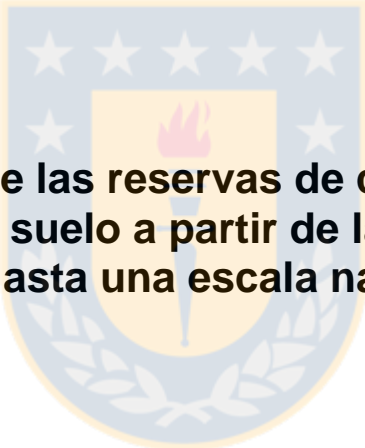Universidad de Concepción
Dirección de Postgrado
Facultad de Agronomía
Programa de Doctorado en Ciencias de la Agronomía

Mapeo digital de las reservas de carbono orgánico del suelo y sobre el suelo a partir de la unidad de muestreo hasta una escala nacional

Tesis para optar al grado de Doctor en Ciencias de la Agronomía

EFRAÍN ALBERTO DUARTE CASTAÑEDA
CHILLÁN-CHILE
2021

Profesor Guía: Erick Zagal Venegas, PhD.
Dpto. de Suelos y Recursos Naturales
Facultad de Agronomía
Universidad de Concepción

Esta tesis ha sido realizada en el Departamento de Suelos y Recursos Naturales de la Facultad de Agronomía, Universidad de Concepción.

Profesor Guía

_____
PhD. Erick Zagal Venegas
Facultad de Agronomía
Universidad de Concepción

Profesor Co-Guía

_____
Dr. Juan Barrera Berrocal
Facultad de Agronomía
Universidad de Concepción

Comisión Evaluadora:

_____
PhD. Francis Dube
Facultad de Ciencias Forestales
Universidad de Concepción

Comisión Evaluadora:

_____
PhD. Alexander Hernández
Agricultural Research Service
United State Department of Agriculture

Directora de Programa

_____
Dr. María Dolores López Belchi
Facultad de Agronomía
Universidad de Concepción

Efraín Alberto Duarte Castañeda, estudiante del Programa Doctorado en Ciencias de la Agronomía, Facultad de Agronomía, Universidad de Concepción, declara ser autor del presente trabajo titulado "**Mapeo digital de las reservas de carbono orgánico del suelo y sobre el suelo a partir de la unidad de muestreo hasta una escala nacional**."

Se permite la reproducción total o parcial, con fines académicos por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento.

DEDICATORIA

A mi madre Edith Alis Castañeda Laínez (†) y mi padre Efraín Duarte Baires.

A mis hijas Claudia Sofía Duarte Garrido y Nohelia Abigaíl Duarte Garrido quienes me inspiraron en este reto y con su inocencia, alegría y amor estuvieron presentes durante todo este proceso; pido a Dios que esta meta que logró su papá las motive en su desarrollo académico.

A mi esposa Claudia Garrido Ruiz por su apoyo, paciencia y motivación.

A mis hermanas, sobrinas y personas que se manifestaron con sus palabras de motivación y apoyo.
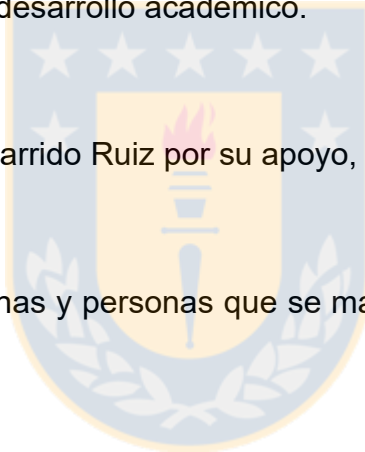
Efraín Alberto Duarte Castañeda

AGRADECIMIENTOS

Agradezco a Dios. por darme el privilegio de alcanzar una meta académica en mi vida.

A mi profesor guía PhD. Erick Zagal Venegas por su confianza, el apoyo, la orientación y supervisión durante todo este proceso de investigación.

A mis asesores el Dr. Juan Barrera Berrocal, el PhD. Francis Dube y el PhD. Alexander Hernández por la orientación y la supervisión en la realización de la tesis. Al MsC. Fabio Casco por su apoyo durante la investigación.

Al Programa de Doctorado de la Facultad de Agronomía de la Universidad de Concepción por el apoyo brindado en el proceso del desarrollo de la pasantía y la publicación de la presente investigación.

TABLA DE CONTENIDO

Índice de tablas

Índice de figuras

Resumen

La construcción de un sistema robusto y transparente para la medición y monitoreo del carbono almacenado en los cinco reservorios considerados por el Grupo Intergubernamental de Expertos sobre el Cambio Climático (IPCC) requiere de la combinación de información proveniente principalmente de dos fuentes: 1) mediciones en terreno, 2) información biofísica derivada de productos satelitales. En la actualidad, la cuantificación de la biomasa de los recursos vegetacionales y el carbono orgánico del suelo (COS) tienen una gran importancia de tipo ecológico y económico, ya que tanto los árboles como el suelo contribuyen significativamente al balance del carbono del planeta Tierra.

Basados en el contexto anterior, en el presente proyecto de investigación se implementaron métodos capaces de medir, monitorear y reportar mediante un mapeo digital del COS y la degradación de los bosques tropicales aplicando enfoques analíticos que integran información biofísica derivada de sensores remotos y datos provenientes de mediciones en terreno, los cuales al ser integrados mediante técnicas de los Sistemas de Información Geográfica (SIG) son capaces de generar modelos espacialmente explícitos del COS y de la degradación del carbono contenido en los bosques.

El enfoque permite a los países implementar metodologías para reportar información transparente y consistente sobre las emisiones y absorciones de carbono en tierras forestales; adicionalmente, estos métodos apoyan a los países en sus reportes nacionales ante la Convención Marco de las Naciones Unidas sobre el Cambio Climático (CMNUCC) y facilita su participación en procesos e iniciativas para la mitigación del cambio climático a través de un mecanismo financiero para la reducción de emisiones causadas por la deforestación y la degradación de los bosques (REDD+).

Abstract

The construction of a robust and transparent system for the measurement and monitoring of the carbon stored in the five reservoirs selected by the Intergovernmental Group of Experts on Climate Change (IPCC) requires the combination of information coming mainly from two sources: 1) field measurements and, 2) biophysical information derived from satellite products. At present, the quantification of the biomass of vegetation resources and soil organic carbon (SOC) are of great ecological and economic importance, since both trees and soil significantly improve the carbon balance of planet Earth.

Based on the above context, we implemented in this research project methods capable of measuring, monitoring and reporting through a digital mapping of the SOC and the degradation of tropical forests applying analytical approaches that integrate biophysical information derived from remote sensors and data from field measurements, which, when integrated by means of Geographic Information Systems (GIS) techniques can generate spatially explicit models of the SOC and the degradation of carbon contained in forests.

The approach allows countries to implement methodologies to report transparent and consistent information on carbon emissions and removals on forest lands; additionally, these methods will support countries in their national

reports to the United Nations Framework Convention on Climate Change (UNFCCC) and facilitate their participation in processes and initiatives for the mitigation of climate change through a financial mechanism for the reduction emissions from deforestation and forest degradation (REDD+).

I. INTRODUCCIÓN GENERAL

1.1 Introducción

Las reservas mundiales del carbono que se encuentra en la biomasa de los bosques y su distribución geográfica están relativamente bien investigadas y se han cuantificado con una certeza razonable en comparación con las reservas de carbono orgánico del suelo (COS: el carbono derivado de la vegetación en descomposición, el crecimiento de hongos y bacterias y las actividades metabólicas de los organismos vivos) (Baccini et al., 2012; Harris et al., 2012; Ruesch & Holly K., 2008; Saatchi et al., 2011). La mayor parte de las políticas nacionales e internacionales para la mitigación del cambio climático se han centrado en conservar el carbono (C) almacenado en los bosques; un ejemplo de esto se observa con el mecanismo desarrollado para proporcionar incentivos financieros denominado REDD+. Sin embargo, además del C de los bosques, el COS es de gran importancia, ya que los suelos almacenan entre dos a tres veces más C en forma orgánica que el C contenido en la atmósfera a nivel mundial (Trumbore, 2009).

Tomando en cuenta la gran cantidad de C almacenado como carbono orgánico en el suelo y, a pesar de una gran cantidad de investigaciones, en la actualidad sigue existiendo una gran incertidumbre sobre la cantidad de las

reservas mundiales de COS, su distribución espacial y las emisiones de C provenientes de los suelos; estas incertidumbres sobre la cantidad de las reservas del COS se debe en gran medida a la poca atención otorgada por parte de los tomadores de decisiones a nivel nacional y a nivel mundial (Gianelle, Oechel, Miglietta, Rodeghiero, & Sottocornola, 2010).

En la actualidad, se identifica una necesidad constante de monitorear los bosques y los suelos, y también se identifica una escasez de datos para un monitoreo eficiente del estado actual de estos recursos (Liang et al., 2016); sumado a esto, el desafío es mucho más complejo cuando las necesidades del monitoreo se trasladan más allá de la biomasa y carbono aéreo de los bosques y éstas trascienden al monitoreo de los cinco reservorios del carbono forestal reconocidos por el IPCC como ser: suelo, raíces, madera muerta, hojarasca y carbono arriba del suelo.

En la actualidad, las tecnologías de teledetección están emergiendo como una herramienta útil para aplicaciones ambientales debido a la disponibilidad de datos históricos, la reducción en el costo de los datos y el aumento de la resolución de las plataformas satelitales. Estas tecnologías junto con datos provenientes de mediciones en terreno serán utilizadas para generar un

modelo para estimar de manera espacial el COS y la degradación de los bosques.

1.1.1 El Carbono Orgánico del Suelo

Después del carbono almacenado en los océanos, el COS es la segunda reserva de carbono más grande en la biosfera. Las estimaciones globales de las reservas de COS reportadas en varios estudios oscilaron entre 504 y 3.000 Pg C (Scharlemann, Tanner, Hiederer, & Kapos, 2014).

En la actualidad, el conocimiento sobre la línea base de COS todavía es bastante limitada, existen estimaciones globales de COS, pero hay una gran variabilidad en los valores informados entre los autores debido a las diferentes fuentes de datos y metodologías utilizadas para realizar las respectivas estimaciones (Henry, Valentini, & Bernoux, 2009; Köchy, Hiederer, & Freibauer, 2015).

La distribución espacial del COS tiene una gran diferencia con respecto al carbono almacenado sobre el suelo, la cual varía con la latitud y las regiones climáticas. La mayor parte del COS se almacena en el hemisferio norte, una

de las regiones más vulnerables al cambio climático. En contraste, las mayores cantidades de carbono forestal (aéreo) se encuentra en los bosques tropicales húmedos y subhúmedos cerca del ecuador (U. Mishra et al., 2013).

A nivel global, las reservas de COS (0 a 30 cm) equivalen a 680 Pg C (Köchy et al., 2015). Con respecto al carbono en las tierras forestales, a nivel mundial las reservas se estiman en 861 ± 66 Pg C, de las cuales 383 ± 30 Pg C (44%) se encuentran en el suelo (hasta 1 m de profundidad), 363 ± 28 Pg C (42%) en la biomasa viva (arriba y abajo del suelo), 73 ± 6 Pg C (8%) en madera muerta, y 43 ± 3 Pg C (5%) en la hojarasca. Geográficamente, 471 ± 93 Pg C (55%) se almacenan en bosques tropicales, 272 ± 23 Pg C (32%) en boques boreales, y 119 ± 6 Pg C (14%) en bosques templados (Pan et al., 2011).

1.1.2 El mapeo del carbono orgánico del suelo

Las tecnologías de teledetección están en continuo desarrollo y nuevos sensores satelitales están surgiendo a un ritmo constante (De Sy et al., 2012). Varios investigadores discuten acerca de los requisitos técnicos para la implementación de sistemas nacionales de monitoreo de suelos y bosques y concluyen sobre la necesidad de utilizar datos satelitales para la observación

de la tierra junto con información proveniente de mediciones realizadas en terreno para construir modelos espacialmente explícitos con altas precisiones (Achard et al., 2007; Martin Herold and Tracy, 2007; Mayaux et al., 2005).

En términos generales, muchos mapas temáticos como el de COS se pueden generar utilizando diferentes métodos de interpolación espacial utilizando como base datos puntuales. En este sentido, los métodos de interpolación espacial proporcionan una herramienta esencial para generar mapas espacialmente explícitos (Lin & Chen, 2004).

Los métodos para predecir mapas basados en covariables como modelo de regresión múltiple, derivan relaciones entre el elemento geográfico objetivo (por ejemplo COS) y covariables como su medio ambiente (morfología del terreno, la vegetación, entre otros) (Ersahin, 2003).

Métodos híbridos, como la regresión kriging y co-kriging, emplean auto correlaciones espaciales y las relacionan entre el elemento geográfico y sus covariables para predecir mapas (Robinson & Metternicht, 2006). En términos generales, muchos coinciden que las exactitudes de los mapas se pueden mejorar al usar técnicas híbridas con información auxiliar tal como litología,

topografía, exposición solar, pendiente etc. (Lin & Chen, 2004; Minasny & McBratney, 2007).

Tomando en cuenta la importancia del suelo como reservorio de C y sumado al auge de los Sistemas de Información Geográfica (SIG) y la facilidad de acceso a datos provenientes de sensores remotos de manera gratuita, en los últimos años, la aplicación de métodos geoespaciales para la construcción de mapas temáticos de COS es diversa, siendo el método híbrido de la Regresión Kriging (RK) uno de los más utilizados (Tabla 1.1). RK: es una técnica de interpolación espacial que combina una regresión de la variable dependiente en variables auxiliares (como parámetros de superficie terrestre, imágenes de sensores remotos y mapas temáticos) (Hengl, Heuvelink, & Rossiter, 2007).

Tabla 1.1 Experiencias internacionales sobre el uso de métodos geoespaciales para la construcción de mapas temáticos de carbono orgánico del suelo.

País	Método utilizado	Autor
Etiopía	Modelo de regresión múltiple	(Nyssen et al., 2008)
Italia	Regresión y Kriging Ordinario	(D'Acqui, Santi, & Maselli, 2007)
Laos, Nigeria	Kriging ordinario (OK), ponderación de distancia inversa (IDW), co-kriging (OCK)	(Phachomphon, Dlamini, & Chaplot, 2010)
Tanzania	Regresión Kriging	(Rossi et al., 2009)
U.S.A.	Espectroscopia de reflectancia de infrarrojo cercano (NIRS), y espectroscopía de reflectancia infrarroja media (MIRS).	(Reeves Iii, Follett, McCarty, & Kimble, 2006)
Kenia	Regresión Kriging	(Mora-Vallejo, Claessens, Stoorvogel, & Heuvelink, 2008)
Trópicos	Funciones de pedotransferencia (OTF)	(Minasny & Hartemink, 2011)
Senegal	Enfoque: Árbol de decisión	(Stoorvogel, Kempen, Heuvelink, & de Bruin, 2009)
Turquía	Regresión Kriging	(Sonmez et al., 2017)
India	Regresión Kriging	(Umakant Mishra et al., 2009)
India	Random Forest y Regresión Kriging	(Hinge, Surampalli, & Goyal, 2018)
Uttar Pradesh, India	Regresión Kriging	(Kumar, Velmurugan, Hamm, & Dadhwal, 2018)
Noreste de China	Regresión Kriging	(Liu et al., 2014)

Fuente: Elaboración propia

Identificar las fuentes de incertidumbre asociadas al mapeo de COS es una tarea muy relevante, en este sentido, se identifica a los diferentes métodos de mapeo como una de las fuentes de incertidumbre, sin embargo, no son la fuente principal de inconsistencia, en muchos de los casos, los mapas producidos con diferentes métodos de mapeo tienen valores comparables y forman una superficie continua de distribución de COS con diferencias aceptables; por lo tanto, se concluye que la fuente principal de incertidumbre e inconsistencias para el mapeo del COS se encuentra en el punto de muestro, especialmente en la distribución y representatividad de las muestras (Guevara et al., 2018).

Uno de los métodos más comunes para obtener datos para estimar la biomasa y el carbono a pequeña o gran escala es mediante un inventario tradicional basado en un muestreo en campo; este tipo de métodos tradicionales de medición en terreno combinado con la teledetección permite obtener información relevante para los inventarios de forma más eficiente; una de las principales ventajas de esta tecnología es la capacidad de obtener datos espacialmente explícitos en grandes áreas de manera oportuna y económica (Aguirre-Salado et al., 2014).

Los inventarios están reconocidos como una fuente de información muy importante, reconoce a los inventarios de carbono como una fotografía del estado actual de cada componente de carbono o compartimiento (troncos, ramas, hojarasca, suelos y raíces) ya que de estos se obtienen mediciones desde el sitio, pero tienen la desventaja de su alto costo de implementación, especialmente cuando estos inventarios se realizan en tierras con difícil acceso como son los terrenos forestales (MacDicken, 1997).

1.1.3 Degradación de los bosques

El mapeo y el monitoreo de la cobertura del suelo han sido ampliamente reconocidos como un objetivo científico importante (Friedl et al., 2002; Hansen, Defries, Townshend, & Sohlberg, 2000; Wulder et al., 2008); sin embargo, en la actualidad el mapeo y monitoreo de la degradación de los bosques son poco estudiados y conocidos, en gran parte porque no hay una definición aceptada internacionalmente sobre degradación forestal y porque los programas internacionales de reducción de emisiones se han centrado en la deforestación debido a que esta es más fácil de detectar y, por lo tanto, más fácil de monitorear. (Pearson, Brown, Murray, & Sidman, 2017).

La degradación de los bosques se puede definir de muchas maneras; la Convención Marco de las Naciones Unidas sobre el Cambio Climático (CMNUCC), define la degradación de los bosques como cualquier pérdida directa y persistente en la densidad de carbono a lo largo del tiempo en las tierras forestales que permanecen como tierras forestales (Penman et al., 2003). La degradación debe considerarse en forma continua dentro de los límites de las definiciones de "bosque"; en este sentido, un desafío clave corresponde en definir primero la reserva de carbono de referencia, contra la cual se puede monitorear el cambio (Guariguata, Nasi, & Kanninen, 2009).

En comparación con los bosques intactos, los bosques degradados suelen ser mucho más bajos en biomasa, ecológicamente menos productivos, fragmentados y con menor cobertura de dosel (Edwards David et al., 2011). La información sobre las estimaciones de las emisiones de Gases de Efecto Invernadero (GEI) provenientes de la degradación de los bosques varía ampliamente y presenta altas incertidumbres como consecuencia de muchos factores como i) diferencias en las definiciones de degradación de bosques y la definición propiamente de bosques, ii) metodologías de estimación e iii) incapacidad de medir y monitorear adecuadamente la degradación de los bosques (Bullock, Woodcock, & Olofsson, 2018).

1.1.4 Medición y monitoreo de la degradación forestal

Las técnicas de detección remota ofrecen beneficios para monitorear la degradación de los bosques en áreas de difícil acceso, como son las regiones de los trópicos (Herold & Skutsch, 2011). La teledetección se ha implementado con éxito para monitorear el cambio de la cubierta forestal y para la contabilidad global del carbono (Hansen et al., 2013; Harris et al., 2012); sin embargo, ninguna metodología disponible ha demostrado ser capaz de cartografiar de forma continua la degradación forestal (Herold et al., 2011). Particularmente en el mecanismo REDD+, la mayoría de los sistemas de monitoreo no cuantifican e informan activamente sobre las emisiones relacionadas con la degradación de los bosques (Hosonuma et al., 2012).

Para el monitoreo de la degradación de los bosques, se requiere una combinación de: (i) la tasa de cambio en la cubierta forestal y (ii) la cantidad de carbono almacenado en el bosque. En este sentido, varios satélites miden la cobertura forestal, la pérdida y alteración del dosel, y las mediciones de la estructura del bosque (Chambers et al., 2007), pero ninguna tecnología satelital puede medir directamente la densidad del carbono (GOFC-GOLD, 2009).

Tomando en cuenta que las tecnologías de teledetección están en continuo desarrollo y nuevos sensores satelitales y aéreos están surgiendo a un ritmo constante (De Sy et al., 2012), especialmente con la apertura en el año 2008 del archivo de datos Landsat del Servicio Geológico de los Estados Unidos (USGS), existe la posibilidad de cuantificar espacialmente el cambio de la superficie terrestre de forma retrospectiva y prospectiva a nivel global (C. Woodcock, 2008). A partir de esta fecha, se han observado aumentos sustanciales en las descargas de imágenes Landsat que llevaron a una rápida expansión de aplicaciones operativas y científicas, al servicio del gobierno, el sector privado y la sociedad civil (Zhu et al., 2019).

Debido a su largo registro de mediciones continuas y su resolución espacial, las imágenes de la serie de satélites Landsat son una de las fuentes de datos más importantes para estudiar diferentes tipos de cambios en la cobertura del suelo, como son la deforestación, la expansión e intensificación de la agricultura, el crecimiento urbano y la pérdida de humedales (Coppin & Bauer, 1996; C. E. Woodcock, Macomber, Pax-Lenney, & Cohen, 2001).

El manejo de una serie de datos derivados de sensores remotos requieren del uso de tecnologías capaces de procesar y analizar diversa magnitud de datos; como ejemplo se puede mencionar a la herramienta Google Earth Engine, la cual es una plataforma basada en la nube para el análisis geoespacial a escala planetaria que aporta las capacidades computacionales masivas de Google a una variedad de problemas sociales de alto impacto que incluyen la deforestación, la sequía, los desastres naturales, las enfermedades, la seguridad alimentaria, la gestión del agua, el monitoreo del clima y protección del medio ambiente (Gorelick et al., 2017).

El monitoreo de los recursos forestales requiere de un análisis consistente para una serie de tiempo, herramientas como TimeSync permiten al intérprete tener en cuenta el contexto espacial, espectral y temporal, así como las imágenes auxiliares de alta resolución de Google Earth al etiquetar píxeles de referencia; TimeSync es una herramienta clave para el análisis de serie de tiempo Landsat porque es casi imposible encontrar datos de referencia con la misma frecuencia temporal, profundidad y cobertura espacial de los datos de Landsat (Kennedy, Yang, & Cohen, 2010).

