 2022) y doce compuestos Sentinel-1 (1 por mes) para 2019 y 2022 cada uno. Se usaron el 60 % de los datos como conjunto de entrenamiento, 20 % de validación y 20 % de prueba. Debido al bajo número de datos, tanto los de 2019 como los de 2022 se usaron para entrenar un solo modelo que hizo las predicciones de AGB en 2019 y 2022. Los resultados mostraron que random forest obtuvo el menor error sobre los datos de prueba (RMSE = 14.10 Mg / ha; rRMSE = 64.4 %) en comparación con U-Net 3D LF (RMSE = 18.10 Mg / ha; rRMSE = 82.5 %). De acuerdo a ambos modelos, el AGB del bosque tropical seco disminuyó entre 2019 y 2022 entre un 6.6 % y 7.9 %. A pesar de que los modelos entrenados obtuvieron un error alto, permitieron identificar la tendencia de pérdida de AGB en el bosque de toda la región de estudio. Ante un conjunto tan pequeño de datos, en conjunto con una relación

señal / ruido entre el AGB y las variables predictoras tan baja, random forest es capaz de obtener mejores resultados que la U-Net 3D LF. Estos resultados indican que los métodos probados permiten conocer de manera general si una región ha sufrido degradación forestal (medida como disminución de AGB), aunque debido a la magnitud del error, no conocer su ubicación espacial.

Palabras clave: pérdida de biomasa aérea, bosque tropical seco, cambio de biomasa aérea, Sentinel-1 y Sentinel-2.

Métodos

Esquema general

El método constó de cuatro pasos principales: 1) el preprocesamiento de las imágenes multispectrales (MS) y radar de apertura sintética (SAR), 2) la creación del conjunto de datos de entrenamiento, validación y prueba, 3) el entrenamiento del algoritmo de aprendizaje profundo (U-Net 3D LF) y el de aprendizaje automatizado (random forest; RF), y su evaluación sobre los datos de validación y prueba, y finalmente, 4) la predicción del AGB del área de estudio por los mejores modelos del algoritmo U-Net 3D LF y RF. A continuación, se describen cada uno de estos.

Área de estudio

La zona de estudio correspondió a la parte occidental de la cuenca del río Ayuquila en Jalisco, la cual se encuentra en los siguientes municipios: Autlán de Navarro, Ejutla, El Grullo, El Limón, Tolimán, Tonaya, Tuxcacuesco y Unión de Tula (Figura 26). Esta región abarca aprox. 3,500 km² y presenta una temperatura media anual entre 18 y 24 °C, así como una precipitación media anual de 1000 mm, concentrada en los meses de junio a octubre (Cuevas and Guzmán, 1998; García, 1998). La altitud de la zona de estudio se encuentra entre los 500 y 3800 m s.n.m. y presenta dos tipos de vegetación, el bosque tropical caducifolio (BTC), que es el más abundante en las partes bajas (500 – 2500 m s.n.m.) y el bosque templado, en las partes más altas de la cuenca (2500 – 3800 m s.n.m.; Borrego and Skutsch, 2019; Cuevas and Guzmán, 1998). El BTC se caracteriza por estar dominado por especies arbóreas de la familia Fabaceae, tales como *Lysiloma* sp., *Acacia* sp. y *Vachellia* sp.

La mayoría del bosque de la cuenca del río Ayuquila se encuentra degradado, debido a su uso como fuente de leña y postes, áreas de pastoreo para el ganado y zonas bajo ciclos de descanso para aprovechamiento agrícola (Borrego and Skutsch, 2019; Morales-Barquero et al., 2015; Salinas-Melgoza et al., 2018). En la zona de estudio también existen programas gubernamentales enfocados al sector agrícola y ganadero que favorecen la degradación del bosque (Morales-Barquero et al., 2015). Por otro lado, en este sitio también existen instrumentos encaminados a la conservación de los bosques como la Reserva de la Biosfera Sierra de Manantlán, el programa de pagos por servicios ambientales, así como su vinculación con el programa REDD+ (Morales-Barquero et al., 2014). Por ello, en el sitio también se encuentran presiones que empujan hacia la degradación del bosque, así como fuerzas que promueven la conservación del bosque.

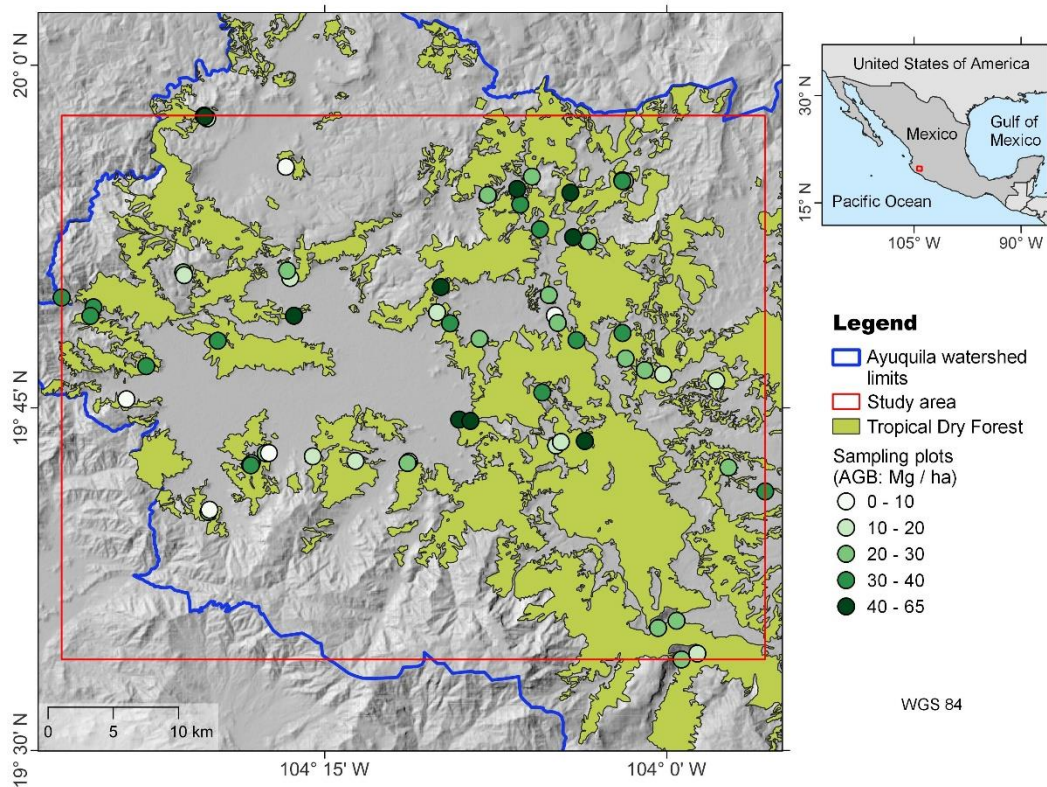


Figura 26. Localización de la cuenca del río Ayuquila y de sitios de muestreo del bosque tropical caducifolio (BTC).

Generación de datos de entrenamiento

Se muestrearon un total de 63 parcelas circulares de 500 m² cada una. De estas, 41 parcelas se muestrearon en mayo – junio del 2019 y 22 en 2022. De las parcelas del 2019, 23 sitios correspondieron a bosque conservado, mientras que los 18 restantes a bosque degradado. Estos últimos se clasificaron como bosque degradado, a partir de una interpretación visual de imágenes Landsat y Google Earth de 2016 – 2018. Los sitios de bosque degradado se identificaron como aquellos que presentaran algún cambio notorio en su textura en el periodo analizado (Jiménez-Rodríguez et al., 2022). Posteriormente, se consultaron los sitios de bosque degradado con la Junta Intermunicipal del Río Ayuquila (JIRA) para asegurar que los sitios elegidos presentaban actividad antropogénica. Por su parte, los datos de 2022 trataron de abarcaron áreas en el espectro completo de biomasa, desde valores muy bajos (aprox. 4 Mg / ha) hasta muy altos (aprox. 65 Mg / ha), sin hacer especial énfasis en la detección de algún cambio entre 2019 y 2022.

En el área central de las parcelas (*i.e.*, 30 m²) se midieron todos los árboles con diámetro a la altura del pecho (DAP) > 2.5 cm, mientras que en el resto del círculo se midieron todos los árboles con DAP > 5 cm. Para cada árbol registrado se midió su DAP y su altura. Posteriormente, la biomasa aérea de cada individuo se calculó utilizando ecuaciones alométricas creadas para la Reserva de la Biosfera Chamela – Cuixmala, Jalisco (ecuación 18; Martínez-Yrizar et al., 1992), que se encuentra a aprox. 130 km del sitio de estudio y utilizando la densidad de madera por especie de la base de datos de (Réjou-Méchain et al., 2017; Zanne et al., 2009).

$$AGB = 10^{(-0.759 + 0.9011 \log_{10}(BA) + 0.5715 \log_{10}(WD) + 0.5654 \log_{10}(H))} \quad (18)$$

donde *AGB* corresponde a la biomasa aérea, *BA* al área basal, *H* a la altura y *WD* a la densidad de madera. El área basal fue calculada como el radio del DAP de cada tallo, al cuadrado, multiplicado por la constante pi (*i.e.*, 3.14159). Debido a que los datos del 2019 no contaron con información taxonómica, la densidad de madera de todos los árboles se calculó como el promedio de la comunidad del muestreo del 2022, a excepción de los cactus, donde se utilizó la densidad de madera de la familia Cactaceae. Después, se calculó

la biomasa aérea como la suma de la biomasa de los individuos en el sitio y se extrapoló a una hectárea (ha), tomando en cuenta las distintas extensiones de muestreo (dos círculos centrales).

Datos de percepción remota

Se utilizaron imágenes Sentinel-1 y Sentinel-2 como insumos para modelar y predecir la AGB del bosque. Sentinel-1 es un sensor de radar de apertura sintética (SAR) en la banda C con dos tipos de transmisión y recepción, vertical – vertical (VV) y vertical – horizontal (VH; Malenovský et al., 2012). Se utilizaron imágenes del coeficiente de retrodispersión (σ^2) en escala lineal, registradas en el modo *Interferometric Wide Swath* con una órbita ascendente (órbita relativa 49) de la colección *Ground Range Detected* (GRD). Para maximizar la concordancia temporal de las imágenes con los datos registrados en campo, se crearon compuestos mensuales para 2019 y 2022. Debido a que estudios previos han recomendado no utilizar filtros espaciales al utilizar imágenes SAR con algoritmos basados en CNN, no se aplicó ningún tipo de filtro espacial para reducir el *speckle* de las imágenes (Ortega et al., 2019). Esto debido a que los algoritmos basados en CNN, al incluir la información espacial de los vecindarios de píxeles a distintas escalas, deberían de ser capaces de identificar y compensar el ruido del *speckle* y relacionar ciertos patrones espaciales con la variable de interés (*i.e.*, biomasa).

Se utilizaron las imágenes Sentinel-2 de la colección de reflectancia debajo de la atmósfera como imágenes multiespectrales (MS). A pesar de que estas imágenes cuentan con trece bandas, sólo se utilizaron las de mayor resolución espacial (*i.e.*, R, G, B y NIR con resolución de 10 m); debido que una mayor resolución espacial se ha asociado con un mayor potencial de modelación usando arquitecturas CNN (Du et al., 2020; Giang et al., 2020; Matosak et al., 2022; Torres et al., 2021; Zhang et al., 2018). Aunque estudios previos sugieren utilizar índices de vegetación o variables de textura como variables para predecir la biomasa de un bosque (Eckert, 2012; Gallardo-Cruz et al., 2012; Ozkan and Demirel, 2021; Singh et al., 2015; Solórzano et al., 2017), al utilizar CNN, esto no es recomendado, ya que en teoría las CNN son capaces de identificar patrones en más de una sola banda (similar a lo que hace un índice) y resumir relaciones espaciales entre píxeles contiguos (similar a lo que hacen las variables de textura; Kattenborn et al., 2021). Al igual

que las imágenes Sentinel-1 se construyeron mosaicos en los años donde se registraron los datos de campo, *i.e.*, 2019 y 2022. Debido a que las imágenes MS son sensibles a nubes y sombras, se hicieron tres compuestos cuatrimestrales para obtener compuestos de alta calidad, *i.e.*, Ene-Abr, May-Ago, Sep-Dic. Estudios previos recomiendan utilizar compuestos más cercanos a la fecha de captura de la información en campo (Réjou-Méchain et al., 2019), sin embargo, en este caso se trató de aprovechar la información en distintos estados fenológicos del bosque para modelar su biomasa. Esto debido a que algunos trabajos han sugerido que contar con información de la época de lluvias y secas permite mejorar los resultados (Jiménez-Rodríguez et al., 2022). En este proceso se enmascararon todas las nubes y sombras utilizando el método s2cloudless (Skakun et al., 2022). En este método, se optimizaron los siguientes umbrales: cobertura de nubes de las imágenes, probabilidad de pertenencia a nube del píxel, reflectancia en el infrarrojo cercano de objetos oscuros. Además, se ajustaron los parámetros de distancia de proyección de sombras de las nubes y tamaño del buffer alrededor de las nubes. Finalmente, con las imágenes enmascaradas se hizo un compuesto de mediana.

Tras obtener los compuestos MS y SAR, se obtuvieron 12 imágenes SAR por año, así como tres imágenes MS por año (*i.e.*, 2019 y 2022). Estas imágenes se estandarizaron mediante la resta de la media y la división por su desviación estándar, antes de ser sometidas como entradas de la U-Net LF y RF. Este proceso tuvo el propósito de facilitar la optimización de los algoritmos utilizados, pero también ayudó a disminuir la variación de reflectancia y retrodispersión entre años. Esto se esperaba que redujera el ruido entre la señal remota y la biomasa del bosque. Por ejemplo, el mismo sitio puede contener una biomasa prácticamente igual en los dos años evaluados (*i.e.*, 2019 y 2022), mientras que las condiciones particulares de precipitación o temperatura de cada año, pueden provocar que la señal de reflectancia o retrodispersión sean distintas. Por lo tanto, al estandarizar las imágenes, se esperaba que el valor de un mismo sitio en los dos años fuera más similar que al comparar sus valores de reflectancia y retrodispersión sin estandarizar.

Datos de entrenamiento, verificación y prueba

Para crear el conjunto de datos de entrenamiento, verificación y prueba se organizaron los datos de campo con sus correspondientes imágenes MS y SAR, *i.e.*, los datos de campo del

2019 con las imágenes del 2019 y viceversa. De esta manera, se maximizó la correspondencia temporal entre los datos de campo y las imágenes. Debido a que los métodos basados en CNN permiten detectar patrones en la dimensión espacial, se crearon áreas de la imagen que fueran lo más cercano a la extensión de las parcelas de muestreo, *i.e.*, 3 x 3 píxeles. Además, la etiqueta objetivo de cada imagen correspondió a vecindarios de 3 x 3 píxeles, donde todos los píxeles tomaron el valor de la AGB calculada para la parcela. Los datos de entrenamiento correspondieron al 60 % de los datos (38 parcelas), mientras que los datos de verificación y prueba correspondieron a un 20% y 19 %, respectivamente (13 y 12 parcelas). Cabe mencionar que cada conjunto de datos mostró una distribución de valores de biomasa similar (ver Figura 29). Debido a que el preprocesamiento de los datos siguió distintos pasos para ser utilizados por la U-Net 3D LF y random forest, estos se detallan junto con la descripción del algoritmo. En U-Net 3D LF se utilizaron cuadrados de 3 x 3 píxeles como entradas (*i.e.*, vecindarios de píxeles), mientras que para random forest se utilizó por separado la información de los píxeles de estos cuadrados (*i.e.*, píxeles).

Algoritmos

U-Net 3D LF

El algoritmo utilizado para realizar la tarea de regresión (*i.e.*, predecir valores de AGB a partir de las imágenes MS y SAR) fue una U-Net 3D con una fusión tardía (U-Net 3D LF; Figura 27). Esta arquitectura se eligió debido a las siguientes razones: 1) permite separar las convoluciones en la dimensión espacial y temporal (*i.e.*, 3D), 2) acepta entradas con un distinto número de observaciones en la dimensión temporal (*i.e.*, MS: 3 y SAR: 12), y 3) incluye una fusión tardía para realizar la predicción basada tanto en la información MS como SAR. Esta arquitectura cuenta con un par de U-Net 3D cuyas salidas se fusionan antes de pasar a la función de activación final y realizar la predicción final. En comparación con la U-Net 3D tradicional, a la U-Net 3D LF se le tuvieron que hacer las siguientes adecuaciones: cambiar las capas de reescalamiento de la información de tomar el valor máximo de una ventana (*max pool*) a tomar el valor promedio (*average pool*), utilizar la función lineal como función de activación para realizar la predicción, en lugar de la softmax, y utilizar el error cuadrático medio (MSE) como métrica de error, en lugar de

alguna métrica que incorporara la entropía cruzada categórica. Estas modificaciones se realizaron para adaptar la arquitectura U-Net a realizar una tarea de regresión y poder hacer la predicción basada en convoluciones con distintas dimensiones temporales (*i.e.*, SAR y MS).

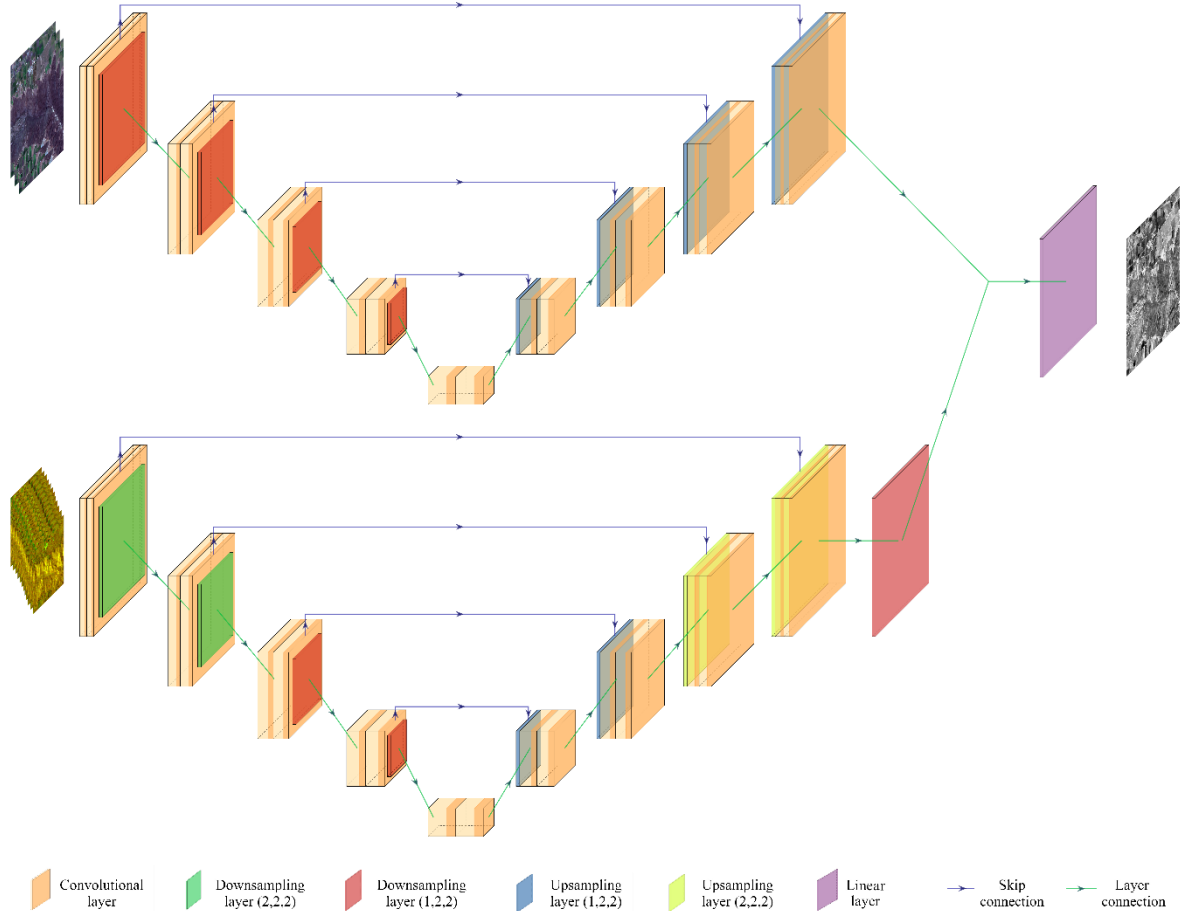


Figura 27. Esquema de la arquitectura U-Net 3D LF.

Los hiperparámetros del número de capas ocultas (3 – 5 capas) y la probabilidad de *dropout* (0.1, 0.3, 0.5) de la U-Net LF se exploraron para detectar la combinación que permitiera obtener la mayor exactitud total y *avgF1-score* sobre los datos de validación. Debido a que estos parámetros determinan la arquitectura a la U-Net, se recomienda probar diferentes combinaciones para identificar la mejor, ya que suelen seguir un comportamiento no lineal. El número de capas ocultas es directamente proporcional a la capacidad de identificar patrones complejos del algoritmo, mientras que el *dropout* es una técnica de

regularización que ayuda a evitar el sobreajuste (Chollet y Allaire 2018; Srivastava et al., 2014). Otros hiperparámetros utilizados fueron una tasa de aprendizaje de 0.0001, 64 filtros en la primera capa y un máximo número de épocas de 50 con una interrupción anticipada (*early stopping*), si el MSE no disminuye por lo menos 4 unidades en 10 épocas.

Para aumentar el tamaño de los datos de entrenamiento y verificación, los datos se aumentaron de manera artificial creando replicas en espejo, tanto en una dirección vertical como horizontal. Tras realizar este procedimiento, los datos de entrenamiento constaron de 114 imágenes, mientras que los de verificación, de 39 imágenes. A los datos de prueba no se les aplicó un aumento artificial de los datos, por lo cual, sólo constaron de 12 imágenes. De esta manera, los datos de entrada MS de la U-Net 3D LF constaron de arreglos con 3 entradas en la dimensión temporal, 3 entradas en las dos dimensiones espaciales y 4 entradas en la dimensión espectral. Por su parte, las SAR constaron de arreglos de 12 entradas en la dimensión temporal, 3 entradas en las dos dimensiones espaciales y 2 en la espectral (Figura 27).

Random forest

El algoritmo random forest es uno frecuentemente utilizados en tareas de aprendizaje automatizado utilizando imágenes de percepción remota. Al igual que muchos otros algoritmos dentro de este grupo realizan clasificaciones o regresiones a nivel de píxel. Aunque en el capítulo 3 se utilizó también un random forest, la principal diferencia con el utilizado en este capítulo fue que este se utilizó para hacer una regresión en lugar de una clasificación. A pesar de ello, el principio del algoritmo es el mismo, es decir, crear árboles de decisión aleatorios (tomando diferentes variables predictivas) y al final, hacer un ensamble para definir el valor predicho como el que recibió más votos por cada uno de los árboles individuales. En este trabajo se construyó el algoritmo random forest con 500 árboles aleatorios y tomando 12 variables para construir cada árbol (*i.e.*, $p/3$, donde $p = 36$, 12 MS y 24 SAR).

Los datos de entrenamiento, verificación y prueba se utilizaron tal cual en el algoritmo random forest, sin un aumento artificial de los datos. Debido a que el algoritmo random forest trabaja con información por píxel, los datos de entrenamiento, verificación y prueba, se reacomodaron para poder ser ingresados al random forest. En este proceso, se

concatenaron los datos de la dimensión temporal sobre la dimensión espectral, de tal manera, que los datos MS contaron con 12 bandas (4 bandas en 3 tiempos) y las SAR 24 bandas (2 bandas en 12 tiempos). Además, los datos de cada píxel se tomaron por separado, en lugar de cuadrados de 3 x 3 como en el caso de la U-Net 3D LF. De esta manera, los datos de entrenamiento constaron de 342 observaciones (3 x 3 x 38), los datos de verificación de 117 observaciones (3 x 3 x 13) y los datos de prueba de 108 observaciones (3 x 3 x 12).

Por último, debido a que RF fue el algoritmo que obtuvo el menor error sobre los datos de prueba, se calcularon los valores de importancia relativa de las variables predictivas incorporadas en el modelo. Esto con el fin de mostrar cuáles variables fueron las más relevantes para predecir la AGB a partir de las variables remotas incorporadas.

Métricas de evaluación

Para evaluar el ajuste de los algoritmos entrenados se utilizó la raíz cuadrada del error cuadrático medio (RMSE; ecuación 19). Además, como otra métrica informativa se calculó la raíz cuadrada del error cuadrático medio expresada en términos relativos (rRMSE; ecuación 20).

$$RMSE = \sqrt{\frac{1}{N} \sum_i^N (\widehat{AGB}_i - AGB_i)^2} \quad (19)$$

$$rRMSE = \frac{RMSE}{\frac{1}{N} \sum_i^N AGB_i} \quad (20)$$

Donde \widehat{AGB} corresponde al AGB predicha por el modelo y AGB al AGB observada (obtenida mediante el muestreo en campo). Esta métrica se evaluó sobre los datos de entrenamiento, validación y prueba. Finalmente, para obtener los intervalos de confianza (IC), estos se calcularon como $1.96 * RMSE$.

Mapa de biomasa

Con las arquitecturas que obtuvieron el menor RMSE sobre los datos de prueba se realizaron predicciones sobre las imágenes completas de 2019 y 2022. Debido a que los algoritmos fueron entrenados con datos de 2019 y 2022, el mismo algoritmo se utilizó para realizar las predicciones sobre las imágenes de 2019 y 2022. Estos mapas se limitaron a realizar predicciones sobre las áreas cubiertas de BTC, debido a que este tipo de bosque era

el contenido en los datos. Por ello, se utilizó una máscara de cubierta de BTC, realizada mediante una clasificación supervisada y una posterior edición mediante interpretación visual. Una vez que los mapas de AGB de 2019 y 2022 fueron obtenidos, se calculó la biomasa para cada uno de los mapas para detectar posibles cambios en la AGB del BTC en toda el área de estudio. Debido al gran error de estas predicciones, no se pudieron detectar áreas de disminución de AGB con certeza; por lo cual, también se realizó una comparación de AGB sobre toda el área de estudio. En este escenario esperábamos que las predicciones no permitieran detectar la ubicación de las áreas degradadas entre 2019 y 2022 (*i.e.*, disminución de AGB), pero sí detectar la tendencia general de cambio de la biomasa entre las dos fechas, ya que el error cometido por el modelo debería sobrestimar y subestimar valores en similares proporciones. Finalmente, debido al gran error de las predicciones, se realizaron dichas predicciones utilizando ambos algoritmos para ver la consistencia en sus evaluaciones. Además, para el algoritmo que permitió las estimaciones con menor error, se analizaron visualmente circunstancias de aciertos y errores para ejemplificar su desempeño.

Pruebas aleatorias

Una vez observado que el algoritmo RF fue el que permitió obtener los resultados con menor error, únicamente sobre este algoritmo se realizaron pruebas posteriores para evaluar el posible efecto de la división de datos sobre el desempeño de este algoritmo. En este proceso se hicieron aleatorizaciones en la división de los datos de entrenamiento y validación de acuerdo a tres técnicas distintas: 1) *10-fold cross validation* (10-fold), 2) *bootstrap* con 25 repeticiones y 3) *leave-one-out cross validation* (LOO). Además, debido a que RF es un algoritmo que realiza una regresión por píxel, se probó dividiendo los datos respetando la pertenencia de cada píxel a una parcela o no (*i.e.*, totalmente aleatorio). Esto quiere decir que en la división de datos por parcela, los nueve píxeles de una parcela tenían que formar parte del mismo conjunto de datos (*i.e.*, entrenamiento o verificación), mientras que en el totalmente aleatorio, cada píxel podía formar parte de un conjunto de datos distinto, sin importar su pertenencia a una parcela (*i.e.*, píxeles de una misma parcela podían formar parte del conjunto de datos de entrenamiento y otros del de verificación). En este proceso, se utilizaron todos los píxeles que se sobreponían con la extensión de la parcela en campo. Posteriormente, para cada método se calculó su RMSE promedio y su

error estándar. Estos análisis ayudaron a esclarecer tres posibilidades: 1) si los resultados obtenidos dependieron en fuerte manera de la división de datos, 2) si el haber elegido otro método de evaluación como una validación cruzada habría obtenido un valor artificial más alto de RMSE, y 3) simular si el haber obtenido un conjunto más grande de datos podría haber mejorado los resultados.

Por último, utilizando la división de datos del procedimiento de *10-fold cross validation* se entrenaron diez modelos RF y se predijo con cada uno de ellos el AGB del área de estudio. Posteriormente con estas diez predicciones se calculó el promedio y la desviación estándar por pixel, así como su coeficiente de variación. Esta última métrica se utilizó para determinar cuáles píxeles presentaron mayor variación y, por lo tanto, consistían de píxeles con menor confiabilidad en su estimado de AGB, y viceversa.

Resultados

Biomasa del bosque

La biomasa del bosque tomó valores desde 2.23 a 65.06 Mg / ha con un promedio de 27.45 Mg / ha (2019: 26.83 Mg / ha; 2022: 28.61 Mg / ha) y con una desviación estándar de 14.32 Mg / ha (2019: 14.33 Mg / ha; 2022: 14.55 Mg / ha). Estos estadísticos descriptivos mostraron que las dos tomas de datos tuvieron una distribución similar de valores (Figura 28).