Los mapas de cobertura actual e histórica de la tierra son los insumos básicos para el establecimiento de sistemas de monitoreo eficientes, la principal fuente de información que requiere la confección de una mapa forestal consiste en la obtención de muestras confiables y representativas de las áreas que se desean mapear; para esto la herramientas como Collect Earth que son de código abierto permiten obtener muestras de cobertura y uso de la tierra para diferentes series de tiempo; esta herramienta es el producto de un esfuerzo de colaboración entre la NASA, la Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO), el Servicio Forestal de los Estados Unidos (USFS) y Google (Saah et al., 2019).

Con base a lo anterior, en la presente investigación, se busca proponer un método confiable e innovador para estimar los contenidos de COS presente en las tierras forestales y la degradación del carbono contenido en la biomasa forestal a través de la combinación de elementos derivados de productos satelitales junto con datos provenientes de inventarios forestales multipropósitos a fin de poder proporcionar información transparente y consistente.

1.2 Hipótesis

Los datos colectados en las unidades de muestreo de los inventarios nacionales forestales en combinación con covariables ambientales derivadas de sensores remotos permiten estimar mediante un enfoque espacial el carbono orgánico del suelo y el carbono sobre el suelo de los bosques tropicales.

1.3 Objetivos

1.3.1 Objetivo general

- Estimar el carbono orgánico del suelo y las pérdidas de carbono por degradación de la biomasa forestal a partir de variables ambientales y fisiográficas obtenidas mediante sensores remotos y variables provenientes de un inventario forestal nacional multipropósito.

1.3.2 Objetivos específicos

- Generar un mapa digital de COS a partir de variables ambientales explicativas espacialmente explícitas y datos provenientes de mediciones realizadas en campo y laboratorio.

- Mapear y estimar la degradación del carbono contenido en grandes áreas de bosque de coníferas mediante un análisis multitemporal para el periodo 1990 - 2018.

- Generar un mapa digital de C forestal (aéreo) a partir de variables ambientales explicativas espacialmente explícitas y datos provenientes de un inventario forestal nacional multipropósito.

Capítulo I: Glosario de abreviaturas, siglas y acrónimos

COS	Carbono Orgánico del Suelo
CMNUCC	Convención Marco de las Naciones Unidas sobre el Cambio Climático
FAO	Organización de las Naciones Unidas para la Agricultura y la Alimentación
GEI	Gases de Efecto Invernadero
IPCC	Grupo Intergubernamental de Expertos sobre el Cambio Climático
REDD+	Reducción de Emisiones causadas por la Deforestación y la Degradación de los bosques
SIG	Sistemas de Información Geográfica
GEE	Google Earth Engine
GIS	Geographic Information Systems
IPCC	Intergovernmental Group of Experts on Climate Change
REDD+	Reduction Emissions from Deforestation and Forest Degradation
RK	Regresión Kriging
SOC	Soil Organic Carbon
UNFCCC	United Nations Framework Convention on Climate Change
USFS	United States Forest Service
USGS	United States Geological Survey
C	Carbono
ha	Hectárea
Pg C	Petagramos de carbono

1.4 Referencias bibliográficas

Achard, F., DeFries, R., Eva, R., Hansen, M., Mayaux, P., & Stibig, H. (2007). Pan-tropical monitoring of deforestation. *Environmental Research Letters, 2*(4), 045022.

Aguirre-Salado, C. A., Treviño-Garza, E. J., Aguirre-Calderón, O. A., Jiménez-Pérez, J., González-Tagle, M. A., Valdéz-Lazalde, J. R., . . . Miranda-Aragón, L. (2014). Mapping aboveground biomass by integrating geospatial and forest inventory data through a k-nearest neighbor strategy in North Central Mexico. *Journal of Arid Land, 6*(1), 80-96. doi:10.1007/s40333-013-0191-x

Angiuli, E., & Trianni, G. (2014). Urban Mapping in Landsat Images Based on Normalized Difference Spectral Vector. *IEEE Geoscience and Remote Sensing Letters, 11*(3), 661-665. doi:10.1109/LGRS.2013.2274327

Baccini, A., Goetz, S. J., Walker, W. S., Laporte, N. T., Sun, M., Sulla-Menashe, D., . . . Houghton, R. A. (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change, 2,* 182. doi:10.1038/nclimate1354https://www.nature.com/articles/nclimate1354#supplementary-information

Badgley, G., Field, C. B., & Berry, J. A. (2017). Canopy near-infrared reflectance and terrestrial photosynthesis. *Science Advances, 3*(3), e1602244. doi:10.1126/sciadv.1602244

Breiman, L. (2001). Random Forests. *Machine Learning, 45*(1), 5-32. doi:10.1023/A:1010933404324

Bullock, E. L., Woodcock, C. E., & Olofsson, P. (2018). Monitoring tropical forest degradation using spectral unmixing and Landsat time series analysis. *Remote Sensing of Environment*, 110968. doi:https://doi.org/10.1016/j.rse.2018.11.011

Chambers, J. Q., Asner, G. P., Morton, D. C., Anderson, L. O., Saatchi, S. S., Espírito-Santo, F. D. B., . . . Souza, C., Jr. (2007). Regional ecosystem structure and function: ecological insights from remote sensing of tropical forests. *Trends in Ecology & Evolution, 22*(8), 414-423. doi:10.1016/j.tree.2007.05.001

Cohen, W. B., Yang, Z., & Kennedy, R. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync — Tools for calibration and validation. *Remote Sensing of Environment, 114*(12), 2911-2924. doi:https://doi.org/10.1016/j.rse.2010.07.010

Coppin, P. R., & Bauer, M. E. (1996). Digital change detection in forest ecosystems with remote sensing imagery. *Remote Sensing Reviews, 13*(3-4), 207-234. doi:10.1080/02757259609532305

D'Acqui, L. P., Santi, C. A., & Maselli, F. (2007). Use of Ecosystem Information to Improve Soil Organic Carbon Mapping of a Mediterranean Island. *Journal of Environmental Quality, 36*(1), 262-271. doi:10.2134/jeq2005.0283

De Sy, V., Herold, M., Achard, F., Asner, G. P., Held, A., Kellndorfer, J., & Verbesselt, J. (2012). Synergies of multiple remote sensing data sources for REDD+ monitoring. *Current Opinion in Environmental Sustainability, 4*(6), 696-706. doi:https://doi.org/10.1016/j.cosust.2012.09.013

Edwards David, P., Larsen Trond, H., Docherty Teegan, D. S., Ansell Felicity, A., Hsu Wayne, W., Derhé Mia, A., . . . Wilcove David, S. (2011). Degraded lands worth protecting: the biological importance of Southeast Asia's repeatedly logged forests. *Proceedings of the Royal Society B: Biological Sciences, 278*(1702), 82-90. doi:10.1098/rspb.2010.1062

Ersahin, S. (2003). Comparing Ordinary Kriging and Cokriging to Estimate Infiltration Rate. *Soil Science Society of America Journal, 67*(6), 1848-1855. doi:10.2136/sssaj2003.1848

Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., . . . Schaaf, C. (2002). Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment, 83*(1), 287-302. doi:https://doi.org/10.1016/S0034-4257(02)00078-0

Gianelle, D., Oechel, W., Miglietta, F., Rodeghiero, M., & Sottocornola, M. (2010). Cataloguing Soil Carbon Stocks. *Science, 330*(6010), 1476. doi:10.1126/science.330.6010.1476-c

GOFC-GOLD. (2009). Reducing greenhouse gas emissions from deforestation and degradation in developing countries: a sourcebook of methods and procedures for monitoring, measuring and reporting GOFC-GOLD Report, Version COP14-2 GOFC-GOLD Project Office, Natural Resources. In. Canada, Alberta, Canada.

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. doi:https://doi.org/10.1016/j.rse.2017.06.031

Guariguata, M. R., Nasi, R., & Kanninen, M. (2009). Forest degradation: it is not a matter of new definitions. *Conservation Letters, 2*(6), 286-287. doi:10.1111/j.1755-263X.2009.00075.x

Guevara, M., Olmedo, G. F., Stell, E., Yigini, Y., Aguilar Duarte, Y., Arellano Hernández, C., Arévalo, G. E., . . . Vargas, R. (2018). No Silver Bullet for Digital Soil Mapping: Country-specific Soil Organic Carbon Estimates across Latin America. In: SOIL Discuss.

Hansen, M. C., Defries, R. S., Townshend, J. R. G., & Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing, 21*(6-7), 1331-1364. doi:10.1080/014311600210209

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., . . . Townshend, J. R. G. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science, 342*(6160), 850. doi:10.1126/science.1244693

Harris, N. L., Brown, S., Hagen, S. C., Saatchi, S. S., Petrova, S., Salas, W., . . . Lotsch, A. (2012). Baseline Map of Carbon Emissions from Deforestation in Tropical Regions. *Science, 336*(6088), 1573. doi:10.1126/science.1217962

Hechenbichler, K., & Schliep, K. (2004). *Weighted k-Nearest-Neighbor Techniques and Ordinal Classification*. Retrieved from https://doi.org/10.5282/UBM/EPUB.1769

Hengl, T., Heuvelink, G. B. M., & Rossiter, D. G. (2007). About regression-kriging: From equations to case studies. *Computers & Geosciences, 33*(10), 1301-1315. doi:https://doi.org/10.1016/j.cageo.2007.05.001

Hengl, T., Heuvelink, G. B. M., & Stein, A. (2004). A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma, 120*(1), 75-93. doi:https://doi.org/10.1016/j.geoderma.2003.08.018

Henry, M., Valentini, R., & Bernoux, M. (2009). Soil carbon stocks in ecoregions of Africa, Biogeosciences Discuss. In (Vol. 6, pp. 797-823): Biogeosciences.

Herold, M., Román-Cuesta, R. M., Mollicone, D., Hirata, Y., Van Laake, P., Asner, G. P., . . . MacDicken, K. (2011). Options for monitoring and estimating historical carbon emissions from forest degradation in the context of REDD+. *Carbon Balance and Management, 6*(1), 13. doi:10.1186/1750-0680-6-13

Herold, M., & Skutsch, M. (2011). Monitoring, reporting and verification for national REDD + programmes: two proposals. *Environmental Research Letters, 6*(1), 014002. doi:10.1088/1748-9326/6/1/014002

Hinge, G., Surampalli, R. Y., & Goyal, M. K. (2018). Prediction of soil organic carbon stock using digital mapping approach in humid India. *Environmental Earth Sciences, 77*(5), 172. doi:10.1007/s12665-018-7374-x

Hosonuma, N., Herold, M., De Sy, V., De Fries, R. S., Brockhaus, M., Verchot, L., . . . Romijn, E. (2012). An assessment of deforestation and forest degradation drivers in developing countries. *Environmental Research Letters, 7*(4), 044009. doi:10.1088/1748-9326/7/4/044009

Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment, 25*(3), 295-309. doi:https://doi.org/10.1016/0034-4257(88)90106-X

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*(3), 440-451. doi:10.1016/S0034-4257(96)00112-5

Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — Temporal segmentation algorithms. *Remote Sensing of*

Environment, 114(12), 2897-2910. doi:https://doi.org/10.1016/j.rse.2010.07.008

Kumar, N., Velmurugan, A., Hamm, N. A. S., & Dadhwal, V. K. (2018). Geospatial Mapping of Soil Organic Carbon Using Regression Kriging and Remote Sensing. *Journal of the Indian Society of Remote Sensing, 46*(5), 705-716. doi:10.1007/s12524-017-0738-y

Köchy, M., Hiederer, R., & Freibauer, A. (2015). Global distribution of soil organic carbon – Part 1: Masses and frequency distributions of SOC stocks for the tropics, permafrost regions, wetlands, and the world. In (Vol. 1, pp. 351-365): SOIL.

Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., . . . Sobrino, J. A. (2013). Satellite-derived land surface temperature: Current status and perspectives. *Remote Sensing of Environment, 131*, 14-37. doi:https://doi.org/10.1016/j.rse.2012.12.008

Liang, X., Kankare, V., Hyyppä, J., Wang, Y., Kukko, A., Haggrén, H., . . . Vastaranta, M. (2016). Terrestrial laser scanning in forest inventories. *ISPRS Journal of Photogrammetry and Remote Sensing, 115*, 63-77. doi:https://doi.org/10.1016/j.isprsjprs.2016.01.006

Lin, G.-F., & Chen, L.-H. (2004). A spatial interpolation method based on radial basis function networks incorporating a semivariogram model. *Journal of Hydrology, 288*(3), 288-298. doi:https://doi.org/10.1016/j.jhydrol.2003.10.008

Liu, L., Wang, H., Dai, W., Lei, X., Yang, X., & Li, X. (2014). Spatial variability of soil organic carbon in the forestlands of northeast China. *Journal of Forestry Research, 25*(4), 867-876. doi:10.1007/s11676-014-0533-3

MacDicken, K. G. (1997). *A Guide to Monitoring Carbon Storage in Forestry and Agroforestry Projects*: Winrock International Institute for Agricultural Development, Forest Carbon Monitoring Program.

MARN. (2017). Manual de Campo Inventario de Biomasa y Carbono en Sistemas No Bosque. In. Santo Domingo, República Dominicana.

Martin Herold and Tracy, J. (2007). Linking requirements with capabilities for deforestation monitoring in the context of the UNFCCC-REDD process. *Environmental Research Letters, 2*(4), 045025.

Masek, J. G., Vermote, E. F., Saleous, N. E., Wolfe, R., Hall, F. G., Huemmrich, K. F., . . . Teng-Kui, L. (2006). A Landsat surface reflectance dataset for North America, 1990-2000. *IEEE Geoscience and Remote Sensing Letters, 3*(1), 68-72. doi:10.1109/LGRS.2005.857030

Mayaux, P., Holmgren, P., Achard, F., Eva, H., Stibig, H.-J., & Branthomme, A. (2005). Tropical forest cover change in the 1990s and options for future monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences, 360*(1454), 373.

Milla, F., Díaz, R., & Emanuelli, P. (2014). Inventario Nacional Forestal Multipropósito de República Dominicana 2014-2015. Elementos de planificación y protocolo para las operaciones de medición. Disponible en: http://www.reddccadgiz.org/documentos/doc_1313366786.pdf. In. El Salvador.

Minasny, B., & Hartemink, A. E. (2011). Predicting soil properties in the tropics. *Earth-Science Reviews, 106*(1), 52-62. doi:https://doi.org/10.1016/j.earscirev.2011.01.005

Minasny, B., & McBratney, A. B. (2007). Spatial prediction of soil properties using EBLUP with the Matérn covariance function. *Geoderma, 140*(4), 324-336. doi:https://doi.org/10.1016/j.geoderma.2007.04.028

Mishra, U., Jastrow, J. D., Matamala, R., Hugelius, G., Koven, C. D., Harden, J. W., . . . Hinzman, L. D. (2013). Empirical estimates to reduce modeling uncertainties of soil organic carbon in permafrost regions: a review of recent progress and remaining challenges. *Environmental Research Letters, 8*(3), 035020. doi:10.1088/1748-9326/8/3/035020

Mishra, U., Lal, R., Slater, B., Calhoun, F., Liu, D., & Van Meirvenne, M. (2009). Predicting Soil Organic Carbon Stock Using Profile Depth Distribution Functions and Ordinary Kriging. *Soil Science Society of America Journal, 73*(2), 614-621. doi:10.2136/sssaj2007.0410

Mora-Vallejo, A., Claessens, L., Stoorvogel, J., & Heuvelink, G. B. M. (2008). Small scale digital soil mapping in Southeastern Kenya. *CATENA, 76*(1), 44-53. doi:https://doi.org/10.1016/j.catena.2008.09.008

Nyssen, J., Temesgen, H., Lemenih, M., Zenebe, A., Haregeweyn, N., & Haile, M. (2008). Spatial and temporal variation of soil organic carbon stocks

in a lake retreat area of the Ethiopian Rift Valley. *Geoderma, 146*(1), 261-268. doi:https://doi.org/10.1016/j.geoderma.2008.06.007

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., . . . Hayes, D. (2011). A Large and Persistent Carbon Sink in the World's Forests. *Science, 333*(6045), 988.

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*(1), 3. doi:10.1186/s13021-017-0072-2

Penman, J., Gytarsky, M., Hiraishi, T., Krug, T., Kruger, D., Pipatti, R., . . . Wagner, F. (2003). *Definitions and methodological options to inventory emissions from direct human-induced degradation of forests and devegetation of other vegetation types*. Kanagawa Prefecture: Institute for Global Environmental Strategies.

Phachomphon, K., Dlamini, P., & Chaplot, V. (2010). Estimating carbon stocks at a regional level using soil information and easily accessible auxiliary variables. *Geoderma, 155*(3), 372-380. doi:https://doi.org/10.1016/j.geoderma.2009.12.020

Reeves Iii, J. B., Follett, R. F., McCarty, G. W., & Kimble, J. M. (2006). Can Near or Mid-Infrared Diffuse Reflectance Spectroscopy Be Used to Determine Soil Carbon Pools? *Communications in Soil Science and Plant Analysis, 37*(15-20), 2307-2325. doi:10.1080/00103620600819461

Robinson, T. P., & Metternicht, G. (2006). Testing the performance of spatial interpolation techniques for mapping soil properties. *Computers and Electronics in Agriculture, 50*(2), 97-108. doi:https://doi.org/10.1016/j.compag.2005.07.003

Rossi, J., Govaerts, A., De Vos, B., Verbist, B., Vervoort, A., Poesen, J., . . . Deckers, J. (2009). Spatial structures of soil organic carbon in tropical forests—A case study of Southeastern Tanzania. *CATENA, 77*(1), 19-27. doi:https://doi.org/10.1016/j.catena.2008.12.003

Ruesch, A., & Holly K., G. (2008). New IPCC Tier-1 Global Biomass Carbon Map For the Year 2000. In. Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Saah, D., Johnson, G., Ashmall, B., Tondapu, G., Tenneson, K., Patterson, M., . . . Chishtie, F. (2019). Collect Earth: An online tool for systematic reference data collection in land cover and use applications. *Environmental Modelling & Software, 118*, 166-171. doi:https://doi.org/10.1016/j.envsoft.2019.05.004

Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T. A., Salas, W., . . . Morel, A. (2011). Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences, 108*(24), 9899. doi:10.1073/pnas.1019576108

Scharlemann, J. P. W., Tanner, E. V. J., Hiederer, R., & Kapos, V. (2014). Global soil carbon: understanding and managing the largest terrestrial carbon pool. *Carbon Management, 5*(1), 81-91. doi:10.4155/cmt.13.77

SEMARENA. (2008). Perfil nacional para evaluar las capacidades nacionales de implementación del Principio 10 de la Declaración de Rio. In (pp. 119). Santo Domingo, República Dominicana.

Sobrino, J. A., & Raissouni, N. (2000). Toward remote sensing methods for land cover dynamic monitoring: Application to Morocco. *International Journal of Remote Sensing, 21*(2), 353-366. doi:10.1080/014311600210876

Sonmez, B., Ozbahce, A., Kececi, M., Akgul, S., Aksoy, E., & Madenoglu, S. (2017). Turkey's national geospatial soil organic carbón information system. In. Rome, Italy.

Souza, C., Siqueira, V. J., Sales, H. M., Fonseca, V. A., Ribeiro, G. J., Numata, I., . . . Barlow, J. (2013). Ten-Year Landsat Classification of Deforestation and Forest Degradation in the Brazilian Amazon. *Remote Sensing, 5*(11). doi:10.3390/rs5115493

Souza, C. M., Roberts, D. A., & Cochrane, M. A. (2005). Combining spectral and spatial information to map canopy damage from selective logging and forest fires. *Remote Sensing of Environment, 98*(2), 329-343. doi:https://doi.org/10.1016/j.rse.2005.07.013

Stoorvogel, J. J., Kempen, B., Heuvelink, G. B. M., & de Bruin, S. (2009). Implementation and evaluation of existing knowledge for digital soil mapping in Senegal. *Geoderma, 149*(1), 161-170. doi:https://doi.org/10.1016/j.geoderma.2008.11.039

Trumbore, S. (2009). Radiocarbon and Soil Carbon Dynamics. *Annual Review of Earth and Planetary Sciences, 37*(1), 47-66. doi:10.1146/annurev.earth.36.031207.124300

Vapnik, V. (1998). The Support Vector Method of Function Estimation. In J. A. K. Suykens & J. Vandewalle (Eds.), *Nonlinear Modeling: Advanced Black-Box Techniques* (pp. 55-85). Boston, MA: Springer US.

Vermote, E., Justice, C., Claverie, M., & Franch, B. (2016). Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sensing of Environment, 185*, 46-56. doi:https://doi.org/10.1016/j.rse.2016.04.008

Viscarra Rossel, R. A., Webster, R., Bui, E. N., & Baldock, J. A. (2014). Baseline map of organic carbon in Australian soil to support national carbon accounting and monitoring under climate change. *Global Change Biology, 20*(9), 2953-2970. doi:10.1111/gcb.12569

Woodcock, C. (2008). 1018 Free Access to Landsat Imagery Teach by the Book Science Education. *Science, 80*(320), 1011-1012.

Woodcock, C. E., Macomber, S. A., Pax-Lenney, M., & Cohen, W. B. (2001). Monitoring large areas for forest change using Landsat: Generalization across space, time and Landsat sensors. *Remote Sensing of Environment, 78*(1), 194-203. doi:https://doi.org/10.1016/S0034-4257(01)00259-0

Wulder, M. A., White, J. C., Goward, S. N., Masek, J. G., Irons, J. R., Herold, M., . . . Woodcock, C. E. (2008). Landsat continuity: Issues and opportunities for land cover monitoring. *Remote Sensing of Environment, 112*(3), 955-969. doi:https://doi.org/10.1016/j.rse.2007.07.004

Xu, H. (2008). A new index for delineating built-up land features in satellite imagery. *International Journal of Remote Sensing, 29*(14), 4269-4276. doi:10.1080/01431160802039957

Zhu, Z., Wulder, M. A., Roy, D. P., Woodcock, C. E., Hansen, M. C., Radeloff, V. C., . . . Scambos, T. A. (2019). Benefits of the free and open Landsat data policy. *Remote Sensing of Environment, 224*, 382-385. doi:https://doi.org/10.1016/j.rse.2019.02.016

II. MONITORING APPROACH FOR TROPICAL CONIFEROUS FOREST DEGRADATION USING REMOTE SENSING AND FIELD DATA

Monitoring Approach for Tropical Coniferous Forest Degradation Using Remote Sensing and Field Data

Efraín Duarte [1,2], Juan A. Barrera [1], Francis Dube [3], Fabio Casco [4], Alexander J. Hernández [5], and Erick Zagal [1,*]

[1] Department of Soils and Natural Resources, Faculty of Agronomy, Universidad de Concepción,
Vicente Méndez 595, Casilla 537, Chillán 3812120, Chile; efrainduarte@udec.cl (E.D.); jbarrera@udec.cl (J.A.B.).

[2] Doctoral Program in Agronomic Sciences, Faculty of Agronomy, Universidad de Concepción,
Vicente Méndez 595, Casilla 537, Chillán 3812120, Chile.

[3] Department of Silviculture, Faculty of Forest Sciences, Universidad de Concepción, Victoria 631,
Casilla 160-C, Concepción 4030000, Chile; fdube@udec.cl

[4] Food and Agriculture Organization (FAO) of the United Nations, Tegucigalpa, Honduras; fabio.cascogutierrez@fao.org

[5] Wildland Resources Department, Utah State University, Logan, UT 84322-5230, USA; alex.hernandez@usu.edu

* Correspondence: ezagal@udec.cl; Phone.: +569-42-2208853

Abstract

Current estimates of CO_2 emissions from forest degradation are generally based on insufficient information and are characterized by high uncertainty, while a global definition of ´forest degradation´ is currently being discussed in the scientific arena. This study proposes an automated approach to monitor degradation using Landsat time series. The methodology was developed using Google Earth Engine (GEE) and applied in a pine forest area of the Dominican Republic. Land cover change mapping was conducted using the Random Forest (RF) algorithm and resulted in a cumulative overall accuracy of 92.8%. Forest degradation was mapped with a 70.7% user accuracy and a 91.3%

producer accuracy. Estimates of the degraded area had a margin of error of 10.8%. A number of 344 Landsat collections, corresponding to the period from 1990 to 2018, were used in the analysis. Additionally, 51 sample plots from a Forest Inventory were used. The carbon stocks and emissions from forest degradation were estimated using the RF algorithm with an R^2 of 0.78.

GEE proved to be an appropriate tool to monitor the degradation of tropical forests, and the methodology developed herein is a robust, reliable, and replicable tool that could be used to estimate forest degradation and improve Monitoring, Reporting, and Verification (MRV) systems under the Reducing Emissions from Deforestation and Forest Degradation (REDD+) mechanism.

Keywords: Forest Degradation; REDD+; Google Earth Engine; Random Forest; Dynamic Land Cover Change; Landsat; Carbon; MRV.
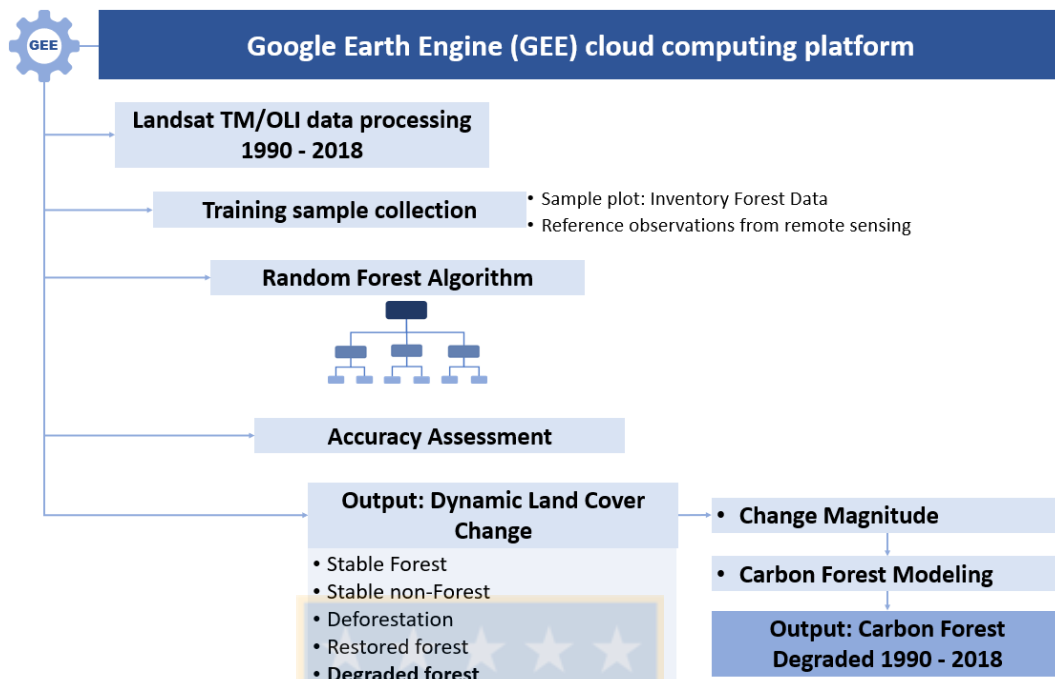
Figure 2.1. Graphical abstract

Fuente: Elaboración propia

2.1. Introduction

Forest monitoring has been an important scientific objective mainly due to the large number of ecosystem services that serve humanity. One of the most efficient methods to monitor them is through geographically explicit and consistent mapping over time [1,2].

Currently, forest monitoring has been focused on quantifying deforestation; spatial representation and the monitoring of forest degradation are poorly studied, mainly because there is no clear, standardized, and

recognized definition of 'forest degradation' globally [3]. Additionally, international initiatives and programs that finance forest emissions reductions have focused on estimating deforestation, which is easier to measure and monitor than forest degradation [4].

The first step in measuring forest degradation is to define key concepts such as i) the forest and ii) forest degradation. These concepts have been widely debated [5], and their definitions vary between institutions and organizations. The Intergovernmental Panel on Climate Change (IPCC) defines forest degradation as *"direct human-induced long-term loss (persisting for X years or more) of at least Y% of forest carbon stocks [and forest values] since time T and not qualifying as deforestation or an elected activity under Article 3.4 of the Kyoto Protocol"* [6]. Thus, defining a carbon (C) stock baseline is the first step to monitor this continual C loss.

Forest degradation, along with deforestation, has been reported as the second most common source (after fuel combustion) of global anthropogenic greenhouse gas (GHG) emissions, comprising over 17% of global CO_2 emissions [7,8]. The assessment and reporting of CO_2 emissions caused by forest degradation is a crucial step to achieve the goals under international policies such as the REDD+ mechanism, which mainly aims to reduce the emissions from deforestation and forest degradation. The REDD+ mechanism also includes (i) the conservation of forest carbon stocks, (ii) the sustainable management of forests, and (iii) the enhancement of forest carbon stocks [9].

To monitor the five REDD+ activities, it is essential to have a robust and transparent system for Measuring, Reporting, and Verifying (MRV) GHG emissions alongside with methods that combine terrestrial and satellite techniques for the measurement and monitoring of emissions and the removal of C from forest resources [10,11].

To estimate and report the GHG emissions and their removal, the principal recommendations of the IPCC are to use activity data and emission factors at a national scale [12]. The most practical method of mapping land-cover changes at a national scale is to use spatially explicit data through remote sensing (RS) [13].

RS methods to monitor deforestation have been successfully used for global C accounting [4,14]. However, unlike deforestation, no available method has been reliable for monitoring degradation [15], thereby restricting C accounting [16]. Forest degradation monitoring requires estimating the rate of change rate for i) the forest cover and ii) the forest C stock. In this sense, satellite imagery is key to monitoring changes in forest cover (density, structure, and composition) but fails to monitor the C stock [17]. Therefore, in addition to using satellite imagery, there is a clear need to employ data from field measurements to achieve more accurate estimates of CO_2 emissions.

Since RS technologies are advancing and new satellite are emerging at a constant pace [18], particularly since the United States Geological Survey

33

(USGS) adopted a free and open Landsat in 2008, it is now possible to spatially quantify changes in the Earth's surface retrospectively and prospectively at a global scale [13]. Due to their long record of continuous measurements and high spatial resolution, Landsat series satellite images are some of the most important information sources for studying the different classes of land cover change [19] and has facilitated the characterization of land change using time series of Landsat images [20].

Applying Landsat time series requires the use of technologies with high capacity to access, storage and tools to perform analysis of large data sets, these special technologies are available at no cost to everyone through the Google Earth Engine (GEE). GEE is a cloud computing platform consisting of a tool for analyzing geospatial information, through which we can analyze the land use and land change use by applying highly interactive algorithms on a global scale with a code editor via the Javascript API [21]. Cloud-based Landsat imagery has been widely used for mapping land cover and especially deforestation [22-26]. However, forest degradation mapping has been rarely investigated using satellite imagery and data sampling using GEE.

The current study developed a method for automated forest degradation measuring and monitoring using field data from a National Forest Inventory and a time series of Landsat images using GGE. The objective of this study was to provide a dynamic land cover change map (included pine forest degradation) in the Dominican Republic and to map C stocks of pine forests to

estimate CO_2 emissions from forest degradation for the 1990–2018 time period. Our processing and mapping algorithm uses Landsat data to characterize the forest cover's extension, loss, and degradation. Our approach aims at determining the magnitude of forest degradation (cover change), the spatial distribution of C stocks in the forest, and the amount of CO_2 emissions from forest degradation per unit area for a given period of analysis.

2.2 Materials and Methods

2.2.1. Study Area

Our research area is located in the Dominican Republic, mainly in two thirds of the so-called Hispaniola Island in the east, which is the second largest island in the Greater Antilles. The territory of the country covers 48,198 km^2 (18°28′35″N and 69°53′36″W) **(Figure 2.2).** The Dominican Republic has diverse bioclimatic and topographic zones, ranging from dry regions, where precipitation reaches 450 mm yr^{-1}, to humid regions, where precipitation reaches 2,500 mm yr^{-1}, at altitudes over 3,000 m.a.s.l. The northwest–southeast trending mountain range includes the highest peaks in the Caribbean, Pico Duarte (3,098 m.a.s.l) [27]. This wide variety of geographic conditions has given rise to diverse ecosystems and habitats, including arid, semi-arid, humid, and tropical sub-humid zones [28].

The study was carried out in the pine forests of the Dominican Republic, which cover 3,287 km^2. Most of this area lies in the Cordillera Central (the highest elevation mountain range on the island) and comprises four large protected areas that were declared national parks (NPs): (i) NP Armando Bermúdez, (ii) NP José del Carmen Ramírez, (iii) NP Valle Nuevo, and (iv) NP Sierra de Bahoruco.
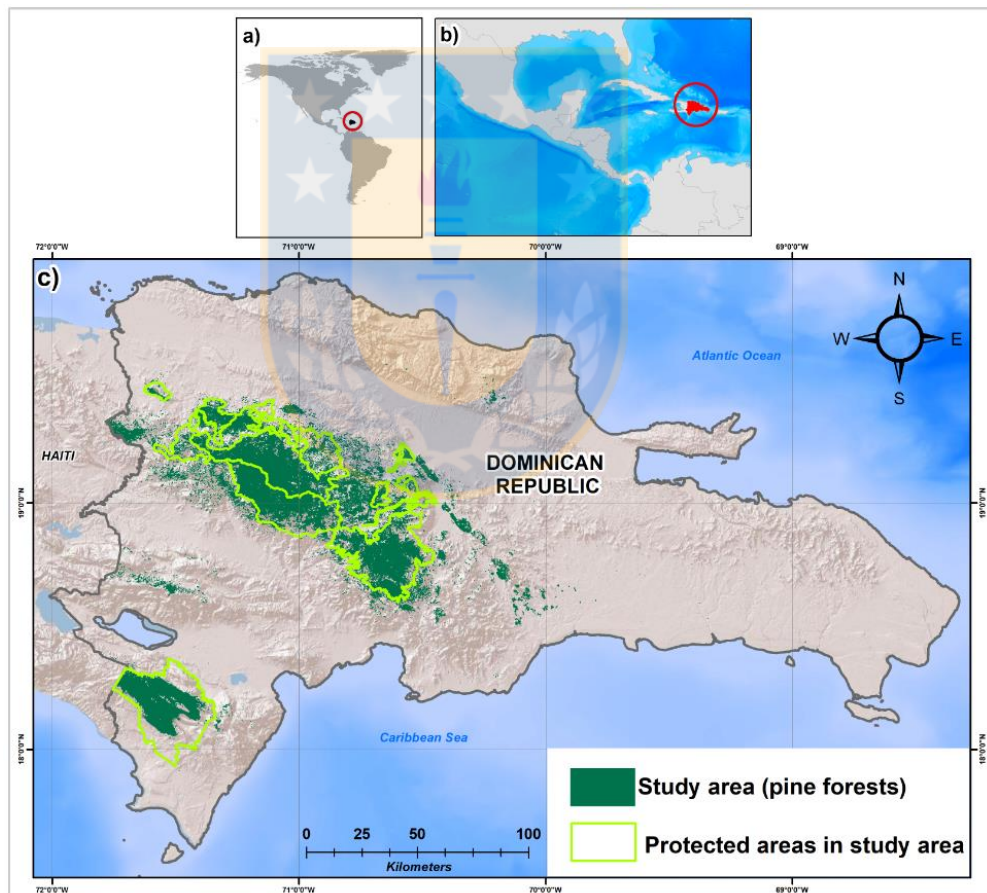
Figure 2.2. Study area (a) General location of the Dominican Republic, (b) regional location, and (c) study area including the protected areas.
Fuente: Elaboración propia

The water service for human consumption and agricultural use in most of the country is one of the main environmental services offered by the Central Cordillera to the Dominican Republic. The vegetation patterns vary mainly due to large-scale climatic factors, such as the direction of the northeast or southeast winds. This mountain range features the principal pine forests of the country. The higher sites of the mountain range include moist broadleaf forests, while the windward side includes forests of West Indies pine (*Pinus occidentalis* Swartz) [29]. This pine is endemic to the island of Hispaniola (19°N, 71°W), although it has been known to scientists for more than 200 years and still covers extensive areas of the Dominican Republic and Haiti [30].

Recent studies conducted under the REDD+ program by the Ministry of the Environment and Natural Resources (MARN) have indicated that the illegal extraction of wood for firewood and charcoal to be used as fuel, timber, and weak management are the principal drivers of pine forest degradation in the Dominican Republic [31].

2.2.2 Forest, deforestation, and degradation definitions

In recent times, the definition of ´forest´ has taken on particular relevance due to the challenges of countries to monitor the CO_2 emissions from the forest sector as part of the objectives to establish robust MRV systems for REDD+. In general, the definitions of 'forest' include references to threshold parameters that include the minimum area of land, minimum tree height, and minimum

canopy cover. Many countries are aligned with the minimum thresholds described by the United Nations Food and Agriculture Organization's (FAO) Global Forest Resource Assessment (FRA). Through the FRA, since 2000, all countries have aligned themselves to adopt a definition of ´forest´ with common parameters, such as (i) a canopy coverage of more than 10%, (ii) trees of 5 m, and (iii) land of at least 0.5 ha [32]. The present study subscribes to the definition of 'forest' adopted by the Dominican Republic, with a focus on pine forests in accordance with the Reference Emission Levels/Forest Reference Levels (FREL/FRL): *"land of at least 0.5 ha covered by pine trees higher than 5 m and with a canopy cover of more than 30%, or by trees able to reach these thresholds, and predominantly under forest land use, this excludes land that is mainly under agricultural or urban land uses"* [33].

Based on the definition of forest described above, in the current study, forest degradation is defined as *"the loss of carbon content in forest lands that remain as forest lands with a decrease in canopy cover that does not qualify as deforestation and that can be caused by anthropogenic activities".* Forest degradation has a human-induced negative impact on carbon stock changes; our operational definition for measuring forest degradation is based on indicators such as forest structure (changes in canopy cover) [34] that affect the ability of the forest to store carbon under natural conditions [35]. These definitions demonstrate that deforestation and forest degradation involve different conditions, processes, and concepts. Deforestation suggests a change in land use from forest to non-forest land use, altering the original

structure and environment of the forest, while degradation occurs in forest lands that are maintained as forest lands but suffer losses in their forest ecosystem functions **(Figure 2.3).**
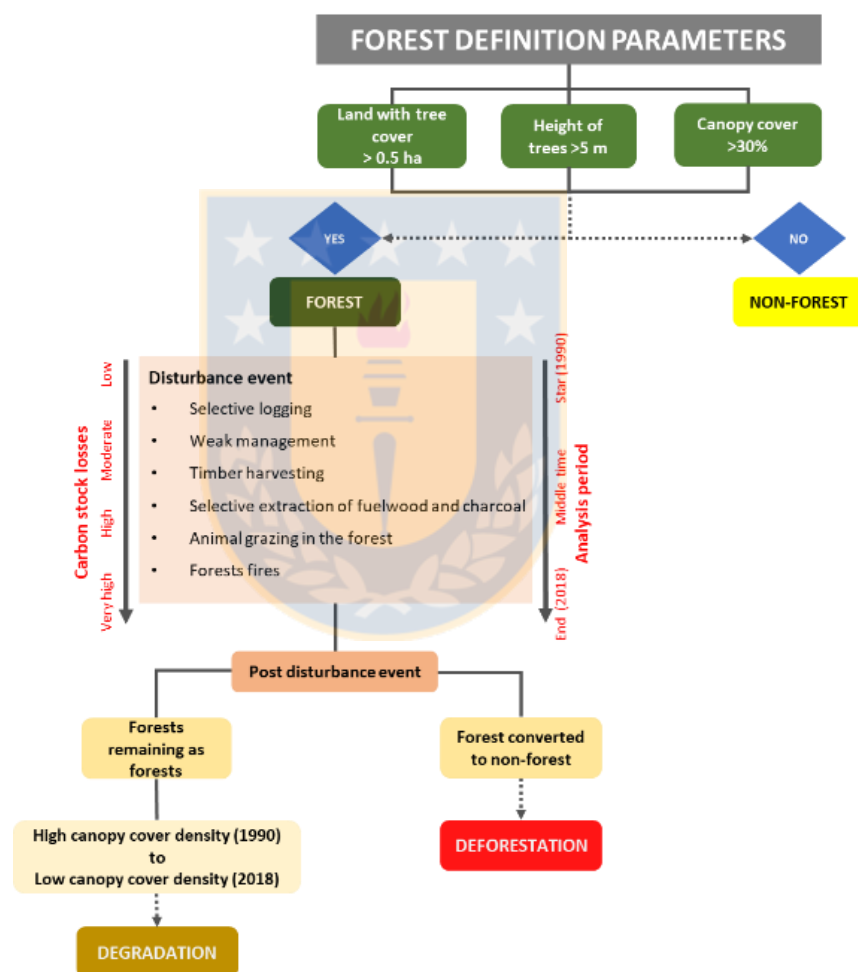
Figure 2.3. Main parameters and elements that interact in forest degradation in the Dominican Republic.

Fuente: Elaboración propia

2.2.3. Method

Once the baseline data were collected and the key concepts (forest, deforestation, and forest degradation) were defined, a method to quantify the degradation of pine forests in the Dominican Republic was developed. The overall structure of the method **(Figure 2.4)** consists of four stages: (i) the preprocessing and selection of Landsat images, (ii) the computation of the spectral indices to map land cover for the 1990–2018 period, (iii) changing the magnitude of mapping, and (iv) mapping the carbon stocks in pine forests. The entire process was accompanied by an accuracy analysis for each step, in which a classification, probability, and regression model was applied using the Smile Random Forest (RF) algorithm. The GEE cloud-based platform was also used in our research.
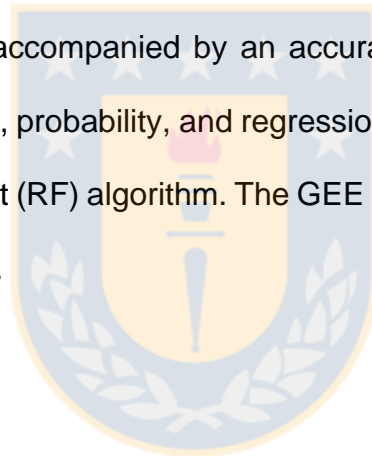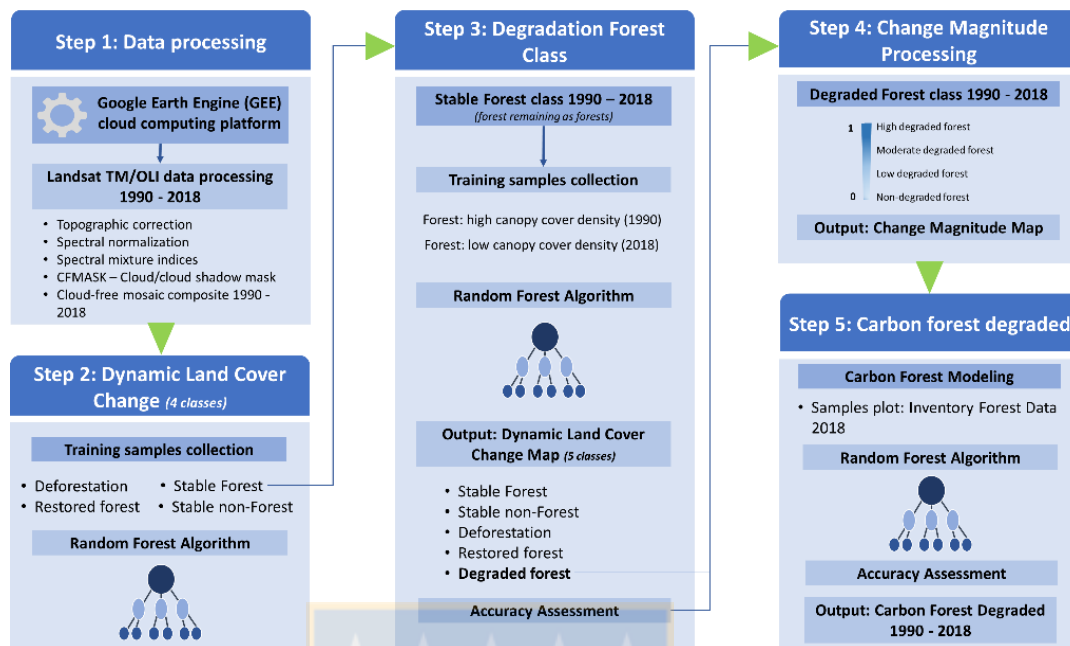
Figure 2.4. Flowchart: Representation of the methodology used in our research.
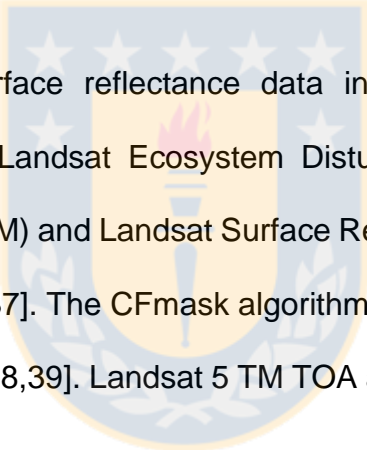
Fuente: Elaboración propia

Our models were developed using Landsat Thematic Mapper (TM) and Operational Land Imager (OLI) applying the RF algorithm in GEE to generate a dynamic land cover change map, degraded forest map, and carbon forest map. The model was trained and validated using sample plots from the forest inventory and satellite images

2.2.4. Reference Data

Landsat TM/OLI Data processing

We used Landsat-5 TM and Landsat 8 OLI surface reflectance data with 16 day and 30 m resolutions (available in the GEE computing platform) [21]. All Landsat-5TM surface reflectance data from year 1990 ± 0.5 (a total of 22 images) and Landsat-8 OLI surface reflectance data from year 2018 ± 0.5 (a total of 322 images) available in GEE were used in this study.