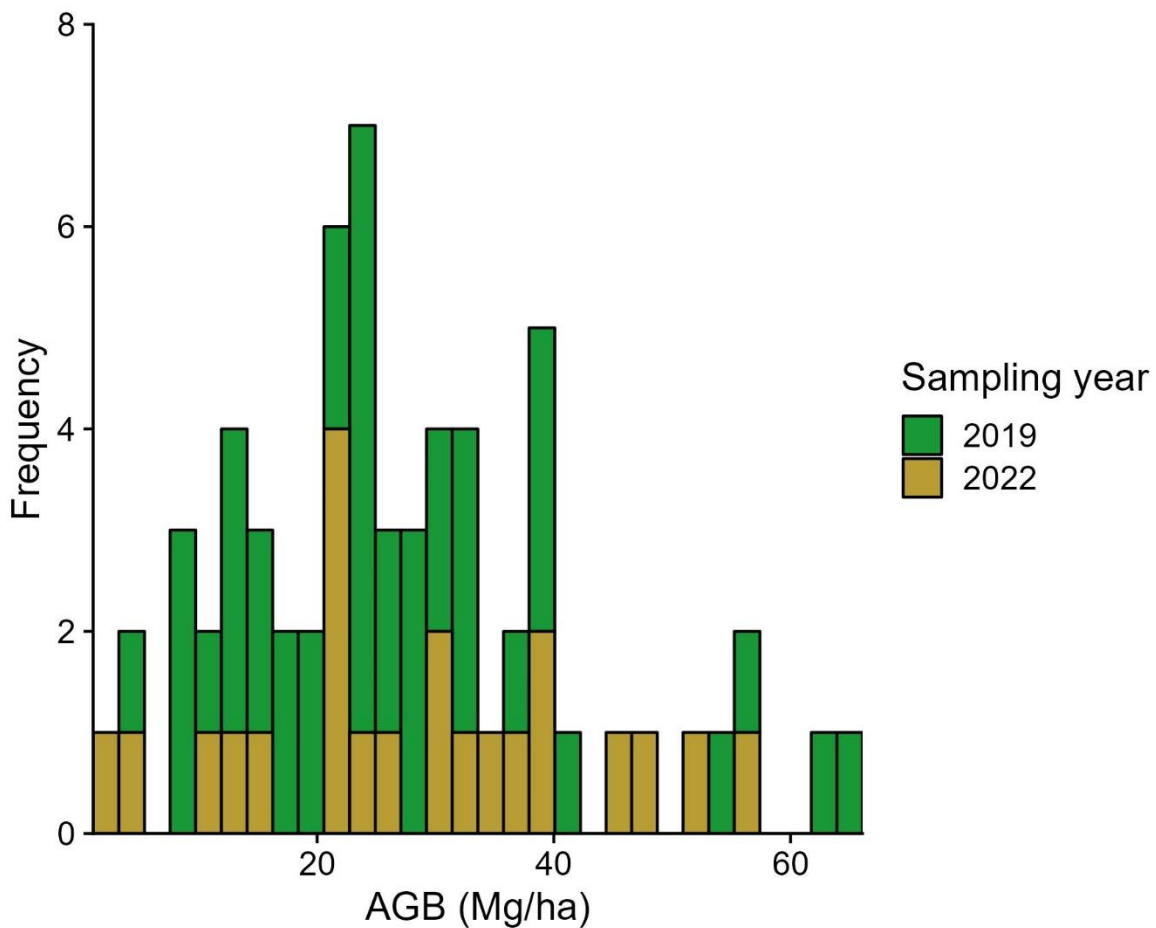


Figura 28. Histograma de biomasa aérea (AGB) de las parcelas muestreadas en este trabajo, divididas por año de registro.

Patrones espectrales y temporales de los datos

Al analizar los patrones espectrales y temporales de los datos de entrenamiento, se observó que estos no presentaron un claro patrón con respecto a los valores de AGB. Sin embargo, sí se detectaron algunos de los patrones esperados como una disminución de la reflectancia en las bandas del espectro visible (R, G, B), así como un aumento de la reflectancia en la banda NIR en el t_3 o t_2 (t_3 : después de la época de lluvias y t_2 : lluvias), en comparación con t_1 (t_1 : temporada seca; Figura 29). Por su parte, al analizar las imágenes SAR no se observó algún patrón esperado del coeficiente de retrodispersión en relación con AGB o con el tiempo (Figura 30). Por otro lado, estos patrones permitieron esclarecer que los patrones de

reflectancia y retrodispersión fueron similares en los dos años de registro de datos (2019 y 2022).

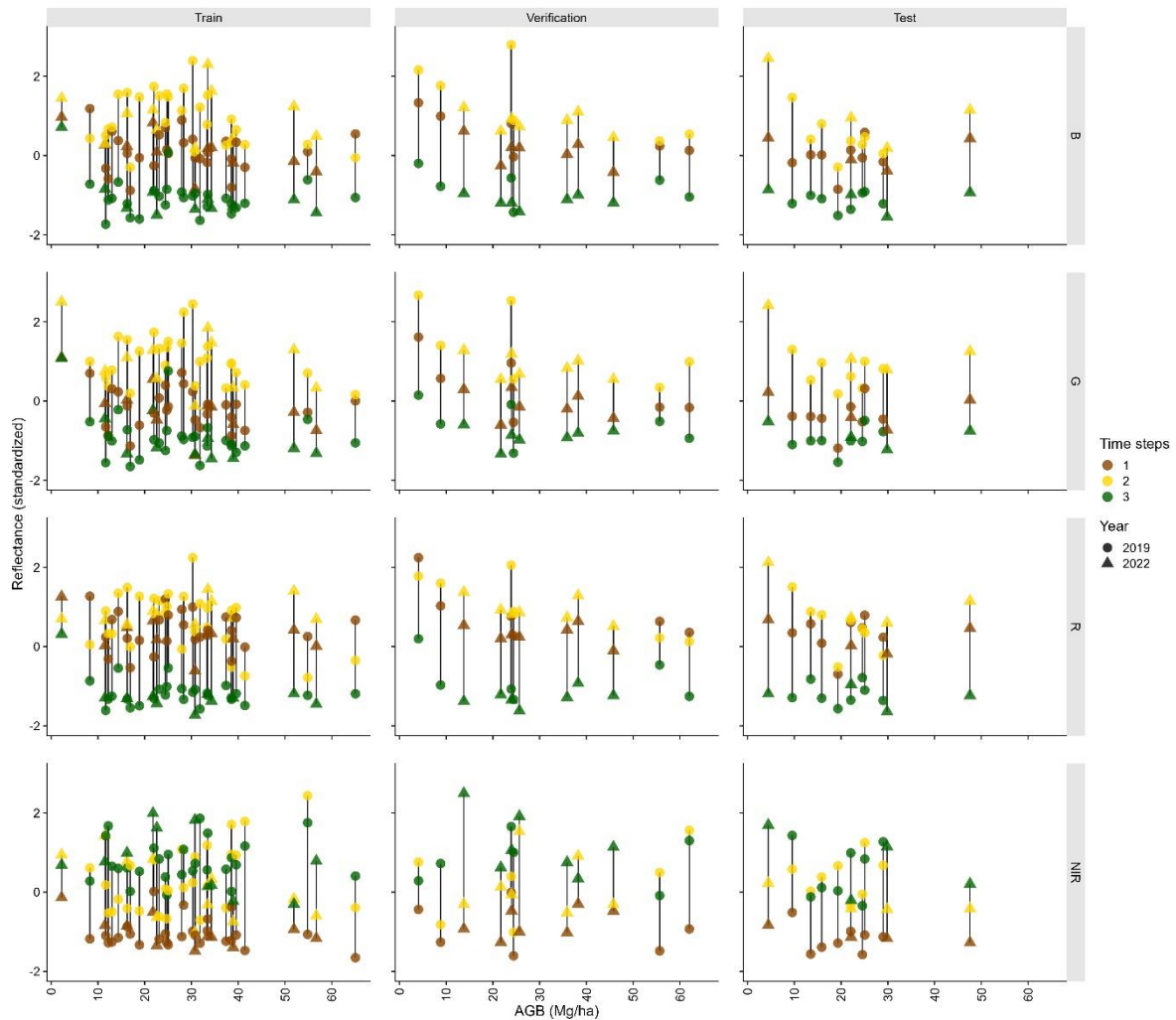


Figura 29. Promedio de la reflectancia de los nueve píxeles por sitio de muestreo en las tres imágenes MS utilizadas para entrenar la U-Net 3D LF (t1: Ene-Abr, t2: May-Ago, t3: Sep-Dic). Además, el símbolo muestra el año del registro de los datos. El gráfico se encuentra dividido por los tres conjuntos de datos utilizados: entrenamiento, verificación y prueba.

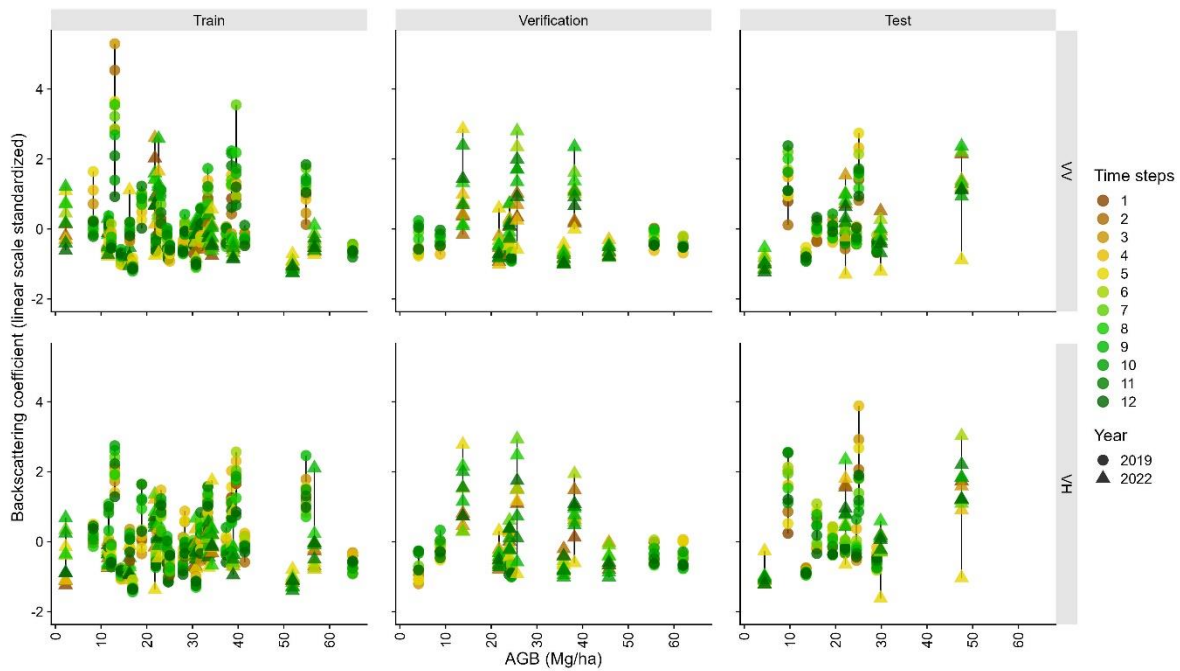


Figura 30. Promedio del coeficiente de retrodispersión de los nueve píxeles por sitio de muestreo en las doce imágenes mensuales SAR utilizadas para entrenar la U-Net 3D LF. Además, el símbolo muestra el año del registro de los datos. Los datos están divididos por los tres conjuntos de datos utilizados: entrenamiento, verificación y prueba.

Evaluación de los algoritmos

U-Net 3D LF

La arquitectura que alcanzó el menor RMSE sobre los datos de validación, *i.e.*, 15.50 Mg / ha, presentó los siguientes hiperparámetros: 4 capas ocultas, *dropout* = 0.3, mientras que esta obtuvo un RMSE de 6.91 Mg / ha en los datos de entrenamiento. La evaluación de las predicciones de esta arquitectura sobre los datos de prueba alcanzó un RMSE = 18.10 Mg / ha, el cual fue un poco mayor al alcanzado sobre los datos de validación (Cuadro 6). Al graficar los datos de AGB observados vs los predichos por la U-Net 3D LF, se observó que la U-Net 3D LF fue capaz de obtener un buen ajuste sobre los datos de entrenamiento; sin embargo, en los datos de validación y de prueba, hizo predicciones alrededor del valor promedio de AGB (Figura 31).

Cuadro 6. Algoritmos que alcanzaron la menor raíz cuadrada del error cuadrático medio (RMSE) sobre los datos de prueba. Además, se muestra el RMSE sobre el conjunto de datos de entrenamiento, validación y prueba sin aumento artificial de los datos.

Algoritmo	RMSE (Mg / ha)			rRMSE (%)		
	Prueba	Entrenamiento	Validación	Prueba	Entrenamiento	Validación
U-Net 3D LF	18.10	6.91	15.50	82.51	24.25	52.43
MS + SAR	14.10	2.38	16.10	64.40	8.35	54.44

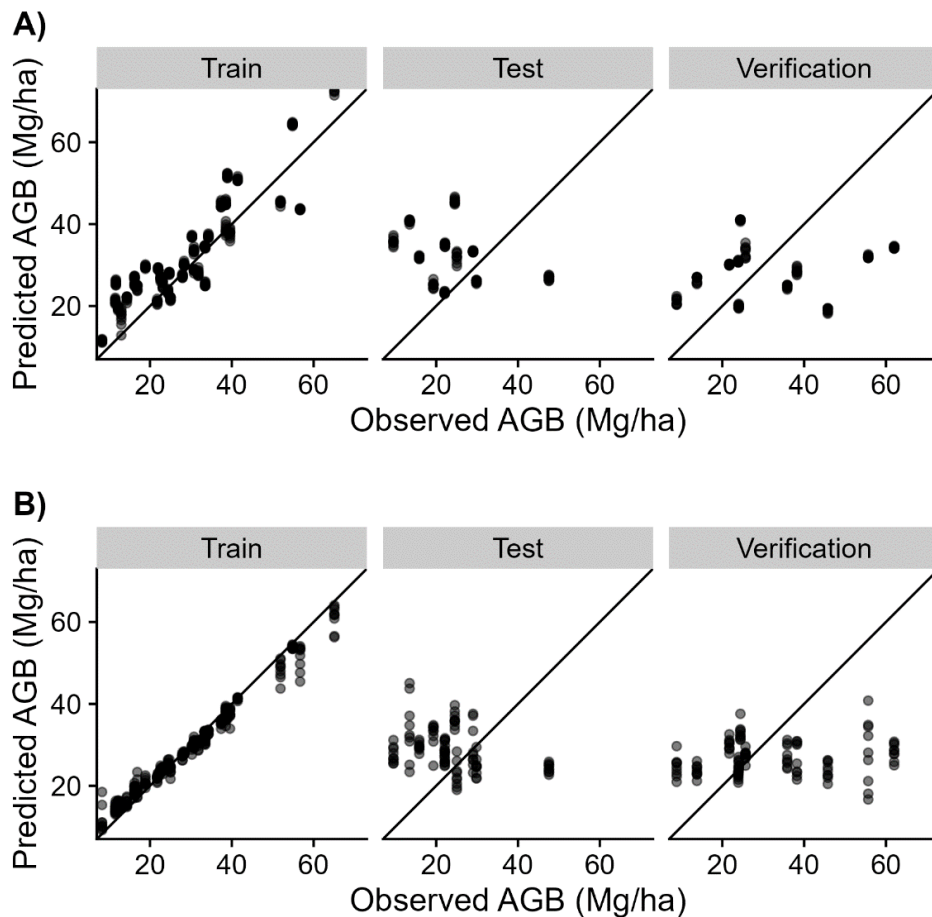


Figura 31. Comparación entre los valores observados y predichos por A) U-Net 3D LF y B) RF en los datos de entrenamiento, prueba y verificación. La línea recta muestra un ajuste perfecto entre los valores observados y predichos.

Random forest

Por su parte, el random forest alcanzó un RMSE = 16.10 Mg / ha sobre los datos de validación, mientras que fue un RMSE = 2.38 sobre los datos de entrenamiento. La evaluación de las predicciones de este algoritmo sobre los datos de prueba alcanzó un RMSE = 14.10 Mg / ha (Cuadro 6). Al graficar los datos observados con los predichos, se observó el mismo patrón que con la U-Net 3D LF, *i.e.*, un buen ajuste sobre los datos de entrenamiento, pero mucho menor sobre los datos de verificación y prueba (Figura 31).

Los resultados del cálculo de la importancia relativa de las variables predictivas del RF muestran que las variables más importantes fueron aquellas provenientes de las imágenes MS, sobre todo las de los tiempos 2 y 3 (Figura 32). Esto quiere decir que entre las variables más importantes para modelar el AGB, se encuentran las bandas R, G y NIR del mosaico de la época de lluvias (mayo a septiembre).

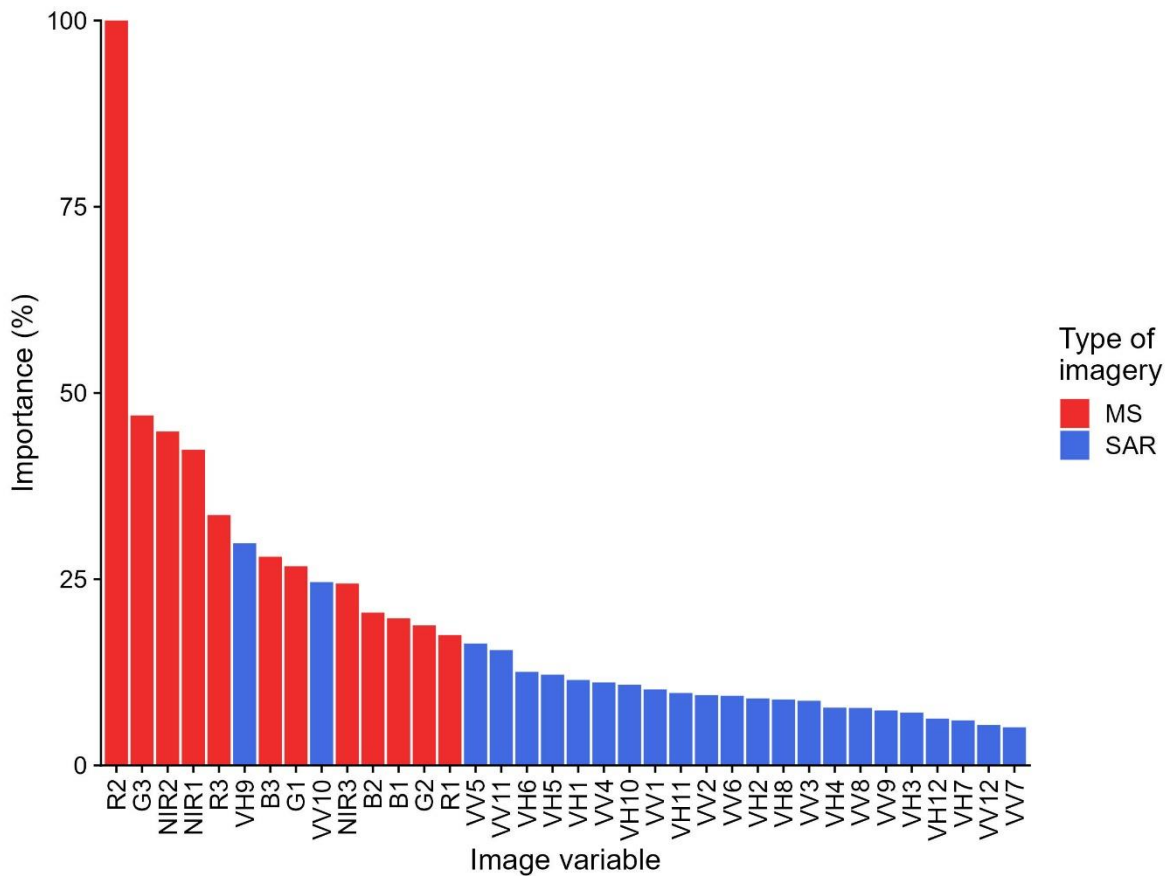


Figura 32. Importancia relativa de las variables predictivas para predecir AGB utilizando random forest.

Mapas de AGB

Los mapas de AGB fueron producidos con los datos de 2019 y 2022, respectivamente, y utilizando la arquitectura U-Net 3D LF y el algoritmo RF, que obtuvo el menor error sobre los datos de prueba. Al observar las predicciones de AGB realizadas por los dos algoritmos se observó que RF tendió a predecir valores más homogéneos de AGB (0 – 60 Mg / ha), mientras que U-Net 3D LF hizo predicciones con un mayor intervalo de valores entre el mínimo y el máximo (0 – 170 Mg / ha; Figura 33 y Figura 34). Por su parte, la comparación de la biomasa presente en el área de estudio entre 2019 y 2022 mostró una disminución de AGB entre ambos años, utilizando los dos algoritmos (Cuadro 7; Figura 35). En esta comparación, se detectaron pérdidas más grandes en las predicciones realizadas por el algoritmo U-Net 3D LF, asociadas a la predicción de valores de AGB muy altos en 2019.

Cuadro 7. Biomasa aérea estimada (AGB) para el bosque tropical caducifolio (BTC) del área de estudio en 2019 y 2022, así como el área promedio por pixel. El área total del BTC es de 81353.24 ha.

Área BTC (ha)	Método	Año	AGB (Mg)	Diferencia (Mg)	Diferencia (%)
80 495 ha	U-Net 3D	2019	2 866	-227 692.69	-7.94
			852.90		
	RF	2022	2 639	-154 143.14	-6.60
			160.21		
	RF	2019	2 336	219.82	076.68
			2 182		

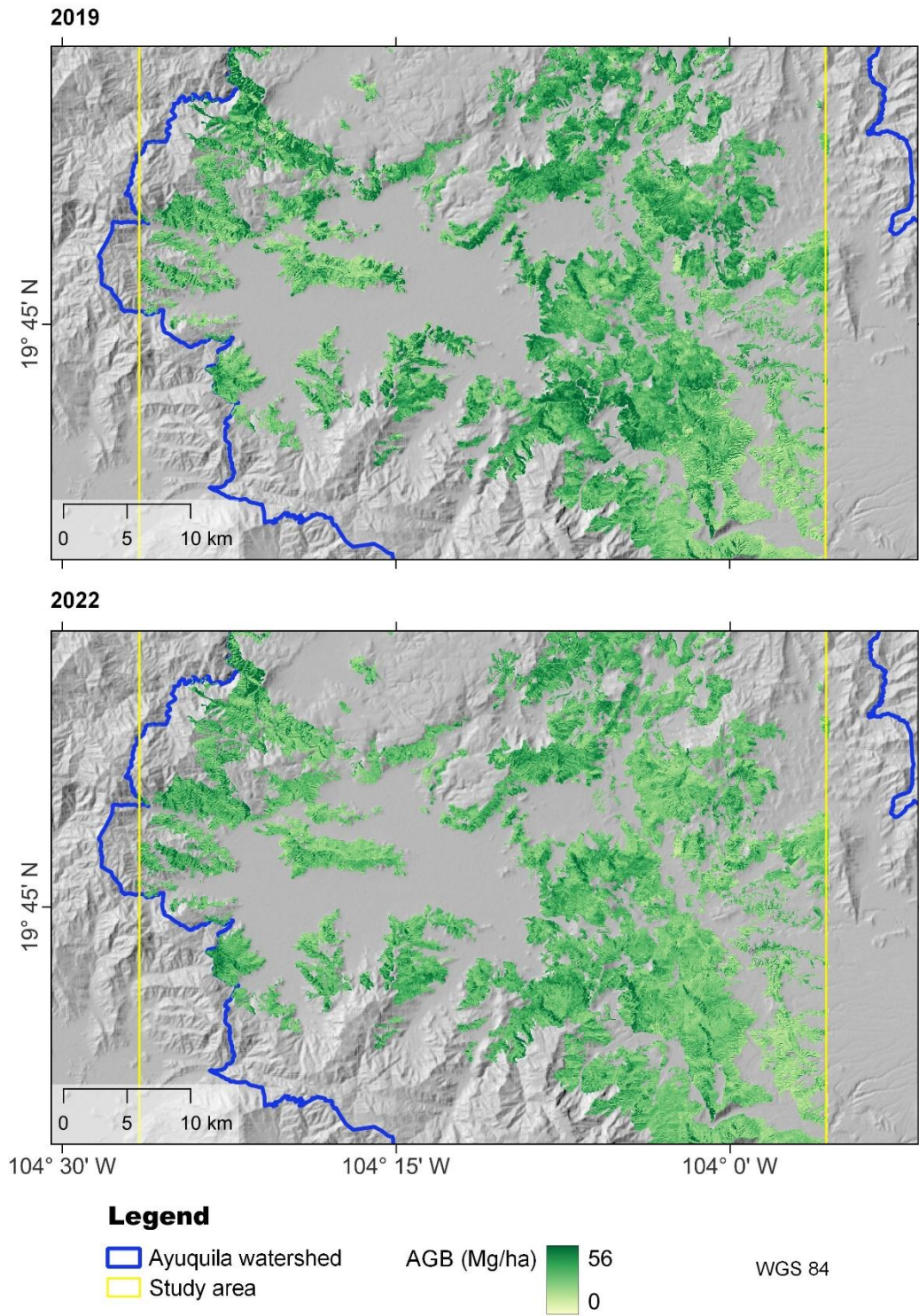


Figura 33. Predicción de biomasa en el BTC realizada por el modelo U-Net 3D LF para 2019 y 2022.

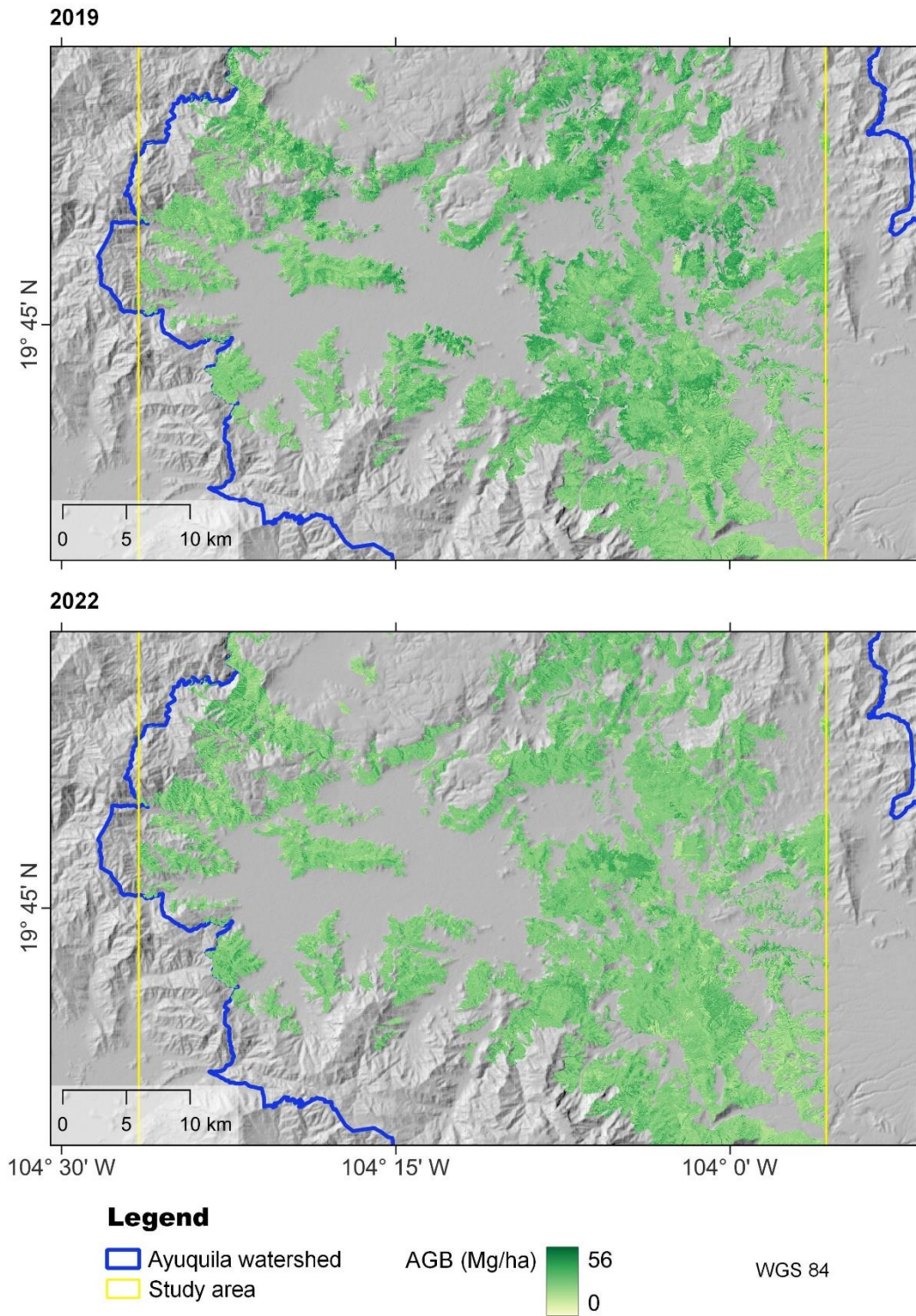
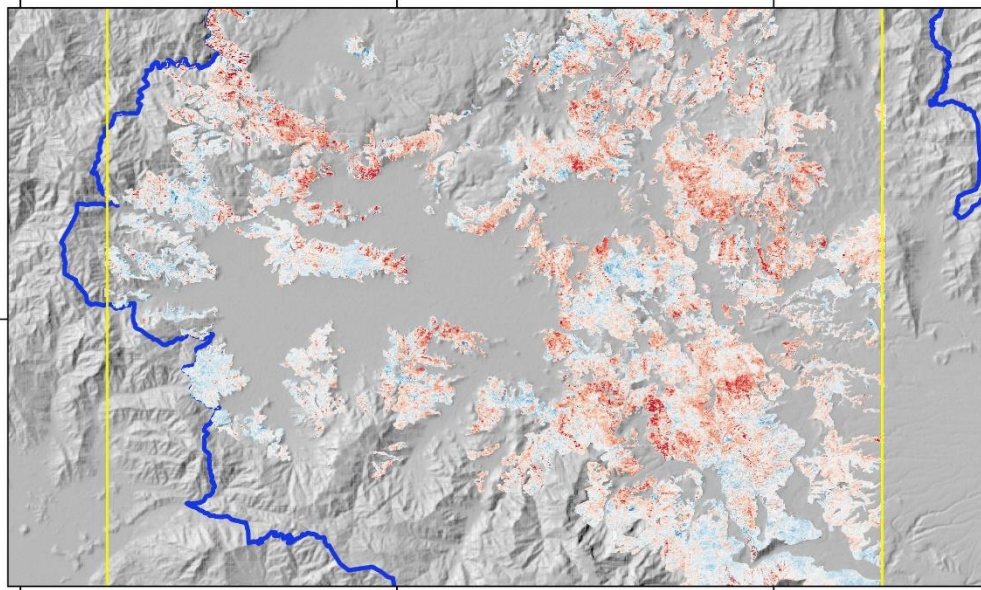
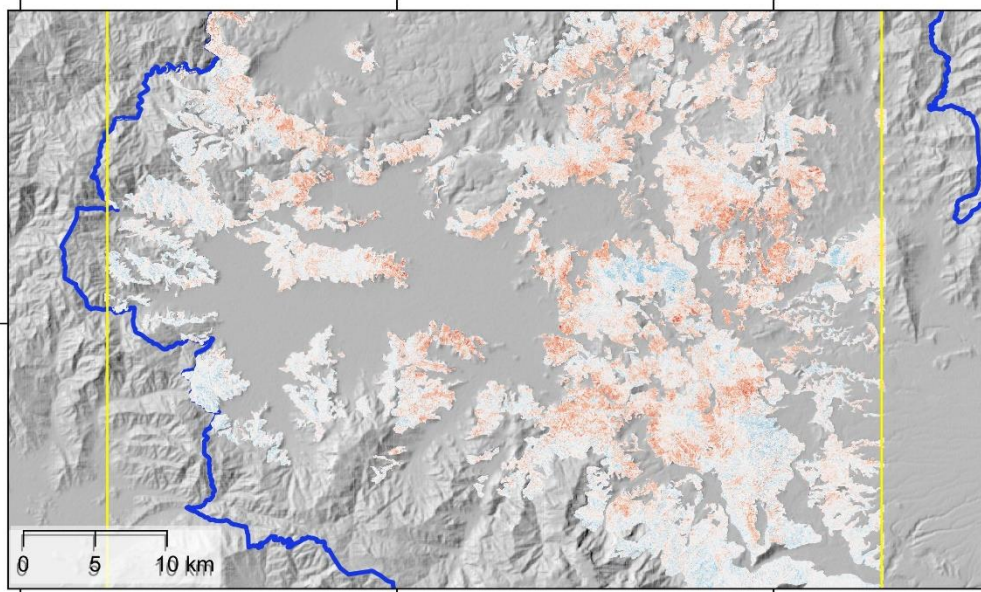


Figura 34. Predicción de biomasa en el BTC realizada por el modelo RF para 2019 y 2022.