The Landsat surface reflectance data in GEE were atmospherically corrected using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) (TM) and Landsat Surface Reflectance Corrected (LaSRC) (OLI) algorithms [36,37]. The CFmask algorithm was used to mask the clouds and cloud shadows [38,39]. Landsat 5 TM TOA and OLI TOA collections were also used [40].

Digital elevation data were obtained from the Shuttle Radar Topography Mission (SRTM) [41] in GEE. These data have a 30 m spatial resolution. SRTM data were used to calculate the topographic slope and elevation. In addition, the empirical Earth rotation model (ERM) was used as a basis to apply a terrain illumination correction algorithm [42], which allowed us to topographically correct each image. For reflectance images, we used the medoid method [43] **(Figure 2.5).**

Once the images were pre-processed, a composite mosaic was developed. This mosaic was formed by combining spatially overlapping images into a single image based on a function of multiple spectral and temporal aggregation ranges [44]. This mosaic (multi-band and multi-date) was built with the images of the years 1990 and 2018.
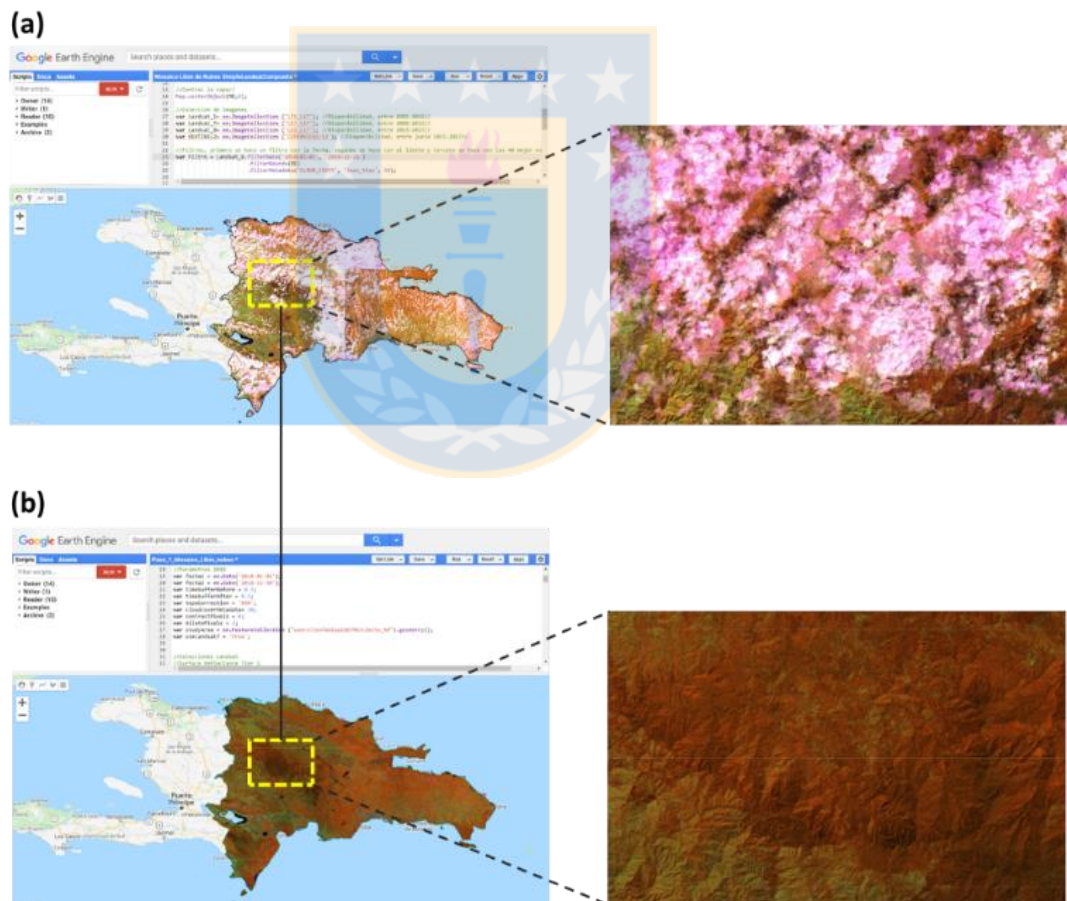
Figure 2.5. Mosaic of seasonal images for the Dominican Republic in 2018. (a) Original composite mosaic, (b) Medoid composite (Shortwave Infrared 1 (SWIR1), Near-Infrared (NIR) and Red band).

Fuente: Elaboración propia

Field Inventory Data

The reference field data included in this study are based on the National Forest Inventory (NFI) collected by the Ministry of the Environment and Natural Resources (MARN) of the Dominican Republic with the support of the REDD/CCAD-GIZ program and the World Bank's Forest Carbon Partnership Facility (FCPF) (https://www.forestcarbonpartnership.org/country/dominican-republic (ERPD document, September, 2019)) [45]. In 2012, the Dominican Republic designed its MRV strategy. This strategy proposed two major lines of monitoring forest resources: (i) satellite monitoring and (ii) terrestrial monitoring. For terrestrial monitoring, the country executed an NFI between 2017 and 2018 with a plan to develop permanent sampling plots to be measured every 5 years according to the action plan of the country's MRV System. The NFI of the Dominican Republic contains 404 sampling units located in the different forest classes, such as moist broadleaf forests (204 plots), subdivided into semi-humid broadleaf forests (117 plots), humid broadleaf forests (76 plots), and broadleaf cloud forest (11 plots); pine forests (59 plots), subdivided into high canopy cover density (19 plots) and low canopy cover density (40 plots); dry forests (71 plots); and mangrove forests (70 plots).

The plot is rectangular with a size of 0.125 hectares (ha) (25 m × 50 m). Different forest characteristics and topographical factors were measured at all plots (tree species, height, diameter at breast height (DBH), soil organic matter, number of trees, geographical coordinates, elevation, and slope). To

design the NFI, the methodology proposed by the REDD/CCAD-GIZ program was used [46,47]. For our study, we used the data from 51 plots located in the pine forest areas. More information about the methodology and the results of the NFI is available in the FREL/FRL submission of the Dominican Republic [33] to REDD+ UNFCCC (**Figure 2.6**) (**Appendix D Table 2.10 and Appendix E Figure 2.15**).
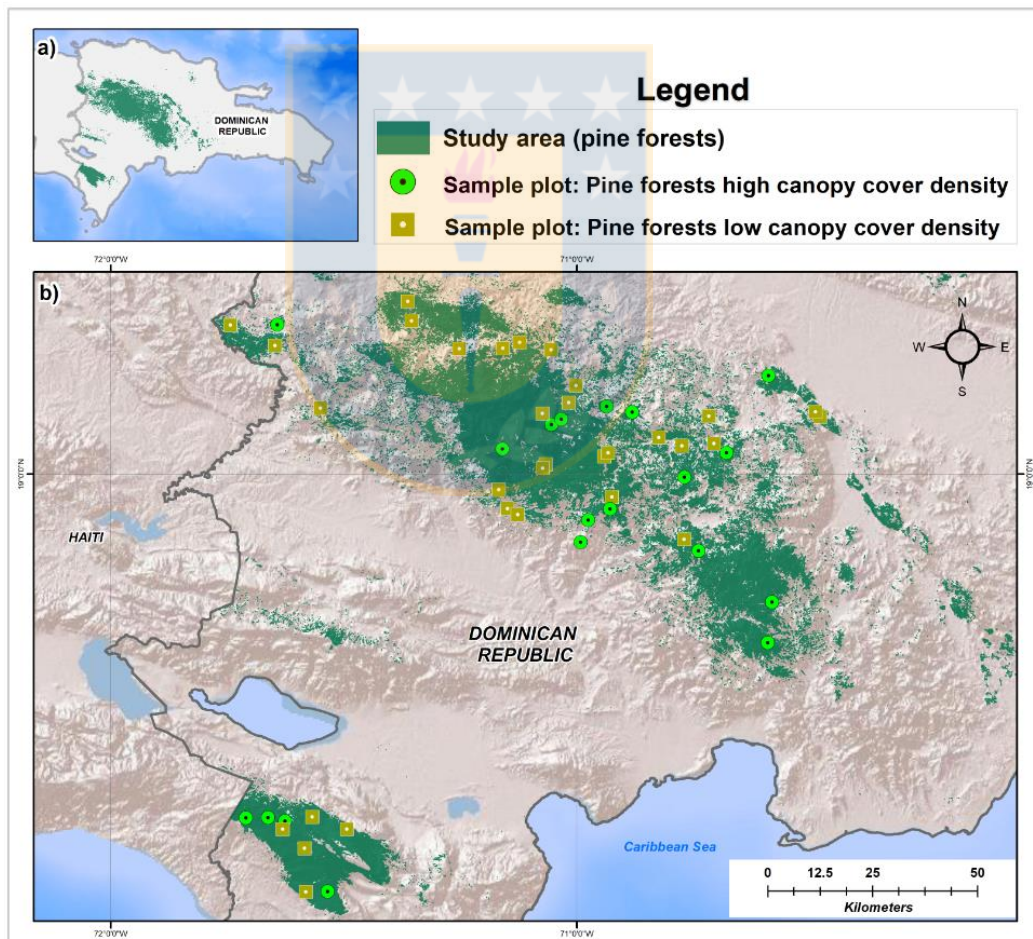
Figure 2.6. (a) General location of the Dominican Republic; (b) location of permanent plots in pine forests with low and high canopy cover density.

Fuente: Elaboración propia

2.2.5 Classification of dynamic land cover change

One of the most relevant tasks for RS is land cover mapping. Different spectral indices are used to improve such mapping techniques [49]. For land cover change mapping, we generated a composite mosaic for the 1990–2018 period, and spectral indices were selected based on the known characteristics of land cover classes. Landsat time series and spectral analyses were used to detect deforestation and degradation forests.

Classes that were determined in the dynamic land cover map correspond to a stable forest, stable non-forest, degradation, deforestation, and restored forest. Seven different vegetation indices were used to monitor the dynamics of forest change during the 1990–2018 period. Among the most relevant indices used are the Enhanced Vegetation Index (EVI), which was designed to enhance the vegetation signal with improved sensitivity in high biomass areas [50]; the Soil Adjust Vegetation Index (SAVI) [51]; and the Normalized Difference Fraction Index (NDFI), which was constructed to highlight degraded or cleared forest areas. The NDFI values in intact forests are expected to be high (i.e., approximately 1) due to the combination of high "Green Vegetation (GV)" and low Non-Photosynthetic Vegetation (NPV) and soil values [52]. The spectral indices used are closely related to the land cover defined in our research; **Appendix C Table 2.9** details each index used.

Land cover change samples

Training samples for land cover change classification were derived from a visual analysis using Landsat from GEE and high spatial resolution images from Google Earth (GE) **(Figure 2.7).**

First, we established four cover change classes: stable forest, stable non-forest, deforestation, and forest restoration. We allocated 97 samples to areas that appeared to be stable forest or stable non-forest, and the remaining 15 samples were allocated to land cover change. The second step was to detect forest degradation based on the stable forest class, commonly called the 'forest mask´, following the criteria established in the definition of a forest in our study. We focused on two classes: non-degraded forest and degraded forest. A total of 90 samples were allocated to areas with disturbances observed in the 1990–2018 period. Reference samples were randomly distributed over each land cover class in the study area, while a single pixel was used as the sample unit (**Figure 2.8**).

The training dataset was used to improve the supervised classification, the per-band pixel values of the stacked composite images were extracted from the training samples, and the resulting data were used to train the RF classifiers [53]. We used the RF algorithm because it is a built-in classifier in GEE and has been widely demonstrated to improve the accuracy of maps by combining random subsets of trees to classify the training samples. In GEE,

the RF algorithm is applied through the following function: *(ee.Classifier.smileRamdomForest)*. Moreover, the algorithm can be configured in three ways *(ee.setOutputMode)* based on the classification mode (class/type maps), regression mode (maps with continuous values predicted), and the probability mode (map with rescaled values between 0 and 1). In the current study, the RF algorithm was applied in the classification mode using GEE to obtain the land cover, in the regression mode to estimate the carbon maps, and in the probability mode to estimate the change magnitude maps.
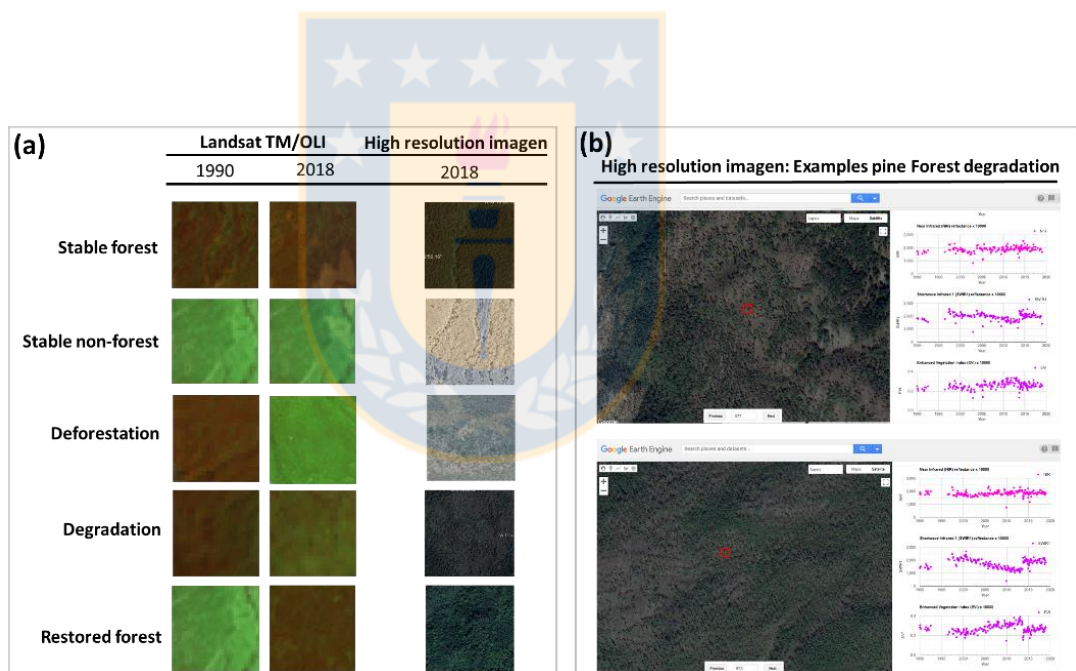
Figure 2.7. Example of the image composites of different land cover dynamic change classes for pine forests of the Dominican Republic. (a) NIR - SWIR1 - RED Landsat Thematic Mapper (TM) and Operational Land Imager (OLI) versus a high-resolution; (b) Example of degraded pine forests observed using Google Earth Engine (GEE).
Fuente: Elaboración propia
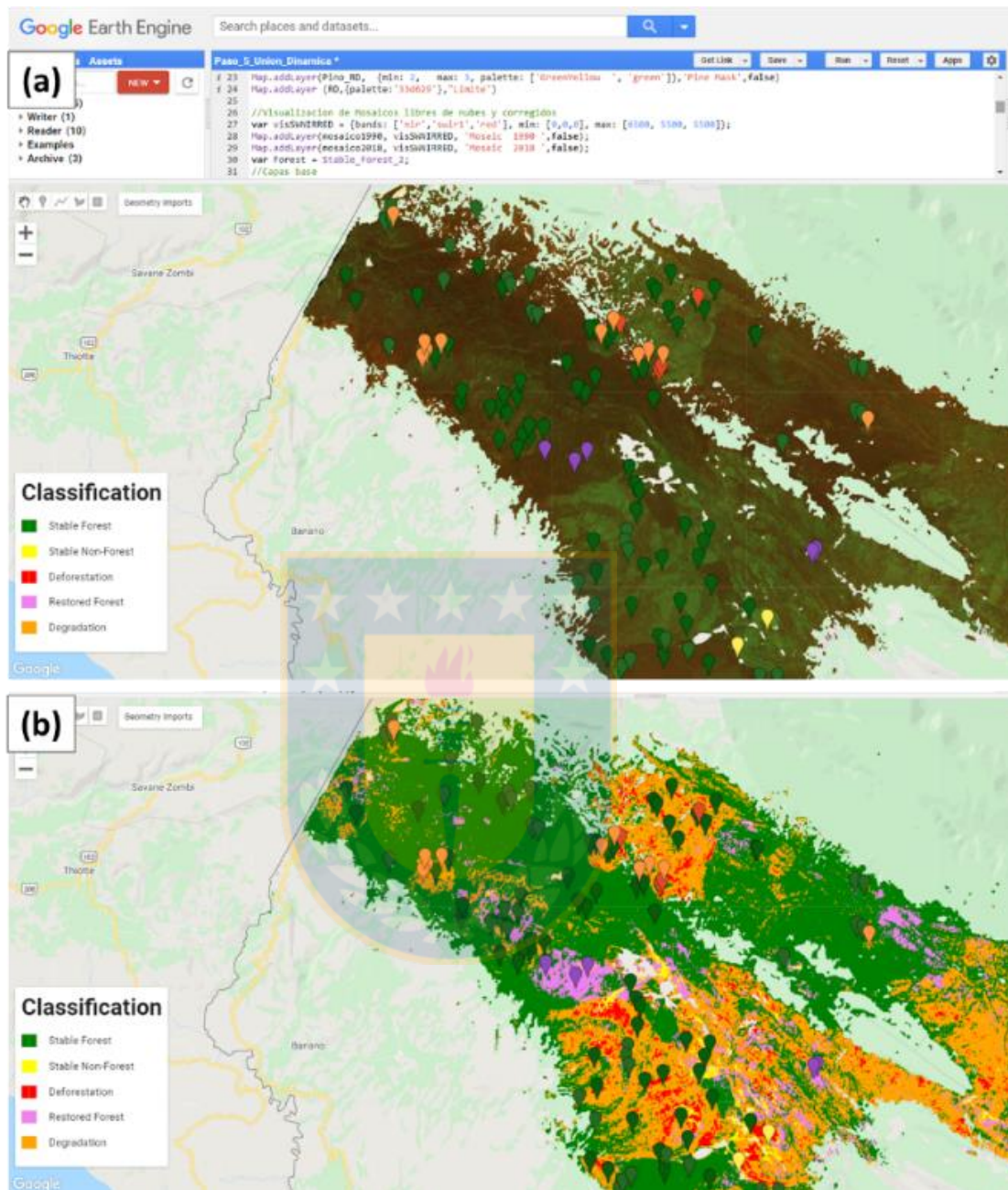
Figure 2.8. Training samples of 5 land cover classes in GEE; (a) samples of land cover change class; (b) dynamics of land cover change classification.
Fuente: Elaboración propia

2.2.6. Carbon stock and change magnitude

Estimating the spatial–temporal distributions of forest carbon stocks subject to land cover changes is critical for estimating and reporting GHG emissions [54]. To spatially represent explicit forest degradation along with degraded carbon, we generated a carbon map of pine forests using 51 sampled plots from the NFI data and Landsat 2018.

The size of each parcel is 0.125 hectares (ha) (25 m × 50 m). Because the size of each plot and the Landsat pixels do not match, we used the geographic coordinates of the centroid of the plot. Thus, we applied the carbon values in units per hectare (t ha^{-1}) to each pixel.

For each plot, four of the five C pools were measured [55]: Aboveground biomass (AGB), belowground biomass (BGB), deadwood (DW), and leaf litter **(Table 2.1)** [47,56,57] (See section: Field Inventory Data; for more details). To convert the biomass to carbon, the IPCC default carbon factor value (0.47) was used. Each pool was modeled independently using spectral responses (vegetation indices) from Landsat applying a regression model with the RF algorithm   in   GEE

(ee.Classifier.randomForest(30).setOutputMode('REGRESSION'). Through this process, each spatial pixel acquired an AGB (t C ha^{-1}) value, and its related standard deviation as a measure of uncertainty. The code generated to model the C stock using the GEE is available in **Appendix B**.

The change magnitude is assumed to be an approximate indicator of the amount of tree removal or canopy damage that occurred due to disturbances [58]. The change magnitude was estimated from the spectral indices for the stable forest class using the composite mosaic for the 1990–2018 period. We fitted an RF probability model in GEE to represent the structural forest changes in each of the spectral indices described previously *(ee.Classifier.randomForest(50).setOutputMode('PROBABILITY')*.

The disturbance monitoring algorithm used to identify the forest changes in each pixel location is closely related to the Continuous Change Detection and Classification algorithm (CCDC) [20,58,59] but was adapted using RF models to predict the change magnitude probabilities for bi-temporal observations. Several studies have successfully applied this algorithm to different sensors and different spectral indices to detect changes [60,61]. This

change detection algorithm operates on the time series of each pixel in the study area.

The decrease in a spectral index caused by a disturbance is recognized by a certain change magnitude. For example, values close to 1 in the NDFI spectral indices indicate high proportions of green vegetation (GV) or stable forests, while values close to 0 in the NDFI imply higher proportions of soil (So). The code generated to estimate the change magnitude using the GEE is available in **Appendix B**.