U-Net 3D LF change



Random forest change



Legend

-  Ayuquila watershed
-  Study area

AGB change (Mg / ha)



WGS 84

Figura 35. Cambios de biomasa en el BTC entre 2019 y 2022 de acuerdo con las predicciones realizadas por los modelos RF y U-Net 3D LF.

Análisis de errores

Al analizar los resultados obtenidos de los cambios detectados con las predicciones del método RF, se detectaron algunos aciertos y errores (Figura 36). Este escenario era esperado debido al alto valor del RMSE presentado en la validación y prueba del algoritmo. A pesar de ello, se detectaron circunstancias donde la disminución de AGB correspondió a una remoción de la cobertura arbórea del BTC de la zona o una aparente ganancia de AGB (Figura 36 A y D). Sin embargo, también se detectaron casos de errores asociados a la presencia de artefactos en las imágenes (Figura 36 B y C).

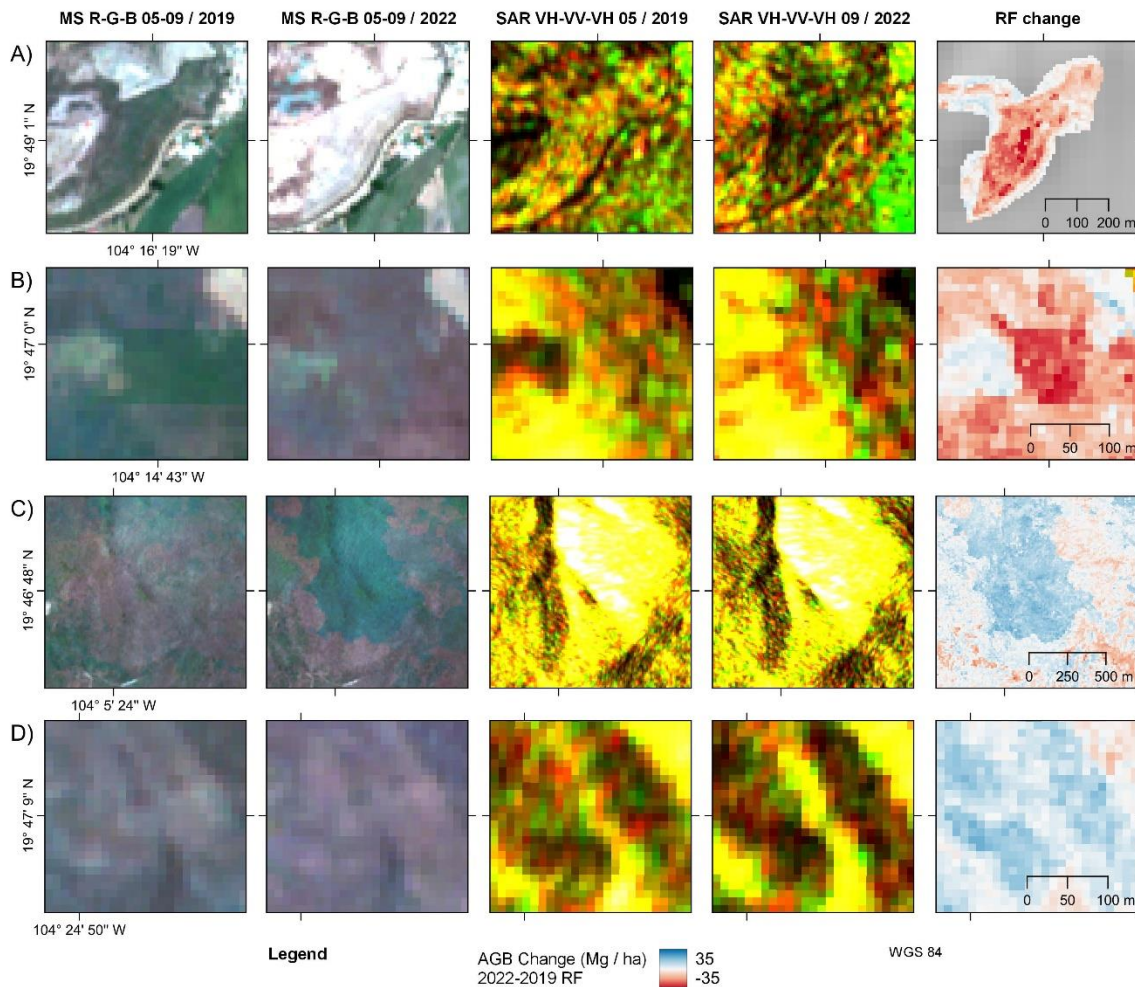


Figura 36. Ejemplos de aciertos o errores detectados en la diferencia del AGB entre los mapas predichos con random forest para 2019 y 2022. A) una pérdida de AGB, B) una falsa detección de pérdida de AGB, C) una falsa detección de ganancia de AGB, D) una posible ganancia de AGB. El cambio de AGB se calculó como al AGB de 2022 menos el AGB de 2019, de tal manera que los valores negativos implican una pérdida y los valores positivos una ganancia.

Evaluaciones aleatorias

Los resultados de las aleatorizaciones del algoritmo RF muestran que el error no se reduce haciendo una validación cruzada en 10 grupos (*10-fold cross validation*), ni utilizando un *bootstrap* (25 divisiones de entre 30 – 40 % de datos para validar, 60 – 70 % de datos para entrenamiento), al respetar las divisiones por sitios (*i.e.*, todos los píxeles de un sitio se van para entrenamiento o verificación). Sin embargo, una validación cruzada LOO sí aumenta un poco el RMSE promedio obtenido (Cuadro 8). Además, al dividir los datos por parcela,

se obtienen valores de error mayores a los obtenidos al dividir los datos de manera totalmente aleatoria. Esto quiere decir que: 1) los resultados obtenidos no dependieron de la selección de datos de entrenamiento, verificación y prueba, 2) el elegir una evaluación con alguno de los métodos aleatorios utilizados tampoco permite obtener estimados del desempeño mayores, a excepción de LOO, y 3) debido a los resultados obtenidos con la distribución totalmente aleatoria, es probable que con un conjunto de datos más grande se habrían obtenido mejores resultados.

Cuadro 8. Resultados obtenidos para el modelo RF aleatorizando la división de datos: 1) tradicional, respetando la división por parcela (9 píxeles por parcela) o 2) totalmente aleatorizado. Se presentan los resultados con tres métodos 10-fold, Bootstrap y leave-one-out cross validation (LOO). Abreviaciones: EE, error estándar.

División de datos	Método	RMSE		
		Promedio	EE	n
Por parcela	10-fold	14.7	1.81	10
	Bootstrap	15.6	0.34	25
	LOO	12.8	1.10	63
Totalmente aleatorio	10-fold	8.03	0.48	10
	Bootstrap	8.97	0.09	25
	LOO	5.75	0.23	567

El resultado de las diez predicciones de AGB de acuerdo con la división de datos de 10-fold indicó que tanto las predicciones de 2019 como las del 2020 presentaron, en general, un bajo coeficiente de variación (CV; entre 0 y 15 %); sin embargo, se logran observar algunas zonas con CV altos (alrededor de 30 %) sobre todo en las zonas noroeste y sureste (Figura 37). Esto quiere decir que al cambiar los datos de entrenamiento, en la mayoría del área de estudio se predicen valores muy similares de AGB, aunque hay sus excepciones. El coeficiente de variación para cada píxel tuvo un intervalo de valores entre 0 y 51 %, lo cual quiere decir que en general los estimados de AGB por píxel no variaron mucho. Para ejemplificar los resultados de las aleatorizaciones, si un píxel obtuvo un valor promedio de AGB = 10 Mg / ha y un porcentaje de variación del 10 %, esto quiere decir que el 95 % de las aleatorizaciones obtuvieron valores entre 8.04 y 11.96 Mg / ha (promedio \pm 1.96 EE).

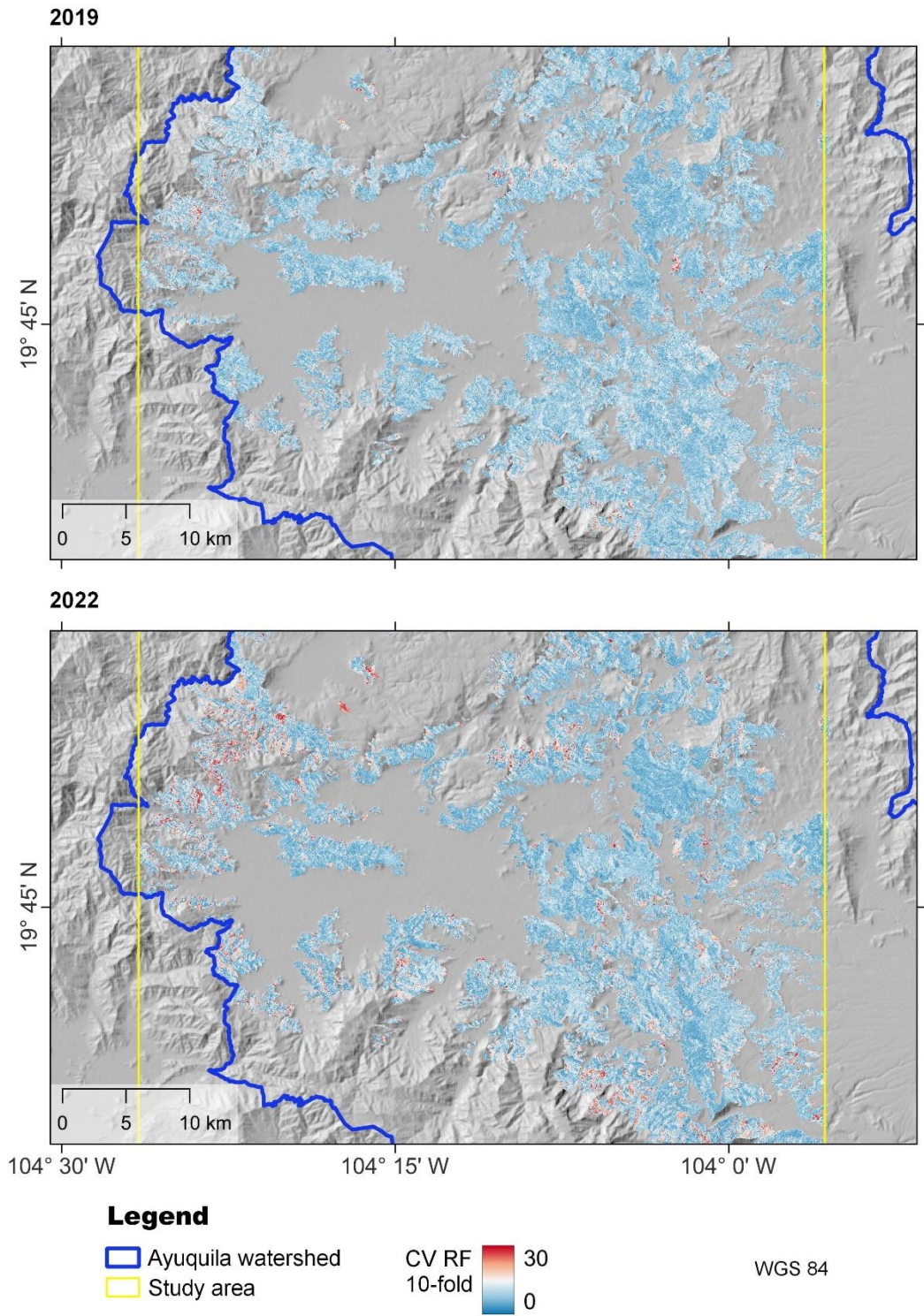


Figura 37. Coeficiente de variación por píxel de acuerdo a las aleatorizaciones realizadas con el algoritmo random forest y usando un 10-fold cross validation para el año 2019 y 2022.

Discusión

Estudio de la AGB en la zona de estudio

Este estudio evaluó el uso de algoritmos de aprendizaje profundo (U-Net 3D LF) y aprendizaje automatizado (random forest) con imágenes SAR y MS para modelar la AGB en un BTC. Estos algoritmos se utilizaron para hacer predicciones sobre el mismo BTC en 2019 y 2022 para poder detectar la degradación forestal como una disminución en AGB, aunque se observó un error relativamente grande, considerando la amplitud del intervalo de confianza ($14 < \text{RMSE} < 18.20 \text{ Mg / ha}$; $27.44 < \text{IC} < 35.67 \text{ Mg / ha}$). Es probable que este error se haya debido a una combinación de factores que incluyen el uso de parcelas pequeñas, errores de geoposicionamiento, errores en el cálculo de la biomasa de los datos de 2019, entre otros (ver sección Limitantes sobre el método de muestreo y el procesamiento de imágenes). Esto se traduce en que pérdidas de biomasa con una magnitud similar correspondieron a eventos de deforestación (*i.e.*, verdaderos positivos), aunque también a errores causados por la presencia de artefactos (*i.e.*, falsos positivos; Figura 36). Por ello, esta información no permite distinguir claramente entre errores y eventos de disminución de biomasa. Además, debido a que la degradación forestal implica una disminución en la biomasa sin un cambio de cubierta, resultó prácticamente imposible distinguir este proceso de cambios causados por artefactos en las imágenes. Por ello, consideramos que los resultados tienen una baja posibilidad de utilizarse para detectar la degradación forestal.

Aunque los mapas obtenidos no son de alta calidad, podrían servir para identificar áreas de reducción de biomasa para después enfocar la interpretación visual en ciertas áreas. Estos servirían como una guía para diferenciar disminuciones verdaderas de AGB y errores, para después asociarlos con algunos de los conductores de la degradación forestal en la zona de estudio son la obtención de leña y postes, áreas de pastoreo para el ganado y zonas bajo ciclos intermedios de descanso para aprovechamiento agrícola (Borrego and Skutsch, 2019; Morales-Barquero et al., 2015; Salinas-Melgoza et al., 2018). Además, este tipo de resultados podrían brindar luz sobre métodos más adecuados para detectar alguna causa de degradación forestal específica (Gao et al., 2020). De acuerdo con nuestro muestreo de campo, sólo potencialmente solo podríamos detectar la degradación forestal

causada por una extracción selectiva de árboles completos (*e.g.*, uso de postes) o zonas bajos ciclos intermitentes de aprovechamiento agrícola; mientras que la degradación forestal causada por una extracción de ramas o por el pastoreo de ganado, sería indetectable con nuestros datos de campo. Para detectar estas últimas causas de la degradación forestal habría que considerar un diferente muestreo de campo que permitiera modelar la pérdida de ramas o de individuos jóvenes y después tratar de relacionarlo con un método de percepción remota.

Debido a los errores frecuentes en los mapas de AGB, se prefirió realizar un análisis a nivel de toda la AGB del BTC en la región, para detectar la tendencia de cambio general de la zona de estudio. Este análisis asume que las estimaciones por píxel pueden contener errores; sin embargo, dichos errores deberían distribuirse aleatoriamente de tal manera que el estimado de toda el área de estudio no debería contener un sesgo particular para alguno de los años estimados. Por un lado, este acercamiento permitió hacer una estimación sobre la tendencia de cambio de toda la biomasa del área de estudio, aunque pierde la información de la localización en el espacio de estos cambios. Este análisis mostró que el BTC disminuyó su AGB entre 2019 y 2022, de acuerdo a ambos modelos. U-Net 3D LF predijo una AGB en 2019 mayor, así como una disminución también mayor (7.94 %), en comparación con la estimada por RF (6.60 %; Cuadro 7). Debido que los modelos predijeron valores alrededor de la media en los datos de validación y prueba, se analizó esta disminución de AGB no había sido causada simplemente porque el promedio de AGB de 2019 era menor al de 2022. Sin embargo, se observó que este no fue el caso, ya que el promedio de las parcelas muestreadas en 2022 fue mayor al de 2019 (2019: 26.83 Mg / ha; 2022: 28.61 Mg / ha). En conjunto, estos hallazgos confirman que el BTC del sitio de estudio se encuentra perdiendo AGB.

Comparación de algoritmos U-Net 3D LF y RF

Las predicciones realizadas por el algoritmo RF se presentaron un menor error sobre los datos de prueba. Por ello, se concluyó que RF permitió obtener resultados más precisos (aunque sobre los datos de validación obtuvo un RMSE más alto). Además, las predicciones de RF se encontraron dentro del intervalo de datos registrado en campo aprox. 0 – 70 Mg / ha, mientras que las predicciones de la U-Net 3D tomaron valores de entre 0 y

170 Mg / ha, lo cual representa una evidente sobrestimación de AGB en un bajo número de píxeles. El hallazgo de que un algoritmo de aprendizaje automática superara a uno de aprendizaje profundo fue contrario a lo encontrado en estudios previos para modelar la AGB a partir variables de percepción remota (Dong et al., 2020; Moradi et al., 2022; Talebiesfandarani and Shamsoddini, 2022; Zhang et al., 2022). Es probable que esta discrepancia se deba a una combinación de: 1) una baja proporción de señal / ruido causada tanto por una alta variación de los datos, como una relación difusa entre la señal remota y los valores de AGB, así como 2) de un conjunto de datos pequeño.

La gran variabilidad presente en los datos creemos que fue causada por la heterogeneidad presente en tres dimensiones: entre sitios, entre temporadas de un mismo año y entre años (2019 y 2022), así como por ciertas limitaciones del estudio (ver sección: Limitantes sobre el método de muestreo y el procesamiento de imágenes). Es muy probable que la variación entre datos respondió a diferentes condiciones topográficas (pendiente, orientación, entre otras). Estudios previos han reportado que las variables topográficas son importantes predictores de la AGB en la región (Mermoz et al., 2014; Réjou-Méchain et al., 2014; Salinas-Melgoza et al., 2018). Por lo tanto, estas variables podrían ser mejores predictores que las variables utilizadas en este estudio (*i.e.*, bandas MS y SAR). Además, debido a que las bandas SAR se registran en un ángulo oblicuo, estas son sensibles a la topografía de una región; y, por lo tanto, la retrodispersión de ciertas zonas se puede ver afectada por la orientación y pendiente de las laderas (Mermoz et al., 2014). Ya que el sitio de estudio presenta topografía muy accidentada, es probable que la interacción de ésta con la geometría del sensor haya opacado alguna posible relación entre la retrodispersión y la vegetación (Flores-Anderson et al., 2019). En este sentido, la inclusión de variables topográficas como términos predictores quizás permitiría mejorar los resultados al incluir información biofísica relevante relacionada con la AGB, así como ayudar a explicar parte de la variación presente en la retrodispersión de las imágenes SAR. Por otro lado, la variación entre sitios se pudo haber visto magnificado debido a las diferencias en la composición de especies de árboles presentes en cada uno de ellos, así como de distintos tipos de suelo (Mermoz et al., 2014).

En el caso de la variación en distintas temporadas de un mismo año, las bandas MS siguieron el patrón esperado, *i.e.*, menor reflectancia en NIR y mayor en el R durante la época de secas (t_1) y viceversa en lluvias (t_3 y t_2 ; Figura 29). En el caso de las bandas SAR se esperaba una mayor retrodispersión en la banda VH en la temporada de lluvias (meses 5 – 10), debido a la presencia de follaje. Aunque este fue el caso en algunos sitios, otros mostraron valores más altos de retrodispersión en la temporada de secas (meses 11 – 4). En este caso, creemos que cambios en la humedad del suelo (sobre todo en la época de secas, donde se esperaría mayor interacción con el suelo), también hayan provocado una mayor variabilidad intranual en la señal de retrodispersión de cada sitio (Bhogapurapu et al., 2022; Joshi et al., 2016). De esta manera, valores muy altos de retrodispersión en la banda VH pudieron haber sido causados por un aumento en la humedad del suelo (Pulliainen et al., 1999). Finalmente, debido a que los sitios mostraron tanto poca como una alta variación en su retrodispersión intra-anual, consideramos que la relación de la geometría del sensor y la topografía quizás fueron más relevantes para determinar su retrodispersión, que la cubierta vegetal. Por lo tanto, esto pudo obscurecer la relación entre la retrodispersión y la AGB de la vegetación arbórea presente en cada sitio.

Aunque la variabilidad interanual también puede ser una fuente de ruido, en respuesta a algunas variaciones climáticas entre los dos años en los que se obtuvieron datos (*e.g.*, precipitación en 2019 y 2022), es probable que esta no tuvo un impacto importante. Por ejemplo, en las Figura 29 y Figura 30 se puede observar que aunque algunos datos de retrodispersión o reflectancia provienen de imágenes del 2019 y otras del 2022, no se observa alguna variación radical asociada únicamente al año. Además, consideramos que la estandarización de las imágenes ayudo a disminuir este efecto y aumentar el tamaño de muestra (Figura 29 y Figura 30). Por ello, creemos que esta fuente de variación fue la menos relevante para explicar la falta de resultados con bajo error. Sin embargo, la combinación de estas tres fuentes de variación provocó una enorme variabilidad en la relación AGB con la reflectancia y el coeficiente de retrodispersión, lo cual dificultó su modelación.

Por otra parte, el hecho de que se haya encontrado un buen ajuste sobre los datos de entrenamiento y no sobre los demás datos implica que hubo cierto sobreajuste de los dos

algoritmos a los datos de entrenamiento. A pesar de que en la U-Net 3D LF se utilizaron técnicas para evitar el sobreajuste, como el *dropout* o el *early stopping*, esto no fue suficiente como para poder generalizar los resultados a los datos de validación y prueba (Ying, 2019). Por su parte, el algoritmo RF presentó un patrón similar, aunque con menor error. Este hallazgo sugiere que el conjunto de datos fue muy pequeño o por lo menos, lo fue para poder encontrar un patrón en los datos que presentaron una alta variabilidad. Estudios previos sugieren que el efecto del tamaño del conjunto de datos de entrenamiento puede afectar en mayor o menor manera los resultados dependiendo de la complejidad de la tarea y la facilidad de reconocimiento de patrones (Adarme et al., 2020; Hamedianfar et al., 2022; Kattenborn et al., 2021). Por ello, se consideró que gracias al tamaño de muestra, RF fue capaz de superar a U-Net 3D LF, ya que es un algoritmo que puede obtener buenos resultados con un conjunto limitado de datos. Por el contrario, el potencial de la U-Net 3D LF no pudo ser observado debido al número reducido de muestras. En este caso, es probable que la alta plasticidad del algoritmo de aprendizaje profundo provocara un mayor sobreajuste que su contraparte RF.

Los resultados de las evaluaciones aleatorias parecen apoyar el hecho de que la principal limitante del análisis realizado fue el número de datos (aunado a su gran variabilidad), ya que estos mostraron que el RMSE prácticamente no mejora al cambiar la división de los datos en los tres conjuntos (*i.e.*, entrenamiento, validación y prueba). Por lo tanto, el desempeño no será muy distinto, aunque se seleccione un conjunto distinto de datos de entrenamiento y validación. Además, estos análisis sugieren que se podría obtener un menor error si se tuviera un conjunto de datos mayor, que abarcara una mayor variación de la relación entre AGB y la reflectancia y retrodispersión de las imágenes. Por ejemplo, los resultados del Cuadro 8 indican que al hacer aleatorizaciones por pixel, sin respetar la pertenencia de un píxel a algún cuadro de muestreo (*i.e.*, píxeles de una parcela de muestreo divididos en datos de entrenamiento y validación), se puede disminuir el error casi a la mitad (Cuadro 8). Aunque estos análisis de aleatorización sólo se realizaron para el algoritmo RF por cuestiones de tiempo de cómputo, se esperaría que el algoritmo U-Net 3D LF mostrara un patrón similar.

Los resultados de las predicciones y su coeficiente de variación mostraron que la gran mayoría del área de estudio mostró predicciones de AGB con baja variación (CV entre 0 y 15 %). Esto implica que aunque se cambie el conjunto de datos de entrenamiento, se obtienen predicciones de AGB relativamente similares para casi toda el área de estudio. Sin embargo, dado que el RMSE fue alto para la verificación y prueba del método RF SAR + MS, esto implica que consistentemente el AGB se predice mal en la mayoría del área de estudio, sin importar si se cambian los datos de entrenamiento. Esto se traduce en que las predicciones son precisas, pero no exactas.

En el caso del uso de arquitecturas basadas en U-Net para la modelación de AGB, el error relativo de la evaluación de este trabajo, en general, fue mayor al reportado en trabajos previos (Buxbaum et al., 2022; Castro et al., 2020; Dong et al., 2020; Ghosh and Behera, 2021; Hamedianfar et al., 2022; Moradi et al., 2022; Talebiesfandarani and Shamsoddini, 2022; Zhang et al., 2022; Zhang et al., 2019b). La mayoría de estos trabajos tratan de evaluar alguna arquitectura basada en CNN para modelar la biomasa de diferentes tipos de bosques, cubiertas vegetales a una escala distinta. En este punto cabe resaltar, que la capacidad de modelación de la biomasa de un sistema también depende del tipo de vegetación y tipo de sistema (*e.g.*, tipo de vegetación o si se trata de bosque maduro o secundario; Eckert, 2012; Solórzano et al., 2021). Esto implica que muchos de estos trabajos no son estrictamente comparables con el nuestro, ya que se realizaron en sistemas más sencillos en comparación al de este estudio (*e.g.*, plantaciones, campos agrícolas; Buxbaum et al., 2022; Castro et al., 2020); aunque hay algunos hechos en bosques templados (Moradi et al., 2022; Zhang et al., 2019b) o áreas cubiertas de bambú (Dong et al., 2020); u obtienen los datos de la AGB de fuentes globales calibradas (pero no de datos de campo tal cual; Talebiesfandarani and Shamsoddini, 2022), o incorporando otro tipo de variables remotas (*e.g.*, LiDAR o coherencia SAR; Ghosh and Behera, 2021; Zhang et al., 2019b).

En el caso del random forest también se encontró que el error relativo reportado en este trabajo fue mayor al de estudios previos (Chen et al., 2018; Ghosh and Behera, 2018; Hernández-Stefanoni et al., 2020; Li et al., 2020; Zhang et al., 2019b; Zhang et al., 2023) o similar (Forkuor et al., 2020; Zhang et al., 2022). De igual manera que en la comparación

con U-Net algunos son realizados sobre plantaciones (Ghosh and Behera, 2018), bosques templados (Zhang et al., 2022; Zhang et al., 2019b; Zhang et al., 2023), bosques subtropicales (Li et al., 2020) o tropicales naturales (Forkuor et al., 2020; Ghosh and Behera, 2018; Hernández-Stefanoni et al., 2020). Además, algunos de estos utilizan otro tipo de información como LiDAR (Hernández-Stefanoni et al., 2020; Zhang et al., 2019b; Zhang et al., 2023), imágenes Landsat (Li et al., 2020; Zhang et al., 2019b) o SAR banda L (Hernández-Stefanoni et al., 2020; Zhang et al., 2022). Por ende, resulta difícil extraer patrones generales y al igual que en el caso de la U-Net, el desempeño de este tipo de algoritmos depende de una combinación de factores que incluyen el sistema estudiado, las métricas incorporadas como predictivas, el número de datos, entre otras.

Finalmente, debido a que el algoritmo que permitió obtener los mejores resultados fue el random forest, sería interesante poner a prueba el desempeño de algunos índices como el NDBI o el índice de vegetación de radar (Almeida-Filho et al., 2009; Mandal et al., 2020). Debido a que esperábamos obtener un mejor resultado con el algoritmo de aprendizaje profundo, no se utilizaron transformaciones de los datos como variables predictivas. Sin embargo, en una futura evaluación sería interesante incorporar algunos de estos índices como parte de las variables predictivas.

Importancia de las bandas

Los resultados de la importancia relativa de cada variable predictiva en el algoritmo RF demostraron que la información contenida en las bandas MS fue la más relevante, en comparación con las SAR. Este patrón de mayor relevancia de las bandas MS que las SAR para modelar la AGB coincide con algunos estudios previos (Forkuor et al., 2020; Ghosh and Behera, 2018; Hirschmugl et al., 2018; Li et al., 2020; Zhang et al., 2022; Zhang et al., 2023; Zhao et al., 2016), pero no con otros (Chen et al., 2018; Laurin et al., 2018; Sinha et al., 2019). A pesar de ello, la mayoría de los estudios considera que la combinación de información MS y SAR brinda las mejores capacidades de monitoreo de la AGB mediante sensores remotos (Deng et al., 2014; Forkuor et al., 2020; Ghosh and Behera, 2018; Hirschmugl et al., 2018; Laurin et al., 2018; Li et al., 2020; Sinha et al., 2019; Zhang et al., 2022; Zhang et al., 2023; Zhao et al., 2016).

En el caso de este estudio, la información más relevante correspondió a los tiempos 2 y 3, que corresponden a la época de lluvias y después de la época de lluvias, respectivamente. Además, entre las bandas más relevante se encontraron R, G y NIR. Esto quiere decir que la información más relevante para modelar la AGB constó de las bandas R y NIR en periodos donde la vegetación cuenta con follaje. Estas bandas están relacionadas con la actividad fotosintética de la vegetación. Por su parte, las bandas SAR no resultaron de mucha importancia, aunque las más importantes fue VH en el mes de septiembre, en el final de la época de lluvias y después VV en octubre. Esto es contrario a lo encontrado en estudios previos donde la información de la temporada de secas resulta más útil (Forkuor et al., 2020; Laurin et al., 2018).