Pixel-based mapping facilitates comparisons and evaluations of changes with direct algebraic calculations [62]. Once the carbon stored in the four pool of the forest was estimated for the year 2018 along with the magnitude of change between 1990–2018, we used **Equation 1** to determine the C stocks in each pixel for the year 1990. Finally, the degraded carbon was estimated as the difference of the C stored in the pine forests between 1990 and 2018 combined with the change magnitude observed in the same period **(Equation 2)**. To estimate forest degradation, only disturbances occurring in the stable forest class were considered for the period analyzed:

$$C_{t1} = \frac{C_{t2}}{1 - CM} \tag{1}$$

where C_{t1} is the C stock in 1990 (Mg ha^{-1}), C_{t2} is the C stock in 2018, and CM is the change magnitude (value > 0 and < 1).

$$CD = C_{t1} - C_{t2} \tag{2}$$

where CD is the carbon degraded for the 1990–2018 period (Mg ha^{-1}), C_{t1} is the C stock in 1990, and C_{t2} is the C stock in 2018.

Finally, to calculate the annual rate of degraded carbon, the stock-difference method (SDM) was used, where changes in carbon stock (ΔCarbon) represent the difference between carbon stocks for a given forest area estimated at two-time points **(Equation 3)**:

$$\Delta Carbon = \frac{Carbon_{t1} - Carbon_{t2}}{t2 - t1} \tag{3}$$

where $\Delta Carbon$ is the annual change in C stocks (Mg·C·ha^{-1}·yr^{-1}), $Carbon_{t1}$ is the C stock in 1990 (Mg·C·ha^{-1}), and $Carbon_{t2}$ is the C stock in 2018 (Mg C·ha^{-1}).

Table 2.1 Definitions and variables used to estimate the carbon stored in the pine forest.

Variable	Unit*	Definition/explanation
AGB	Mg C ha^{-1}	Aboveground biomass: all living and standing dead trees with a diameter at breast height (DBH) equal to or greater than 2 cm.
BGB	Mg C ha^{-1}	Belowground live biomass: roots.
DW	Mg C ha^{-1}	Deadwood: All pieces of wood with a diameter greater than 2 cm lying on the surface of the ground or intermixed with dead leaves.
Litter	Mg C ha^{-1}	Non-woody biomass is recorded, which includes dead leaves (dead biomass) and herbaceous vegetation (living non-woody biomass on the ground). The maximum diameter for woody material to be considered is 2 cm.

Note: 1 Mg ha^{-1} = 1 ton ha^{-1}

Note: To convert biomass to carbon, the IPCC default carbon factor value (0.47) was used.

Fuente: Elaboración propia

2.2.7. Model evaluation: carbon stock

Cross-validation (CV) is one of the most commonly used techniques to evaluate the efficiency of a machine learning (ML) technique; this is due to its wide application in the scientific arena and its efficiency in detecting a model's overfitting problems [63]. To evaluate the performance of the machine learning model applied to map the change magnitude and forest carbon, the following functions were used: coefficient of determination (R^2), mean square error (MSE), root-mean-square error (RMSE), mean absolute deviation (MAD), cumulated forecast error (CFE), and mean absolute percentage error (MAPE):

$$R^2 = \frac{\left(\sum_{i=1}^{n}(y_i - \bar{y}_i)(f_i - \bar{f}_i)\right)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2 \sum_{i=1}^{n}(f_i - \bar{f}_i)^2} \tag{4}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - f_i)^2 \qquad (5)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - f_i)^2} \qquad (6)$$

$$MAD = \frac{\sum |y_i - f_i|}{n} \qquad (7)$$

$$CFE_t = \sum_{i=1}^{t}(y_i - f_t) \qquad (8)$$

$$MAPE = \frac{\sum |\frac{y_i - f_i}{y_i}|}{n} \times 100 \qquad (9)$$

where $n\,(i = 1, 2, \dots, n)$ is the number of samples used for the machine learning model, y_i is the value observed (C stock), $\bar{y}_i$ is the corresponding mean value, f_i is the predicted value (C stock), and $\bar{f}_i$ is the mean value.

2.2.8. Accuracy assessment and analysis

We used the confusion matrix statistical accuracy assessment method to evaluate the dynamic land cover change classification. The overall accuracy (OA), user's accuracy (UA), and producer's accuracy (PA) were applied to each class, and the Kappa coefficient was used to assess the class map and determine the level of agreement between two raters. The standard deviations and the confidence intervals (at a 95% significance level) were also estimated.

Since land change classes (degradation, deforestation, and forest restoration) tend to cover only a small portion of the study objectives compared to stable areas (stable forest and stable non-forest), it is recommended to

stratify the study based on a map that represents the classes of principal interest to ensure an effective statistical sample representation in land change classes such as degradation, deforestation, and forest restoration [64].

An accuracy assessment of the dynamic land cover change map (for the 1990–2018 period), generated through a sampling-based approach to estimate the area of forest degradation in the Dominican Republic (**Figure 6**), was performed on the following land cover change classes:

Stable forest: Pine forests that remain pine forests without disturbance. This forest contains over 30% canopy cover.

Stable non-Forest: Other non-forest lands, such as agriculture, wetlands, grasslands.

Deforestation: Elimination of the forest canopy cover that exceeds 30%. Results in a land-use change.

Degradation: This entails any disturbance that changes the canopy cover density between 100% and 30%. Does not result in a land-use change.

Forest restoration: Conversion of non-forested land to forest: This includes forest restoration with a canopy cover greater than 30% (through natural and artificial means) on deforested land.

For the accuracy assessment, a total of 1,124 spatial sampling points were established for the study area using the stratified random sampling approach following best practices [64] **(Equation 10)**. It was necessary to modify the minimum sample size to determine the objective standard error of the

degradation area, rather than the OA of the map, and thereby ensure that sample size would be large enough to produce sufficiently accurate estimates [65]. The stratified area estimator-design tools hosted in SEPAL were used to generate random spatial points **(Table 2.2)**. SEPAL is a cloud-based computing platform developed by FAO, which uses the GEE and OpenForis Geospatial Toolkits [66].

$$n = \frac{(\sum W_i\,S_i)^2}{[S(\hat{O})]^2 + (1\,/\,N)\sum W_i\,S_i^2} \approx \left(\frac{\sum W_i S_i}{S(\hat{O})}\right)^2 \tag{10}$$

where n = number of points in the study area, $S(\hat{O})$ is the standard error of the estimated OA, W_i is the mapped proportion of the class area i, and S_i is the standard deviation of land cover classes i.

We performed an analysis following best practices to assess the accuracy of the map classification, and the area of change was estimated using a classification error matrix. For details on the matrix nomenclature, refer to *Olofsson et al. (2014)* [64].

Table 2.2 Strata area, sample allocation for the stratified random sample, and weights for the study period (1990–2018).

	Stable forest	Stable non-Forest	Deforestation	Restored forest	Degradation
Area (ha)	252,408	2,527	2,856	23,452	47,534
Wi (%)	76.77	0.77	0.87	7.13	14.46
Samples	800	50	50	74	150

Fuente: Elaboración propia

Collecting reference observations of forest degradation is a complex task, primarily because degradation is a continuous process that must be observed over a long period. In this sense, satellite images with high spectral resolutions have become a key tool, but they are not sufficient to reconstruct the landscape's historical dynamics. Therefore, this study required the use of Landsat observation time series supported by very high spatial resolution (VHRS) imagery. Independent stratified validation samples were visually interpreted from the VHRS time-series images of Collect Earth (CE). We built a survey in CE that helped us access multiple satellite images, including archives including VHSR imagery (Google Earth, Bing Maps) and a set of satellite images from the GEE catalog, along with their derived spectral indices [67] **(Figure 2.9)**.

To facilitate the historical collection of reference data, other GEE assessment tools were adapted and used, such as the Accuracy and Area Estimation Toolbox (AREA2) developed by Bullock and Olofsson (2018) (see github.com/bullocke/AREA2). The sample interpretation tool allowed us to determine reference labels for the 1,124 samples collected. This algorithm helped estimate the map accuracy and disturbance area and visualize the time series trends of each sample using a dataset created using the Continuous Degradation Detection (CODED) methodology in GEE [20] (see **Appendix B** for the code developed in this study using the GEE).
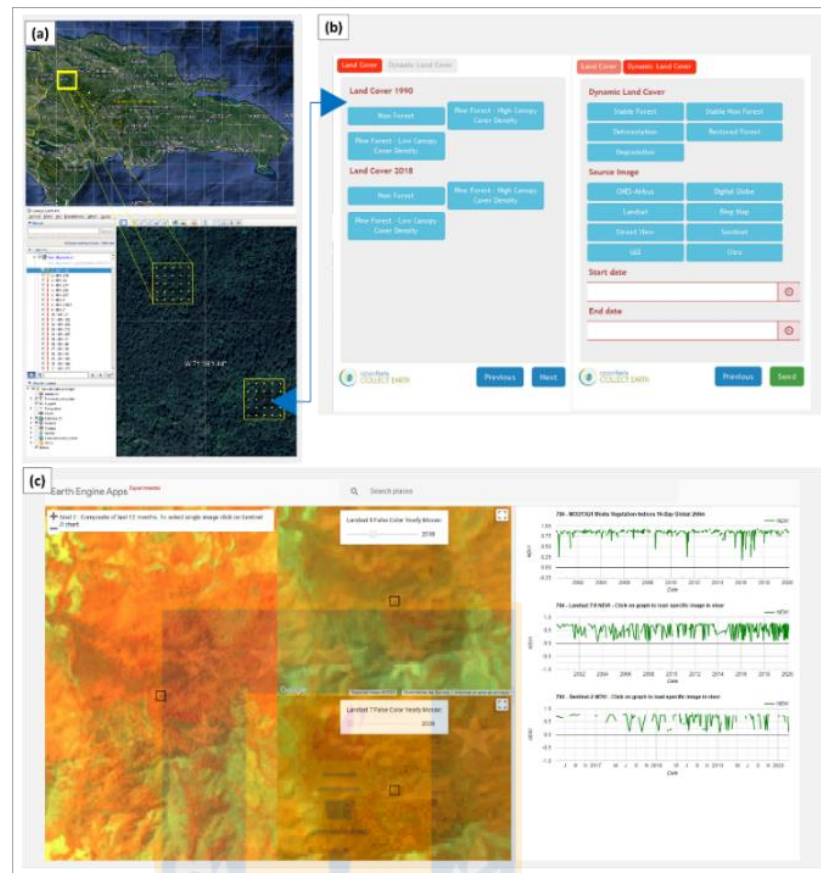
Figure 2.9. Tools used to collect the reference observations of forest degradation. (a) Collect Earth interface in Google Earth Pro; (b) Collect Earth survey; (c) time series tools online viewer.
Fuente: Elaboración propia

To review and assign reference labels to each of the 1,124 selected special sample units, three trained interpreters were delegated. These interpreters did not know the classes of the assigned samples. An additional interpreter reviewed all samples with low or medium confidence. At least three interpreters reviewed all units labeled as degradation. The decision matrix and labels assigned to the samples evaluated in the time series are shown in **Table 2.**

Table 2.3 The decision matrix for the validation samples interpreted from high-resolution images and Landsat time series using Collect Earth (CE) and GEE, respectively.

Land Cover 1990	Land Cover 2018	Reference Class
Non-forest	Non-forest	Stable non-forest
Pine forest: high canopy cover density	Pine forest: low canopy cover density	Degradation
Pine forest: high canopy cover density	Non-forest	Deforestation
Pine forest: low canopy cover density	Non-forest	Deforestation
Non-forest	Pine forest: low canopy cover density	Restored forest
Non-forest	Pine forest: high canopy cover density	Restored forest
Pine forest: high canopy cover density	Pine forest: high canopy cover density	Stable forest
Pine forest: low canopy cover density	Pine forest: high canopy cover density	Stable forest
Pine forest: low canopy cover density	Pine forest: low canopy cover density	Stable forest

Fuente: Elaboración propia

2.3. Results

2.3.1. Dynamic land cover changes from 1990 to 2018

The total study area corresponded to 328,777 ha of pine forests in the Dominican Republic. The results showed that degraded forests accounted for 11% ± 1.21% (95% confidence interval) of the total study area between 1990 and 2018, while 79% ± 1.28% remained stable and did not suffer any disturbances; further, 2% ± 0.61% were deforested. In total, we estimated that 36,808 ± 446 ha of pine forests was degraded.

The margin of error of the area estimate for forest degradation was 10.8% (95% CI). The user's accuracy was 70.7%, and the producer's accuracy for

forest degradation was 91.3%. The overall accuracy of the dynamic land cover change map was 92.8%. The main results corresponding to the accuracy assessment are shown in **Table 2.4**.

Table 2.4 Confusion matrix—sample counts, area proportions, area estimates, and accuracy measures for stable forest, stable non-forest, deforestation, restored forest, and forest degradation.

Confusion matrix, random sample counts

| | | Reference | | | | | | | | |
		Stable forest	Stable non-forest	Deforestation	Restored forest	Degradation	Total	Pixels	W_i	Ha
Map	Stable forest	**787**	0	2	2	9	800	2,804,538	0.768	252,408
	Stable non-forest	0	**48**	0	2	0	50	28,074	0.008	2,527
	Deforestation	0	2	**47**	0	1	50	31,729	0.009	2,856
	Restored forest	11	5	0	**57**	1	74	260,578	0.071	23,452
	Degradation	29	6	9	0	**106**	150	528,158	0.145	47,534
	Total	*827*	*61*	*58*	*61*	*117*	*1,124*	*3,653,077*	*1*	*328,777*

Confusion matrix, area proportions

| | Reference | | | | |
	Stable forest	Stable non-forest	Deforestation	Restored forest	Degradation
Stable forest	0.7552	0.0000	0.0019	0.0019	0.0086
Stable non-forest	0.0000	0.0074	0.0000	0.0003	0.0000
Deforestation	0.0000	0.0003	0.0082	0.0000	0.0002
Restored forest	0.0106	0.0048	0.0000	0.0549	0.0010
Degradation	0.0280	0.0058	0.0087	0.0000	0.1022
Total	0.7938	0.0183	0.0188	0.0572	0.1119

Accuracy and area estimates

Area [pix]	2,899,809	66,953	68,526	208,850	408,939
Area [ha]	260,983	6,026	6,167	18,796	36,804
S(Area)	0.0065	0.0031	0.0031	0.0038	0.0062
S(Area) [ha]	2,143	1,034	1,031	1,240	2,033
95% CI [ha]	4,201	2,026	2,021	2,430	3,985
Margin of error [%]	1.61	33.62	32.77	12.93	10.83
User's acc (%)	98.38	96.00	94.00	77.03	70.67
Producer's acc (%)	95.14	40.25	43.52	96.11	91.27
Overall	92.8%				
Kappa	0.85				

Fuente: Elaboración propia

We analyzed the dynamics of land cover change as they relate to the country's protected areas and identified that 71% of the degraded pine forests exist within the protected area. Among these, the main forest belongs to Sierra de Bahoruco National Park (NP), with 14,166 ha (30% of the total degraded area), followed by Valle Nuevo NP and José del Carmen Ramírez NP, with 8,736 ha (18% of the total degraded area) and 6,462 ha (16% of the total degraded area), respectively. **Table 2.5** shows the locations of the protected areas with the degraded pine forests from 1990 to 2018. A geographic representation of the dynamic land cover change map obtained in our study is also provided (**Figure 2.10).** Detailed results on land cover change are available via a dashboard called "Accuracy assessment and analysis tools" available in **Appendix A**.

Table 2.5 Dynamics of land cover change associated with the different classes of protected areas in the Dominican Republic.

Protected Area Category	Deforestation (ha)	Degradation (ha)	Restored forest (ha)	Stable forest (ha)	Stable non-forest (ha)
Natural Monument	0	5	19	337	0
Natural reserve	71	1,800	1,293	12,350	21
National Park	2,151	31,779	9,867	175,081	1,580
Protected Landscape	6	151	48	3,083	3
Strict Protection Area	3	98	30	820	7
Habitat/Species Management Area	0	0	1	16	0
Non-Protected Area	625	13,701	12,193	60,722	916

Fuente: Elaboración propia
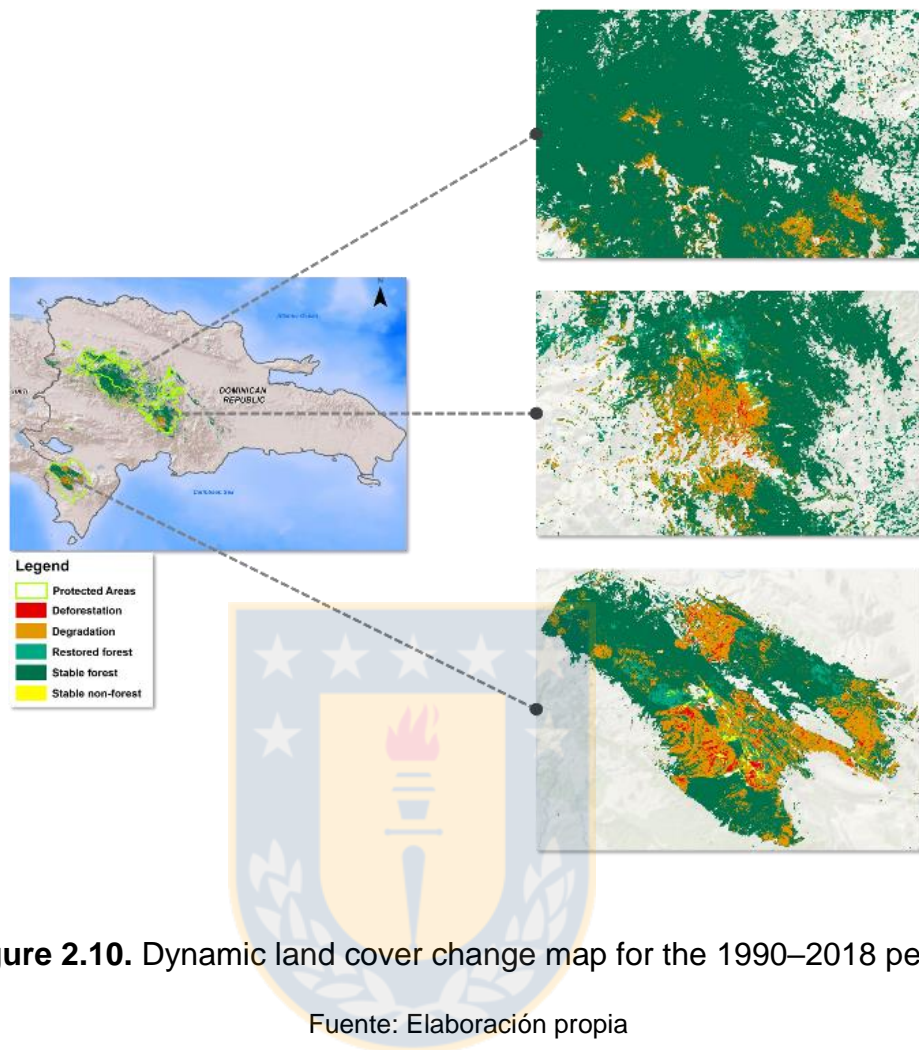
Figure 2.10. Dynamic land cover change map for the 1990–2018 period.

Fuente: Elaboración propia

2.3.2. Carbon stock

The total carbon stock in the pine forest area analyzed was composed of AGB C, BGB C, DW C, and litter C pools. The results for the total carbon analysis are presented in **Table 2.6, Figure 2.11, and Figure 2.12.** The analysis shows that the total carbon stock in 2018 was approximately 19,002,000 Mg C, with an average of 64.4 Mg C ha^{-1}. The RMSE of the model was 13.4 Mg C ha^{-1}, the R^2 was 0.78, the CFE was 0.35, and the MAPE

reached 21%. Of the total carbon stock stored, 75.8% (14,410,609 Mg C) was in the National Park, while 18.4% (3,498,042 Mg C) was outside protected areas, and 4.3% (824,182 Mg C) was stored in natural reserves **(Table 2.7)**. Detailed results on the carbon stored from the different pools and protected area categories are available via a dashboard provided in **Appendix A**.

Table 2.6 Results of the accumulated carbon stock model, carbon stock for the different pools estimated, and their error measures based on random forest modeling.

Pool	N	Mg C	Mg C ha^{-1}	R^2 (%)	MSE	RMSE (Mg C ha^{-1})	MAD	CFE	MAPE (%)
Total	51	19,002,000	66.9	78.1%	179.09	13.38	10.85	0.35	21.1%
AGB	51	12,098,753	43.3	75.5%	96.99	9.85	8.09	- 7.83	24.8%
BGB	51	3,638,370	13.2	75.8%	7.41	2.72	2.08	- 1.82	20.9%
DW	42	1,289,859	3.53	80.1%	12.33	3.51	1.97	16.11	175.0%
Litter	50	548,420	2.2	79.3%	1.96	1.40	0.90	- 10.10	86.0%

Fuente: Elaboración propia
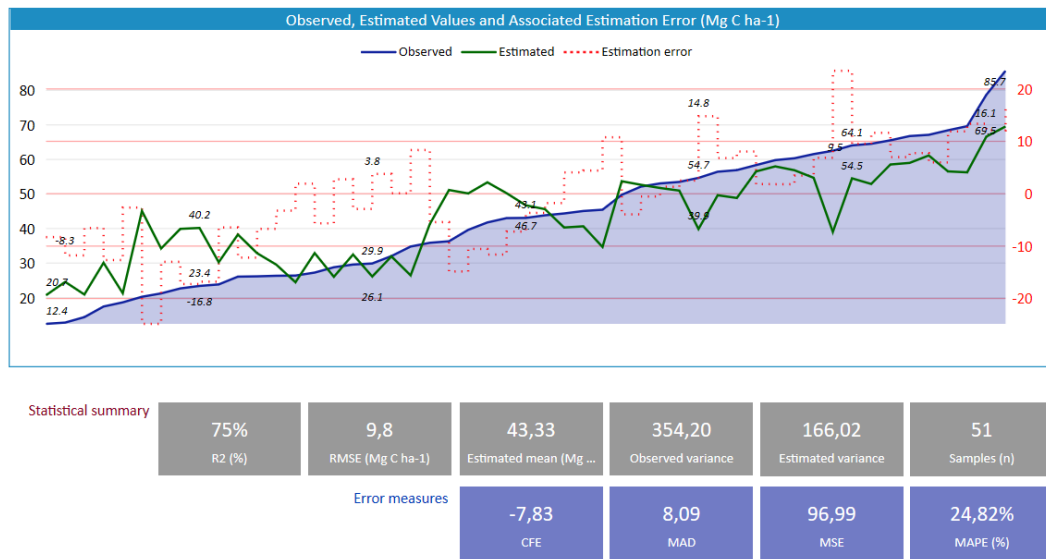
Figure 2.11. Carbon stored in the pine forests in 2018—a statistical summary and the error measures used to evaluate the performance of the model via random forest modeling.