Estudios previos sugieren que el mayor potencial para estudiar la biomasa mediante sensores SAR, es mediante aquellos de menor frecuencia (*i.e.*, mayor longitud de onda con las bandas L o P), ya que permiten penetrar más el follaje e interactuar con los troncos (Ghasemi et al., 2011; Hernández-Stefanoni et al., 2020; Hirschmugl et al., 2020; Lu, 2006; Mermoz et al., 2014; Zhao et al., 2016). A pesar de ello, algunos estudios previos han reportado buenas capacidades de modelar la AGB a partir de variables de SAR banda C, sobre todo en sistemas de baja AGB, como era el caso de este estudio (Chen et al., 2018; Laurin et al., 2018; Sinha et al., 2019). En este caso, pensábamos que el SAR banda C podría sobrellevar este aspecto, debido a que durante la época de secas la mayoría del bosque no presenta follaje, y, por lo tanto, las microondas podrían interactuar directamente con los troncos. Esto también puede implicar que las microondas interactúen con estructuras más pequeñas como ramas (que se encuentran presentes en la época de secas) o inclusive con la rugosidad del suelo o su humedad (Flores-Anderson et al., 2019). La interacción sobre todo con la humedad del suelo permitiría explicar la enorme variación detectada en la retrodispersión de algunos sitios (Figura 30; Pulliainen et al., 1999). Además, cabe mencionar que la información SAR utilizada en este estudio correspondió únicamente a la información de retrodispersión; sin embargo, en un futuro sería interesante explorar el uso también de la información de fase (la cual no se encuentra actualmente disponible en GEE).

Limitantes sobre el método de muestreo y el procesamiento de imágenes

La idea que subyace al uso de los algoritmos basados en CNN para modelar la AGB se encuentra en el potencial de métricas de textura para modelar dicho atributo. Se ha propuesto que esto se debe a que estas métricas permiten sobrepasar el problema de la saturación de la relación de las variables espectrales y la biomasa, así como registrar patrones espaciales de los tamaños de las copas de los árboles (Couteron et al., 2005; Eckert, 2012; Gallardo-Cruz et al., 2012). Varios estudios previos han reportado buenas capacidades para modelar la AGB a partir de la textura de imágenes multiespectrales (Ozdemir and Karnieli, 2011; Ozkan and Demirel, 2021; Singh et al., 2015; Solórzano et al., 2017), radar (Berninger et al., 2018; Ghosh and Behera, 2018; Hernández-Stefanoni et al., 2020; Jiménez-Rodríguez et al., 2022; Liao et al., 2020) e inclusive LiDAR (Hernández-Stefanoni et al., 2020) en bosques tropicales. Estos estudios sugieren que las relaciones espaciales entre los píxeles de un área determinada permiten describir y predecir los valores de AGB presentes en el bosque. De hecho, reportes previos mencionan que algunos algoritmos de aprendizaje automatizado se benefician en gran medida del uso de algunas variables de textura (Dong et al., 2020; Zhang et al., 2022). Algunos estudios han sugerido que esto puede estar relacionado con la presencia de sombras entre copas contiguas o directamente entre la señal remota de las copas (Gerard and North; Proisy et al., 2011; Sarker and Nichol, 2011; Solórzano et al., 2017). Por esta razón, se hipotetizaba que el uso de las CNN como método predictor sería prometedor, ya que permitiría captar estas relaciones tono-espaciales sin necesidad de predefinir las métricas calculadas. Sin embargo, los resultados mostraron lo contrario.

Tanto utilizando métricas de textura como de cualquier CNN, la resolución de las imágenes determina el detalle de la información espacial que estos tipos de algoritmos pueden identificar. Por ello, es probable que la resolución de las imágenes utilizadas en este estudio fuera demasiado gruesa para proveer información de gran relevancia (Lu and Batistella, 2005; Matosak et al., 2022; Solórzano et al., 2017; Torres et al., 2021). Además, es posible que al contar con un tamaño de ventana tan reducido (3 x 3 píxeles), las capacidades de resumir información espacial fueran muy bajas o no permitieran identificar patrones espaciales, relevantes para modelar la AGB (Dong et al., 2020; Hamedianfar et al.,

2022). Por lo tanto, la resolución espacial de las imágenes pudo jugar en contra del potencial de modelación de la AGB usando la U-Net 3D LF. Asimismo, esto provocó que la información proveniente de las convoluciones espaciales no fuera tan útil, y, por lo tanto, esto pudo contribuir a obtener mejores resultados con RF (*i.e.*, en un método por píxel).

Otro aspecto relacionado con la resolución espacial de las imágenes y la riqueza de la información que se puede extraer de una imagen, es la extensión de la parcela de muestreo. En este caso, se eligió continuar con un esquema de muestreo de áreas de 500 m² para tener un conjunto de datos mayor (aunados a los datos de 2019). Estudios previos sugieren utilizar tamaños de parcelas grandes para caracterizar la AGB, debido a la variación que esta puede presentar (Hernández-Stefanoni et al., 2018; Réjou-Méchain et al., 2019; Réjou-Méchain et al., 2014; Saatchi et al., 2011). Por ejemplo, varios estudios sugieren utilizar parcelas de muestreo mayores a 0.25 ha, debido a la autocorrelación negativa encontrada entre parcelas de muestreo muy cercanas (relacionados a claros y otras variaciones locales) o a la relación inversa entre la variación de los atributos del bosque y el tamaño de la parcela (Chave et al., 2004; Pinto et al., 2021; Wagner et al., 2010). Por otro lado, tanto otros estudios sugieren utilizar parcelas de muestreo mayores (aprox. 1000 m²) para obtener mediciones con menor error de AGB, ya que disminuye la importancia del error entre la posición la registrada con el GPS y de la imagen (Hernández-Stefanoni et al., 2018; Saatchi et al., 2011). En conclusión, el tener parcelas de muestreo mayores a 0.25 ha permite mejorar la modelación de la AGB, ya que disminuyen la influencia de las condiciones muy locales sobre el estimado de AGB, así como el error entre la imagen y las coordenadas registradas en campo. Este aspecto pudo jugar en contra de la capacidad de los algoritmos para modelar la AGB.

A pesar de que es deseable utilizar parcelas de mayor tamaño para estimar algunos atributos del bosque, existe un compromiso entre el tamaño y el número de parcelas de muestreo. Usualmente ante un escenario de recursos limitados, el utilizar parcelas de muestreo de mayor tamaño provocará una reducción en el número de parcelas muestreadas. Debido a que en los modelos cada parcela es contada como una observación, tampoco es deseable registrar muy pocas parcelas de gran tamaño. Por lo tanto, ante esta circunstancia, el mejor escenario sería el de utilizar parcelas de muestreo lo suficientemente grandes para

que no sean tan sensibles a condiciones muy locales y permitir contar con un número estadísticamente representativo para ajustar un modelo ($n \approx 30$). En el caso de este estudio, parece ser que el tamaño de las parcelas de muestreo fue muy pequeño, lo cual se puede considerar como una fuente más de error o ruido en los datos. Una alternativa prometedora para resolver este tipo de compromisos podría ser utilizar información de otras fuentes remotas de mayor detalle (*i.e.*, GEDI, LiDAR) para tratar de generar un mayor conjunto de datos de menor calidad. Creemos que esta alternativa podría ser prometedora y debería explorarse en el futuro, aunque también conllevará sus propias dificultades (*e.g.*, modelos de biomasa basados en GEDI calibrados a nivel global). De hecho, en este estudio, algunos sitios de muestreo (*i.e.*, 12) coincidieron con el paso del sensor GEDI.

Adicionalmente, debido a que los datos de 2019 no contaron con información taxonómica, se utilizó una densidad de madera promedio calculada a partir del muestreo de 2022. Esto es otra fuente de error que contribuyó a que se obtuvieran modelos con errores relativamente altos, ya que este acercamiento ignora la variación de densidad de madera entre especies.

Por otro lado, el uso de las convoluciones en la dimensión temporal se basó en el uso de este tipo de operaciones para realizar alguna tarea de clasificación o regresión (Pelletier et al., 2019). Las convoluciones temporales permiten resumir patrones temporales relevantes para la tarea en cuestión, similar a lo que pueden realizar otras arquitecturas como las redes neuronales recursivas (RNN; Hamedianfar et al., 2022; Pelletier et al., 2019). Además, Otros trabajos han sugerido que el uso por separado de las dimensiones temporales, espectrales y espaciales permiten mejorar los resultados (Giannopoulos et al., 2022). En este caso, se construyó una arquitectura que aprovecha las convoluciones en la dimensión temporal y espacial. A pesar de ello, los resultados no fueron muy prometedores.

Esté capítulo parece indicar que no se requieren de técnicas más avanzadas de aprendizaje, como los algoritmos de aprendizaje profundo para modelar la AGB mediante información proveniente de sensores remotos, sino más bien en un aumento del conjunto de datos para entrenar este tipo de algoritmos, en su calidad (parcelas de mayor tamaño) o en el detalle de la información de percepción remota (otras fuentes como GEDI). Por un lado,

es probable que estos algoritmos permitan mejorar dicha modelación únicamente utilizando un conjunto de datos más grande y con parcelas de muestreo mayores. Por otro lado, el mayor potencial de este tipo de estudios puede encontrarse en el uso con información de mayor detalle espacial o espectral, o inclusive con datos de otra naturaleza (*e.g.*, datos LiDAR o datos SAR de mayor longitud de onda y corregidos por topografía), ya que los resultados muestran una débil relación entre el AGB y los patrones espectrales y espaciotemporales de las imágenes. Por lo tanto, creemos que el futuro del monitoreo de la degradación forestal puede requerir de análisis multifuente y de grandes conjuntos de datos.

Finalmente, la incorporación de información topográfica o del clima como variables predictivas en los modelos podrían ayudar a mejorar los resultados. Estas variables han demostrado ser relevantes para explicar la variación de la biomasa en determinados paisajes (Hernández-Stefanoni et al., 2020; Réjou-Méchain et al., 2019; Salinas-Melgoza et al., 2018).

Conclusiones

Los algoritmos de aprendizaje profundo han permitido mejorar las capacidades para monitorear la superficie terrestre; sin embargo, estos requieren de un conjunto de datos de entrenamiento grande para poder superar la precisión alcanzada con algoritmos previamente disponibles como los de aprendizaje automatizado. Este estudio mostró para modelar la AGB a partir de información de percepción remota, el algoritmo random forest obtuvo menor error que la U-Net 3D LF. Por lo tanto, consideramos que dicho resultado es explicado por el mayor sobreajuste del algoritmo U-Net 3D LF que el random forest, ante un conjunto de datos pequeño y muy variable. Además, los resultados demuestran la dificultad de monitorear la degradación forestal como un proceso de reducción de AGB entre dos tiempos determinados, a partir de información de percepción remota. Aunque los resultados no permitieron identificar sitios degradados, permitieron establecer la potencial reducción del AGB del bosque tropical seco de toda la región entre 2019 y 2022.

Finalmente, el desarrollo de este tipo de trabajos brindará herramientas para establecer métodos que permitan incrementar las capacidades para detectar la degradación forestal y monitorearla.

Capítulo 6. Discusión general

Principales hallazgos

El objetivo principal de la tesis fue evaluar si los algoritmos de aprendizaje profundo son capaces de mejorar la precisión obtenida para describir y monitorear la superficie terrestre. Debido a que estos algoritmos incorporan la dimensión espacial y temporal, aparte de la espectral, han permitido obtener resultados más precisos y ganado popularidad en los estudios de percepción remota (Isaienkov et al., 2020; Pelletier et al., 2019; Zhang et al., 2016; Zhang et al., 2019c). Este trabajo mostró que estos algoritmos permiten ayudar a mejorar las capacidades para algunas tareas (*e.g.*, clasificación de tipos de vegetación y usos del suelo o detección de la deforestación); sin embargo, no para todas (*e.g.*, detección de la degradación forestal como disminución de la biomasa arbórea).

Los capítulos 3 y 4 mostraron que este tipo de algoritmos son capaces de mejorar los resultados para describir distintos tipos de cubiertas y usos del suelo (*e.g.*, bosques maduros, secundario o degradados y plantaciones), así como para detectar cambios de algunas de estas clases (bosque vs no bosque). Por el contrario, el capítulo 5 mostró que para modelar la biomasa área a partir de imágenes satelitales, estos algoritmos no permitieron obtener un mejor resultado que opciones más sencillas (*i.e.*, random forest). Estos resultados corresponden con los reportes previos en los que se menciona que la ventaja de estos algoritmos depende fuertemente de la tarea a realizar y del tamaño de datos disponibles para entrenar a los algoritmos (Chollet and Allaire, 2018; LeCun et al., 2015).

Los algoritmos de aprendizaje profundo, en particular aquellos basados en CNN como la U-Net, parecen ser mejores para resolver tareas de clasificación, aun cuando se cuente con conjunto de datos relativamente pequeño. De igual manera, la complejidad del esquema de clasificación parece influir en las capacidades de discriminación (un mayor o menor número de clases; Solórzano et al., 2021). Estos algoritmos parecen requerir de un mayor número de datos para resolver tareas más complejas, o presentar una relación señal / ruido más pequeña, como fue el caso del capítulo 5. Por ejemplo, cuando hay una relación señal / pequeña es difícil diferenciar entre una señal de la cubierta o proceso de interés y el ruido; mientras que al ser ésta grande, resulta fácil diferenciar entre la señal de interés (verdaderos

positivos) y el ruido (falsos positivos). Por ello, es probable que otro tipo de métodos que requieran menos datos de entrenamiento y que suponen cierto tipo de respuesta entre las variables predictivas y de respuesta como los modelos lineales o algoritmos de aprendizaje automatizado permitan un mejor resultado en dichas circunstancias, como fue el caso de lo reportado en el capítulo 5.

En los capítulos 3 al 5 se comparó un algoritmo basado en la U-Net con alguno de los algoritmos de aprendizaje automatizado más populares para aplicaciones de percepción remota como random forest y SVM (Dabija et al., 2021; Sheykhmousa et al., 2020). En los primeros dos capítulos se observó que los algoritmos de aprendizaje profundo (U-Net) superaron a los algoritmos de ML, similar a lo reportado en trabajos similares (Adarme et al., 2020; de Bem et al., 2020; Du et al., 2020; Giang et al., 2020; Ienco et al., 2019; Ji et al., 2018; Kussul et al., 2017; Matosak et al., 2022; Robinson et al., 2019; Stoian et al., 2019). En ambos casos, el algoritmo DL superó al ML en *avgF1-score* 0.15 – 0.38 y *OA* 0.01 - 0.23. Por su parte, un patrón similar se detectó al comparar los *F1-score* por clase. Esto quiere decir que la integración de información en la dimensión espacial (clasificación), espectral y temporal (deforestación) trajo beneficios para el trabajo de clasificación. Además, los resultados de este trabajo coinciden con reportes previos de que los algoritmos DL permiten obtener resultados con menor ruido y, por lo tanto, prácticamente no requieren de un post procesamiento (Adarme et al., 2020; de Bem et al., 2020).

Por el contrario, en el capítulo 5 se encontró que dadas las condiciones del estudio, el algoritmo ML fue superior al basado en la U-Net para predecir la AGB del bosque y, por lo tanto, podría tener un mayor potencial de detectar cambios en el AGB del bosque para identificar sitios que sufrieron una degradación. Este resultado fue contrario al reportado en estudios similares para modelar el AGB a partir de imágenes de percepción remota (Buxbaum et al., 2022; Castro et al., 2020; Dong et al., 2020; Ghosh and Behera, 2021; Hamedianfar et al., 2022; Moradi et al., 2022; Talebiesfandarani and Shamsoddini, 2022; Zhang et al., 2022; Zhang et al., 2019b). Sin embargo, el único otro artículo que reportó un mejor desempeño con un algoritmo ML que uno CNN, justifica dicho resultado por el reducido tamaño de los datos de entrenamiento y la alta plasticidad de las CNN que provocó un sobreajuste mayor que el algoritmo ML (Tamiminia et al., 2021). Nuestra

conclusión sobre este capítulo fue la misma. Este tipo de resultados también resaltan la necesidad de comparar este tipo de algoritmos con otros métodos con un mismo conjunto de datos, ya que difícilmente se puede conocer el resultado *a priori*.

Estudios previos han reportado que las CNN con un número más grande de capas ocultas o filtros son capaces de detectar patrones más complejos con mejor precisión (Buduma, 2017; LeCun et al., 2015; Li et al., 2018). Sin embargo, cuando se tiene un conjunto de datos de entrenamiento limitados, como el caso de este estudio, el principal problema con arquitecturas más complejas radica en que tienden a sobreajustarse a los datos de entrenamiento. Quizás esta razón permite explicar por qué los capítulos 3 y 4 obtuvieron mejores resultados al utilizar un algoritmo DL, mientras que el capítulo 5, no. Aunque en todos los casos, los datos de entrenamiento eran limitados, en términos del número de píxeles usados, el menor conjunto de datos fue el del capítulo 5. Por ello, este fue el único capítulo donde se observó un buen ajuste a los datos de entrenamiento, pero bajo ajuste para los datos de validación y prueba.

Cabe resaltar que aunque los algoritmos de aprendizaje profundo permiten obtener mejores resultados que otros algoritmos de aprendizaje automatizado, en algunos casos, no permiten entender cuáles características de los datos son relevantes para desarrollar la tarea con éxito (Hamedianfar et al., 2022). Esto implica que este tipo de métodos serán de gran ayuda para cuando el enfoque del estudio sea clasificar o modelar alguna variable de interés con la mayor precisión posible. Sin embargo, si el propósito es entender cuáles factores son determinantes para que ocurra cierta actividad o transición, este tipo de algoritmos no serán de gran utilidad. En ese caso, es probable que resulte mejor utilizar modelos más sencillos, como regresiones logísticas o lineales, que obtengan menores valores de precisión o ajuste, pero que permitan identificar relaciones significativas (o al menos estadísticamente distintas a un efecto aleatorio). En su defecto, algunos algoritmos de aprendizaje automatizado (*e.g.*, random forest) también permiten detectar cuáles variables son más relevantes, como se presentó en el capítulo 5.

Por último, aunque este estudio no comparó las capacidades de clasificación de la U-Net (capítulo 3) con la U-Net 3D (capítulo 4), trabajos similares han reportado que el utilizar la dimensión temporal por separado ayuda a aumentar, aunque sea ligeramente, la

precisión obtenida en distintas tareas de clasificación o detección de cambios (Giannopoulos et al., 2022; Isaienkov et al., 2020; Ji et al., 2018; Masolele et al., 2021). Es probable que esta separación de las dimensiones permita explicar la obtención de un menor *avgF1-score* por un algoritmo ML, SVM, en el capítulo 4. Aunque este algoritmo incorporó la dimensión temporal dentro de la espectral, es probable que su menor capacidad de identificación de sitios deforestadas tenga que ver con que éste no separa la dimensión temporal de la espectral. A pesar de ello, cabe mencionar que varios trabajos han concatenado la dimensión temporal en la espectral y obtenido buenos resultados para resolver diferentes tareas con la U-Net 2D (Adarme et al., 2020; Bragagnolo et al., 2021a, b; Isaienkov et al., 2020; Torres et al., 2021).

Implicaciones para el monitoreo de las cubiertas arbóreas

El desarrollar métodos automatizados, como el uso de algoritmos de aprendizaje profundo, para monitorear los tipos de cubiertas, usos del suelo, su dinámica, y su contenido de biomasa permite obtener evaluaciones con un menor tiempo de procesamiento y recursos. Esto implica que pueden ser técnicas más costo eficientes que minimicen los pasos que requieran de interpretación visual o el registro de datos de campo. Además, este tipo de métodos se pueden incorporar en cadenas de trabajo donde se revisen los datos en un esquema activo para optimizar los recursos de interpretación visual (Adarme et al., 2020; Bragagnolo et al., 2021a). Por ejemplo, algunos de estos trabajos mencionan que las U-Net pueden detectar áreas omitidas en los datos de entrenamiento por error humano (Bragagnolo et al., 2021a), lo cual permite detectar dichas omisiones y posteriormente, corregir los datos de entrenamiento, volver a entrenar el algoritmo y aumentar la precisión para el monitoreo. Por otro lado, estos métodos pueden utilizar datos antiguos e incorporar datos nuevos, lo cual representa una opción atractiva para plantear programas de monitoreo basados en información preexistente y la incorporación de nuevos datos. Por ello, es probable que el futuro de los sistemas de monitoreo de la superficie terrestre se base en este tipo de algoritmos (Bragagnolo et al., 2021b). Sin embargo, como con cualquier método automatizado, su potencial dependerá de la cantidad y calidad de datos con la que se puede alimentar este tipo de métodos.

A pesar de que la diferenciación entre tipos de cubiertas arbóreas permite obtener mayor detalle de los procesos que ocurren en la superficie terrestre, esta discriminación ha sido muy difícil de lograr mediante sensores remotos y algoritmos previamente disponibles. El interés principal por diferenciar entre bosque maduro, secundario y plantaciones radicó en las diferencias que presentan en cuanto a características como biodiversidad, carbono almacenado y servicios ecosistémicos provistos, así como prácticas de aprovechamiento de los recursos naturales (Fernández-Montes de Oca et al., 2021; Gibson et al., 2011; Gutiérrez-Vélez et al., 2011; Putz and Redford, 2010; Tropek et al., 2014; Wright, 2005).

Por ello, en los capítulos 3 y 4, realizados en parte de la Selva Lacandona, se trató de separar las clases de cubierta arbórea, por lo menos en dos categorías. Esto permite describir los tipos de cubierta arbórea y sus transiciones con mayor detalle, lo cual puede ser crítico para determinar el efecto de algunas presiones internacionales o nacionales (demanda de ciertos cultivos o implementación de incentivos). En muchas ocasiones, el cambio de uso de suelo de bosque a plantaciones puede ir acompañado de diversas afectaciones tanto sociales como ambientales que incluyen el despojo de tierras, la explotación laboral, la alteración del ciclo hidrológico, la degradación de los suelos, la pérdida de hábitat, entre otras (Isaac-Márquez et al., 2016; Meijaard et al., 2018; Romero and Albuquerque, 2018). Por ello, estos capítulos no solo trataron de demostrar que los algoritmos DL permiten mejorar los resultados en términos de precisión, sino también mejorar el detalle de algunos esquemas de clasificación. Esta información será de vital importancia para estimar de manera indirecta potenciales afectaciones en los territorios estudiados.

Además, el diferenciar entre bosque maduro y secundario, permite conocer si dichos cambios implican una expansión de la frontera agrícola, o más bien a una sustitución de cultivos. Por otro lado, esta discriminación entre cubiertas arbóreas permite brindar mayor detalle sobre el uso del territorio, por ejemplo, si algunas áreas se encuentran bajo un uso intensivo, muchas veces asociados a una degradación de suelos (*e.g.*, plantaciones de palma), de otras zonas dominadas por árboles, pero bajo prácticas con un menor impacto ambiental (*e.g.*, áreas con ciclos de descanso agrícola; Isaac-Márquez et al., 2016; Ochoa-Gaona et al., 2007). Por último, esta información ayudará a estimaciones futuras sobre las

emisiones de carbono asociadas a la deforestación y la degradación forestal, así como del cálculo de tasas de pérdida de hábitat natural o cambios en la calidad de servicios ecosistémicos (Chi et al., 2016).

En el caso de la zona de estudio de Ayuquila, esta es una zona de acciones tempranas de REDD+, el programa de reducción de emisiones asociadas a la deforestación y degradación de los bosques. Por ello, en esta zona el estudio se enfocó a tratar de cuantificar la degradación forestal como una reducción en la biomasa del bosque. Estudios previos han reportado que la probabilidad de que un bosque se degrade está asociada con un índice de pobreza y el tamaño poblacional de las localidades cercanas a los bosques (Borrego and Skutsch, 2019; Morales-Barquero et al., 2015). Por otro lado, el porcentaje de bosque degradado a nivel municipal en la región está relacionado de manera negativa con la distancia a carreteras y el porcentaje de tierras parceladas (Jiménez-Rodríguez et al., 2022). Sin embargo, estos estudios habían trabajado con la detección del bosque degradado como un problema de clasificación (*i.e.*, bosque degradado vs bosque conservado) y no como un proceso, es decir como una disminución de la biomasa. Aunque este enfoque resulta mucho más complicado, tiene la ventaja de que considera la variación espacial natural del AGB del bosque, en respuesta a algunas variables biofísicas como la pendiente, el tipo de suelo, entre otras. Aunque el capítulo 5 trató de abordar esto, los resultados no fueron tan prometedores como en el caso de los dos capítulos anteriores. A pesar de ello, se logró obtener un estimado del carbono almacenado en el bosque tropical caducifolio de la zona. Esta predicción no fue muy distinta de lo que se hubiera obtenido al multiplicar el área de BTC por el promedio de AGB contenido en una ha. Sin embargo, nuestros resultados permitieron obtener mapas de AGB con una variación espacial que puede ser más realista. En este sentido, consideramos que futuros métodos deberían estar encaminados en tratar de detectar la degradación forestal como un proceso, a pesar de que eso suponga un reto mucho mayor.

Adicionalmente, el contar con evaluaciones de clasificaciones de tipos de vegetación y usos del suelo, deforestación y degradación forestal de alta resolución permiten caracterizar a la superficie con un mayor detalle y precisión espacial (*i.e.*, 10 m). Tanto en el caso de los tipos de cubiertas como en el de deforestación, el obtener estimaciones con mayor

resolución, permite detectar áreas más pequeñas que permiten obtener estimaciones más realistas. En el caso particular de la deforestación, esto resulta crítico para detectar áreas pequeñas, ya que esta diferencia de resolución puede resultar en un aumento significativo de las estimaciones de áreas, sobre todo de la deforestación (Reiche et al., 2021). En el caso del capítulo 5, el realizar estimaciones con las imágenes Sentinel-1 y Sentinel-2 permiten obtener estimaciones de biomasa también con mayor resolución. Aunque en este caso la ventaja no es tan evidente como en el caso de los capítulos 3 y 4, debido a que los estimados de biomasa normalmente se estandarizan a una ha. A pesar de ello, la ventaja de evaluar la biomasa con imágenes de mayor resolución es que permite captar variaciones espaciales con mayor detalle espacial.

Evolución de los capítulos de la tesis

Los últimos tres capítulos de la tesis analizaron la superficie terrestre y su dinámica desde tres diferentes perspectivas. La primera constó de la clasificación de la cubierta terrestre, la segunda, de la detección de áreas de pérdida de cubierta arbórea y la última, de la estimación en el contenido de carbono en la cubierta forestal y su cambio. Por un lado, estos estudios evaluaron el uso de algoritmos de aprendizaje profundo en diferentes escalas temporales, ya sea para una descripción de un solo tiempo (clasificación de la cubierta) o el cambio entre dos periodos (ya sea de cubiertas o de contenido de carbono). Por otro lado, estos enfoques también estudiaron a la superficie terrestre en dos niveles de estudio, ya sea a nivel de tipo de cubierta (tipo de cubierta y cambio de cubierta) o de una característica específica de un tipo de cubierta forestal (contenido de carbono). De esta manera, cada uno de estos estudios evaluó estas herramientas para estudiar distintos procesos, escalas temporales y niveles de estudio.

Visto desde otro panorama, cada uno de estos trabajos constaron de una tarea distinta, en los primeros dos casos (capítulos 3 y 4) se trató de una tarea de clasificación, ya que se deseaba obtener un mapa temático de la ubicación y extensión espacial, ya fuera de un tipo de cubierta o de determinados cambios. Por su parte, el último capítulo constó de una regresión, ya que a partir de las imágenes utilizadas se obtuvo un mapa de la cantidad de carbono secuestrado en un bosque. Por ello, este conjunto de estudios mostró la flexibilidad

y adaptabilidad de los algoritmos de aprendizaje profundo para realizar distintos tipos de tareas.

Además, el avance de los capítulos constó de una evolución en la complejidad de dicho algoritmo. Por ejemplo, el uso de la U-Net para crear una clasificación de tipos de vegetación y usos del suelo constó de la utilización de la dimensión espectral y espacial, mientras que en los últimos dos capítulos se ampliaron las capacidades de dicho algoritmo para utilizar también la dimensión temporal, la U-Net 3D. En el primer caso, la incorporación de la dimensión espacial resultó importante sobre todo para diferenciar entre tipos de cubiertas arbóreas. En el segundo caso, la incorporación de la dimensión temporal fue vital para detectar los cambios de pérdida de cubierta arbórea entre dos fechas. En el último capítulo, aunque también se incorporó la dimensión temporal, la arquitectura de la U-Net se modificó para aprovechar al máximo las capacidades de las imágenes utilizadas: MS y SAR. Gracias a que las imágenes SAR no son sensibles (o casi no) a la presencia de nubes, potencialmente podrían tener muchas más observaciones sin artefactos en comparación con MS. Por ello, en el último capítulo se decidió aumentar el número de observaciones en la dimensión temporal de las imágenes SAR. Sin embargo, esto supuso un cambio en la arquitectura de la U-Net para poder trabajar con insumos con distinto número de observaciones en la dimensión temporal. Tras estas modificaciones propusimos una U-Net con fusión tardía.

Finalmente, el orden en el que se realizaron los capítulos también representó un incremento en la complejidad de las imágenes utilizadas como entrada a los algoritmos. En el capítulo 3 se utilizó una imagen de un solo tiempo, lo cual minimizó la presencia de artefactos en la imagen, pero aumentó el área para la cual no se pudo obtener una clasificación por la presencia de nubes. Por su parte, los capítulos 4 y 5 utilizaron mosaicos multitemporales que permitieron describir un área completa de interés, pero a costa de aumentar la presencia de artefactos en las imágenes. Esta evolución pone en evidencia la disyuntiva que suele existir entre obtener imágenes sin nubes para toda un área de estudio y la frecuencia de artefactos en ellas. Finalmente, tanto en los capítulos 4 y 5, varios de los errores detectados estuvieron asociados con la presencia de estos artefactos.

Comparación MS + SAR, MS y SAR

Además de poner a prueba distintos algoritmos, este trabajo contó con un segundo objetivo, el cual constó en evaluar estas herramientas con imágenes multiespectrales y radar, y evaluar las posibles ventajas de utilizarlas en conjunto. Cada tipo de información o imagen generada mediante un sensor remoto tiene sus propias ventajas y desventajas. Por ejemplo, las imágenes multiespectrales permiten obtener información detallada de los tipos de cubiertas de una región; sin embargo, también son sensibles a la presencia de nubes y sombras, los cuales crean artefactos que pueden hacer más difícil trabajar con ellas. Las imágenes radar prácticamente no son sensibles a nubes, por lo tanto, permiten obtener información sin artefactos meteorológicos (aunque sí presentan otro tipo de artefactos como los topográficos), pero no tienen la riqueza ni el detalle de las imágenes multiespectrales (Flores-Anderson et al., 2019; Heckel et al., 2020; Hirschmugl et al., 2020; Zhao et al., 2016; Zhu et al., 2017a). A pesar de ello, estudios previos han reportado mejores resultados con MS + SAR para realizar clasificaciones de tipos de vegetación y usos del suelo (Heckel et al., 2020; Hirschmugl et al., 2020; Khan et al., 2020; Tavares et al., 2019; Van Tricht et al., 2018) o detección de deforestación (Joshi et al., 2016; Reiche et al., 2018).