Fuente: Elaboración propia

Table 2.7 Results of the accumulated carbon stocks and carbon stock models for the different pools in each protected area category.

Protected Area Category	Mg C total	Litter (Mg C)	AGB (Mg C)	DW (Mg C)	BGB (Mg C)
Natural Monument	22,507	833	14,872	643	4,453
Natural Reserve	824,182	27,840	531,320	38,284	160,987
National Park	14,410,609	395,917	9,182,340	1,022,543	2,754,263
Protected Landscape	180,743	7,942	120,722	6,917	36,218
Strict Protection Area	64,857	2,075	42,030	2,454	12,592
Habitat/Species Management Area	1,058	35	692	24	205
Non-Protected Area	3,498,043	113,777	2,206,776	218,994	669,653

Fuente: Elaboración propia
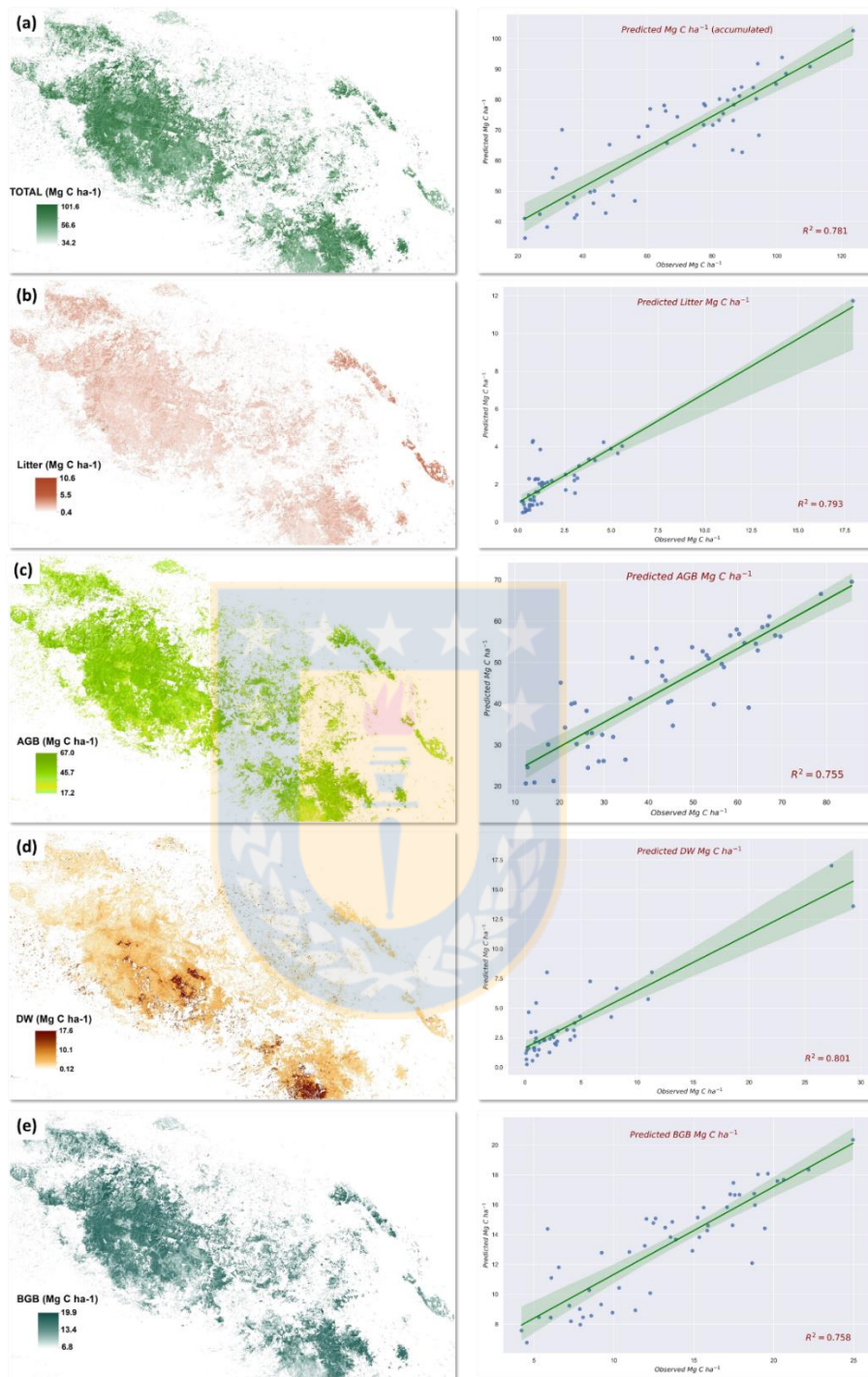
Figure 2.12. Spatial distribution of the average predicted carbon stocks in pine forests (zoomed in image of the national parks with the highest density of pine forests) and the carbon prediction graph with a 95% confidence interval: (a) total carbon; (b) litter C; (c) AGB C; (d) downed dead C; (e) BGB C (Note: 1 Mg ha−1 = 1 ton ha−1).

Fuente: Elaboración propia

2.3.3. Carbon degraded in the 1990–2018 period

The total carbon degraded in the pine forest area analyzed was 3,479,159 Mg C. Converting this degraded carbon into emission- and climate-related units of CO_2-equivalent emissions (Metric tons CO_2-equivalent units), the emissions caused by the degradation of the pine forests in the period 1990–2018 were 12,756,916 tCO_2eq, with an annual average of 2.6 Mg C ha^{-1} yr^{-1} (9.5 tCO_2eq ha^{-1} yr^{-1}). Of the total degraded C stock, 73.9% (2,570,081 Mg C) was found in National Park, while 2.9% (102,401 Mg C) and 22.8% (792,048 Mg C) of C were degraded in natural reserves and non-protected areas, respectively **(Table 2.8)**. Detailed results on the degraded carbon in pine forests in the 1990–2018 period for the different pools and protected area categories are provided in **Appendix A**.

Table 2.8 Results of degraded carbon for different pools per protected area category.

Protected Area Category	Total Carbon (Mg)	Carbon (Mg) Litter	Carbon (Mg) AGB	Carbon (Mg) DW	Carbon (Mg) BGB
Natural Monument	242	9	135	43	41
Natural reserve	102,401	2,584	55,559	18,834	17,320
National Park	2,570,081	61,449	1,404,486	595,785	423,726
Protected Landscape	8,512	297	4,743	1,440	1,482
Strict Protection Area	5,873	141	3,167	1,152	970
Habitat/Species Management Area	2	0	1	0	0
Non-Protected Area	792,048	20,407	431,030	162,352	132,047

Fuente: Elaboración propia
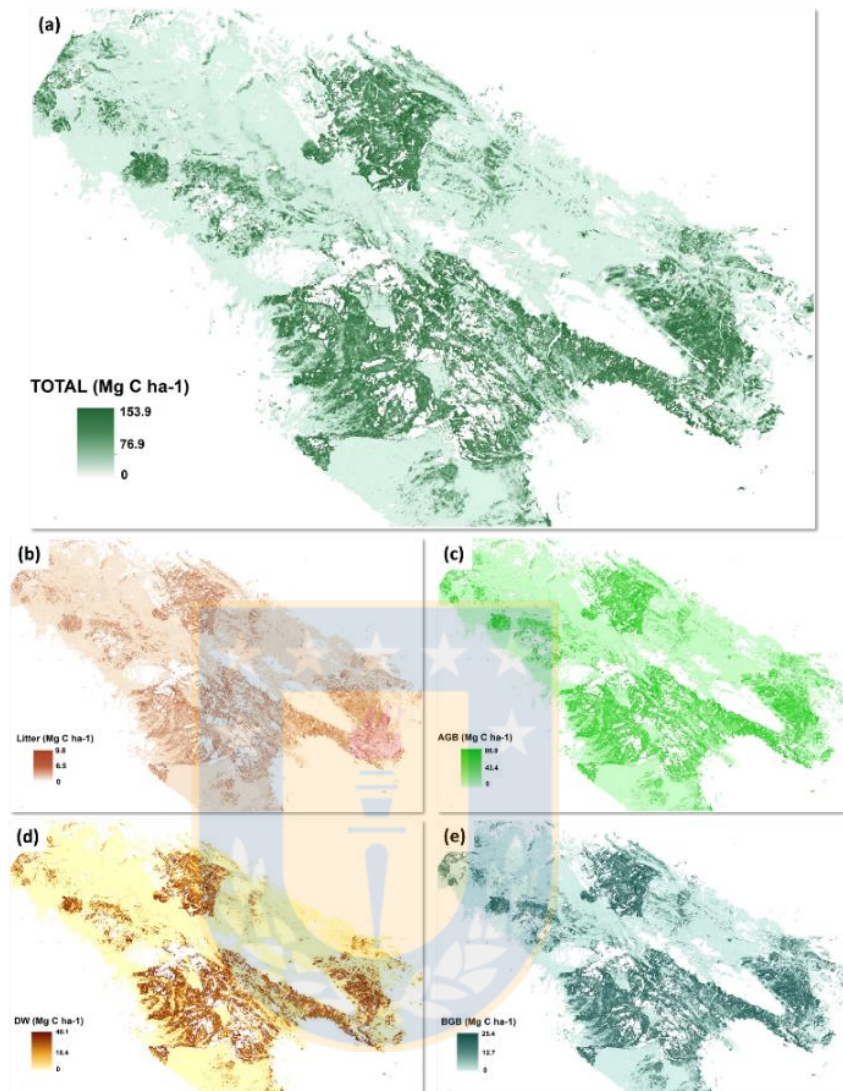
Figure 2.13. Spatial distribution of degraded carbon (zoomed-in image of the Sierra de Bahoruco National Park protected area with the highest degradation): (a) total carbon; (b) AGB C; (c) BGB C; (d) DW C; (e) and litter C in the pine forests of the Dominican Republic.

Fuente: Elaboración propia

2.4. Discussion

2.4.1. Validation of dynamic land cover change map from the 1990–2018 period.

Most of the countries that are part of the REDD+ mechanism do not quantify or report their emissions caused by forest degradation [16]. Efforts to find such a method have been great, and the challenge of obtaining accurate estimates remains under investigation and debate in the scientific arena. One of the main agreements in the measurement and monitoring of forest degradation is that a long series of temporal and spatio–temporal observations are required to detect disturbances in forest cover, which is why satellite images like those from Landsat are key inputs to establish more robust MRV systems for REDD+.

Using Landsat data in GEE, we developed a methodology to monitor pine forest degradation in the tropics. This approach was determined to be precise, with an overall 92.9% and 91% producer's accuracy in the degraded forest class of the dynamic land use change map of pine forests in the Dominican Republic. Our estimates of forest degradation are compatible with those of other studies on a sub-national scale. For example, the OA obtained in degradation and deforestation mapping in Rondônia, Brazil, was 91%, while the producer's accuracy reached 68% in the forest degradation class [58]. In the forests of the Brazilian Amazon, the OA in degradation and deforestation

mapping was 92%, while the producer's accuracy was 80% in the forest degradation class [68]. Another study on forest degradation estimation in southeast Asia reported an OA of 78% using high-resolution satellite images [80], while a study using SPOT images with spectral mixing models in the eastern Amazon showed results that also indicated good agreement (86% OA) [69].

The dynamic mapping analysis determined an efficient stratification in the study area and allowed for an impartial estimation. Margins of error of 10.8% were obtained when mapping forest degradation at a 95% confidence level. Although the mapping of forest degradation in tropical forests is scarce in the literature, we observed some consistency between our results and those of other studies in the temporal scale, spatial and spectral resolution of the images used, accuracy, and the use of vegetation index analysis as a method to evaluate and map tropical forest disturbances.

Historical data collection on forest changes is a challenging task because such data are not readily available everywhere, and temporal change data are not detailed enough for the validation of time series maps. The current study used the AREA2 algorithm developed by Bullock and Olofsson (2018) by applying the Time Series Viewer. This is a sample interpretation tool used to determine the reference labels derived from a mapped dataset. However, a new challenge involved assessing the changes detected by the model. For this

process, independent datasets were selected and assessed using CE [70]. The combination of these tools proved to be efficient in our study.

Using spectral index measures to validate the dynamic land cover change map in our study allowed us to extend tele-interpretation techniques and facilitated the visual detection of historical change processes, especially for degraded forest detection. In other research, the NDFI was used to map degradation; ultimately, the NDFI was found to be more sensitive to disturbances from tropical forests than other spectral indices [71]. We use different spectral indices to improve our estimates of forest degradation and performed a regression analysis with random forest to determine the importance of the indices in the constructed model. We found that the NDFI and EVI were the main variables able to explain the model and thereby classify areas with forest degradation for the 1990–2018 period **(Figure 2.14a and 2.14b).**
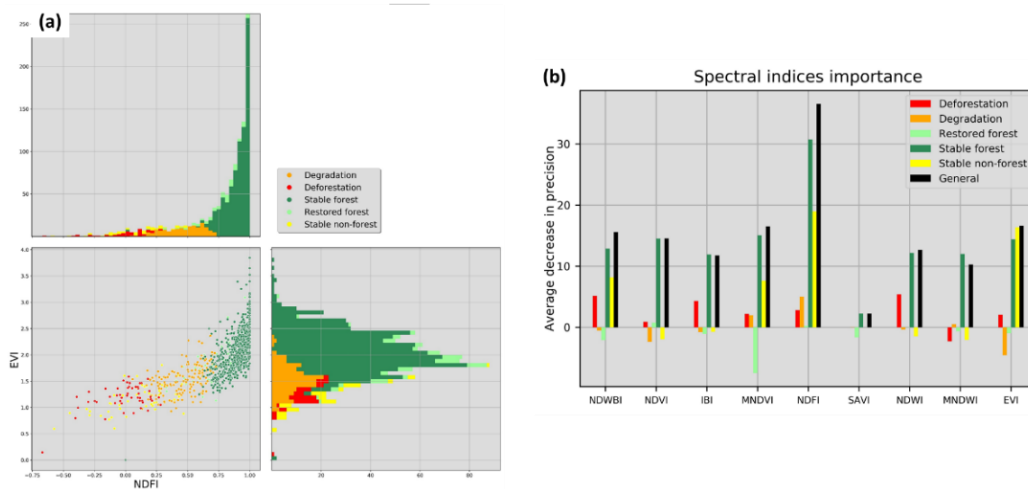
Figure 2.14. Scatterplot of two spectral indices (EVI and NDFI); the points represent the values of the most important variables used for land cover change classification; (b) spectral index importance analysis used for dynamic land cover change classification.

Fuente: Elaboración propia

While the Dominican Republic has experienced a decrease in deforestation in recent times, forest degradation has been on the rise. In our study, we found that an area equivalent to 14% of pine forests of the Dominican Republic was degraded between 1990 and 2018. This is a critical element that must be considered for the development of a REDD+ program in the country. The Dominican Republic established a FREL/FRL to obtain results-based payments under the REDD+ mechanism supported by the FCPF. This includes emissions and removal in the remaining forest land (emissions from forest degradation) for the 2005–2015 period [45]. In this sense, the methodology proposed, and the results obtained in the current study, contribute directly to monitoring and quantifying CO_2, emissions and removal. Furthermore, this
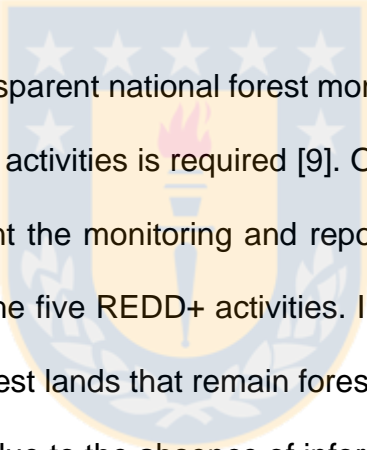
methodology is a valuable tool that can be used in other tropical countries to monitor forest degradation.

2.4.2 Carbon model assessment

Forest inventory is an important source of information for a variety of strategic purposes in forest management. Based on 51 carbon samples obtained in pine forests, we generated a predictive regression model of the carbon stored in AGB, BGB, DW, and litter. Satisfactory results were obtained when applying the RF algorithm to estimate the total carbon stock in the study area, obtaining an R^2 of 78.1% and an RMSE of 13.38 Mg C ha^{-1}. It was determined that 19,002,000 Mg C is stored in the pine forests, of which 3,479,159 Mg C was degraded (18% of C stock), which is equivalent to 124,255 Mg C yr^{-1} for the 1990–2018 study period.

We found some differences between our results and previous estimates; for example, the MARN of the Dominican Republic estimated the local emissions from forest degradation at a level of 182,937 Mg C yr^{-1}, including all forest classes, and at a level of 46,591 Mg C yr^{-1} specifically for pine forests for the 2005–2015 reference period [48]. The differences between both estimates are mainly due to the scale of monitoring (sub-national vs. national), the methodology, the reference period, the algorithms used for image processing and classification, and especially the differences in the accuracy obtained between both estimates.

These differences also suggest that between 1990 and 2005, the degradation rate in the pine forest was much higher compared than that during the period 2005–2015. This provides a new research opportunity to understand the drivers of degradation that have decreased in the country. However, we believe that the degradation estimates found in both studies differ from each other due to the previously mentioned technical and methodological factors.

A robust and transparent national forest monitoring system to monitor and reporting five REDD+ activities is required [9]. Often, capabilities and national circumstances prevent the monitoring and reporting of CO_{2eq} emissions and their removal under the five REDD+ activities. In our study, the forest carbon stock increases in forest lands that remain forest lands were not estimated for the reference period due to the absence of information on the average annual increase in biomass in the studied forest. This offers a new research opportunity for a "step-wise" approach that will allow estimate net CO_{2eq} emissions as part of one of the five REDD+ activities called "enhancement of forest carbon stocks".

2.4.3 Google Earth Engine Platform

The processing and analysis of the data from our study was accomplished using the JavaScript API via the code editor of the GEE platform. Using GEE,

the processing time efficiency increased, and satisfactory results were obtained (see the GEE code developed for this study in Appendix B). GEE is considered a multidisciplinary tool; since 2017, its application has increased notably, especially in land-use mapping and water resources. Landsat images (82%), the RF algorithm (52%), and the NDVI spectral indices are the most frequently used methods in recent studies related to vegetation [72].

In our study, we faced a series of complexities in the detection of forest degradation, especially since a very detailed approach is required. This approach involves detecting the reduction or modifications in the forest structure that are not considered a total loss over an extended period of time and at a large scale. Using the characteristics of the GEE platform by extracting spectral characteristics from satellite images, it was possible to satisfactorily solve the main challenges we encountered.

The methodology proposed here was able to detect the dynamics of change in land use and forest degradation. While there is room to improve the methodology, the use of GEE with its computing capacity and the availability of free satellite images provide powerful support for mapping forest degradation. Additionally, GEE allowed us to quantify the carbon stored until 2018 and the degraded carbon in four pine forest pools. Thus, this method could become a key monitoring tool (at the sub-national and national levels). Moreover, this tool will allow authorities to monitor forest degradation according to indicator 15.3.1 of the Sustainable Development Goals (SDGs),

which defines the area of degraded land and will serve as an essential element for developing national MRV systems for REDD+ strategy implementation and supporting the report on GHG emissions of the United Nations Framework Convention on Climate Change (UNFCCC) in an efficient, robust, and transparent manner.

2.5. Conclusions

The current study developed and applied a methodology for forest degradation mapping based on available data from the Open Access Landsat and the GEE platforms. Additionally, GEE was used in combination with tools such as SEPAL and CE by the FAO. The main objective was to estimate the degraded pine forest area and quantify the degraded carbon for a period of 28 years (1990–2018) by applying machine learning models such as Random Forest.

The methodology applied in this study shows new possibilities for forest degradation monitoring and estimating CO_2 emissions from forest degradation using spectral information derived from Landsat archives and data from the forest inventory; combining both sources of information can also help improve the MRV systems for the REDD+ mechanism.

The model assessment revealed a dynamic land change map with a cumulative overall accuracy of 92%, in relevant classes (such as forest

degradation) with a UA of 70.7%, and a PA of 91.2%. A carbon stock model was also developed with an R^2 of 79% to estimate degradation in terms of the Mg C ha^{-1}. The applied models were built, trained, and validated to demonstrate the efficiency of the methodology. The results obtained indicate that this methodology can be an especially useful tool for time series processing to map forest degradation by applying technologies such as GEE.

GEE has excellent potential for the "wall to wall" forest degradation mapping of tropical pine forest ecosystems. More research is still required to assess the ability of GEE to map degradation in broadleaf forests and dry forest ecosystems by applying Machine Learning techniques combined with spatial data and field measurements. The approaches presented herein could become a key tool for measuring and monitoring emissions from forest degradation in the tropics.

Author Contributions: Conceptualization, E.D., F.D., F.C., and E.Z.; methodology, E.D., J.A.B., F.C., A.J.H.; software, E.D. and F.C.; validation, E.D. and F.C.; analysis, E.D., F.C., J.A.B., F.D., A.J.H., and E.Z.; investigation, E.D., F.C., J.A.B., F.D., A.J.H., and E.Z.; writing—original draft, E.D., F.C., J.A.B., F.D., A.J.H., and E.Z.; supervision, E.D., F.C., J.A.B., F.D., A.J.H., and E.Z.

Appendix A

A consolidation of the results obtained is available on a dashboard with free online access called "Accuracy assessment and analysis tools". This dashboard presents assessments of the dynamic land cover change, the spectral mixture analyses from Landsat, the error matrix, the carbon model, and all the results obtained. The dashboard allows users to pose questions and filter graphic and alphanumeric data with a geographic viewer. The dashboard is available at the following link: [Accuracy assessment and analysis tools](#)

Appendix B

The codes developed in this study using the GEE cloud-based computing platform are available at the following link: [Degradation code](#)

Appendix C

Table 2.9 Spectral indices used for forest degradation mapping.

Spectral Indices	Equation	
Normalized Difference Vegetation Index (NDVI) [73]	$NDVI = \dfrac{(NIR - \text{Red})}{(NIR + Red)}$	(14)
Normalized Difference Spectral Vector (NDSV) [74]	$NDSV_{ij} = \dfrac{B_i - B_j}{B_i + B_j}$ Where B_i and B_j are two generic bands.	(15)
Enhanced Vegetation Index (EVI) [50]	$EVI = G \; x \; \dfrac{(NIR + Red)}{(NIR + C1 \; x \; Red - C2 \; x \; Blue + L)}$	(16)
Soil Adjust Vegetation Index (SAVI) [51]	$SAVI = \dfrac{(NIR - Red)}{(NIR + Red + L)} * (1 + L)$	(17)
Index-Based built-up Index (IBI) [52]	$NDBI = \dfrac{(MIR - NIR)}{(MIR + NIR)}$ $SAVI = \dfrac{(NIR - Red)(1 + l)}{(NIR + Red + l)}$ $MNDWI = \dfrac{(Green - MIR)}{(Green + MIR)}$	(18)
Near-infrared reflectance of vegetation (NIRv) [75]	$NIRv = \left(NDVI_{median\,monthly} - 0.08\right) x \; NIR_{median\,monthly}$	(|9)
Normalized Difference Fraction Index (NDFI) [52]	$NDFI = \dfrac{GV_{shade} - (NPV + Soil)}{GV_{sahde} + NPV + Soil}$ *Where:* $GV_{shade} = \dfrac{GV}{100 - Shade}$	(20)

Fuente: Elaboración propia

Appendix D

Table 2.10 Reference field: National Forest Inventory (NFI) collected by the Ministry of the Environment and Natural Resources (MARN) of the Dominican Republic.

ID	Class	AGB (Mg C ha⁻¹)	BGB (Mg C ha⁻¹)	Litter (Mg C ha⁻¹)	DW (Mg C ha⁻¹)	Total accumulated (Mg C ha⁻¹)	Long	Lat
1	Pine forest: low canopy cover density	18.6	7.9	0.4	0.0	37.5	-71.363	19.371
2	Pine forest: high canopy cover density	26.2	7.2	1.3	0.0	35.2	-71.642	19.321
3	Pine forest: low canopy cover density	12.4	4.6	0.5	0.3	22.2	-71.743	19.321
4	Pine forest: low canopy cover density	17.4	5.3	0.6	1.2	26.8	-71.354	19.330
5	Pine forest: low canopy cover density	34.8	11.3	0.6	2.2	56.1	-71.647	19.277
6	Pine forest: low canopy cover density	85.7	25.0	3.3	2.9	123.6	-71.158	19.272
7	Pine forest: low canopy cover density	56.4	15.3	2.6	2.7	77.4	-71.055	19.268
8	Pine forest: low canopy cover density	49.8	15.6	1.8	2.2	77.5	-71.122	19.284
9	Pine forest: low canopy cover density	41.8	12.0	0.9	7.7	65.2	-71.252	19.270
10	Pine forest: high canopy cover density	43.0	13.2	4.6	2.5	69.3	-70.589	19.212
11	Pine forest: low canopy cover density	29.6	10.3	0.4	0.1	49.1	-71.001	19.191
12	Pine forest: high canopy cover density	58.4	17.5	4.1	0.5	86.8	-71.054	19.107
13	Pine forest: high canopy cover density	27.3	8.5	3.1	0.8	43.7	-70.936	19.146
14	Pine forest: high canopy cover density	64.1	20.2	3.8	0.8	99.8	-70.881	19.134
15	Pine forest: high canopy cover density	68.4	20.6	1.4	4.4	102.8	-71.033	19.119
16	Pine forest: low canopy cover density	60.3	17.9	5.0	3.7	92.7	-70.478	19.124
17	Pine forest: low canopy cover density	59.8	17.3	1.2	2.4	84.8	-70.488	19.134
18	Pine forest: low canopy cover density	23.4	9.2	0.9	4.0	48.3	-71.073	19.131
19	Pine forest: low canopy cover density	45.4	12.3	1.3	27.4	86.4	-70.717	19.125
20	Pine forest: low canopy cover density	29.9	8.1	0.3	0.0	38.2	-71.550	19.143
21	Pine forest: low canopy cover density	52.1	15.9	1.1	4.4	80.2	-71.017	19.154
22	Pine forest: high canopy cover density	53.5	15.2	1.3	4.9	77.9	-71.159	19.055
23	Pine forest: high canopy cover density	45.1	14.9	1.6	2.8	74.5	-70.679	19.046
24	Pine forest: low canopy cover density	39.6	12.6	0.8	0.0	60.1	-71.067	19.022
25	Pine forest: low canopy cover density	32.0	9.2	5.6	0.7	49.5	-70.706	19.066
26	Pine forest: low canopy cover density	78.7	22.2	3.0	2.9	110.2	-70.824	19.079
27	Pine forest: low canopy cover density	26.3	7.3	3.2	5.8	43.4	-70.941	19.039
28	Pine forest: low canopy cover density	35.9	11.0	2.6	29.3	83.4	-70.933	19.046
29	Pine forest: low canopy cover density	20.3	5.9	5.4	0.8	33.7	-70.775	19.061

30	Pine forest: high canopy cover density	43.8	13.7	17.9	0.1	82.3	-70.769	18.994
31	Pine forest: low canopy cover density	43.1	11.9	0.9	0.3	57.3	-71.073	19.013
32	Pine forest: low canopy cover density	61.5	17.1	0.6	8.2	89.1	-71.167	18.967
33	Pine forest: low canopy cover density	21.2	6.1	0.6	1.6	30.8	-70.925	18.952
34	Pine forest: high canopy cover density	69.6	18.8	0.7	4.4	93.6	-70.975	18.902
35	Pine forest: high canopy cover density	56.9	15.8	0.7	11.3	86.5	-70.929	18.926
36	Pine forest: low canopy cover density	28.8	9.9	0.4	0.1	47.2	-71.148	18.927
37	Pine forest: low canopy cover density	64.5	17.4	0.2	0.0	82.1	-71.126	18.913
38	Pine forest: high canopy cover density	62.5	18.6	0.7	0.9	89.3	-70.738	18.836
39	Pine forest: high canopy cover density	66.8	18.8	0.0	0.2	88.5	-70.992	18.855
40	Pine forest: low canopy cover density	23.8	7.9	0.3	0.0	37.3	-70.770	18.861
41	Pine forest: high canopy cover density	22.7	6.5	1.0	0.0	31.8	-70.581	18.726
42	Pine forest: high canopy cover density	36.3	12.5	1.0	1.3	60.9	-70.590	18.639
43	Pine forest: high canopy cover density	67.1	19.0	1.2	11.0	101.6	-71.709	18.263
44	Pine forest: high canopy cover density	53.1	17.6	3.0	1.1	86.7	-71.662	18.263
45	Pine forest: low canopy cover density	26.4	8.6	1.0	1.0	42.4	-71.568	18.265
46	Pine forest: high canopy cover density	65.5	19.6	0.8	1.0	94.1	-71.625	18.256
47	Pine forest: low canopy cover density	14.4	6.0	0.6	0.0	29.1	-71.493	18.238
48	Pine forest: low canopy cover density	12.8	4.2	0.3	2.0	22.1	-71.584	18.197
49	Pine forest: low canopy cover density	26.1	13.5	1.1	0.9	65.7	-71.631	18.239
50	Pine forest: high canopy cover density	54.7	19.4	1.3	1.7	94.5	-71.534	18.105
51	Pine forest: low canopy cover density	44.4	13.9	0.8	0.0	66.1	-71.581	18.104

Fuente: Elaboración propia

Appendix E
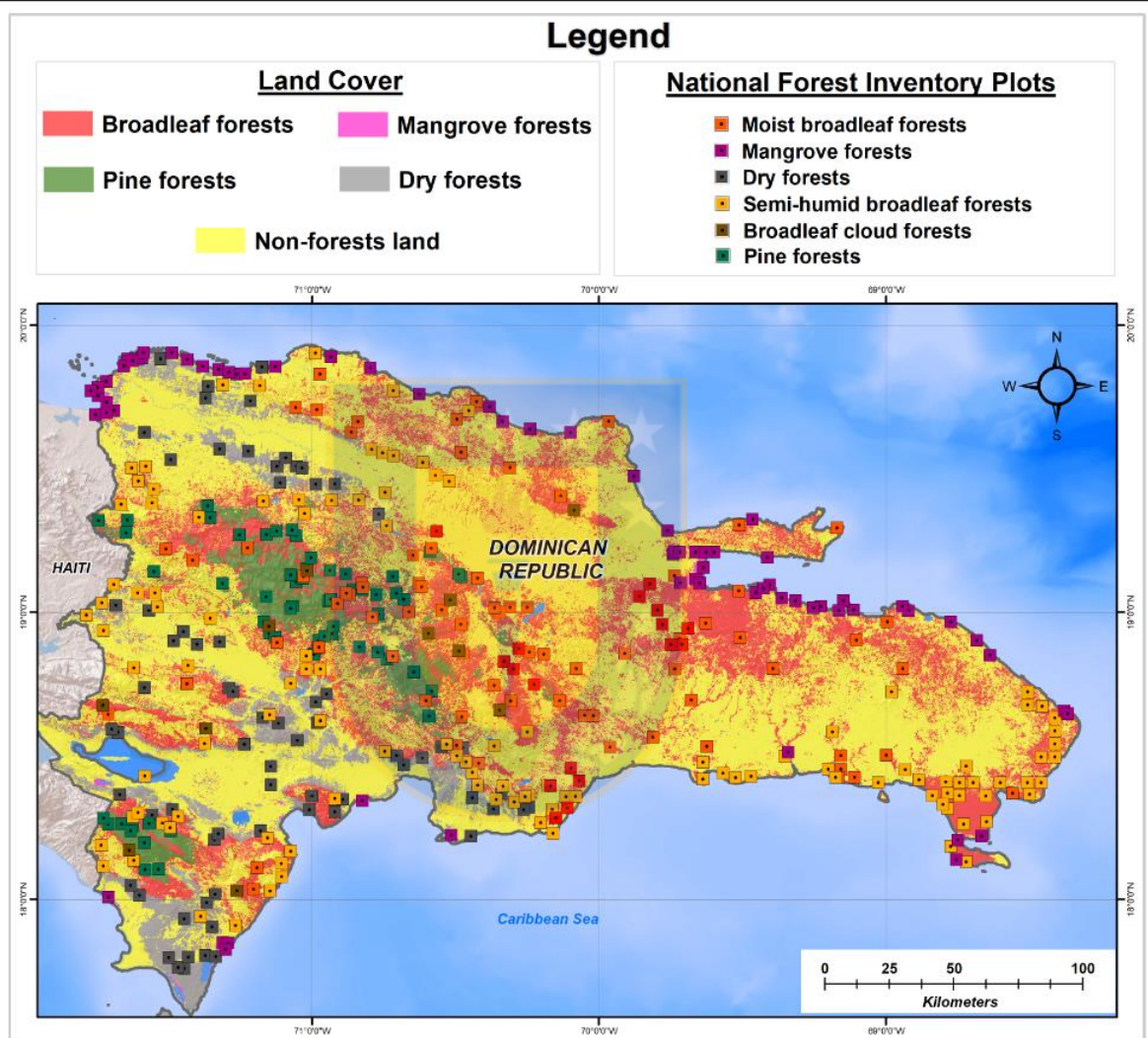
Figure 2.15. Reference map of the National Forest Inventory (NFI) collected by the Ministry of the Environment and Natural Resources (MARN) of the Dominican Republic.

Fuente: Elaboración propia

Capítulo II: Glosario de abreviaturas, siglas y acrónimos

AGB	Aboveground biomass
BGB	Belowground biomass
CCDC	Continuous Change Detection and Classification algorithm
CE	Collect Earth
CFE	Cumulated Forecast Error
CODED	Continuous Degradation Detection
CO_2	Carbon dioxide
CV	Cross-validation
DBH	Diameter at Breast Height
DW	Deadwood
ERPD	Emission Reduction Program Document
EVI	Enhanced Vegetation Index
FAO	United Nations Food and Agriculture Organization
FCPF	Forest Carbon Partnership Facility
FRA	Global Forest Resource Assessment
FREL/FRL	Reference Emission Levels/Forest Reference Levels
GE	Google Earth
GEE	Google Earth Engine
GHG	Greenhouse Gas Emissions
GV	Green Vegetation
IPCC	Intergovernmental Panel on Climate Change
LEDAPS	Landsat Ecosystem Disturbance Adaptive Processing System
RF	Random Forest
MAD	Mean Absolute Deviation
MAPE	Mean Absolute Percentage Error
MARN	Ministry of the Environment and Natural Resources
m.a.s.l	Meters Above Sea Level
Mg C	Megagramme Carbon

ML	Machine Learning
MRV	Monitoring, Reporting, and Verification
MSE	Mean Square Error
NDFI	Normalized Difference Fraction Index
NFI	National Forest Inventory
NPs	National Parks
NPV	lNon-Photosynthetic Vegetation
OA	Overall Accuracy
OLI	Operational Land Imager
PA	Producer's Accuracy
REDD+	Reducing Emissions from Deforestation and Forest Degradation
RMSE	Root-Mean-Square Error
RS	Remote Sensing
SAVI	Soil Adjust Vegetation Index
SDM	Stock-difference method
SRTM	Shuttle Radar Topography Mission
SOC	Soil Organic Carbon
TM	Thematic Mapper
UA	User's Accuracy
USGS	United States Geological Survey

2.6 References

1.   Friedl MA, McIver DK, Hodges JCF, et al. Global land cover mapping from MODIS: algorithms and early results. Remote Sensing of Environment. 2002 2002/11/01/;83(1):287-302.

2.   Wulder MA, White JC, Goward SN, et al. Landsat continuity: Issues and opportunities for land cover monitoring. Remote Sensing of Environment. 2008 2008/03/18/;112(3):955-969.

3.   Pearson TRH, Brown S, Murray L, et al. Greenhouse gas emissions from tropical forest degradation: an underestimated source. Carbon Balance and Management. 2017 2017/02/14;12(1):3.

4.   Harris NL, Brown S, Hagen SC, et al. Baseline Map of Carbon Emissions from Deforestation in Tropical Regions. Science. 2012;336(6088):1573.

5.   Chazdon RL, Brancalion PHS, Laestadius L, et al. When is a forest a forest? Forest concepts and definitions in the era of forest and landscape restoration. Ambio. 2016 2016/09/01;45(5):538-550.

6.   Arana Pardo JI, Birdsey R, Boehm M, et al. IPCC Report on Definitions and Methodological Options to Inventory Emissions from Direct Human-induced Degradation of Forests and Devegetation of Other Vegetation Types. Eggleston HS, Buendia L, Miwa K, Ngara T, Tanabe K ed. Kanagawa, Japan: Institute for Global Environmental Strategies (IGES); 2003.

7.   van der Werf GR, Morton DC, DeFries RS, et al. CO2 emissions from forest loss. Nature Geoscience. 2009 11/01/online;2:737.

8.   Metz B, Davidson R, Bosch R, et al. Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. (Cambridge: Cambridge University Press); 2007.

9.   UNFCCC. Report of the Conference of the Parties on Its Sixteenth Session, Held in Cancun From 29 November to 10 December 2010, Addendum Part Two: Action Taken by the Conference of the Parties at Its Sixteenth Session. 2011; Cancún, México.

10.   Ochieng RM, Visseren-Hamakers IJ, Arts B, et al. Institutional effectiveness of REDD+ MRV: Countries progress in implementing

technical guidelines and good governance requirements. Environmental Science & Policy. 2016 2016/07/01/;61:42-52.

11. Goetz SJ, Hansen M, Houghton RA, et al. Measurement and monitoring needs, capabilities and potential for addressing reduced emissions from deforestation and forest degradation under REDD+. Environmental Research Letters. 2015 2015/12/01;10(12):123001.

12. Milne R, Jallow P. Basis for Consistent Representation of Land Areas. IPCC Good Practice Guidance for LULUCF2005.

13. Woodcock C. Free Access to Landsat Imagery Teach by the Book Science Education. Science. 2008;80(320):1011-1012.

14. Hansen MC, Potapov PV, Moore R, et al. High-Resolution Global Maps of 21st-Century Forest Cover Change. Science. 2013;342(6160):850.

15. Herold M, Hirata Y, Laake PV, et al. A review of methods to measure and monitor historical forest degradation. Unasylva. 2011;62:1-31.

16. Hosonuma N, Herold M, De Sy V, et al. An assessment of deforestation and forest degradation drivers in developing countries. Environmental Research Letters. 2012 2012/10/08;7(4):044009.

17. Chambers JQ, Asner GP, Morton DC, et al. Regional ecosystem structure and function: ecological insights from remote sensing of tropical forests. Trends in Ecology & Evolution. 2007;22(8):414-423.

18. De Sy V, Herold M, Achard F, et al. Synergies of multiple remote sensing data sources for REDD+ monitoring. Current Opinion in Environmental Sustainability. 2012 2012/12/01/;4(6):696-706.

19. Woodcock CE, Macomber SA, Pax-Lenney M, et al. Monitoring large areas for forest change using Landsat: Generalization across space, time and Landsat sensors. Remote Sensing of Environment. 2001 2001/10/01/;78(1):194-203.

20. Zhu Z, Woodcock CE. Continuous change detection and classification of land cover using all available Landsat data. Remote Sensing of Environment. 2014 2014/03/25/;144:152-171.

21. Gorelick N, Hancher M, Dixon M, et al. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment. 2017 2017/12/01/;202:18-27.

22. Azzari G, Lobell DB. Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring. Remote Sensing of Environment. 2017 2017/12/01;202:64-74.

23. Chen B, Xiao X, Li X, et al. A mangrove forest map of China in 2015: Analysis of time series Landsat 7/8 and Sentinel-1A imagery in Google Earth Engine cloud computing platform. ISPRS Journal of Photogrammetry and Remote Sensing. 2017 2017/09/01;131:104-120.

24. Dong J, Xiao X, Menarguez MA, et al. Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. Remote Sensing of Environment. 2016 2016/11/01;185:142-154.

25. Goldblatt R, Stuhlmacher MF, Tellman B, et al. Using Landsat and nighttime lights for supervised pixel-based image classification of urban land cover. Remote Sensing of Environment. 2018 2018/02/01;205:253-275.

26. Huang H, Chen Y, Clinton N, et al. Mapping major land cover dynamics in Beijing using all Landsat images in Google Earth Engine. Remote Sensing of Environment. 2017 2017/12/01;202:166-176.

27. Orvis KH. The highest mountain in the Caribbean: Controversy and resolution via GPS [Article]. Caribbean Journal of Science. 2003;39(3):378-380.

28. SEMARENA. Perfil nacional para evaluar las capacidades nacionales de implementación del Principio 10 de la Declaración de Rio. Santo Domingo, República Dominicana2008. p. 119.

29. Kennedy LM, Horn SP, Orvis KH. Modern pollen spectra from the highlands of the Cordillera Central, Dominican Republic. Review of Palaeobotany and Palynology. 2005 2005/11/01;137(1):51-68.

30. Darrow WK, Zanoni T. HISPANIOLAN PINE (PINUS OCCIDENTALIS SWARTZ) A LITTLE KNOWN SUB-TROPICAL PINE OF ECONOMIC POTENTIAL. The Commonwealth Forestry Review. 1990;69(2 (219)):133-146.

31. MARN. Analysis of the Direct and Indirect Drivers of Deforestation and Forest Degradation (DD) in the Dominican Republic. Santo Domingo, Dominican Republic2018. p. 161.

32. FAO. Global Forest Resources Assessment 2015. Rome: UN Food and Agriculture Organization; 2015.

33. MARN, Feliz K, Rodríguez L, et al. Dominican Republic Reference Emissions Levels/Forest Reference Levels Dominican Republic 2019. p. 71.

34. Romero-Sanchez ME, Ponce-Hernandez R. Assessing and Monitoring Forest Degradation in a Deciduous Tropical Forest in Mexico via Remote Sensing Indicators. Forests. 2017;8(9).

35. Thompson ID, Guariguata MR, Okabe K, et al. An Operational Framework for Defining and Monitoring Forest Degradation. Ecology and Society. 2013;18(2).

36. Masek JG, Vermote EF, Saleous N, et al. LEDAPS Calibration, Reflectance, Atmospheric Correction Preprocessing Code, Version 2. ORNL Distributed Active Archive Center; 2013.

37. Vermote E, Justice C, Claverie M, et al. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. Remote Sensing of Environment. 2016 2016/11/01/;185:46-56.

38. Zhu Z, Woodcock CE. Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sensing of Environment. 2012 2012/03/15/;118:83-94.

39. Foga S, Scaramuzza PL, Guo S, et al. Cloud detection algorithm comparison and validation for operational Landsat data products. Remote Sensing of Environment. 2017 2017/06/01/;194:379-390.

40. Goldblatt R, Rivera Ballesteros A, Burney J. High Spatial Resolution Visual Band Imagery Outperforms Medium Resolution Spectral Imagery for Ecosystem Assessment in the Semi-Arid Brazilian Sertão. Remote Sensing. 2017;9(12).

41. Farr TG, Rosen PA, Caro E, et al. The Shuttle Radar Topography Mission. Reviews of Geophysics. 2007 2007/06/01;45(2).

42. Tan B, Masek JG, Wolfe R, et al. Improved forest change detection with terrain illumination corrected Landsat images. Remote Sensing of Environment. 2013 2013/09/01/;136:469-483.

43. Flood N. Seasonal Composite Landsat TM/ETM+ Images Using the Medoid (a Multi-Dimensional Median). Remote Sensing. 2013;5(12).

44.	developers G. Compositing and Mosaicking | Google Earth Engine | Google Developers. 2019.

45.	MARN. Emission Reductions Program Document (ER-PD), Dominican Republic. Santo Domingo, Dominican Republic2019. p. 368.

46.	Milla F, Díaz R, Emanuelli P. National Multipurpose Forest Inventory of the Dominican Republic 2014-2015. Planning and protocol elements for measurement operations. Available : http://www.reddccadgiz.org/documentos/doc_1313366786.pdf. El Salvador2014.

47.	Emanuelli P, Gonzales J, Nuñez J, et al. National Forest Inventory of the Dominican Republic 2018. Final report (ESP). Available : http://www.reddccadgiz.org/documentos/doc_1984105887.pdf. Santo Domingo, Dominican Republic2018. p. 389.