En este trabajo se utilizaron imágenes Sentinel-1 y Sentinel-2, las cuales son imágenes de distribución libre generadas por la agencia espacial europea. Sentinel-1 es un sensor en la banda C que no tiene tanta penetración como las bandas L y P (Deus, 2016; Flores-Anderson et al., 2019; Joshi et al., 2016; Pompeu Pavanelli et al., 2018), las cuales usualmente se mencionan como las ideales para estudiar sistemas arbóreos. Por lo tanto, la banda C interactúa más con ramas y hojas, que con los troncos, como lo hace la banda P o L (Flores-Anderson et al., 2019). Quizás esta es la principal razón para explicar por qué no resultaron de mucha utilidad para los capítulos 4 y 5. Sin embargo, la información proveída por Sentinel-1 ayudó a diferenciar distintos tipos de cubierta arbórea (capítulo 3) y aunque las imágenes banda L o P podrían ser el insumo óptimo, estas no son de distribución libre o sólo se pueden encontrar mosaicos anuales (*e.g.*, ALOS PALSAR, banda L). Quizás la razón por la que las imágenes Sentinel-1 fueron útiles para el trabajo del capítulo 3 fue su sensibilidad a arreglos geométricos y humedad del suelo.

Por su parte, las imágenes Sentinel-2 son producidas por un sensor multiespectral con 13 bandas. A pesar de su alta resolución espectral, sólo cuatro bandas presentan una resolución de 10 m. Debido a que estudios previos sugieren mejores resultados con mayor resolución espacial, se utilizaron solo esas. Sin embargo, en un futuro sería interesante explorar el potencial de usar otras bandas de menor resolución espacial como las bandas de SWIR y RE. En este sentido, Sentinel-2 representa uno de los tipos de imágenes gratuitos de mayor resolución espacial. Sin embargo, mejores resultados quizás se podrían obtener con imágenes no gratuitas de mayor resolución espacial (*e.g.*, Clark and McKechnie, 2020; Flood et al., 2019; Zhang et al., 2018). Este aspecto también representa un compromiso entre precisión de los algoritmos entrenados y su costo, relacionado con la resolución espacial de las imágenes.

A partir de nuestros resultados y estudios previos, la ventaja de utilizar MS + SAR parece variar de la aplicación, el área de estudio o el sistema de estudio, el sistema de clasificación, la resolución de las imágenes utilizadas, la correspondencia temporal entre las bandas SAR y MS y el método utilizado para analizar los datos (Adarme et al., 2020; Joshi et al., 2016; Reiche et al., 2018). Por ello, aunque en principio se podría suponer que al usar las imágenes MS + SAR se obtendrían mejores resultados, creemos que difícilmente se podrá dar una conclusión *a priori*, por lo tanto, recomendamos probar con ambos tipos de imágenes. Sin embargo, sí se quiere simplificar el conjunto de variables predictivas, las bandas MS deberían brindar mejores resultados que solo utilizar las SAR.

Por ejemplo, algunos estudios previos sugieren que al modelar características del bosque mediante percepción remota, se pueden obtener resultados más prometedores al trabajar en sistemas sencillos, como plantaciones o cultivos, en comparación con sistemas complejos como bosques naturales, debido a su relativa simplicidad y homogeneidad (Couteron et al., 2005; Gallardo-Cruz et al., 2012; Proisy et al., 2007; Solórzano et al., 2017). Este factor está relacionado con la razón de señal / ruido, ya que si se tiene un sistema muy sencillo, cualquier diferencia o alteración del mismo sistema se podrá relacionar con la clase, característica o proceso de interés. Por el contrario, si se estudia un sistema complejo, al tener muchas variables no controladas (*e.g.*, diversidad de especies, diversidad de alturas, formas de copas, etc.), la relación señal / ruido será menor y cualquier

algoritmo tenderá a poder diferenciar en menor medida la clase, característica o proceso de interés (Gerard and North; Proisy et al., 2011). A pesar de ellos, otros estudios ponen en duda tal aseveración (Trisasongko and Paull, 2020).

Por otro lado, tanto en nuestro estudio como en anteriores, el trabajar con sistemas de clasificación más sencillos parece facilitar la obtención de resultados con alta precisión; mientras que sistemas muy detallados y con un mayor número de clases, parecen brindar resultados con menor precisión (Buxbaum et al., 2022; Clark and McKechnie, 2020; Dong et al., 2020; Wagner et al., 2020a). Esto está relacionado con la dificultad de la tarea a realizar, así como, la potencial confusión entre clases muy detalladas.

La resolución de las imágenes ha sido frecuentemente mencionada como un aspecto esencial para determinar el desempeño al utilizar algoritmos, como las CNN, que consideran la dimensión espacial (Du et al., 2020; Giang et al., 2020; Hamdi et al., 2019; Huang et al., 2018; Neves et al., 2020; Wagner et al., 2019; Zhang et al., 2018). Esto se debe a que la información espacial de este tipo de información permite diferenciar tamaños de copas de los árboles o arreglos regulares típicos de plantaciones (Clark and McKechnie, 2020; Datcu et al., 2020; Flood et al., 2019; Hooser et al., 2020; Wagner et al., 2020a; Wagner et al., 2020b). Por ello, es muy probable que el utilizar imágenes de muy alta resolución permita obtener mejores resultados; sin embargo, éstas no son de distribución libre. En este estudio, se eligió utilizar imágenes de menor resolución espacial, pero que fueran de distribución libre. Aunque esta estrategia probablemente permita obtener una menor exactitud total y *F1-score* promedio que al utilizar imágenes de mayor resolución espacial, es un método aplicable para cualquier parte del mundo y de muy bajo costo. Por ello, esta técnica podría ser una de fácil aplicación para realizar tareas de monitoreo de la superficie terrestre.

Otros estudios han utilizado imágenes gratuitas de alta a mediana resolución espacial, como Landsat-8 o Sentinel-2, para realizar clasificaciones o detección de deforestación con resultados similares a los nuestros o mejores (Adarme et al., 2020; Bragagnolo et al., 2021a; de Bem et al., 2020; Giannopoulos et al., 2022; Isaienkov et al., 2020; Maretto et al., 2021; Matosak et al., 2022; Pelletier et al., 2019). Interesantemente, los pocos estudios que han comparado el desempeño de este tipo de algoritmos con imágenes Landsat-8 y

Sentinel-2, reportan mejores resultados con las de mayor resolución (Matosak et al., 2022; Torres et al., 2021). Esto confirma que la resolución espacial de las imágenes, aun utilizando imágenes de mediana resolución, es un factor clave para determinar la precisión de un método basado en CNN para detectar la deforestación.

La ventaja de usar MS + SAR dependió también de la correspondencia temporal entre los dos tipos de imágenes. Por ejemplo, en el capítulo 3 su ventaja fue evidente probablemente porque se usaron imágenes MS de un solo tiempo y un compuesto multitemporal de SAR de una ventana muy corta de tiempo (1 mes). Esto permitió maximizar la correspondencia temporal entre los dos tipos de bandas. Además, es probable que la información SAR haya sido de vital importancia para distinguir entre tipos de cubierta arbórea donde su reflectancia es bastante similar y sobre todo para la detección de plantaciones (Danylo et al., 2021; Monsalve-Tellez et al., 2022; Solórzano et al., 2021). Sin embargo, en este caso, al ser la U-Net MS + SAR la que obtuvo la mayor precisión y *F1-score* promedio, no se pudo obtener una cuantificación sobre toda el área de estudio, debido a la presencia de nubes en las bandas MS.

Por su parte, en el capítulo 4 se usaron compuestos multitemporales MS para evitar la presencia de nubes. Sin embargo, esto supuso la inclusión de ciertos artefactos en las imágenes MS utilizadas. Esto pudo haber sido la razón por la cual la inclusión de las bandas SAR no aumentaron el *F1-score* promedio. Además, es probable que dicha situación se viera exacerbada debido a que la mayoría de pérdidas en la clase bosque secundario / plantaciones se debiera a la primera clase, donde la información SAR no ayuda mucho.

Por otro lado, en el capítulo 5 ya no se evaluó la ventaja de utilizar MS, SAR o MS + SAR. Debido a la arquitectura usada, era probable que si las bandas SAR no resultaran ventajosas, simplemente esta información la desecharía el algoritmo (debido a la fusión tardía). A pesar de ello, estudios previos sugerían que los datos más informativos serían MS, seguidos de los SAR para modelar la AGB (Zhang et al., 2022). Por otro lado, esta arquitectura evitó limitar al número de observaciones a través del tiempo disponibles para las imágenes SAR, lo cual suponíamos brindaría más información para modelar la AGB. Pensando en que este estudio se realizó en un bosque tropical caducifolio, esto permitiría

contar con más observaciones en la época de lluvias y secas; mientras que para MS, estuvo limitado a la obtención de un compuesto sin nubes.

Finalmente, cabe mencionar, que en este trabajo se evaluaron únicamente algoritmos de aprendizaje profundo y algoritmos de aprendizaje automatizado; sin embargo, en la literatura se pueden encontrar algoritmos basados en series de tiempo que suelen obtener mejores resultados al combinar información MS + SAR. Esto probablemente se deba a que este tipo de algoritmos maximizan la correspondencia temporal entre las bandas MS + SAR. Aunque estos algoritmos también representan una alternativa atractiva para realizar cualquiera de las tareas realizadas en este estudio, en éste se optó por trabajar con algoritmos que permitieran detectar patrones espaciales y temporales, en lugar de solo estos últimos (como en las series de tiempo).

Comparaciones con estudios similares

Al comparar el *F1-score* y *OA* de la clasificación obtenida en el capítulo 3 con otras realizadas utilizando la U-Net, se observó que este trabajo se encontró entre los valores más bajos (Du et al., 2020; Flood et al., 2019; Giang et al., 2020; Hamdi et al., 2019; Huang et al., 2018; Isaienkov et al., 2020; Neves et al., 2020; Robinson et al., 2019; Stoian et al., 2019; Ulmas and Liiv, 2020; Wagner et al., 2020a; Wagner et al., 2020b; Wagner et al., 2019; Yi et al., 2019; Zhang et al., 2018). Sin embargo, el *avgF1-score* y *OA* del capítulo 4 se encontró dentro de los valores reportados para detectar deforestación utilizando alguna arquitectura basada en la U-Net (Adarme et al., 2020; Bragagnolo et al., 2021a, b; Isaienkov et al., 2020; Maretto et al., 2021; Matosak et al., 2022; Torres et al., 2021). En el caso del capítulo 5, el error de la modelación de AGB mediante el uso de una arquitectura basada en la U-Net, fue mayor al reportado en trabajos previos utilizando arquitecturas basadas en CNN (Buxbaum et al., 2022; Castro et al., 2020; Dong et al., 2020; Ghosh and Behera, 2021; Hamedianfar et al., 2022; Huy et al., 2022; Moradi et al., 2022; Pyo et al., 2019; Talebiesfandarani and Shamsoddini, 2022; Tamiminia et al., 2021; Zhang et al., 2022; Zhang et al., 2019b). Sin embargo, hay que destacar que muchos de los trabajos con los que se compararon el capítulo 3 y 5, se podrían considerar con sistemas de clasificación más sencillos o sistemas más simples en comparación con los abordados en este trabajo.

Aspectos metodológicos

En este estudio se utilizaron dos métodos que permitieron mejorar las estimaciones de área de las clases de interés. La primera fue seguir las recomendaciones de Olofsson et al. (2014; Cochran, 1977; Olofsson et al., 2014) para verificar mapas con clases muy dispares para obtener estimados corregidos del área de cada clase, utilizando los resultados de la matriz de confusión. Este método permite obtener estimados corregidos de área por clase y con un error asociado. El segundo fue seguir las recomendaciones de Olofsson et al. (2020) al utilizar un estrato de buffer para potencialmente contener los errores de omisión de alguna clase rara. Esto provoca que el área por la que se multiplica un error de omisión de una clase rara en una clase dominante sea menor, lo cual implica que el intervalo de confianza de la clase rara se reduce para evitar que sea estadísticamente igual a 0. Un aspecto interesante es que aunque este tipo de métodos son comúnmente utilizados para describir la precisión de un mapa, en el contexto del uso de algoritmos de aprendizaje profundo raramente son utilizados (pero ver Clark and McKechnie, 2020; Du et al., 2020).

Por último, cabe mencionar que sería interesante combinar métodos para monitorear la deforestación y degradación forestal. Por ejemplo, los métodos basados en series de tiempo permiten detectar un evento de este tipo dentro de una ventana temporal corta, mientras que los algoritmos de aprendizaje profundo utilizados en este trabajo tienen un retraso más grande. Por el otro lado, los algoritmos de aprendizaje profundo permiten incorporar la información espacial de los píxeles para detectar tanto la degradación forestal como la deforestación. Por ello, un método que conjuntara las ventajas de ambos tipos de métodos resultaría prometedor; sin embargo, no resulta muy claro cómo se podría hacer dicha fusión. Trabajos futuros deberían estar encaminados en esta dirección.

Lecciones aprendidas y recomendaciones para estudios futuros

La labor de interpretación visual realizada en este estudio tomó un gran esfuerzo de tiempo. Por ejemplo, para generar los datos del capítulo tres se invirtió aproximadamente un mes, en la delineación de los polígonos y revisión de los mismos. Debido a que la generación de datos de este tipo requiere de mucho tiempo, el uso de datos previamente disponibles consta de una alternativa más rápida. Por ejemplo, en el caso de estudios realizados en

México, se podría utilizar la información generada por instituciones como el INEGI o CONAFOR (*e.g.*, series de tipos de vegetación y usos del suelo o áreas deforestadas), hacer los ajustes necesarios y utilizarlos como datos de entrenamiento. Frecuentemente, este tipo de información contiene errores de clasificación, por lo que algunos estudios han propuesto utilizar algunos análisis de reducción de dimensiones, como los mapas autoorganizativos o algún método de ordenación, para analizar la calidad de los datos de entrenamiento y detectar muestras erróneas o con mucho ruido (Bragagnolo et al., 2021b; Simoes et al., 2021). Este tipo de métodos resultan muy atractivos para automatizar la generación de datos de entrenamiento, basándose en esfuerzos cartográficos previos e incluyendo una fase de curación de los datos.

Por otro lado, hay que reconocer que la delineación de polígonos temáticos para generar los datos de entrenamiento, validación y prueba de los capítulos tres y cuatro no se encuentran libres de errores. Aunque dichos datos se revisaron varias veces, hubo situaciones donde la clase de un determinado polígono no fue totalmente clara. Además, debido a que se utilizaron los mismos insumos para generar la clasificación y los datos de entrenamiento, es probable que el efecto de algunos artefactos haya tenido impacto en la delimitación y clasificación manual de polígonos. Aunque habría sido deseable usar otros insumos de mayor resolución espacial para generar los datos de entrenamiento (*e.g.*, Google Earth), debido a que no se encontraron imágenes para las fechas y área de estudio, resultó imposible de realizar. Por otro lado, estudios previos han sugerido que varios interpretes revisen los datos de entrenamiento, validación y prueba para minimizar el error en ellos y contar con una medida del error en los datos. Esto permite contar con un valor de referencia de cuál podría ser la exactitud máxima a alcanzar por algún algoritmo de aprendizaje automatizado o profundo.

Por su parte, para la generación de los datos de entrenamiento, validación y prueba del capítulo cinco, se pueden hacer las siguientes recomendaciones: 1) registrar información sobre la identidad taxonómica de los individuos, 2) utilizar parcelas de muestreo de por lo menos 1000 m², 3) registrar la posición de las parcelas con un GPS de alta precisión. El contar con información taxonómica de los individuos muestreados permite realizar los cálculos de la biomasa considerando la variación de la densidad de madera entre especies,

lo cual se traduce en un estimado más preciso. El utilizar parcelas de muestreo de por lo menos 1000 m² ayuda a reducir la sensibilidad a condiciones locales, obtener estimados con menor error de AGB y reducir el efecto del error de la posición del GPS. Por último, el utilizar un GPS de alta precisión permite reducir el desfase entre la posición de la parcela de muestreo y su posición en la imagen satelital.

A partir de esta tesis fue difícil determinar un tamaño mínimo de datos necesario para entrenar de manera exitosa algunos algoritmos de aprendizaje profundo. Sin embargo, los resultados sugieren que un número mínimo de datos podría encontrarse en el orden de las centenas. Además, de acuerdo con este trabajo, la comparación de los resultados obtenidos entre un algoritmo de aprendizaje profundo y uno automatizado brinda herramientas para esclarecer este aspecto. Por ejemplo, en esta tesis, es muy probable que en el capítulo 5, el mejor desempeño del algoritmo de aprendizaje automatizado, en comparación con el de aprendizaje profundo, se haya debido al número reducido de datos. Estudios previos coinciden con este tipo de conclusiones.

Estudios futuros deberían explorar el uso de imágenes SAR de una longitud de onda más larga (*i.e.*, como ALOS-PALSAR 2, banda L) para el estudio de los bosques tropicales, ya sea tanto en un problema de clasificación como de regresión (*i.e.*, para modelar la biomasa). Las imágenes de SAR de onda más larga permiten una mayor penetración en la cobertura forestal, por lo que usualmente se reportan mejores resultados usando estas imágenes. Este estudio prefirió utilizar imágenes Sentinel-1, banda C, debido a que estas imágenes tienen un tiempo de revisita menor al de ALOS-PALSAR 2, por lo que se pueden obtener mosaicos en ventanas de tiempo más cortas. Además, se esperaba que la diferencia entre usar ALOS-PALSAR 2 y Sentinel-1 fuera más relevante al modelar la AGB de un bosque. Sin embargo, debido a que dicho ejercicio se realizó en un bosque tropical seco, se hipotetizó que la diferencia sería mínima debido a que: 1) el estudio se realizó en un bosque de baja biomasa, por lo que probablemente los dos tipos de sensores podrían ser capaces de penetrar el dosel y registrar información de la estructura del bosque y 2) se utilizaron compuestos de la época de secas, donde los árboles no presentan follaje y por lo tanto, podrían detectar la estructura del bosque sin interferencia del dosel. Aunque los resultados no se muestran en la tesis, se hizo una prueba con imágenes ALOS-PALSAR 2 en el

capítulo 5, sin embargo, no se obtuvieron mejores resultados con estas imágenes. A pesar de ello, una en un futuro sería interesante realizar una exploración más formal y detallada.

Conclusiones

Este trabajo demostró que los algoritmos de aprendizaje profundo basados en CNN son una alternativa atractiva para mejorar las capacidades para clasificar y monitorear la cubierta terrestre; sin embargo, su potencial estará limitado por la complejidad de la tarea a resolver, así como del tamaño de los datos de entrenamiento. En particular, este trabajo demostró que para crear una clasificación de tipos de vegetación y usos del suelo, así como para la detección de la deforestación, las arquitecturas basadas en CNN fueron capaces de obtener resultados más precisos que su contraparte ML. Sin embargo, para la detección de la degradación forestal como una disminución de AGB, se observó lo contrario.

Por su parte, aunque en general el uso de las bandas MS y SAR fue beneficioso, no en todas las evaluaciones se obtuvieron mejores resultados utilizando ambas bandas en conjunto. Por ello, consideramos que el beneficio de utilizar ambos sensores depende de la aplicación y la tarea a realizar. A pesar de ello, por sí solas, las bandas MS tendieron a contribuir con información más útil que la SAR.

El desarrollo de este tipo de trabajos brindará herramientas para establecer métodos que incorporen el uso de algoritmos de aprendizaje profundo en su flujo de trabajo y que permitan aumentar las capacidades de descripción y monitoreo de la superficie terrestre. Esto permitirá establecer líneas base más detalladas y precisas, así como determinar tendencias temporales más certeras sobre los procesos de deforestación y degradación forestal. Esto brindará información para determinar el avance en las metas nacionalmente determinadas, así como de algunos objetivos de desarrollo sostenible.

Referencias

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zheng, X., 2016. TensorFlow: A System for Large-Scale Machine Learning, Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation. OSDI, Savannah, EUA, pp. 265–283.
- Adarme, M.O., Feitosa, R.Q., Happ, P.N., Almeida, C.A.D., Gomes, A.R., 2020. Evaluation of deep learning techniques for deforestation detection in the brazilian amazon and cerrado biomes from remote sensing imagery. *Remote Sensing* 12, 910. <https://doi.org/10.3390/rs12060910>
- Agapiou, A., Alexakis, D.D., Sarris, A., Hadjimitsis, D.G., 2014. Evaluating the potentials of sentinel-2 for archaeological perspective. *Remote Sensing* 6, 2176-2194. <https://doi.org/10.3390/rs6032176>
- Ali, I., Greifeneder, F., Stamenkovic, J., Neumann, M., Notarnicola, C., 2015. Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data. *Remote Sensing* 7, 16398-16421. <https://doi.org/10.3390/rs71215841>
- Allaire, J., Chollet, F., 2018. keras: R Interface to 'Keras'.
- Allaire, J.J., 2018. tfruns: Training Run Tools for 'TensorFlow'.
- Allaire, J.J., Chollet, F., 2020. keras: R Interface to 'Keras'.
- Almeida-Filho R, Shimabukuro YE, Rosenqvist A, Sánchez GA., 2009. Using dualpolarized ALOS PALSAR data for detecting new fronts of deforestation in the Brazilian Amazônia. *International Journal of Remote Sensing* 30, 3735-43. <https://doi.org/10.1080/01431160902777175>
- Alom, M.Z., Taha, T.M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M.S., Hasan, M., Van Essen, B.C., Awwal, A.A.S., Asari, V.K., 2019. A state-of-the-art survey on deep learning theory and architectures. *Electronics (Switzerland)* 8, 292. <https://doi.org/10.3390/electronics8030292>
- Aplin, P., 2004. Remote sensing: land cover. *Progress in Physical Geography* 28, 283-293. <https://doi.org/10.1191/0309133304pp413pr>
- Arévalo, P., Olofsson, P., Woodcock, C.E., 2020. Continuous monitoring of land change activities and post-disturbance dynamics from Landsat time series: A test methodology for REDD+ reporting. *Remote Sensing of Environment* 238. <https://doi.org/10.1016/j.rse.2019.01.013>
- Awange, J.L., Kyalo Kiema, J.B., 2013. Fundamentals of Remote Sensing, in: Awange, J.L., Kyalo Kiema, J.B. (Eds.), *Environmental Geoinformatics: Monitoring and Management*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 111-118.
- Badjana, H.M., Olofsson, P., Woodcock, C.E., Helmschrot, J., Wala, K., Akpagana, K., 2017. Mapping and estimating land change between 2001 and 2013 in a heterogeneous landscape in West Africa: Loss of forestlands and capacity building opportunities. *International Journal of Applied Earth Observation and Geoinformation* 63, 15-23. <https://doi.org/10.1016/j.jag.2017.07.006>

- Baek, J., Kim, J.W., Lim, G.J., Lee, D.-C., 2011. Electromagnetic land surface classification through integration of optical and radar remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing* 49, 1214-1222. <https://doi.org/10.1109/TGRS.2010.2096513>
- Baer, H.A., 2011. Global Capitalism and Climate Change, *Handbook on International Political Economy*. WORLD SCIENTIFIC, pp. 395-414.
- Bai, Z.G., Dent, D.L., Olsson, L., Schaepman, M.E., 2008. Proxy global assessment of land degradation. *Soil Use and Management* 24, 223-234. <https://doi.org/10.1111/j.1475-2743.2008.00169.x>
- Ball, J.E., Anderson, D.T., Chan, C.S., 2017. Comprehensive survey of deep learning in remote sensing: theories, tools, and challenges for the community. *Journal of Applied Remote Sensing* 11, 042609. <https://doi.org/10.1117/1.jrs.11.042609>
- Banskota, A., Kayastha, N., Falkowski, M.J., Wulder, M.A., Froese, R.E., White, J.C., 2014. Forest Monitoring Using Landsat Time Series Data: A Review. *Canadian Journal of Remote Sensing* 40, 362-384. <https://doi.org/10.1080/07038992.2014.987376>
- Berger, M., Moreno, J., Johannessen, J.A., Levelt, P.F., Hanssen, R.F., 2012. ESA's sentinel missions in support of Earth system science. *Remote Sensing of Environment* 120, 84-90. <https://doi.org/10.1016/j.rse.2011.07.023>
- Berget, C., Verschoor, G., García-Frapolli, E., Mondragón-Vázquez, E., Bongers, F., 2021. Landscapes on the move: Land-use change history in a mexican agroforest frontier. *Land* 10. <https://doi.org/10.3390/land10101066>
- Berninger, A., Lohberger, S., Stängel, M., Siegert, F., 2018. SAR-based estimation of above-ground biomass and its changes in tropical forests of Kalimantan using L- and C-band. *Remote Sensing* 10. <https://doi.org/10.3390/rs10060831>
- Bhogapurapu, N., Dey, S., Homayouni, S., Bhattacharya, A., Rao, Y.S., 2022. Field-scale soil moisture estimation using sentinel-1 GRD SAR data. *Advances in Space Research* 70, 3845-3858. <https://doi.org/https://doi.org/10.1016/j.asr.2022.03.019>
- Borrego, A., Skutsch, M., 2019. How Socio-Economic Differences Between Farmers Affect Forest Degradation in Western Mexico. *Forests* 10, 893. <https://doi.org/10.3390/f10100893>
- Boser, B., Guyon, I., Vapnik, V., 1992. A training algorithm for optimal margin classifiers, *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, Pittsburgh, PA, USA, pp. 144–152.
- Boyd, D.S., Danson, F.M., 2005. Satellite remote sensing of forest resources: three decades of research development. *Progress in Physical Geography* 29, 1-26. <https://doi.org/10.1191/0309133305pp432ra>
- Bragagnolo, L., da Silva, R.V., Grzybowski, J.M.V., 2021a. Amazon forest cover change mapping based on semantic segmentation by U-Nets. *Ecological Informatics* 62, 101279. <https://doi.org/10.1016/j.ecoinf.2021.101279>
- Bragagnolo, L., da Silva, R.V., Grzybowski, J.M.V., 2021b. Towards the automatic monitoring of deforestation in Brazilian rainforest. *Ecological Informatics* 62, 101279. <https://doi.org/10.1016/j.ecoinf.2021.101279>
- Breiman, L., 2001. Random Forests. *Machine Learning* 45, 5-32. <https://doi.org/10.1023/A:1010933404324>

- Brown, W.M., 1967. Synthetic Aperture Radar. *IEEE Transactions on Aerospace and Electronic Systems* AES-3, 217-229. <https://doi.org/10.1109/TAES.1967.5408745>
- Buduma, N., 2017. *Fundamentals of Deep Learning*. O'Reilly, Boston, USA.
- Bullock, E.L., Nolte, C., Reboredo Segovia, A.L., Woodcock, C.E., 2020. Ongoing forest disturbance in Guatemala's protected areas. *Remote Sensing in Ecology and Conservation* 6, 141-152. <https://doi.org/https://doi.org/10.1002/rse2.130>
- Bullock, E.L., Woodcock, C.E., Olofsson, P., 2018. Remote Sensing of Environment Monitoring tropical forest degradation using spectral unmixing and Landsat time series analysis. *Remote Sensing of Environment*, 0-1. <https://doi.org/10.1016/j.rse.2018.11.011>
- Burivalova, Z., Allnutt, T.F., Rademacher, D., Schlemm, A., Wilcove, D.S., Butler, R.A., 2019. What works in tropical forest conservation, and what does not: Effectiveness of four strategies in terms of environmental, social, and economic outcomes. *Conservation Science and Practice* 1, e28. <https://doi.org/https://doi.org/10.1111/csp2.28>
- Buxbaum, N., Lieth, J.H., Earles, M., 2022. Non-destructive Plant Biomass Monitoring With High Spatio-Temporal Resolution via Proximal RGB-D Imagery and End-to-End Deep Learning. *Frontiers in Plant Science* 13. <https://doi.org/10.3389/fpls.2022.758818>
- Cano Castellanos, I.J., 2014. Entre sueños agrarios y discurso ecologista. Las encrucijadas contemporáneas de la colonización de la Selva Lacandona (Chiapas, México). *EntreDiversidades* 1, 101-143. <https://doi.org/10.31644/ed.2.2014.a04>
- Cano Castellanos, I.J., 2018. De montaña a "reserva forestal". Colonización, sentido de comunidad y conservación en la selva Lacandona. Universidad Nacional Autónoma de México, Instituto de Investigaciones Sociales, México.
- Carabias, J., De la Maza, J., Cadena, R., 2015. El escenario natural y social, in: Carabias, J., De la Maza, J., Cadena, R. (Eds.), *Conservación y desarrollo sustentable en la Selva Lacandona. 25 años de actividades y experiencias*. Natura y Ecosistemas Mexicanos A.C, México, pp. 16-18.
- Card, D.H., 1982. Using Known Map Category Marginal Frequencies To Improve Estimates of Thematic Map Accuracy. *Photogrammetric Engineering and Remote Sensing* 48, 431-439.
- Cassol, H.L., Shimabukuro, Y.E., Beuchle, R., Aragão, L.E.O.C., 2019. Sentinel-1 Time-Series Analysis For Detection Of Forest Degradation By Selective Logging, *Anais do XIX Simpósio Brasileiro de Sensoriamento Remoto*, March ed, São José dos Campos, Brazil, pp. 1-4.
- Castillo-Santiago, M.A., Hellier, A., Tipper, R., De Jong, B.H.J., 2007. Carbon emissions from land-use change: An analysis of causal factors in Chiapas, Mexico. *Mitigation and Adaptation Strategies for Global Change* 12, 1213-1235. <https://doi.org/10.1007/s11027-006-9060-7>
- Castro, W., Junior, J.M., Polidoro, C., Osco, L.P., Gonçalves, W., Rodrigues, L., Santos, M., Jank, L., Barrios, S., Valle, C., Simeão, R., Carromeu, C., Silveira, E., Jorge, L.A.d.C., Matsubara, E., 2020. Deep learning applied to phenotyping of biomass in forages with uav-based rgb imagery, *Sensors (Switzerland)*. MDPI AG, pp. 1-18. <https://doi.org/10.3390/s20174802>

- CBD (Convention on Biological Diversity), 2021. Indicative definitions taken from the Report of the ad hoc technical expert group on forest biological diversity. CBD.
- Centro Nacional de Investigaciones Agrarias, 1982. Determinación de estrategias de desarrollo incorporando criterios ecológicos en el noreste de Chiapas, México.
- Chauvenet, A.L.M., Reise, J., Kümpel, N.F., Pettorelli, N., 2015. Satellite-based Remote Sensing for Measuring the Earth's Natural Capital and Ecosystem Services. <https://doi.org/10.13140/RG.2.2.10383.59043>
- Chave, J., Condit, R., Aguilar, S., Hernandez, A., Lao, S., Perez, R., 2004. Error propagation and scaling for tropical forest biomass estimates. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 359, 409-420. <https://doi.org/10.1098/rstb.2003.1425>
- Chazdon, R.L., 2008. Beyond Deforestation : Restoring Forests and Ecosystem Services on Degraded Lands. *Science* 320, 1458-1458. <https://doi.org/10.1126/science.1155365>
- Chen, L., Ren, C., Zhang, B., Wang, Z., Xi, Y., 2018. Estimation of forest above-ground biomass by geographically weighted regression and machine learning with sentinel imagery. *Forests* 9, 1-20. <https://doi.org/10.3390/f9100582>
- Cheng, G., Wang, Y., Xu, S., Wang, H., Xiang, S., Pan, C., 2017. Automatic Road Detection and Centerline Extraction via Cascaded End-to-End Convolutional Neural Network. *IEEE Transactions on Geoscience and Remote Sensing* 55, 3322-3337. <https://doi.org/10.1109/TGRS.2017.2669341>
- Chi, M., Plaza, A., Benediktsson, J.A., Sun, Z., Shen, J., Zhu, Y., 2016. Big Data for Remote Sensing: Challenges and Opportunities. *Proceedings of the IEEE* 104, 2207–2219. <https://doi.org/10.1109/JPROC.2016.2598228>
- Chollet, F., Allaire, J.J., 2018. *Deep Learning with R*. Manning Publications Co., Shelter Island, USA.
- Chuvienco, E., 2008. *Teledetección Ambiental*. Grupo Planeta, Madrid, España.
- Chuvienco, E., Ventura, G., Martín, M.P., 2005. AVHRR multitemporal compositing techniques for burned land mapping. *International Journal of Remote Sensing* 26, 1013-1018. <https://doi.org/10.1080/01431160412331299235>
- CICC (Comisión Intersecretarial de Cambio Climático), 2017. *Estrategia Nacional para REDD+ 2017-2030 ENAREDD+*. CONAFOR, México.
- Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O., 2016. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation, in: Ourselin, S., Joskowicz, L., Sabuncu, M., Unal, G., Wells, W. (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*. *MICCAI 2016. Lecture Notes in Computer Science* 9901. Springer, Athens, Greece, pp. 424-432. https://doi.org/10.1007/978-3-319-46723-8_49
- Cireşan, D.C., Meier, U., Masci, J., Gambardella, L.M., 2003. Flexible, High Performance Convolutional Neural Networks for Image Classification. *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*, 1237-1242. <https://doi.org/10.5591/978-1-57735-516-8/ijcai11-210>
- Clark, A., McKechnie, J., 2020. Detecting banana plantations in the wet tropics, Australia, using aerial photography and U-net. *Applied Sciences (Switzerland)* 10, 1-15. <https://doi.org/10.3390/app10062017>

- Clerici, N., Valbuena Calderón, C.A., Posada, J.M., 2017. Fusion of sentinel-1a and sentinel-2A data for land cover mapping: A case study in the lower Magdalena region, Colombia. *Journal of Maps* 13, 718-726.
<https://doi.org/10.1080/17445647.2017.1372316>
- Cochran, W.G., 1977. *Sampling techniques*, 3rd ed. John Wiley & Sons, New York.
- Colombo, R., Bellingeri, D., Fasolini, D., Marino, C., 2003. Retrieval of leaf area index in different vegetation types using high resolution satellite data. *Remote Sensing of Environment* 86, 120-131.
- Coluzzi, R., Imbrenda, V., Lanfredi, M., Simoniello, T., 2018. A first assessment of the Sentinel-2 Level 1-C cloud mask product to support informed surface analyses. *Remote Sensing of Environment* 217, 426-443.
<https://doi.org/10.1016/j.rse.2018.08.009>
- Comber, A., Wulder, M.A., 2019. Considering spatiotemporal processes in big data analysis: Insights from remote sensing of land cover and land use. *Transactions in GIS* 23, 879-891. <https://doi.org/10.1111/tgis.12559>
- CONAFOR (Comisión Nacional Forestal), 2020. Mapa de zonas críticas de deforestación, periodo 2001-2018. CONAFOR, México.
- CONAFOR (Comisión Nacional Forestal), 2022. Distribución de la superficie deforestada por ecorregiones y su transición de tierras forestales hacia otras categorías de usos de la tierra. CONAFOR, Ciudad de México.
- CONAFOR (Comisión Nacional Forestal), SEMARNAT (Secretaría de Recursos Naturales y Medio Ambiente), 2019. Glosario.
- Couteron, P., Pelissier, R., Nicolini, E.A., Paget, D., 2005. Predicting tropical forest stand structure parameters from Fourier transform of very high-resolution remotely sensed canopy images. *Journal of Applied Ecology* 42, 1121-1128.
<https://doi.org/10.1111/j.1365-2664.2005.01097.x>
- Crowley, M.A., Cardille, J.A., White, J.C., Wulder, M.A., 2019. Multi-sensor, multi-scale, Bayesian data synthesis for mapping within-year wildfire progression. *Remote Sensing Letters* 10, 302-311. <https://doi.org/10.1080/2150704X.2018.1536300>
- Cuevas, R.N.N., Guzmán, H.L., 1998. El Bosque Tropical Caducifolio En La Reserva de La Biosfera Sierra Manantlan, Jalisco-Colima, Mexico. *Bol. IBUG* 5, 445-491.
- Dabija, A., Kluczek, M., Zagajewski, B., Raczko, E., Kycko, M., Al-Sulttani, A.H., Tardà, A., Pineda, L., Corbera, J., 2021. Comparison of support vector machines and random forests for corine land cover mapping. *Remote Sensing* 13, 777.
<https://doi.org/10.3390/rs13040777>
- Danylo, O., Pirker, J., Lemoine, G., Ceccherini, G., See, L., McCallum, I., Hadi, Kraxner, F., Achard, F., Fritz, S., 2021. A map of the extent and year of detection of oil palm plantations in Indonesia, Malaysia and Thailand. *Scientific Data* 8, 96.
<https://doi.org/10.1038/s41597-021-00867-1>
- Datcu, M., Schwarz, G., Dumitru, C.O., 2020. Deep Learning Training and Benchmarks for Earth Observation Images: Data Sets, Features, and Procedures, in: Sadollah, A. (Ed.), *Recent Trends in Artificial Neural Networks. From Training to Prediction*. InTech Open, London, UK.

- de Bem, P.P., de Carvalho, O.A., Guimarães, R.F., Gomes, R.A.T., 2020. Change detection of deforestation in the Brazilian Amazon using Landsat data and convolutional neural networks. *Remote Sensing* 12, 901. <https://doi.org/10.3390/rs12060901>
- de Vos, J., Marion, M.-O., 2015. Colonización de la Selva Lacandona, in: Carabias, J., De la Maza, J., Cadena, R. (Eds.), *Conservación y desarrollo sustentable en la Selva Lacandona. 25 años de actividades y experiencias*. Natura y Ecosistemas Mexicanos A.C, México, pp. 35-44.
- Delgado-Aguilar, M.J., Fassnacht, F.E., Peralvo, M., Gross, C.P., Schmitt, C.B., 2017. Potential of TerraSAR-X and Sentinel 1 imagery to map deforested areas and derive degradation status in complex rain forests of Ecuador. *International Forestry Review* 19, 102-118. <https://doi.org/10.1505/146554817820888636>
- Deng, S., Katoh, M., Guan, Q., Yin, N., Li, M., 2014. Estimating forest aboveground biomass by combining ALOS PALSAR and WorldView-2 data: A case study at Purple Mountain National Park, Nanjing, China. *Remote Sensing* 6, 7878-7910. <https://doi.org/10.3390/rs6097878>
- Deschamps Ramírez, P., Larson, A.M., 2017. The politics of REDD+ MRV in Mexico: The interplay of the national and subnational levels. CIFOR, CGIAR, Bogor. <https://doi.org/https://doi.org/10.17528/cifor/006568>
- Deus, D., 2016. Integration of ALOS PALSAR and Landsat Data for Land Cover and Forest Mapping in Northern Tanzania, *Land*, p. 43. <https://doi.org/10.3390/land5040043>
- Diakogiannis, F.I., Waldner, F., Caccetta, P., Wu, C., 2020. ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. *ISPRS Journal of Photogrammetry and Remote Sensing* 162, 94-114. <https://doi.org/10.1016/j.isprsjprs.2020.01.013>
- Díaz, S., Demissew, S., Carabias, J., Joly, C., Lonsdale, M., Ash, N., Larigauderie, A., Adhikari, J.R., Arico, S., B?ldi, A.s., Bartuska, A., Baste, I.A., Bilgin, A., Brondizio, E., Chan, K.M.A., Figueroa, V.E., Duraiappah, A., Fischer, M., Hill, R., Koetz, T., Leadley, P., Lyver, P., Mace, G.M., Martin-Lopez, B., Okumura, M., Pacheco, D., Pascual, U., P??rez, E.S., Reyers, B., Roth, E., Saito, O., Scholes, R.J., Sharma, N., Tallis, H., Thaman, R., Watson, R., Yahara, T., Hamid, Z.A., Akosim, C., Al-Hafedh, Y., Allahverdiyev, R., Amankwah, E., Asah, T.S., Asfaw, Z., Bartus, G., Brooks, A.L., Caillaux, J., Dalle, G., Darnaedi, D., Driver, A., Erpul, G., Escobar-Eyzaguirre, P., Failler, P., Fouda, A.M.M., Fu, B., Gundimeda, H., Hashimoto, S., Homer, F., Lavorel, S., Lichtenstein, G., Mala, W.A., Mandivenyi, W., Matczak, P., Mbizvo, C., Mehrdadi, M., Metzger, J.P., Mikissa, J.B., Moller, H., Mooney, H.A., Mumby, P., Nagendra, H., Nesshover, C., Oteng-Yeboah, A.A., Pataki, G.r., Rou, M., Rubis, J., Schultz, M., Smith, P., Sumaila, R., Takeuchi, K., Thomas, S., Verma, M., Yeo-Chang, Y., Zlatanova, D., 2015. The IPBES Conceptual Framework - connecting nature and people. *Current Opinion in Environmental Sustainability* 14, 1-16. <https://doi.org/10.1016/j.cosust.2014.11.002>
- DOF (Diario Oficial de la Federación), 2017. Reglas de Operación del Programa Apoyos para el Desarrollo Forestal Sustentable 2018. DOF, México.
- DOF (Diario Oficial de la Federación), 2018. Ley General de Cambio Climático. DOF, México.

- Dong, L., Du, H., Han, N., Li, X., Zhu, D.e., Mao, F., Zhang, M., Zheng, J., Liu, H., Huang, Z., He, S., 2020. Application of convolutional neural network on lei bamboo above-ground-biomass (AGB) estimation using Worldview-2. *Remote Sensing* 12. <https://doi.org/10.3390/rs12060958>
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P., 2012. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment* 120, 25-36. <https://doi.org/10.1016/j.rse.2011.11.026>
- Du, L., McCarty, G.W., Zhang, X., Lang, M.W., Vanderhoof, M.K., Li, X., Huang, C., Lee, S., Zou, Z., 2020. Mapping Forested Wetland Inundation in the Neural Networks. *Remote Sensing* 12, 644. <https://doi.org/10.3390/rs12040644>
- Eckert, S., 2012. Improved Forest Biomass and Carbon Estimations Using Texture Measures from WorldView-2 Satellite Data. *Remote Sensing* 4, 810-829. <https://doi.org/10.3390/rs4040810>
- Ellis, E.A., Navarro-Martínez, A., García-Ortega, M., 2021. Drivers of forest cover transitions in the Selva Maya, Mexico: Integrating regional and community scales for landscape assessment. *Land Degradation & Development* 32, 3122-3141. <https://doi.org/https://doi.org/10.1002/ldr.3972>
- ESA (European Space Agency), 2016. Sentinel Data Access Annual Report. Copernicus, European Commission, ESA, SERCO.
- ESA (European Space Agency), 2018. Sentinel Data Access Annual Report. Copernicus, European Commission, ESA, SERCO.
- ESA (European Space Agency), 2019. About Copernicus: Overview.
- ESA (European Space Agency), 2020a. Sentinel-1 SAR Technical Guide. European Space Agency (ESA).
- ESA (European Space Agency), 2020b. Sentinel-2 MSI Technical Guide. European Space Agency (ESA).
- ESA (European Space Agency), 2020c. Sentinel-2 MSI User Guide.
- ESA (European Space Agency), 2020d. SNAP.
- Espinosa, D., Ocegueda, S., Aguilar, C., Flores, O., Llorente-Bousquets, J., 2008. El conocimiento biogeográfico de las especies y su regionalización natural, *Capital natural de México*, vol. I: Conocimiento actual de la biodiversidad. Conabio, México, pp. 33-65.
- Falbel, D., Zak, K., 2020. U-Net: Convolutional Networks for Biomedical Image Segmentation.
- Fan, J., Ma, C., Zhong, Y., 2019. A Selective Overview of Deep Learning. 1-37.
- FAO (Food and Agriculture Organization), 2009. Towards Defining Forest Degradation : Comparative Analysis of Existing definition.
- FAO (Food and Agriculture Organization), 2011. Assessing forest degradation. Towards the development of globally applicable guidelines. FAO, Roma.
- FAO (Food and Agriculture Organization), 2017. Openforis Accuracy Assessment Tool.
- FAO (Food and Agriculture Organization), 2018. FRA 2020. Terms and Definitions. Food and Agriculture of the United Nations, Rome.

- FAO (Food and Agriculture Organization), 2020. Global Forest Resources Assessment: Main Report. FAO, Rome, Italy.
- Fernández-Montes de Oca, A., Gallardo-Cruz, J.A., Martínez, M., 2015. Deforestación en la región Selva Lacandona, in: Carabias, J., De la Maza, J., Cadena, R. (Eds.), Conservación y desarrollo sustentable en la Selva Lacandona. 25 años de actividades y experiencias. Natura y Ecosistemas Mexicanos A.C., México, D.F., México, pp. 61-67.
- Fernández-Montes de Oca, A.I., Gallardo-Cruz, A., Ghilardi, A., Kauffer, E., Solórzano, J.V., Sánchez-Cordero, V., 2021. An integrated framework for harmonizing definitions of deforestation. *Environmental Science and Policy* 115, 71-78.
<https://doi.org/10.1016/j.envsci.2020.10.007>
- Filipponi, F., 2018. BAIS2: Burned Area Index for Sentinel-2. *Proceedings* 2, 364.
<https://doi.org/10.3390/ecrs-2-05177>
- Fitch, J.P., 1988. Synthetic Aperture Radar. Springer-Verlag, New York.
- Flood, N., Watson, F., Collett, L., 2019. Using a U-net convolutional neural network to map woody vegetation extent from high resolution satellite imagery across Queensland, Australia. *International Journal of Applied Earth Observation and Geoinformation* 82, 101897. <https://doi.org/10.1016/j.jag.2019.101897>
- Flores-Anderson, A.I., Herndon, K.E., Thapa, R.B., Cherrington, E., 2019. The SAR Handbook. Comprehensive Methodologies for Forest Monitoring and Biomass Estimation. NASA, Huntsville, USA. <https://doi.org/10.25966/nr2c-s697>
- Forkuor, G., Zougrana, J.-B.B., Dimobe, K., Ouattara, B., Prasad, K.V., Tondoh, J.E., 2020. Above Ground Biomass Mapping in West African Dryland Forest Using Sentinel-1 and 2 Datasets - A Case Study. *Remote Sensing of Environment* 236.
- Franklin, S.E., Maudie, A.J., Lavigne, M.B., 2001a. Using spatial co-occurrence texture to increase forest structure and species composition classification accuracy. *Photogrammetric Engineering and ...*, 849-855.
- Franklin, S.E., Wulder, M.A., Gerylo, G.R., 2001b. Texture analysis of IKONOS panchromatic data for Douglas-fir forest age class separability in British Columbia. *International Journal of Remote Sensing* 22, 2627-2632.
- Gallardo-Cruz, J.A., Meave, J.A., González, E.J., Lebrija-trejos, E., Romero-Romero, M.A., Pérez-García, E.A., Gallardo-Cruz, R., Hernández-Stefanoni, J.L., Martorell, C., 2012. Predicting tropical dry forest successional attributes from space: is the key hidden in image texture? *PloS one* 7, e30506-e30506.
<https://doi.org/10.1371/journal.pone.0030506>
- Gallardo-Cruz, J.A., Peralta-Carreta, C., Solórzano, J.V., Fernández-Montes de Oca, A.I., Nava, L.F., Kauffer, E., Carabias, J., 2021. Deforestation and trends of change in protected areas of the Usumacinta River basin (2000 – 2018), Mexico and Guatemala. *Regional Environmental Change* 21, 97-97. <https://doi.org/10.1007/s10113-021-01833-8>
- Gao, Y., Skutsch, M., Paneque-Gálvez, J., Ghilardi, A., 2020. Remote sensing of forest degradation: a review. *Environmental Research Letters*. <https://doi.org/10.1088/1748-9326/abaad7>
- García, E., 1998. Isotermas Medias Anuales. CONABIO, México.

- Gargiulo, M., Dell'aglio, D.A.G., Iodice, A., Riccio, D., Ruello, G., 2020. Integration of sentinel-1 and sentinel-2 data for land cover mapping using w-net. *Sensors (Switzerland)* 20, 2969. <https://doi.org/10.3390/s20102969>
- Gatti, L.V., Basso, L.S., Miller, J.B., Gloor, M., Gatti Domingues, L., Cassol, H.L.G., Tejada, G., Aragão, L.E.O.C., Nobre, C., Peters, W., Marani, L., Arai, E., Sanches, A.H., Corrêa, S.M., Anderson, L., Von Randow, C., Correia, C.S.C., Crispim, S.P., Neves, R.A.L., 2021. Amazonia as a carbon source linked to deforestation and climate change. *Nature* 595, 388-393. <https://doi.org/10.1038/s41586-021-03629-6>
- Gebhardt, S., Wehrmann, T., Ruiz, M.A.M., Maeda, P., Bishop, J., Schramm, M., Kopeinig, R., Cartus, O., Kellndorfer, J., Ressl, R., Santos, L.A., Schmidt, M., 2014. MAD-MEX: Automatic wall-to-wall land cover monitoring for the mexican REDD-MRV program using all landsat data. *Remote Sensing* 6, 3923-3943. <https://doi.org/10.3390/rs6053923>
- Gerard, F.F., North, P.R.J., Analyzing the Effect of Structural Variability and Canopy Gaps on Forest BRDF Using a Geometric-Optical Model.
- Ghasemi, N., Sahebi, M.R., Mohammadzadeh, A., 2011. A review on biomass estimation methods using synthetic aperture radar data. *International Journal of Geomatics and Geosciences* 1, 776-788. <https://doi.org/10.1109/VETECEF.2011.6093256>
- Ghatak, A., 2019. *Deep Learning with R*. Springer, Singapore. <https://doi.org/10.1007/978-981-13-5850-0>
- Ghosh, S.M., Behera, M.D., 2018. Aboveground biomass estimation using multi-sensor data synergy and machine learning algorithms in a dense tropical forest. *Applied Geography* 96, 29-40. <https://doi.org/10.1016/j.apgeog.2018.05.011>
- Ghosh, S.M., Behera, M.D., 2021. Aboveground biomass estimates of tropical mangrove forest using Sentinel-1 SAR coherence data - The superiority of deep learning over a semi-empirical model. *Computers and Geosciences* 150. <https://doi.org/10.1016/j.cageo.2021.104737>
- Giam, X., 2017. Global biodiversity loss from tropical deforestation. *Proceedings of the National Academy of Sciences* 114, 5775-5777. <https://doi.org/10.1073/pnas.1706264114>
- Giang, T.L., Dang, K.B., Toan Le, Q., Nguyen, V.G., Tong, S.S., Pham, V.-M., 2020. U-Net Convolutional Networks for Mining Land Cover Classification Based on High-Resolution UAV Imagery. *IEEE Access* 8, 186257-186273. <https://doi.org/10.1109/access.2020.3030112>
- Giannopoulos, M., Tsagkatakis, G., Tsakalides, P., 2022. 4D U-Nets for Multi-Temporal Remote Sensing Data Classification. *Remote Sensing* 14, 634-634. <https://doi.org/10.3390/rs14030634>
- Gibson, L., Lee, T.M., Koh, L.P., Brook, B.W., Gardner, T.A., Barlow, J., Peres, C.A., Bradshaw, C.J.A., Laurance, W.F., Lovejoy, T.E., Sodhi, N.S., 2011. Primary forests are irreplaceable for sustaining tropical biodiversity. *Nature* 478, 378-381. <https://doi.org/10.1038/nature10425>
- Giri, C.P., 2020. *Remote Sensing of Land Use and Land Cover. Principles and Applications*. CRC Press, Boca Raton, USA.

- Gobierno de México, 2020. Contribución determinada a nivel nacional: México. Versión actualizada 2020. SEMARNAT (Secretaría de Medio Ambiente y Recursos Naturales), INECC (Instituto Nacional de Ecología y Cambio Climático), México.
- Godinho Cassol, H.L., Shimabukuro, Y.E., Beuchle, R., Aragão, L.E.O.C., 2019. Sentinel-1 Time-Series Analysis For Detection Of Forest Degradation By Selective Logging. *Xix Sbsr*, 4-4.
- Gómez, C., White, J.C., Wulder, M.A., 2018. Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing* 10, 55-72. <https://doi.org/10.3390/rs10071119>
- Goodfellow, I., Bengio, Y., Courville, A., 2016. *Deep learning*. MIT Press, Cambridge, USA.
- Google Earth Engine, 2019. Sentinel-2 MSI: MultiSpectral Instrument, Level-2A. Google Earth Engine.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* 202, 18-27. <https://doi.org/10.1016/j.rse.2017.06.031>
- Goutte, C., Gaussier, E., 2005. A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation. *Lecture Notes in Computer Science* 3408, 345-359. https://doi.org/10.1007/978-3-540-31865-1_25
- Gutiérrez-Vélez, V.H., DeFries, R., Pinedo-Vásquez, M., Uriarte, M., Padoch, C., Baethgen, W., Fernandes, K., Lim, Y., 2011. High-yield oil palm expansion spares land at the expense of forests in the Peruvian Amazon. *Environmental Research Letters* 6, 044029-044029. <https://doi.org/10.1088/1748-9326/6/4/044029>
- Hamdi, Z.M., Brandmeier, M., Straub, C., 2019. Forest Damage Assessment Using Deep Learning on High Resolution Remote Sensing Data. *Remote Sensing* 11, 1976. <https://doi.org/10.3390/rs11171976>
- Hamedianfar, A., Mohamedou, C., Kangas, A., Vauhkonen, J., 2022. Deep learning for forest inventory and planning: a critical review on the remote sensing approaches so far and prospects for further applications, *Forestry*. Oxford University Press, pp. 451-465. <https://doi.org/10.1093/forestry/cpac002>
- Hansen, M.C., Krylov, A., Tyukavina, A., Potapov, P.V., Turubanova, S., Zutta, B., Ifo, S., Margono, B., Stolle, F., Moore, R., 2016. Humid tropical forest disturbance alerts using Landsat data. *Environmental Research Letters* 11, 034008-034008. <https://doi.org/10.1088/1748-9326/11/3/034008>
- Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* 342, 850-853. <https://doi.org/10.1126/science.1244693>
- Hanski, I., Habitat loss, the dynamics of biodiversity, and a perspective on conservation.
- He, H., Yang, D., Wang, S., Wang, S., Li, Y., 2019. Road Extraction by Using Atrous Spatial Pyramid Pooling Integrated Encoder-Decoder Network and Structural Similarity Loss. *Remote Sensing* 11, 1015. <https://doi.org/10.3390/rs11091015>
- He, K., Zhang, X., Ren, S., Sun, J., 2015. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification, 2015 IEEE International Conference

- on Computer Vision (ICCV). IEEE Xplore, Santiago, Chile, pp. 1026-1034.
<https://doi.org/10.1109/ICCV.2015.123>
- Heckel, K., Urban, M., Schratz, P., Mahecha, M.D., Schullius, C., 2020. Predicting Forest Cover in Distinct Ecosystems: The Potential of Multi-Source Sentinel-1 and -2 Data Fusion. *Remote sensing* 12, 302. <https://doi.org/10.3390/rs12020302>
- Hernández-Stefanoni, J.L., Castillo-Santiago, M.Á., Mas, J.F., Wheeler, C.E., Andres-Mauricio, J., Tun-Dzul, F., George-Chacón, S.P., Reyes-Palomeque, G., Castellanos-Basto, B., Vaca, R., Dupuy, J.M., 2020. Improving aboveground biomass maps of tropical dry forests by integrating LiDAR, ALOS PALSAR, climate and field data. *Carbon Balance and Management* 15. <https://doi.org/10.1186/s13021-020-00151-6>
- Hernández-Stefanoni, J.L., Reyes-Palomeque, G., Castillo-Santiago, M.Á., George-Chacón, S.P., Huechacóna-Ruiz, A.H., Tun-Dzul, F., Rondon-Rivera, D., Dupuy, J.M., 2018. Effects of sample plot size and GPS location errors on aboveground biomass estimates from LiDAR in tropical dry forests. *Remote Sensing* 10. <https://doi.org/10.3390/rs10101586>
- Hijmans, R.J., 2020. raster: Geographic Data Analysis and Modeling.
- Hirschmugl, M., Deutscher, J., Sobe, C., Bouvet, A., Mermoz, S., Schardt, M., 2020. Use of SAR and Optical Time Series for Tropical Forest Disturbance Mapping. *Remote Sensing* 12, 727. <https://doi.org/10.3390/RS12040727>
- Hirschmugl, M., Gallaun, H., Dees, M., Datta, P., Deutscher, J., Koutsias, N., Schardt, M., 2017. Methods for Mapping Forest Disturbance and Degradation from Optical Earth Observation Data: a Review. *Current Forestry Reports* 3, 32-45. <https://doi.org/10.1007/s40725-017-0047-2>
- Hirschmugl, M., Sobe, C., Deutscher, J., Schardt, M., 2018. Combined use of optical and synthetic aperture radar data for REDD+ applications in Malawi. *Land* 7, 116. <https://doi.org/10.3390/land7040116>
- Hodgdon, B., Hughell, D., Ramos, V., McNab, R., 2015. Deforestation trends in the Maya Biosphere Reserve, Guatemala 2000–2013. Rainforest Alliance, New York.
- Hoeser, T., Bachofer, F., Kuenzer, C., 2020. Object Detection and Image Segmentation with Deep Learning on Earth Observation Data: A Review—Part II: Applications. *Remote Sensing* 12, 3053. <https://doi.org/10.3390/rs12183053>
- Hoeser, T., Kuenzer, C., 2020. Object Detection and Image Segmentation with Deep Learning on Earth Observation Data: A Review-Part I : Evolution and Recent Trends. *Remote Sensing* 12, 1667. <https://doi.org/10.3390/rs12101667>
- Hojas-Gascon, L., Cerutti, P.O., Eva, H., Nasi, R., Martius, C., 2015. Monitoring deforestation and forest degradation in the context of REDD+: Lessons from Tanzania. CIFOR Infobrief no. 124. 8-8. <https://doi.org/10.17528/cifor/005642>
- Hostert, P., Griffiths, P., van der Linden, S., Pflugmacher, D., 2015. Time Series Analyses in a New Era of Optical Satellite Data, in: Kuenzer, C., Dech, S., Wagner, W. (Eds.), *Remote Sensing Time Series: Revealing Land Surface Dynamics*. Springer International Publishing, Cham, pp. 25-41.
- Houghton, R.A., 1990. The global effects of tropical deforestation. *Environmental Science and Technology* 24, 414-422. <https://doi.org/10.1021/es00074a001>
- Houghton, R.A., 1991. Tropical Deforestation and Atmospheric Carbon Dioxide. *Climatic Change* 19, 99-118.

- Huang, B., Lu, K., Audebert, N., Khalel, A., Tarabalka, Y., Malof, J., Boulch, A., Saux, B.L., Collins, L., Bradbury, K., Lefèvre, S., El-Saban, M., 2018. Large-scale semantic classification: Outcome of the first year of inria aerial image labeling benchmark, International Geoscience and Remote Sensing Symposium (IGARSS) - IEEE International Geoscience and Remote Sensing Symposium. HAL, Valencia, Spain, pp. 1-4. <https://doi.org/10.1109/IGARSS.2018.8518525>
- Huete, A.R., Liu, H.Q., Batchily, K., Van Leeuwen, W., 1997. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing of Environment* 59, 440-451.
- Huy, B., Truong, N.Q., Khiem, N.Q., Poudel, K.P., Temesgen, H., 2022. Deep learning models for improved reliability of tree aboveground biomass prediction in the tropical evergreen broadleaf forests. *Forest Ecology and Management* 508. <https://doi.org/10.1016/j.foreco.2022.120031>
- Ienco, D., Interdonato, R., Gaetano, R., Ho Tong Minh, D., 2019. Combining Sentinel-1 and Sentinel-2 Satellite Image Time Series for land cover mapping via a multi-source deep learning architecture. *ISPRS Journal of Photogrammetry and Remote Sensing* 158, 11-22. <https://doi.org/10.1016/j.isprsjprs.2019.09.016>
- INEGI (Instituto Nacional de Estadística), CONABIO (Comisión Nacional para el Conocimiento y Uso de la Biodiversidad), INE (Instituto Nacional de Ecología). 2008. *Ecorregiones terrestres de México* 1:1000000.
- IPCC (Intergovernmental Panel on Climate Change), 2003. *Definitions and Methodological Options to Inventory Emissions from 3 Direct Human-induced Degradation of Forests and Devegetation of Other Vegetation*, Milan.
- IPCC (Intergovernmental Panel on Climate Change), 2014. *Climate Change 2014 Synthesis Report Summary for Policymakers Summary for Policymakers*.
- Irvin, J., Sheng, H., Ramachandran, N., Johnson-Yu, S., Zhou, S., Story, K., Rustowicz, R., Elsworth, C., Austin, K., Ng, A.Y., 2020. ForestNet: Classifying Drivers of Deforestation in Indonesia using Deep Learning on Satellite Imagery. *arXiv*, 1-10. <https://doi.org/10.48550/arXiv.2011.05479>
- Isaac-Márquez, R., Valladares, J.L.S., Spencer, A.E., Arcipreste, M.E.A., Aguilar, M.A.A., Márquez, A.P.I., González, M.C.S., 2016. Impactos Sociales y Ambientales de la Palma de Aceite: Perspectiva de los Campesinos en Campeche, México. *Journal of Latin American Geography* 15, 123-146. <https://doi.org/10.1353/lag.2016.0023>
- Isaienkov, K., Yushchuk, M., Khramtsov, V., Seliverstov, O., 2020. Deep Learning for Regular Change Detection in Ukrainian Forest Ecosystem With Sentinel-2. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 14, 1-15. <https://doi.org/10.1109/jstars.2020.3034186>
- Izquierdo-Tort, S., Corbera, E., Barceinas Cruz, A., Naime, J., Angélica Vázquez-Cisneros, P., Carabias Lillo, J., Castro-Tovar, E., Ortiz Rosas, F., Rubio, N., Torres Knoop, L., Dupras, J., 2021. Local responses to design changes in payments for ecosystem services in Chiapas, Mexico. *Ecosystem Services* 50. <https://doi.org/10.1016/j.ecoser.2021.101305>
- Ji, S., Zhang, C., Xu, A., Shi, Y., Duan, Y., 2018. 3D convolutional neural networks for crop classification with multi-temporal remote sensing images. *Remote Sensing* 10, 75. <https://doi.org/10.3390/rs10010075>

- Jiménez-Rodríguez, D.L., Gao, Y., Solórzano, J.V., Skutsch, M., Pérez-Salicrup, D.R., Salinas-Melgoza, M.A., Farfán, M., 2022. Mapping Forest Degradation and Contributing Factors in a Tropical Dry Forest. *Frontiers in Environmental Science* 10. <https://doi.org/10.3389/fenvs.2022.912873>
- Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M.R., Kuemmerle, T., Meyfroidt, P., Mitchard, E.T.A., Reiche, J., Ryan, C.M., Waske, B., 2016. A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring. *Remote Sensing* 8, 70. <https://doi.org/10.3390/rs8010070>
- Kalamandeen, M., Gloor, E., Mitchard, E., Quincey, D., Ziv, G., Spracklen, D., Spracklen, B., Adami, M., Aragañ, L.E.O.C., Galbraith, D., 2018. Pervasive Rise of Small-scale Deforestation in Amazonia. *Scientific Reports* 8, 1-10. <https://doi.org/10.1038/s41598-018-19358-2>
- Karatzoglou, A., Smola, A., Hornik, K., Zeileis, A., 2004. kernlab - An S4 Package for Kernel Methods in R. *Journal of Statistical Software* 11, 1-20.
- Kattenborn, T., Leitloff, J., Schiefer, F., Hinz, S., 2021. Review on Convolutional Neural Networks (CNN) in Vegetation Remote Sensing. *ISPRS Journal of Photogrammetry and Remote Sensing* 173, 24-49. <https://doi.org/10.1016/j.isprsjprs.2020.12.010>
- Kattenborn, T., Schiefer, F., Frey, J., Feilhauer, H., Mahecha, M.D., Dormann, C.F., 2022. Spatially autocorrelated training and validation samples inflate performance assessment of convolutional neural networks. *ISPRS Open Journal of Photogrammetry and Remote Sensing* 5, 100018-100018. <https://doi.org/10.1016/j.ophoto.2022.100018>
- Kemker, R., Salvaggio, C., Kanan, C., 2018. Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning. *ISPRS Journal of Photogrammetry and Remote Sensing* 145, 60-77. <https://doi.org/10.1016/j.isprsjprs.2018.04.014>
- Kerr, J.T., Ostrovsky, M., 2003. From space to species: ecological applications for remote sensing. *Trends in Ecology & Evolution* 18, 299-305. [https://doi.org/10.1016/S0169-5347\(03\)00071-5](https://doi.org/10.1016/S0169-5347(03)00071-5)
- Khan, A., Govil, H., Kumar, G., Dave, R., 2020. Synergistic use of Sentinel-1 and Sentinel-2 for improved LULC mapping with special reference to bad land class: a case study for Yamuna River floodplain, India. *Spatial Information Research* 28, 669-681. <https://doi.org/10.1007/s41324-020-00325-x>
- Kindermann, G., Obersteiner, M., Sohngen, B., Sathaye, J., Andrasko, K., Rametsteiner, E., Schlamadinger, B., Wunder, S., Beach, R., 2008. Global cost estimates of reducing carbon emissions through avoided deforestation. *Proceedings of the National Academy of Sciences of the United States of America* 105, 10302-10307. <https://doi.org/10.1073/pnas.0710616105>
- Kingma, D.P., Ba, J., 2015. Adam: A Method for Stochastic Optimization, *International Conference on Learning Representations*, San Diego, USA, pp. 1-13.
- Koch, M., 2012. *Capitalism and Climate Change: Theoretical Discussion, Historical Development and Policy Responses*. Palgrave MacMillan, London.
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM* 60, 84-90. <https://doi.org/10.1145/3065386>

- Kruid, S., Macedo, M.N., Gorelik, S.R., Walker, W., Moutinho, P., Brando, P.M., Castanho, A., Alencar, A., Baccini, A., Coe, M.T., 2021. Beyond Deforestation: Carbon Emissions From Land Grabbing and Forest Degradation in the Brazilian Amazon. *Frontiers in Forests and Global Change* 4. <https://doi.org/10.3389/ffgc.2021.645282>
- Kuhn, M., Vaughan, D., 2021. yardstick: Tidy Characterizations of Model Performance.
- Kuhn, M., Wickham, H., 2020. Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles.
- Kussul, N., Lavreniuk, M., Skakun, S., Shelestov, A., 2017. Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. *IEEE Geoscience and Remote Sensing Letters* 14, 778-782. <https://doi.org/10.1109/LGRS.2017.2681128>
- Lary, D.J., Alavi, A.H., Gandomi, A.H., Walker, A.L., 2016. Machine learning in geosciences and remote sensing. *Geoscience Frontiers* 7, 3-10. <https://doi.org/10.1016/j.gsf.2015.07.003>
- Latam, M., 2020. México: la invasión de la palma africana en la Selva Lacandona, Mongabay Latam. Mongabay, México.
- Laurin, G.V., Papale, D., Balling, J., Truckenbrodt, J., Urban, M., Corona, P., Mattioli, W., Puletti, N., Rizzo, M., 2018. Above-ground biomass prediction by Sentinel-1 multitemporal data in central Italy with integration of ALOS2 and Sentinel-2 data. *Journal of Applied Remote Sensing* 12. <https://doi.org/10.1117/1.JRS.12.016008>
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521, 436-444. <https://doi.org/10.1038/nature14539>
- Leyva-Ovalle, Á., Valdez-Lazalde, J.R., de los Santos-Posadas, H.M., Martínez-Trinidad, T., Herrera-Corredor, J.A., Lugo-Espinosa, O., García-Nava, J.R., 2017. Monitoreo de la degradación forestal en México con base en el inventario nacional forestal y de suelos (Infys). *Madera Bosques* 23, 69-83. <https://doi.org/10.21829/myb.2017.2321431>
- Li, J., Roy, D.P., 2017. A global analysis of Sentinel-2a, Sentinel-2b and Landsat-8 data revisit intervals and implications for terrestrial monitoring. *Remote Sensing* 9, 902. <https://doi.org/10.3390/rs9090902>
- Li, M., Im, J., Beier, C., 2013. Machine learning approaches for forest classification and change analysis using multi-temporal Landsat TM images over Huntington Wildlife Forest. *GIScience and Remote Sensing* 50, 361-384. <https://doi.org/10.1080/15481603.2013.819161>
- Li, Y., Li, M., Li, C., Liu, Z., 2020. Forest aboveground biomass estimation using Landsat 8 and Sentinel-1A data with machine learning algorithms. *Scientific Reports* 10. <https://doi.org/10.1038/s41598-020-67024-3>
- Li, Y., Zhang, H., Xue, X., Jiang, Y., Shen, Q., 2018. Deep learning for remote sensing image classification: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 8, 1-17. <https://doi.org/10.1002/widm.1264>
- Liao, Z., He, B., Quan, X., 2020. Potential of texture from SAR tomographic images for forest aboveground biomass estimation. *International Journal of Applied Earth Observation and Geoinformation* 88. <https://doi.org/10.1016/j.jag.2020.102049>
- Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest, R news, pp. 18-22.

- Lillesand, T.M., Kiefer, R.W., 2000. Remote sensing and image interpretation, 4 ed. John Wiley & Sons, Nueva York, EUA.
- Liu, J., Herzberger, A., Kapsar, K., Carlson, A.K., Connor, T., 2019. What Is Telecoupling?, Telecoupling. Springer International Publishing, pp. 19-48.
- Lu, D., 2006. The potential and challenge of remote sensing-based biomass estimation. *International Journal of Remote Sensing* 27, 1297-1328. <https://doi.org/10.1080/01431160500486732>
- Lu, D., Batistella, M., 2005. Exploring TM image texture and its relationships with biomass estimation in Rondônia, Brazilian Amazon. *Acta Amazonica* 35, 249-257.
- Lu, D., Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing* 28, 823-870. <https://doi.org/10.1080/01431160600746456>
- Lund, H.G., 2009. What is a degraded forest? White Paper on Forest Degradation Definitions Prepared for FAO, 42-42.
- Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., Johnson, B.A., 2019. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing* 152, 166-177. <https://doi.org/10.1016/j.isprsjprs.2019.04.015>
- Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., Gascon, F., 2017. Sen2Cor for Sentinel-2, Image and Signal Processing for Remote Sensing XXIII. Proceedings of SPIE, Varsovia, Polonia, pp. 1042704-1042701. <https://doi.org/10.1117/12.2278218>
- Malenovský, Z., Rott, H., Cihlar, J., Schaepman, M.E., García-Santos, G., Fernandes, R., Berger, M., 2012. Sentinels for science: Potential of Sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land. *Remote Sensing of Environment* 120, 91-101. <https://doi.org/10.1016/j.rse.2011.09.026>
- Mandal, D., Kumar, V., Ratha, D., Dey, S., Bhattacharya, A., Lopez-Sanchez, J.M., McNairn, H., Rao, Y.S., 2020. Dual polarimetric radar vegetation index for crop growth monitoring using sentinel-1 SAR data. *Remote Sensing of Environment* 247, 111954. <https://doi.org/10.1016/j.rse.2020.111954>
- Maretto, R.V., Fonseca, L.M.G., Jacobs, N., Körting, T.S., Bendini, H.N., Parente, L.L., 2021. Spatio-Temporal Deep Learning Approach to Map Deforestation in Amazon Rainforest. *IEEE Geoscience and Remote Sensing Letters* 18, 771-775. <https://doi.org/10.1109/LGRS.2020.2986407>
- Margono, B.A., Turubanova, S., Zhuravleva, I., Potapov, P., Tyukavina, A., Baccini, A., Goetz, S., Hansen, M.C., 2012. Mapping and monitoring deforestation and forest degradation in Sumatra (Indonesia) using Landsat time series data sets from 1990 to 2010. *Environmental Research Letters* 7. <https://doi.org/10.1088/1748-9326/7/3/034010>
- Martinez-Yrizar, A., Sarukhan, J., Perez-Jimenez, A., Rincon, E., Maass, J., Solis-Magallanes, A., Cervantes, L., 1992. Above-ground phytomass of a tropical deciduous forest on the coast of Jalisco, México. *Journal of Tropical Ecology* 8, 87-96. <https://doi.org/10.1017/S0266467400006131>
- Mas, J.-F., Couturier, S., Paneque-Gálvez, J., Skutsch, M.M., Pérez-Vega, A., Castillo-Santiago, M., Gerardo, B., 2016. Comment on Gebhardt et al. MAD-MEX: Automatic

- Wall-to-Wall Land Cover Monitoring for the Mexican REDD-MRV Program Using All Landsat Data. *Remote Sensing* 8, 533. <https://doi.org/10.3390/rs8070533>
- Mas, J.-F., Velázquez, A., Couturier, S., 2009. La evaluación de los cambios de cobertura/uso del suelo en la República Mexicana. *Investigación ambiental Ciencia y política pública* 1, 23-39.
- Masolele, R.N., De Sy, V., Herold, M., Marcos Gonzalez, D., Verbesselt, J., Gieseke, F., Mullissa, A.G., Martius, C., 2021. Spatial and temporal deep learning methods for deriving land-use following deforestation: A pan-tropical case study using Landsat time series. *Remote Sensing of Environment* 264, 112600. <https://doi.org/10.1016/j.rse.2021.112600>
- Matosak, B.M., Fonseca, L.M.G., Taquary, E.C., Maretto, R.V., Bendini, H.D.N., Adami, M., 2022. Mapping Deforestation in Cerrado Based on Hybrid Deep Learning Architecture and Medium Spatial Resolution Satellite Time Series. *Remote Sensing* 14, 209. <https://doi.org/10.3390/rs14010209>
- Maxwell, A.E., Warner, T.A., Guillén, L.A., 2021. Accuracy assessment in convolutional neural network-based deep learning remote sensing studies—part 2: Recommendations and best practices. *Remote Sensing* 13. <https://doi.org/10.3390/rs13132591>
- McFarland, B.J., 2018. The Context of Tropical Rainforest Deforestation and Degradation. https://doi.org/10.1007/978-3-319-63236-0_2
- McRoberts, R.E., Stehman, S.V., Liknes, G.C., Næsset, E., Sannier, C., Walters, B.F., 2018. The effects of imperfect reference data on remote sensing-assisted estimators of land cover class proportions. *ISPRS Journal of Photogrammetry and Remote Sensing* 142, 292-300. <https://doi.org/10.1016/j.isprsjprs.2018.06.002>
- Meijaard, E., García-Ulloa, J., Sheil, D., Wich, S.A., Carlson, K.M., Juffe-Bignoli, D., Brooks, T.M., 2018. Oil Palm and Biodiversity: A Situation Analysis by the IUCN Oil Palm Task Force. IUCN (International Union for Conservation of Nature).
- Melesse, A.M., Weng, Q., S.Thenkabail, P., Senay, G.B., 2007a. Remote Sensing Sensors and Applications in Environmental Resources Mapping and Modelling. *Sensors* 7, 3209-3241. <https://doi.org/10.3390/s7123209>
- Melesse, A.M., Weng, Q., Thenkabail, P.S., Senay, G.B., 2007b. Remote Sensing Sensors and Applications in Environmental Resources Mapping and Modelling. *Sensors* 7, 3209-3241. <https://doi.org/10.3390/s7123209>
- Mendoza, E., Dirzo, R., 1999. Deforestation in Lacandonia (southeast Mexico): Evidence for the declaration of the northernmost tropical hot-spot. *Biodiversity and Conservation* 8, 1621-1641.
- Mercier, A., Betbeder, J., Rumiano, F., Baudry, J., Gond, V., Blanc, L., Bourgoïn, C., Cornu, G., Ciudad, C., Marchamalo, M., Pocard-Chapuis, R., Hubert-Moy, L., 2019. Evaluation of Sentinel-1 and 2 Time Series for Land Cover Classification of Forest-Agriculture Mosaics in Temperate and Tropical Landscapes. *Remote Sensing* 11, 979. <https://doi.org/10.3390/rs11080979>
- Mermoz, S., Le Toan, T., Villard, L., Réjou-Méchain, M., Seifert-Granzin, J., 2014. Biomass assessment in the Cameroon savanna using ALOS PALSAR data. *Remote Sensing of Environment* 155, 109-119. <https://doi.org/10.1016/j.rse.2014.01.029>

- Meyer, V., Saatchi, S., Ferraz, A., Xu, L., Duque, A., García, M., Chave, J., 2019. Forest degradation and biomass loss along the Chocó region of Colombia. *Carbon Balance and Management* 14, 1-15. <https://doi.org/10.1186/s13021-019-0117-9>
- Milodowski, D.T., Mitchard, E.T.A., Williams, M., 2017. Forest loss maps from regional satellite monitoring systematically underestimate deforestation in two rapidly changing parts of the Amazon. *Environmental Research Letters* 12, 094003. <https://doi.org/https://doi.org/10.1088/1748-9326/aa7e1e>
- Mitchell, A.L., Rosenqvist, A., Mora, B., 2017. Current remote sensing approaches to monitoring forest degradation in support of countries measurement, reporting and verification (MRV) systems for REDD+. *Carbon Balance and Management* 12, 9. <https://doi.org/10.1186/s13021-017-0078-9>
- Monsalve-Tellez, J.M., Torres-León, J.L., Garcés-Gómez, Y.A., 2022. Evaluation of SAR and Optical Image Fusion Methods in Oil Palm Crop Cover Classification Using the Random Forest Algorithm. *Agriculture (Switzerland)* 12, 955. <https://doi.org/10.3390/agriculture12070955>
- Montes de Oca, R.E., Castro, E., Ramírez-Martínez, C., Naime, J., Carabias, J., 2015. Características socioeconómicas del municipio de Marqués de Comillas, in: Carabias, J., De la Maza, J., Cadena, R. (Eds.), *Conservación y desarrollo sustentable en la Selva Lacandona. 25 años de actividades y experiencias.* . Natura y Ecosistemas Mexicanos A.C., México, pp. 219-243.
- Moradi, F., Sadeghi, S.M.M., Heidarlou, H.B., Deljouei, A., Boshkar, E., Borz, S.A., 2022. Above-ground biomass estimation in a Mediterranean sparse coppice oak forest using Sentinel-2 data. *Annals of Forest Research* 65, 165-182. <https://doi.org/10.15287/afr.2022.2390>
- Morales-Barquero, L., Borrego, A., Skutsch, M., Kleinn, C., Healey, J.R., 2015. Identification and quantification of drivers of forest degradation in tropical dry forests: A case study in Western Mexico. *Land Use Policy* 49, 296-309. <https://doi.org/10.1016/j.landusepol.2015.07.006>
- Morales-Barquero, L., Skutsch, M., Jardel-Peláez, E.J., Ghilardi, A., Kleinn, C., Healey, J.R., 2014. Operationalizing the definition of forest degradation for REDD+, with application to Mexico. *Forests* 5, 1653-1681. <https://doi.org/10.3390/f5071653>
- Moreira, A., Prats-Iraola, P., Younis, M., Krieger, G., Hajnsek, I., Papathanassiou, K.P., 2013. A tutorial on synthetic aperture radar. *IEEE Geoscience and Remote Sensing Magazine* 1, 6-43. <https://doi.org/10.1109/MGRS.2013.2248301>
- Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing* 66, 247-259. <https://doi.org/10.1016/j.isprsjprs.2010.11.001>
- Mountrakis, G., Li, J., Lu, X., Hellwich, O., 2018. Deep learning for remotely sensed data. *ISPRS Journal of Photogrammetry and Remote Sensing* 145, 1-2. <https://doi.org/10.1016/j.isprsjprs.2018.08.011>
- Mueller, J.P., Massaron, L., 2019. *Deep Learning for Dummies.* John Wiley & Sons, Hoboken.
- Murray, J.P., Grenyer, R., Wunder, S., Raes, N., Jones, J.P.G., 2015. Spatial patterns of carbon, biodiversity, deforestation threat, and REDD+ projects in Indonesia. *Conservation Biology* 29, 1434-1445. <https://doi.org/10.1111/cobi.12500>

- Mwampamba, T.H., Salinas, M., Larrazabal, A., Skutsch, M., 2015. Opciones en REDD+ para ejidos en la cuenca de Ayuquila, Jalisco. CIGA - UNAM, University of Twente, JIRA, CIEco, Morelia.
- Neves, A.K., Körting, T.S., Fonseca, L.M.G., Neto, C.D.G., Wittich, D., Costa, G.A.O.P., Heipke, C., 2020. Semantic segmentation of Brazilian savanna vegetation using high spatial resolution satellite data and U-net, 2020 XXIV ISPRS Congress (2020 edition). ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Nice, France, pp. 505-511.
- Nielsen, A.A., 2007. The regularized iteratively reweighted MAD method for change detection in multi- and hyperspectral data. *IEEE Transactions on Image Processing* 16, 463-478. <https://doi.org/10.1109/TIP.2006.888195>
- Ochoa-Gaona, S., González-Espinosa, M., 2000. Land use and deforestation in the highlands of Chiapas, Mexico. *Applied Geography* 20, 17-42. [https://doi.org/https://doi.org/10.1016/S0143-6228\(99\)00017-X](https://doi.org/https://doi.org/10.1016/S0143-6228(99)00017-X)
- Ochoa-Gaona, S., Hernández-Vázquez, F., De Jong, B.H.J., Gurri-García, F.D., 2007. Pérdida de diversidad florística ante un gradiente de intensificación del sistema agrícola de Roza-Tumba-Quema: un estudio de caso en la Selva Lacandona, Chiapas, México. *Boletín de la Sociedad Botánica de México* 81, 65-80.
- Oliver, C., Quegan, S., 2004. *Understanding Synthetic Aperture Radar Images*. SciTech, Raleigh.
- Olofsson, P., Arévalo, P., Espejo, A.B., Green, C., Lindquist, E., McRoberts, R.E., Sanz, M.J., 2020. Mitigating the effects of omission errors on area and area change estimates. *Remote Sensing of Environment* 236, 111492. <https://doi.org/10.1016/j.rse.2019.111492>
- 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. <https://doi.org/10.1016/j.rse.2014.02.015>
- Ortega Adarme, M., Doblas Prieto, J., Queiroz Feitosa, R., De Almeida, C.A., 2022. Improving Deforestation Detection on Tropical Rainforests Using Sentinel-1 Data and Convolutional Neural Networks. *Remote Sensing* 14, 3290. <https://doi.org/10.3390/rs14143290>
- Ortega, M.X., Bermudez, J.D., Happ, P.N., Gomes, A., Feitosa, R.Q., 2019. Evaluation of Deep Learning Techniques for Deforestation Detection in the Amazon Forest, PIA19+MRSS19 – Photogrammetric Image Analysis & Munich Remote Sensing Symposium. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Munich, Germany, pp. 18-20. <https://doi.org/10.5194/isprs-annals-IV-2-W7-121-2019>
- Ortiz-espejel, B., Toledo, V.M., 1998. Tendencias en la deforestación de la selva lacandona (Chiapas, México): El caso de las cañadas. *Interciencia* 23, 318-327. <https://doi.org/0378-1844/98/06/318-10>
- Otukei, J.R., Blaschke, T., 2010. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *International Journal of Applied Earth Observation and Geoinformation* 12, 27-31. <https://doi.org/10.1016/j.jag.2009.11.002>

- Ozdemir, I., Karnieli, A., 2011. Predicting forest structural parameters using the image texture derived from WorldView-2 multispectral imagery in a dryland forest, Israel. *International Journal of Applied Earth Observation and Geoinformation* 13, 701-710. <https://doi.org/10.1016/j.jag.2011.05.006>
- Ozkan, U.Y., Demirel, T., 2021. The influence of window size on remote sensing-based prediction of forest structural variables. *Ecological Processes* 10, 60-60. <https://doi.org/10.1186/s13717-021-00330-4>
- Park, J.T., 2015. *Climate Change and Capitalism*, pp. 189-206.
- Pearson, T.R.H., Brown, S., Murray, L., Sidman, G., 2017. Greenhouse gas emissions from tropical forest degradation: An underestimated source. *Carbon Balance and Management* 12. <https://doi.org/10.1186/s13021-017-0072-2>
- Pebesma, E., 2018. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal* 10, 439-446. <https://doi.org/10.32614/RJ-2018-009>
- Pelletier, C., Webb, G.I., Petitjean, F., 2019. Temporal convolutional neural network for the classification of satellite image time series. *Remote Sensing* 11, 523. <https://doi.org/10.3390/rs11050523>
- Pettorelli, N., Schulte to Bühne, H., Tulloch, A., Dubois, G., Macinnis-Ng, C., Queirós, A.M., Keith, D.A., Wegmann, M., Schrodt, F., Stellmes, M., Sonnenschein, R., Geller, G.N., Roy, S., Somers, B., Murray, N., Bland, L., Geijzendorffer, I., Kerr, J.T., Broszeit, S., Leitão, P.J., Duncan, C., El Serafy, G., He, K.S., Blanchard, J.L., Lucas, R., Mairota, P., Webb, T.J., Nicholson, E., 2018. Satellite remote sensing of ecosystem functions: opportunities, challenges and way forward. *Remote Sensing in Ecology and Conservation* 4, 71-93. <https://doi.org/10.1002/rse2.59>
- Pimentel, D., 1997. The value of forests to world food security. *Human Ecology* 25, 91-120. <https://doi.org/10.1023/A:1021987920278>
- Pinto, L.O.R., Souza, C.R.d., Terra, M.d.C.N.S., Mello, J.M.d., Calegário, N., Arcebi Júnior, F.W., 2021. Optimal plot size for carbon-diversity sampling in tropical vegetation. *Forest Ecology and Management* 482. <https://doi.org/10.1016/j.foreco.2020.118778>
- Planet Team, 2017. *Planet Application Program Interface: In Space for Life on Earth.*, San Francisco, CA.
- Pompeu Pavanelli, J., dos Santos, J., Soares Galvão, L., Xaud, M., Xaud, H., 2018. Palsar-2/ALOS-2 and Oli/Landsat-8 data integration for land use and land cover mapping in northern Brazilian Amazon. *Boletim de Ciencias Geodesicas* 24, 250-269. <https://doi.org/10.1590/S1982-21702018000200017>
- Pontius, R.G., Millones, M., 2011. Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing* 32, 4407-4429. <https://doi.org/10.1080/01431161.2011.552923>
- Poortinga, A., Tenneson, K., Shapiro, A., Nquyen, Q., Aung, K.S., Chishtie, F., Saah, D., 2019. Mapping plantations in Myanmar by fusing Landsat-8, Sentinel-2 and Sentinel-1 data along with systematic error quantification. *Remote Sensing* 11, 831. <https://doi.org/10.3390/rs11070831>
- Praticò, S., Solano, F., Di Fazio, S., Modica, G., 2021. Machine learning classification of mediterranean forest habitats in google earth engine based on seasonal sentinel-2 time-

- series and input image composition optimisation. *Remote Sensing* 13, 1-28.
<https://doi.org/10.3390/rs13040586>
- Proisy, C., Barbier, N., Guérault, M., 2011. Biomass prediction in tropical forests: the canopy grain approach, *Remote sensing of biomass: principles and applications*, pp. 1-18.
- Proisy, C., Couteron, P., Fromard, F., 2007. Predicting and mapping mangrove biomass from canopy grain analysis using Fourier-based textural ordination of IKONOS images. *Remote Sensing of Environment* 109, 379-392.
<https://doi.org/10.1016/j.rse.2007.01.009>
- Puissant, A., Hirsch, J., Weber, C., 2005. The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery. *International Journal of Remote Sensing* 26, 733-745.
<https://doi.org/10.1080/01431160512331316838>
- Pulliainen, J.T., Kurvonen, L., Hallikainen, M.T., 1999. Multitemporal behavior of L- and C-band SAR observations of boreal forests. *IEEE Transactions on Geoscience and Remote Sensing* 37, 927-937. <https://doi.org/10.1109/36.752211>
- Putz, F.E., Redford, K.H., 2010. The Importance of Defining 'Forest': Tropical Forest Degradation, Deforestation, Long-Term Phase Shifts, and Further Transitions. *Biotropica* 42, 10-20. <https://doi.org/10.1111/j.1744-7429.2009.00567.x>
- Pyo, J.C., Duan, H., Baek, S., Kim, M.S., Jeon, T., Kwon, Y.S., Lee, H., Cho, K.H., 2019. A convolutional neural network regression for quantifying cyanobacteria using hyperspectral imagery. *Remote Sensing of Environment* 233.
<https://doi.org/10.1016/j.rse.2019.111350>
- Qin, Y., Xiao, X., Wigneron, J.P., Ciais, P., Brandt, M., Fan, L., Li, X., Crowell, S., Wu, X., Doughty, R., Zhang, Y., Liu, F., Sitch, S., Moore, B., 2021. Carbon loss from forest degradation exceeds that from deforestation in the Brazilian Amazon. *Nature Climate Change* 11, 442-448. <https://doi.org/10.1038/s41558-021-01026-5>
- R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.
- R Core Team, 2021. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Viena.
- Ramoelo, A., Cho, M., Mathieu, R., Skidmore, A.K., 2015. Potential of Sentinel-2 spectral configuration to assess rangeland quality. *Journal of Applied Remote Sensing* 9, 094096. <https://doi.org/10.1117/1.jrs.9.094096>
- Reiche, J., de Bruin, S., Hoekman, D., Verbesselt, J., Herold, M., 2015. A Bayesian approach to combine landsat and ALOS PALSAR time series for near real-time deforestation detection. *Remote Sensing* 7, 4973-4996.
<https://doi.org/10.3390/rs70504973>
- Reiche, J., Hamunyela, E., Verbesselt, J., Hoekman, D., Herold, M., 2018. Improving near-real time deforestation monitoring in tropical dry forests by combining dense Sentinel-1 time series with Landsat and ALOS-2 PALSAR-2. *Remote Sensing of Environment* 204, 147-161. <https://doi.org/10.1016/j.rse.2017.10.034>
- Reiche, J., Mullissa, A., Slagter, B., Gou, Y., Tsendbazar, N.E., Odongo-Braun, C., Vollrath, A., Weisse, M.J., Stolle, F., Pickens, A., Donchyts, G., Clinton, N., Gorelick, N., Herold, M., 2021. Forest disturbance alerts for the Congo Basin using Sentinel-1.

- Environmental Research Letters 16, 024005. <https://doi.org/10.1088/1748-9326/abd0a8>
- Réjou-Méchain, M., Barbier, N., Couteron, P., Ploton, P., Vincent, G., Herold, M., Mermoz, S., Saatchi, S., Chave, J., de Boissieu, F., Féret, J.B., Takoudjou, S.M., Péliissier, R., 2019. Upscaling Forest Biomass from Field to Satellite Measurements: Sources of Errors and Ways to Reduce Them, *Surveys in Geophysics*. Springer Netherlands, pp. 881-911. <https://doi.org/10.1007/s10712-019-09532-0>
- Réjou-Méchain, M., Muller-Landau, H.C., Detto, M., Thomas, S.C., Le Toan, T., Saatchi, S.S., Barreto-Silva, J.S., Bourg, N.A., Bunyavejchewin, S., Butt, N., Brockelman, W.Y., Cao, M., Cárdenas, D., Chiang, J.M., Chuyong, G.B., Clay, K., Condit, R., Dattaraja, H.S., Davies, S.J., Duque, A., Esufali, S., Ewango, C., Fernando, R.H.S., Fletcher, C.D., N. Gunatilleke, I.A.U., Hao, Z., Harms, K.E., Hart, T.B., Hérault, B., Howe, R.W., Hubbell, S.P., Johnson, D.J., Kenfack, D., Larson, A.J., Lin, L., Lin, Y., Lutz, J.A., Makana, J.R., Malhi, Y., Marthews, T.R., McEwan, R.W., McMahan, S.M., McShea, W.J., Muscarella, R., Nathalang, A., Noor, N.S.M., Nytch, C.J., Oliveira, A.A., Phillips, R.P., Pongpattananurak, N., Punchi-Manage, R., Salim, R., Schurman, J., Sukumar, R., Suresh, H.S., Suwanvecho, U., Thomas, D.W., Thompson, J., Uriarte, M., Valencia, R., Vicentini, A., Wolf, A.T., Yap, S., Yuan, Z., Zartman, C.E., Zimmerman, J.K., Chave, J., 2014. Local spatial structure of forest biomass and its consequences for remote sensing of carbon stocks. *Biogeosciences* 11, 6827-6840. <https://doi.org/10.5194/bg-11-6827-2014>
- Réjou-Méchain, M., Tanguy, A., Piponiot, C., Chave, J., Hérault, B., 2017. biomass: an r package for estimating above-ground biomass and its uncertainty in tropical forests. *Methods in Ecology and Evolution* 8, 1163-1167. <https://doi.org/10.1111/2041-210X.12753>
- Rignot, E., Salas, W.A., Skole, D.L., 1997. Mapping deforestation and secondary growth in Rondonia, Brazil, using imaging radar and thematic mapper data. *Remote Sensing of Environment* 59, 167-179. [https://doi.org/10.1016/S0034-4257\(96\)00150-2](https://doi.org/10.1016/S0034-4257(96)00150-2)
- Robinson, C., Hou, L., Malkin, K., Soobitsky, R., Czawlytko, J., Dilkina, B., Jojic, N., 2019. Large scale high-resolution land cover mapping with multi-resolution data. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2019-June*, 12718-12727. <https://doi.org/10.1109/CVPR.2019.01301>
- Rockstrom, J., Steffen, W., Noone, K., Persson, A., Chapin Iii, F.S., Lambin, E.F., Lenton, T.M., Scheffer, M., Folk, C., Schellnhuber, H.J., Nykvist, B., de Wit, C.A., Hughes, T., van der Leeuw, S., Rodhe, H., Sorlin, S., Snyder, P.K., Constanza, R., Svedin, U., Falkenmark, M., Karlber, L., Corell, R.W., Fabry, V.J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P., Foley, J.A., 2009. A safe operating space for humanity. *Nature* 461.
- Rogan, J., Franklin, J., Stow, D., Miller, J., Woodcock, C., Roberts, D., 2008. Mapping land-cover modifications over large areas: A comparison of machine learning algorithms. *Remote Sensing of Environment* 112, 2272-2283. <https://doi.org/10.1016/j.rse.2007.10.004>

- Rogan, J.E., Lacher, T.E., 2018. Impacts of Habitat Loss and Fragmentation on Terrestrial Biodiversity, Reference Module in Earth Systems and Environmental Sciences. Elsevier.
- Romero, A.Á., Albuquerque, J., 2018. IMPACTOS SOCIOAMBIENTALES DEL CULTIVO DE PALMA AFRICANA: LOS CASOS MEXICANO Y BRASILEÑO. *Economía & Sociedad* 23, 62-83.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation, International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 234-241. https://doi.org/10.1007/978-3-319-24574-4_28
- Rosete-Vergés, F.A., Pérez-Damián, J.L., Villalobos-Delgado, M., Navarro-Salas, E.N., Salinas-Chávez, E., Remond-Noa, R., 2014. El avance de la deforestación en México 1976-2007. *Madera y Bosques* 20, 21-35.
- Ross, N., 2020. fasterize: Fast Polygon to Raster Conversion, R package version 1.0.3 ed. .
- Roteta, E., Bastarrika, A., Padilla, M., Storm, T., Chuvieco, E., 2019. Development of a Sentinel-2 burned area algorithm: Generation of a small fire database for sub-Saharan Africa. *Remote Sensing of Environment* 222, 1-17. <https://doi.org/10.1016/j.rse.2018.12.011>
- Runyan, C., D'Odorico, P., 2016. Global Deforestation. Cambridge University Press. <https://doi.org/10.1017/CBO9781316471548>
- Ryan, C.M., Hill, T., Woollen, E., Ghee, C., Mitchard, E., Cassells, G., Grace, J., Woodhouse, I.H., Williams, M., 2012. Quantifying small-scale deforestation and forest degradation in African woodlands using radar imagery. *Global Change Biology* 18, 243-257. <https://doi.org/10.1111/j.1365-2486.2011.02551.x>
- Saatchi, S., Marlier, M., Chazdon, R.L., Clark, D.B., Russell, A.E., 2011. Impact of spatial variability of tropical forest structure on radar estimation of aboveground biomass. *Remote Sensing of Environment* 115, 2836-2849. <https://doi.org/10.1016/j.rse.2010.07.015>
- Salinas-Melgoza, M.A., Skutsch, M., Lovett, J.C., 2018. Predicting aboveground forest biomass with topographic variables in human-impacted tropical dry forest landscapes. *Ecosphere* 9. <https://doi.org/10.1002/ecs2.2063>
- Sánchez Colón, S., Flores Martínez, A., Cruz-Leyva, I.A., Velázquez, A., 2009. Estado y transformación de los ecosistemas terrestres por causas humanas. , *Capital natural de México*, vol. II: Estado de conservación y tendencias de cambio, México: Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, pp. 75-129.
- Sandler, T., 1993. Tropical Deforestation: Markets and Market Failures. *Land Economics* 69, 225-233. <https://doi.org/10.2307/3146589>
- Sannier, C., McRoberts, R.E., Fichet, L.-V., 2016. Suitability of Global Forest Change data to report forest cover estimates at national level in Gabon. *Remote Sensing of Environment Journal* 173, 326-338. <https://doi.org/http://doi.org/10.1016/j.rse.2015.10.032>
- Sarker, L.R., Nichol, J.E., 2011. Improved forest biomass estimates using ALOS AVNIR-2 texture indices. *Remote Sensing of Environment* 115, 968-977. <https://doi.org/10.1016/j.rse.2010.11.010>

- Sasaki, N., Putz, F.E., 2009. Critical need for new definitions of “forest” and “forest degradation” in global climate change agreements. *Conservation Letters* 2, 226-232. <https://doi.org/10.1111/j.1755-263x.2009.00067.x>
- Sawangarrearak, S., Thanathamthee, P., 2020. Random forest with sampling techniques for handling imbalanced prediction of university student depression. *Information (Switzerland)* 11, 519. <https://doi.org/10.3390/info11110519>
- Scarpa, G., Gargiulo, M., Mazza, A., Gaetano, R., 2018. A CNN-based fusion method for feature extraction from sentinel data. *Remote Sensing* 10, 1-20. <https://doi.org/10.3390/rs10020236>
- Schmidt, M., Gebhardt, S., Wehrmann, T., Ressler, R., Ruiz, M.M., Tovar, C.M., Morfin, J., Rodríguez, R., Serrano, E., Santos, L., Espinoza, J.A., Elemen, C., Victoria, A., Ornelas, J.L., 2016. Reply to Mas et al.: Comment on Gebhardt et al. MAD-MEX: Automatic wall-to-wall land cover monitoring for the mexican REDD-MRV program using all landsat data. *Remote Sens.* 2014, 6, 3923-3943. *Remote Sensing* 8. <https://doi.org/10.3390/rs8070534>
- SEMARNAP (Secretaría de Medio Ambiente, Recursos Naturales y Pesca), 2000a. Programa de Manejo Reserva de la Biosfera Montes Azules. CONANP, México.
- SEMARNAP (Secretaría de Medio Ambiente, Recursos Naturales y Pesca), 2000b. Programa de Manejo Reserva de la Biosfera Montes Azules. SEMARNAP, INE, México.
- SEMARNAT (Secretaría de Medio Ambiente y Recursos Naturales), 2022. Áreas Naturales Protegidas decretadas. SEMARNAT, México.
- SEMARNAT (Secretaría de Medio Ambiente y Recursos Naturales), 2016. Informe de la Situación del medio ambiente en México. Compendio de estadísticas ambientales, indicadores clave y de desempeño ambiental. Edición 2015., 2015 ed. SEMARNAT (Secretaría de Medio Ambiente y Recursos Naturales), México.
- Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghamisi, P., Homayouni, S., 2020. Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 13, 6308-6325. <https://doi.org/10.1109/JSTARS.2020.3026724>
- Simoës, R., Camara, G., Queiroz, G., Souza, F., Andrade, P.R., Santos, L., Carvalho, A., Ferreira, K., 2021. Satellite Image Time Series Analysis for Big Earth Observation Data. *Remote Sensing* 13, 2428.
- Simula, M., 2009. Towards defining forest degradation: comparative analysis of existing definitions. *Forest Resources Assessment Programme working Paper*, 62-62.
- Singh, M., Evans, D., Chevance, J.B., Tan, B.S., Wiggins, N., Kong, L., Sakhoëun, S., 2018. Evaluating the ability of community-protected forests in Cambodia to prevent deforestation and degradation using temporal remote sensing data. *Ecology and Evolution* 8, 10175-10191. <https://doi.org/10.1002/ece3.4492>
- Singh, M., Evans, D., Friess, D.A., Tan, B.S., Nin, C.S., 2015. Mapping above-ground biomass in a tropical forest in Cambodia using canopy textures derived from Google Earth. *Remote Sensing* 7, 5057-5076. <https://doi.org/10.3390/rs70505057>

- Sinha, S., Jeganathan, C., Sharma, L.K., Nathawat, M.S., 2015. A review of radar remote sensing for biomass estimation. *International Journal of Environmental Science and Technology* 12, 1779-1792. <https://doi.org/10.1007/s13762-015-0750-0>
- Sinha, S., Mohan, S., Das, A.K., Sharma, L.K., Jeganathan, C., Santra, A., Santra Mitra, S., Nathawat, M.S., 2019. Multi-sensor approach integrating optical and multi-frequency synthetic aperture radar for carbon stock estimation over a tropical deciduous forest in India. *Carbon Management* 11, 39-55. <https://doi.org/10.1080/17583004.2019.1686931>
- Skakun, S., Wevers, J., Brockmann, C., Doxani, G., Aleksandrov, M., Batič, M., Frantz, D., Gascon, F., Gómez-Chova, L., Hagolle, O., López-Puigdollers, D., Louis, J., Lubej, M., Mateo-García, G., Osman, J., Peressutti, D., Pflug, B., Puc, J., Richter, R., Roger, J.C., Scaramuzza, P., Vermote, E., Vesel, N., Zupanc, A., Žust, L., 2022. Cloud Mask Intercomparison eXercise (CMIX): An evaluation of cloud masking algorithms for Landsat 8 and Sentinel-2. *Remote Sensing of Environment* 274, 112990. <https://doi.org/10.1016/j.rse.2022.112990>
- Small, D., 2011. Flattening gamma: Radiometric terrain correction for SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing* 49, 3081-3093. <https://doi.org/10.1109/TGRS.2011.2120616>
- Solórzano, J.V., Gallardo-Cruz, J.A., Peralta-Carreta, C., 2020. Potencial del acervo de imágenes Landsat disponible en Google Earth Engine para el estudio del territorio mexicano. *Investigaciones Geograficas* 101, e59821. <https://doi.org/dx.doi.org/10.14350/rig.59821>
- Solórzano, J.V., Mas, J.F., Gao, Y., Gallardo-Cruz, J.A., 2021. Land Use Land Cover Classification with U-Net : Advantages of Combining Sentinel-1 and Sentinel-2 Imagery. *Remote Sensing* 13, 3600. <https://doi.org/10.3390/rs13183600>
- Solórzano, J.V., Meave, J.A., Gallardo-Cruz, J.A., González, E.J., Hernández-Stefanoni, J.L., 2017. Predicting old-growth tropical forest attributes from very high resolution (VHR)-derived surface metrics. *International Journal of Remote Sensing* 38, 492-513. <https://doi.org/10.1080/01431161.2016.1266108>
- Solórzano, J.V., Peralta-Carreta, C., Gallardo-Cruz, J.A., 2018. Identificación y cuantificación del cambio de uso de suelo en las áreas naturales protegidas de la Selva Lacandona y en el municipio de Marqués de Comillas, México.
- Song, J., Gao, S., Zhu, Y., Ma, C., 2019. A survey of remote sensing image classification based on CNNs. *Big Earth Data* 3, 232-254. <https://doi.org/10.1080/20964471.2019.1657720>
- Song, X.P., Hansen, M.C., Stehman, S.V., Potapov, P.V., Tyukavina, A., Vermote, E.F., Townshend, J.R., 2018. Global land change from 1982 to 2016. *Nature* 560, 639-643. <https://doi.org/10.1038/s41586-018-0411-9>
- Souza, C.M., Siqueira, J.V., Sales, M.H., Fonseca, A.V., Ribeiro, J.G., Numata, I., Cochrane, M.A., Barber, C.P., Roberts, D.A., Barlow, J., 2013. Ten-year landsat classification of deforestation and forest degradation in the brazilian amazon. *Remote Sensing* 5, 5493-5513. <https://doi.org/10.3390/rs5115493>
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research* 15, 1929-1958. [https://doi.org/10.1016/0370-2693\(93\)90272-J](https://doi.org/10.1016/0370-2693(93)90272-J)

- Steffen, W., Broadgate, W., Deutsch, L., Gaffney, O., Ludwig, C., 2015. The trajectory of the Anthropocene: The Great Acceleration. *The Anthropocene Review* 2, 81-98. <https://doi.org/10.1177/2053019614564785>
- Steffen, W., Rockström, J., Richardson, K., Lenton, T.M., Folke, C., Liverman, D., Summerhayes, C.P., Barnosky, A.D., Cornell, S.E., Crucifix, M., Donges, J.F., Fetzer, I., Lade, S.J., Scheffer, M., Winkelmann, R., Schellnhuber, H.J., 2018. Trajectories of the Earth System in the Anthropocene. *Proceedings of the National Academy of Sciences of the United States of America* 115, 8252-8259. <https://doi.org/10.1073/pnas.1810141115>
- Steffen, W., Sanderson, R., Tyson, P., Jäger, J., Matson, P., Moore, B., Oldfield, F., Richardson, K., Schellnhuber, H., Turner, B., Wasson, R., 2003. *Global Change and the Earth System: A Planet Under Pressure*.
- Stehman, S.V., 1997. Selecting and Interpreting Measures of Thematic Classification Accuracy. *Remote Sensing of Environment* 62, 77-89.
- Steltzer, H., Welker, J., 2006. Modeling the effect of photosynthetic vegetation properties on the NDVI-LAI relationship. *Ecology and Evolution* 87, 2765-2772.
- Stoian, A., Poulain, V., Inglada, J., Poughon, V., Derksen, D., 2019. Land cover maps production with high resolution satellite image time series and convolutional neural networks: Adaptations and limits for operational systems. *Remote Sensing* 11, 1986. <https://doi.org/10.3390/rs11171986>
- Sudmanns, M., Tiede, D., Augustin, H., Lang, S., 2019. Assessing global Sentinel-2 coverage dynamics and data availability for operational Earth observation (EO) applications using the EO-Compass. *International Journal of Digital Earth*, 1-17. <https://doi.org/10.1080/17538947.2019.1572799>
- Talebiesfandarani, S., Shamsoddini, A., 2022. Global-scale biomass estimation based on machine learning and deep learning methods. *Remote Sensing Applications: Society and Environment* 28. <https://doi.org/10.1016/j.rsase.2022.100868>
- Talukdar, S., Singha, P., Mahato, S., Shahfahad, Pal, S., Liou, Y.A., Rahman, A., 2020. Land-use land-cover classification by machine learning classifiers for satellite observations. A review. *Remote Sensing* 12, 1135. <https://doi.org/10.3390/rs12071135>
- Tamiminia, H., Salehi, B., Mahdianpari, M., Beier, C.M., Klimkowski, D.J., Volk, T.A., 2021. Random Forest Outperformed Convolutional Neural Networks for Shrub Willow Above Ground Biomass Estimation Using Multi-spectral UAS Imagery, *International Geoscience and Remote Sensing Symposium (IGARSS)*. Institute of Electrical and Electronics Engineers Inc., pp. 8269-8272. <https://doi.org/10.1109/IGARSS47720.2021.9553515>
- Tavares, P.A., Beltrão, N.E.S., Guimarães, U.S., Teodoro, A.C., 2019. Integration of sentinel-1 and sentinel-2 for classification and LULC mapping in the urban area of Belém, eastern Brazilian Amazon. *Sensors (Switzerland)* 19. <https://doi.org/10.3390/s19051140>
- Thompson, I.D., Guariguata, M.R., Okabe, K., Bahamondez, C., Nasi, R., Heymell, V., Sabogal, C., 2013. An operational framework for defining and monitoring forest degradation. *Ecology and Society* 18.

- Torres, D.L., Turnes, J.N., Juan, P., Vega, S., Feitosa, R.Q., Silva, D.E., Junior, J.M., Almeida, C., 2021. Deforestation Detection with Fully Convolutional Networks in the Amazon Forest from Landsat-8 and Sentinel-2 Images. *Remote Sensing* 13, 5084.
- Trask, A.W., 2019. *Grokking Deep Learning*. Manning Shelter Island, New York.
- Trebits, R.N., 1987. Synthetic Aperture Radar, in: Eaves, J.L., Reedy, E.K. (Eds.), *Principles of Modern Radar*. Springer US, Boston, MA, pp. 502-537.
- Treitz, P., Rogan, J., 2004. Remote sensing for mapping and monitoring land-cover and land-use change-an introduction. *Progress in Planning* 61, 269-279. [https://doi.org/10.1016/S0305-9006\(03\)00064-3](https://doi.org/10.1016/S0305-9006(03)00064-3)
- Trisasonkko, B.H., Paull, D., 2020. A review of remote sensing applications in tropical forestry with a particular emphasis in the plantation sector, *Geocarto International*. Taylor and Francis Ltd., pp. 317-339. <https://doi.org/10.1080/10106049.2018.1516245>
- Tropek, R., Sedláček, O., Beck, J., Keil, P., Musilová, Z., Šímová, I., Storch, D., 2014. Comment on "High-resolution global maps of 21st-century forest cover change". *Science* 344, 981. <https://doi.org/10.1126/science.1248753>
- Tucker, C., Sellers, P., 1986. Satellite remote sensing of primary production. *International Journal of Remote Sensing* 7, 1395–1416.
- Ulmas, P., Liiv, I., 2020. Segmentation of satellite imagery using U-Net models for land cover classification. *arXiv*, 2003.02899.
- UN (United Nations), 2015. *Transforming our world: The 2030 agenda for sustainable development*. UN, United Nations.
- UN (United Nations), 2016. *Paris Agreement*. United Nations, Paris.
- UNEP (United Nations Environment Programme), CBD (Convention on Biological Diversity), 2010. *Strategic Plan for biodiversity 2011-2020*. Conference to the Parties to the Convention on Biological Diversity (COP) 10. UNEP, CBD, Nagoya.
- UNFCCC (United Nations Framework Convention on Climate Change), 2011. *Consideration of further commitments for Annex I Parties under the Kyoto Protocol*. UNFCCC, Durban, Sudáfrica.
- UNFCCC (United Nations Framework Convention on Climate Change), 2001. *Report of the Conference of the Parties on its seventh session, Bonn*, pp. 1-69. <https://doi.org/10.1503/cmaj.109-3944>
- Ushey, K., Allaire, J.J., Tang, Y., 2020. *reticulate: Interface to 'Python'*.
- Vaca, R.A., Golicher, D.J., Cayuela, L., Hewson, J., Steininger, M., 2012. Evidence of incipient forest transition in Southern Mexico. *PLoS ONE* 7, e42309. <https://doi.org/10.1371/journal.pone.0042309>
- Vaca, R.A., Golicher, D.J., Rodiles-Hernández, R., Castillo-Santiago, M.Á., Bejarano, M., Navarrete-Gutiérrez, D.A., 2019. Drivers of deforestation in the basin of the Usumacinta River: Inference on process from pattern analysis using generalised additive models. *PLoS ONE* 14, 1-21. <https://doi.org/10.1371/journal.pone.0222908>
- Van der Meer, F.D., van der Werff, H.M.A., van Ruitenbeek, F.J.A., 2014. Potential of ESA's Sentinel-2 for geological applications. *Remote Sensing of Environment* 148, 124-133. <https://doi.org/10.1016/j.rse.2014.03.022>
- Van Tricht, K., Gobin, A., Gilliams, S., Piccard, I., 2018. Synergistic use of radar sentinel-1 and optical sentinel-2 imagery for crop mapping: A case study for Belgium. *Remote Sensing* 10, 1642. <https://doi.org/10.3390/rs10101642>

- Vásquez-Grandón, A., Donoso, P.J., Gerding, V., 2018. Forest degradation: When is a forest degraded? *Forests* 9, 1-13. <https://doi.org/10.3390/f9110726>
- Vaughan, D., 2020. rray: Simple Arrays.
- Verbesselt, J., Zeileis, A., Herold, M., 2012. Near real-time disturbance detection using satellite image time series. *Remote Sensing of Environment* 123, 98-108. <https://doi.org/10.1016/j.rse.2012.02.022>
- Vidal, O., López-García, J., Rendón-Salinas, E., 2014. Trends in Deforestation and Forest Degradation after a Decade of Monitoring in the Monarch Butterfly Biosphere Reserve in Mexico. *Conservation Biology* 28, 177-186. <https://doi.org/10.1111/cobi.12138>
- Wagner, F., Rutishauser, E., Blanc, L., Hérault, B., 2010. Effects of plot size and census interval on descriptors of forest structure and dynamics. *Biotropica* 42, 664-671. <https://doi.org/10.1111/j.1744-7429.2010.00644.x>
- Wagner, F.H., Dalagnol, R., Casapia, X.T., Streher, A.S., Phillips, O.L., Gloor, E., Aragão, L.E.O.C., 2020a. Regional Mapping and Spatial Distribution Analysis of Canopy Palms in an Amazon Forest Using Deep Learning and VHR Images. *Remote Sensing* 23, 2225. <https://doi.org/10.3390/rs12142225>
- Wagner, F.H., Sanchez, A., Aidar, M.P.M., Rochelle, A.L.C., Tarabalka, Y., Fonseca, M.G., Phillips, O.L., Gloor, E., Aragão, L.E.O.C., 2020b. Mapping Atlantic rainforest degradation and regeneration history with indicator species using convolutional network. *PloS one* 15, e0229448. <https://doi.org/10.1371/journal.pone.0229448>
- Wagner, F.H., Sanchez, A., Tarabalka, Y., Lotte, R.G., Ferreira, M.P., Aidar, M.P.M., Gloor, E., Phillips, O.L., Aragão, L.E.O.C., 2019. Using the U-net convolutional network to map forest types and disturbance in the Atlantic rainforest with very high resolution images. *Remote Sensing in Ecology and Conservation* 5, 360-375. <https://doi.org/10.1002/rse2.111>
- Whitcraft, A.K., Becker-Reshef, I., Justice, C.O., 2015. A framework for defining spatially explicit earth observation requirements for a global agricultural monitoring initiative (GEOGLAM). *Remote Sensing* 7, 1461-1481. <https://doi.org/10.3390/rs70201461>
- White, J.C., Wulder, M.A., 2014. The landsat observation record of Canada: 1972-2012. *Canadian Journal of Remote Sensing* 39, 455-467. <https://doi.org/10.5589/m13-053>
- Whittle, M., Quegan, S., Uryu, Y., Stüewe, M., Yulianto, K., 2012. Detection of tropical deforestation using ALOS-PALSAR: A Sumatran case study. *Remote Sensing of Environment* 124, 83-98. <https://doi.org/10.1016/j.rse.2012.04.027>
- Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., François, R., Golemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Milton Bache, S., Müller, K., Ooms, J., Robinson, D., Paige Seidel, D., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Tidyverse. *Journal of Open Source Software* 4, 1686. <https://doi.org/10.21105/joss.01686>
- Wilson, A.M., Jetz, W., 2016. Remotely Sensed High-Resolution Global Cloud Dynamics for Predicting Ecosystem and Biodiversity Distributions. *PLoS Biology* 14, e1002415. <https://doi.org/10.1371/journal.pbio.1002415>
- Wright, S.J., 2005. Tropical forests in a changing environment. *Trends in ecology & evolution* 20, 553-560. <https://doi.org/10.1016/j.tree.2005.07.009>
- Wulder, M.A., White, J.C., Loveland, T.R., Woodcock, C.E., Belward, A.S., Cohen, W.B., Fosnight, E.A., Shaw, J., Masek, J.G., Roy, D.P., 2016. The global Landsat archive:

- Status, consolidation, and direction. *Remote Sensing of Environment* 185, 271-283. <https://doi.org/10.1016/j.rse.2015.11.032>
- WWF (World Wildlife Fund), 2020. Palm Oil Buyers Scorecard.
- Xiao, J., Chevallier, F., Gomez, C., Guanter, L., Hicke, J.A., Huete, A.R., Ichii, K., Ni, W., Pang, Y., Rahman, A.F., Sun, G., Yuan, W., Zhang, L., Zhang, X., 2019. Remote sensing of the terrestrial carbon cycle: A review of advances over 50 years. *Remote Sensing of Environment* 233, 111383. <https://doi.org/10.1016/j.rse.2019.111383>
- Xie, Y., Sha, Z., Yu, M., 2008. Remote sensing imagery in vegetation mapping: a review. *Journal of Plant Ecology* 1, 9-23. <https://doi.org/10.1093/jpe/rtm005>
- Ye, Y., Yang, C., Zhu, B., Zhou, L., He, Y., Jia, H., 2021. Improving co-registration for sentinel-1 sar and sentinel-2 optical images. *Remote Sensing* 13, 928. <https://doi.org/10.3390/rs13050928>
- Yi, Y., Zhang, Z., Zhang, W., Zhang, C., Li, W., Zhao, T., 2019. Semantic Segmentation of Urban Buildings from VHR Remote Sensing Imagery Using a Deep Convolutional Neural Network. *Remote Sensing* 11, 1774. <https://doi.org/10.3390/rs11151774>
- Ying, X., 2019. An Overview of Overfitting and its Solutions, *Journal of Physics: Conference Series*, 2 ed. Institute of Physics Publishing. <https://doi.org/10.1088/1742-6596/1168/2/022022>
- Yu, X., Wu, X., Luo, C., Ren, P., 2017. Deep learning in remote sensing scene classification: a data augmentation enhanced convolutional neural network framework. *GIScience and Remote Sensing* 54, 741-758. <https://doi.org/10.1080/15481603.2017.1323377>
- Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., Xu, H., Tan, W., Yang, Q., Wang, J., Gao, J., Zhang, L., 2020. Deep learning in environmental remote sensing: Achievements and challenges. *Remote Sensing of Environment* 241, 111716. <https://doi.org/10.1016/j.rse.2020.111716>
- Yuan, X., Shi, J., Gu, L., 2021. A review of deep learning methods for semantic segmentation of remote sensing imagery. *Expert Systems with Applications* 169, 114417. <https://doi.org/10.1016/j.eswa.2020.114417>
- Zanne, A.E., Lopez-Gonzalez, G., Coomes, D.A., Ilic, J., Jansen, S., Lewis, S.L., Miller, R.B., Swenson, N.G., Wiemann, M.C., Chave, J., 2009. Towards a worldwide wood economics spectrum. *Dryad*. <https://doi.org/10.5061/dryad.234>
- Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J., Atkinson, P.M., 2019a. Joint Deep Learning for land cover and land use classification. *Remote Sensing of Environment* 221, 173-187. <https://doi.org/10.1016/j.rse.2018.11.014>
- Zhang, F., Tian, X., Zhang, H., Jiang, M., 2022. Estimation of Aboveground Carbon Density of Forests Using Deep Learning and Multisource Remote Sensing. *Remote Sensing* 14. <https://doi.org/10.3390/rs14133022>
- Zhang, L., Shao, Z., Liu, J., Cheng, Q., 2019b. Deep learning based retrieval of forest aboveground biomass from combined LiDAR and landsat 8 data. *Remote Sensing* 11. <https://doi.org/10.3390/rs11121459>
- Zhang, L., Zhang, L., Du, B., 2016. Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience and Remote Sensing Magazine* 4, 22-40. <https://doi.org/10.1109/MGRS.2016.2540798>

- Zhang, L., Zhang, X., Shao, Z., Jiang, W., Gao, H., 2023. Integrating Sentinel-1 and 2 with LiDAR data to estimate aboveground biomass of subtropical forests in northeast Guangdong, China. *International Journal of Digital Earth* 16, 158-182.
<https://doi.org/10.1080/17538947.2023.2165180>
- Zhang, P., Ke, Y., Zhang, Z., Wang, M., Li, P., Zhang, S., 2018. Urban land use and land cover classification using novel deep learning models based on high spatial resolution satellite imagery. *Sensors (Switzerland)* 18, 3717. <https://doi.org/10.3390/s18113717>
- Zhang, Y., Ling, F., Foody, G.M., Ge, Y., Boyd, D.S., Li, X., Du, Y., Atkinson, P.M., 2019c. Mapping annual forest cover by fusing PALSAR/PALSAR-2 and MODIS NDVI during 2007–2016. *Remote Sensing of Environment* 224, 74-91.
<https://doi.org/10.1016/j.rse.2019.01.038>
- Zhao, H., Liu, F., Zhang, H., Liang, Z., 2019. Convolutional neural network based heterogeneous transfer learning for remote-sensing scene classification. *International Journal of Remote Sensing* 40, 8506-8527.
<https://doi.org/10.1080/01431161.2019.1615652>
- Zhao, J., Gong, M., Liu, J., Jiao, L., 2014. Deep learning to classify difference image for image change detection. *Proceedings of the International Joint Conference on Neural Networks*, 411-417. <https://doi.org/10.1109/IJCNN.2014.6889510>
- Zhao, P., Lu, D., Wang, G., Liu, L., Li, D., Zhu, J., Yu, S., 2016. Forest aboveground biomass estimation in Zhejiang Province using the integration of Landsat TM and ALOS PALSAR data. *International Journal of Applied Earth Observation and Geoinformation* 53, 1-15. <https://doi.org/10.1016/j.jag.2016.08.007>
- Zhou, Y., Dong, J., Liu, J., Metternicht, G., Shen, W., You, N., Zhao, G., Xiao, X., 2019. Are there sufficient Landsat observations for retrospective and continuous monitoring of land cover changes in China? *Remote Sensing* 11, 1808.
<https://doi.org/10.3390/rs11151808>
- Zhu, C., Yang, X., 1998. Study of remote sensing image texture analysis and classification using wavelet. *International Journal of Remote Sensing* 19, 3197-3203.
- Zhu, L., Suomalainen, J., Liu, J., Hyypä, J., Kaartinen, H., Haggren, H., 2017a. A Review: Remote Sensing Sensors, in: Rustamov, R.B., Hasanova, S., Zeynalova, M.H. (Eds.), *Multi-purposeful Application of Geospatial Data*. InTechOpen, pp. 19-42.
- Zhu, X.X., Tuia, D., Mou, L., Xia, G.S., Zhang, L., Xu, F., Fraundorfer, F., 2017b. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience and Remote Sensing Magazine* 5, 8-36.
<https://doi.org/10.1109/MGRS.2017.2762307>
- Zhu, Z., Woodcock, C.E., 2014. Continuous change detection and classification of land cover using all available Landsat data. *Remote sensing of Environment* 144, 152-171.
- Zou, K.H., Warfield, S.K., Bharatha, A., Tempany, C.M., Kaus, M.R., Haker, S.J., Wells, W.M., Jolesz, F.A., Kikinis, R., 2004. Statistical validation of image segmentation quality based on a spatial overlap index. *Academic radiology* 11, 178–189.
[https://doi.org/https://doi.org/10.1016/s1076-6332\(03\)00671-8](https://doi.org/https://doi.org/10.1016/s1076-6332(03)00671-8)