48.	MARN, Vargas O. Dominican Republic Reference Emissions Levels/Forest Reference Levels Dominican Republic 2018. p. 71.

49.	Hansen MC, Defries RS, Townshend JRG, et al. Global land cover classification at 1 km spatial resolution using a classification tree approach. International Journal of Remote Sensing. 2000 2000/01/01;21(6-7):1331-1364.

50.	Huete AR, Liu HQ, Batchily K, et al. A comparison of vegetation indices over a global set of TM images for EOS-MODIS [Article]. Remote Sensing of Environment. 1997;59(3):440-451.

51.	Huete AR. A soil-adjusted vegetation index (SAVI). Remote Sensing of Environment. 1988 1988/08/01/;25(3):295-309.

52.	Souza CM, Roberts DA, Cochrane MA. Combining spectral and spatial information to map canopy damage from selective logging and forest fires. Remote Sensing of Environment. 2005 2005/10/15/;98(2):329-343.

53.	Breiman L. Random Forests. Machine Learning. 2001 2001/10/01;45(1):5-32.

54.	Ma W, Domke GM, Woodall CW, et al. Contemporary forest carbon dynamics in the northern U.S. associated with land cover changes. Ecological Indicators. 2020 2020/03/01/;110:105901.

55.	Swingland IR, Bettelheim EC, Grace J, et al. Measuring, monitoring, and verification of carbon benefits for forest–based projects. Philosophical

Transactions of the Royal Society of London Series A: Mathematical, Physical and Engineering Sciences. 2002 2002/08/15;360(1797):1669-1683.

56.   MARN. Manual de Campo Inventario de Biomasa y Carbono en Sistemas No Bosque. Santo Domingo, Dominican Republic2017.

57.   Feliz K, Rodríguez L, Galán M, et al. Dominican Republic Reference Emissions Levels/Forest Reference Levels Dominican Republic 2019. p. 71.

58.   Bullock EL, Woodcock CE, Olofsson P. Monitoring tropical forest degradation using spectral unmixing and Landsat time series analysis. Remote Sensing of Environment. 2018 2018/11/28/:110968.

59.   Zhu Z, Woodcock CE, Olofsson P. Continuous monitoring of forest disturbance using all available Landsat imagery. Remote Sensing of Environment. 2012 2012/07/01/;122:75-91.

60.   Pasquarella VJ, Bradley BA, Woodcock CE. Near-Real-Time Monitoring of Insect Defoliation Using Landsat Time Series. Forests. 2017;8(8).

61.   Tang X, Bullock EL, Olofsson P, et al. Near real-time monitoring of tropical forest disturbance: New algorithms and assessment framework. Remote Sensing of Environment. 2019 2019/04/01/;224:202-218.

62.   Gómez C, White JC, Wulder MA, et al. Historical forest biomass dynamics modelled with Landsat spectral trajectories. ISPRS Journal of Photogrammetry and Remote Sensing. 2014 2014/07/01/;93:14-28.

63.   Arlot S, Celisse A. A survey of cross-validation procedures for model selection. Statist Surv. 2010 2010;4:40-79.

64.   Olofsson P, Foody GM, Herold M, et al. Good practices for estimating area and assessing accuracy of land change. Remote Sensing of Environment. 2014 2014/05/25/;148:42-57.

65.   GFOI. Integration of remote-sensing and ground-based observations for estimation of emissions and removals of greenhouse gases in forests: Methods and Guidance from the Global Forest Observations Initiative. Edition 2.0 ed. Rome: Food and Agriculture Organization; 2016. p. 226.

66.   Tondapu G, Markert K, Lindquist EJ, et al. A SERVIR FAO Open Source Partnership: Co-development of Open Source Web Technologies using

Earth Observation for Land Cover Mapping. AGU Fall Meeting Abstracts; December 01, 20182018. p. IN21B-27.

67.   Bey A, Sánchez-Paus Díaz A, Maniatis D, et al. Collect Earth: Land Use and Land Cover Assessment through Augmented Visual Interpretation. Remote Sensing. 2016;8(10).

68.   Souza C, Siqueira VJ, Sales HM, et al. Ten-Year Landsat Classification of Deforestation and Forest Degradation in the Brazilian Amazon. Remote Sensing. 2013;5(11).

69.   Souza C, Firestone L, Silva LM, et al. Mapping forest degradation in the Eastern Amazon from SPOT 4 through spectral mixture models. Remote Sensing of Environment. 2003 2003/11/15/;87(4):494-506.

70.   Saah D, Johnson G, Ashmall B, et al. Collect Earth: An online tool for systematic reference data collection in land cover and use applications. Environmental Modelling & Software. 2019 2019/08/01/;118:166-171.

71.   Schultz M, Clevers JGPW, Carter S, et al. Performance of vegetation indices from Landsat time series in deforestation monitoring. International Journal of Applied Earth Observation and Geoinformation. 2016 2016/10/01/;52:318-327.

72.   Tamiminia H, Salehi B, Mahdianpari M, et al. Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. ISPRS Journal of Photogrammetry and Remote Sensing. 2020 2020/06/01/;164:152-170.

73.   Sobrino JA, Raissouni N. Toward remote sensing methods for land cover dynamic monitoring: Application to Morocco. International Journal of Remote Sensing. 2000 2000/01/01;21(2):353-366.

74.   Trianni G, Lisini G, Angiuli E, et al. Scaling up to National/Regional Urban Extent Mapping Using Landsat Data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2015;8(7):3710-3719.

75.   Badgley G, Field CB, Berry JA. Canopy near-infrared reflectance and terrestrial photosynthesis. Science Advances. 2017;3(3):e1602244.

III. DIGITAL MAPPING OF SOIL ORGANIC CARBON STOCKS IN THE FOREST LANDS OF DOMINICAN REPUBLIC

Digital mapping of Soil Organic Carbon Stocks in the forest lands of Dominican Republic

Efraín Duarte [a,b], Erick Zagal [a,*], Juan A. Barrera [a], Francis Dube [c], Fabio Casco [d] and Alexander J. Hernández [e].

[a] *Department of Soils and Natural Resources, Faculty of Agronomy, University of Concepcion, Vicente Méndez 595, Casilla 537, Chillán 3812120, Chile; efrainduarte@udec.cl (E.D.); jbarrera@udec.cl (J.A.B.).*

[b] *Faculty of Agronomy, University of Concepcion, Vicente Méndez 595, Casilla 537, Chillán 3812120, Chile.*

[c] *Department of Silviculture, Faculty of Forest Sciences, University of Concepcion, Victoria 631, Casilla 160-C, Concepción 4030000, Chile; fdube@udec.cl*

[d] *Food and Agriculture Organization (FAO) of the United Nations, Tegucigalpa, Honduras; fabio.cascogutierrez@fao.org*

[e] *United States Department of Agriculture, Agricultural Research Service, Utah State University, Logan, UT 84322-6300, USA; alexander.hernandez@usda.gov*

* Correspondence: ezagal@udec.cl; Phone.: +569-42-2208853.

Abstract

Mapping the spatial distribution of soil organic carbon (SOC) in lands covered by tropical forests is important to understand the relationship and dynamics of SOC in this type of ecosystem. In this study, the Random Forest (RF) algorithm was used to map SOC stocks of topsoil (0-15 cm) in forest lands of the Dominican Republic. The methodology was developed using geospatial datasets available in the Google Earth Engine (GEE) platform combined with a set of 268 soil samples. Twenty environmental covariates were analyzed, including climate, topography, and vegetation. The results indicate that Model A (combining all 20 covariates) was only marginally better than Model B (combining topographic and

climatic covariates), and Model C (only combining multispectral remote sensing data derived from Landsat 8 OLI images). Model A and Model B yielded SOC mean values of 110.35 and 110.87 Mg C ha^{-1}, respectively. Model A reported the lowest prediction error and uncertainty with an R^2 of 0.83, an RMSE of 35.02 Mg C ha^{-1}. There was a strong dependence of SOC stocks on multispectral remote sensing data. Therefore, multispectral remote sensing proved accurate to map SOC stocks in forest ecosystems in the region.

Keywords: Random Forest; Landsat; Machine Learning; Tropical Forest; Environmental Covariates, Google Earth Engine.
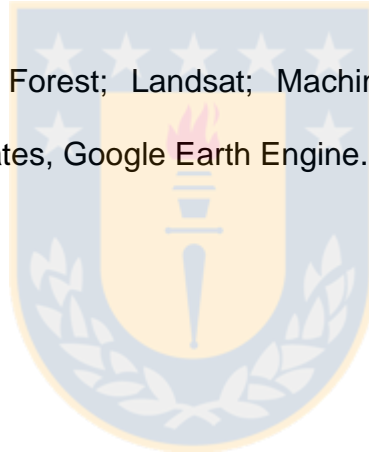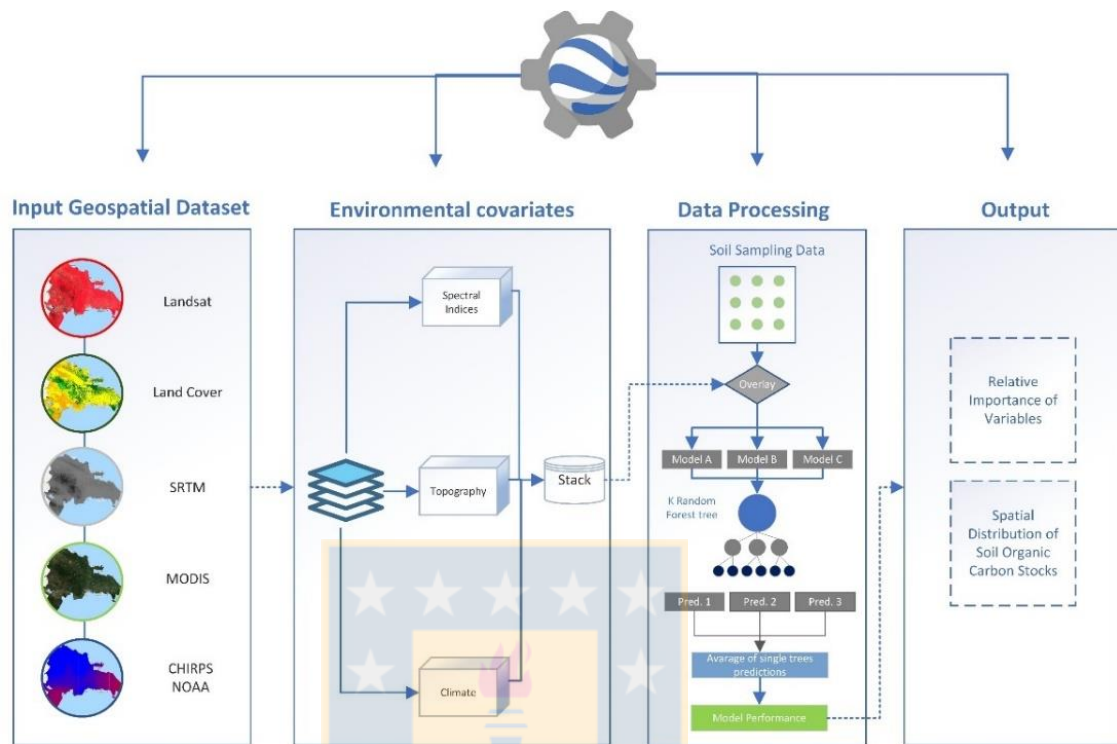
Figure 3.1 Graphical abstract

Fuente: Elaboración propia

3.1 Introduction

Soils hold the largest carbon (C) pool on Earth after the oceans, with an estimated total of 1,500-2,400 Pg C up to 1 m depth (Scharlemann et al., 2014; Tifafi et al., 2018). Soil organic carbon (SOC) directly influences the physicochemical properties nutrient retention capacity and infiltration rate of the soil (Scholten et al., 2017; Viscarra Rossel et al., 2016). In addition, SOC has the potential to mitigate the adverse impacts of current and future climate change (Edenhofer et al., 2014) and help improve the primary productivity of the biosphere (Grinand et al., 2017).

The Global Climate Observing System (GCOS) has identified 54 Essential Climate Variables (ECVs), including components of the cryosphere, biosphere, and hydrosphere that are "needed to understand and predict the evolution of climate, to guide mitigation and adaptation measures, to assess risks and enable attribution of climate events to underlying causes, and to underpin climate services" (WMO, 2020). The present work focuses on one of these ECVs, SOC stocks, particularly those stored in lands covered by tropical forests.

The global C stocks in forest biomass and their spatial distribution are relatively well documented and have been estimated with reasonable accuracy compared to SOC stocks (Baccini et al., 2012; Harris et al., 2012; Ruesch, A., Gibbs, 2008; Saatchi et al., 2011). Most local and international policies for climate change mitigation have focused on conserving and studying the C stored in forests. In addition to this, SOC is of great importance because the Earth's soils store two-to-three times as much carbon in organic form as there is C in the atmosphere globally (Trumbore, 2009). In this sense, the construction of a robust and transparent system to measure, report and verify (MRV) SOC changes represents a key tool to support compliance with the Sustainable Development Goals (SDG), specifically the SDG indicator 15.3.1 "Proportion of land that is degraded over the total land area" (FAO, 2020a; Jan & Jeffrey, 2018).

At present, there is still uncertainty about the amount of global SOC stocks and their spatial distribution, mainly due to the little attention given by decision makers at the local, national and international levels (Gianelle et al., 2010). In an analysis and review of 27 studies that estimated SOC globally, it was found that the SOC mean value is approximately 1,460.5 Pg C, ranging from 504 to 3,000 Pg C (Scharlemann et al., 2014). One of the main reasons for the uncertainties found in these estimates is the large number of factors that interfere in SOC dynamics combined with all the uncertainties leading to error propagation associated with the difficulty in assessing C and soil bulk density (Köchy et al., 2015).

Even though there is scientific interest in monitoring forests and soils, there is a lack of data to carry out efficient monitoring and determine the current state of these resources (Liang et al., 2016). In 2017, the Food and Agriculture Organization (FAO), the Intergovernmental Technical Panel on Soils (ITPS), the Intergovernmental Panel on Climate Change (IPCC), the Global Soil Partnership (GSP), the Science-Policy Interface of the United Nations Convention to Combat Desertification (UNCCD-SPI), and the World Meteorological Organization (WMO) jointly organized a Global Symposium on Soil Organic Carbon. This symposium provided guidelines for developing efficient systems and protocols for measuring SOC with higher accuracy (FAO, 2020a). In the last decade, digital soil mapping (DSM) approaches have focused on mapping SOC using remote sensing

techniques as the main emerging tool to improve spatial estimates of SOC (Mahmoudzadeh et al., 2020; Padarian et al., 2019).

DSM allows quantifying the spatial variation of SOC stocks using environmental covariates (Zhang et al., 2017), which describe the relationship of a soil attribute and its spatially implicit forming factors (Jenny, 1941). The environmental auxiliary variables of SOC can be obtained from digital elevation models (DEM) (Farr et al., 2007; B. Wang et al., 2018), remote sensing data (Duarte et al., 2020; Wulder et al., 2016; Xiao et al., 2019) and climatic data (Ermida et al., 2020; Veronesi & Schillaci, 2019). The easy accessibility of satellite images combined with Machine Learning (ML) techniques has significantly improved the accuracy of SOC mapping. In a review of 120 studies on SOC mapping, in which different ML techniques were applied, it was found that the Random Forest (RF) algorithm has optimum performance in the selection of environmental covariates for SOC mapping. At the same time, it also behaves better than other ML techniques and Multiple Linear Regression (MLR) (Lamichhane et al., 2019).

There are few studies on SOC mapping of Dominican Republic lands. The main report comes from the Global Soil Organic Carbon Map (GSOCmap) launched by FAO. In fact, GSOCmap represents the first global estimation of SOC content carried out with a participatory approach to compile all the available data

on soils at the national level (FAO & ITPS, 2020). With regard to the tropics, SOC estimates are very limited on forest lands since most of the research has focused on estimating SOC from agricultural lands.

The objective of this study was to estimate SOC content in forest lands of the Dominican Republic and their spatial distribution by applying ML techniques, a dataset of environmental covariates obtained from remote sensing (RS) and field data. We compared the influence of three groups of predictive variables for SOC mapping: (1) multispectral remote sensing variables, (2) topographic variables, and (3) climatic variables. The performance of the ML model was also evaluated. Our model was implemented in the Google Earth Engine (GEE) cloud-based computing platform (Gorelick et al., 2017).

3.2. Materials and methods

The overall structure of the method (Figure 3) consisted of five stages such as: (i) selection of a geospatial dataset; (ii) data pre-processing; (iii) model building/development; (iv) evaluation of the model performance; and (v) mapping of SOC.

3.2.1. Study area

The study was carried out in the island of Hispaniola (central region of the Caribbean), Dominican Republic. It corresponds to forest lands located between 17°36′ and 19°58′ latitude north, and 68°19′ and 72°01′ longitude west, belonging to the Greater Antilles. The territory of the country covers 48,198 km^2 (**Figure 3.2**).

The Dominican Republic has a tropical climate, which is altered only by the Alyssian winds of the Atlantic and topographical factors of the island. The average annual temperature is 25° C, with August being the hottest month and January the coldest. Precipitation is distributed in two seasons: a rainy season, which goes from April to June and from September to November, with precipitation of 2,500 mm yr^{-1}; and a dry season, which goes from December to March, with precipitation of 450 mm yr^{-1}. On the island, the areas with the highest humidity are in the north because they are influenced by the Atlantic Ocean, while the driest areas are found to the south, along the Caribbean coast (Cano-Ortiz et al., 2015).
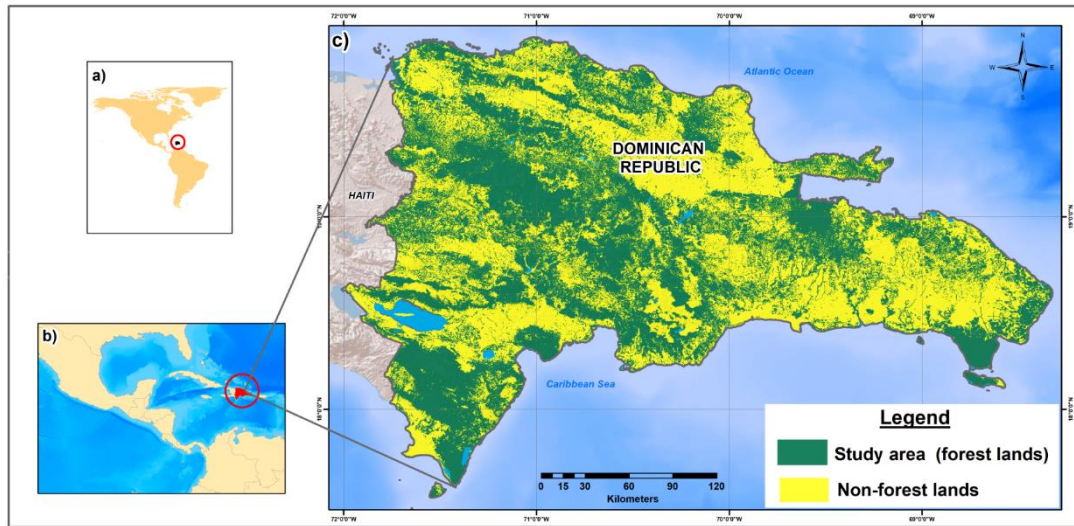
Figure 3.2 Study area: (a) General location of the Dominican Republic, (b) regional location, (c) study area (forest lands).
Fuente: Elaboración propia

The country's native forests (**Figure 3.3**) include pine, broadleaf, dry and mangrove forests. Pine forests are primarily made up of *Pinus occidentalis*, a species endemic to the island (Kennedy et al., 2005). The composition of the broadleaf forests is diverse, with species such as *Swietenia mahagoni*, *Ocotea spp., Sloanea berteriana, Didymopanax tremulus,* and *Clusia rosea*. Dry forests include various species such as *Guaiacum officinale, Phylostilum braziliensis,* and *Prosopis julifora*, while mangrove forests are composed of *Avicennia germinans, Laguncularia racemosa, Conocarpus erectus* and *Rizophora mangle* (MARN, 2019). Shrubland and herbaceous vegetation can also be found (Martin & Fahey, 2006).

The soils are divided into 10 classes based on characteristics such as depth, slope, and drainage. Soil classes correspond to savannah, non-calcareous clay and calcareous soils as well as soils derived from igneous rocks, soils of volcanic and metamorphic origins, recent alluvial soils, organic soils, wetlands, coastal beaches and dunes (MARN, 2012). Our study focused on soils covered by pine, broadleaf, dry and mangrove forests, covering an area of 14,499 km^2.

3.2.2. Reference data

3.2.2.1. Forest mask

To estimate SOC stocks, the study area was defined with a forest mask. For this, we used the land cover map prepared in 2015 by the Ministry of the Environment and Natural Resources (MARN) of the Dominican Republic. On this map, the concept of forest was applied in accordance with the FAO Forest Resources Assessment (FRA-FAO) defined by the country as *"Natural or planted ecosystem of at least 0.5 hectares covered by trees higher than 5 meters and with a canopy cover of more than 40%"* (FAO, 2020b). Forest types **(Figure 3.3)** that are part of this study are defined as follows (MARN, 2012):

Broadleaf forests: represented by trees where the combination of broad-leaved species predominates; it comprises plant communities from semi-humid in transition to cloudy. It is the type of forest with the largest existence in the country. It is classified as cloudy broadleaf forest, located in areas with elevations from

600 to 2,300 m above sea level (m.a.s.l.); moist broadleaf forest in areas with elevations from 300 to 1,500 m.a.s.l. and semi-humid broadleaf forest located in areas with elevations up to 900 meters above sea level, or m.a.s.l.

Dry forests: mostly secondary forest; they are composed of semi-deciduous trees that develop at elevations below 500 m.a.s.l. The greatest presence of these forests is located in the lowlands, both south-southwest and northwest of the country.

Pine Forests: composed of pine species. The authentic Dominican pine forest is located especially in the highlands and the dominant species is *Pinus occidentalis*, which is found in the large mountain ranges with elevations above 2,000 m.a.s.l. The pine forest has two types of cover: high and low canopy cover density.

Mangrove forests: Coastal and wet ecosystems found in swampy and flooded regions; they mainly belong to the Rhizophoraceae family, with exposed supporting roots. This forest is located at elevations below 20 m.a.s.l.
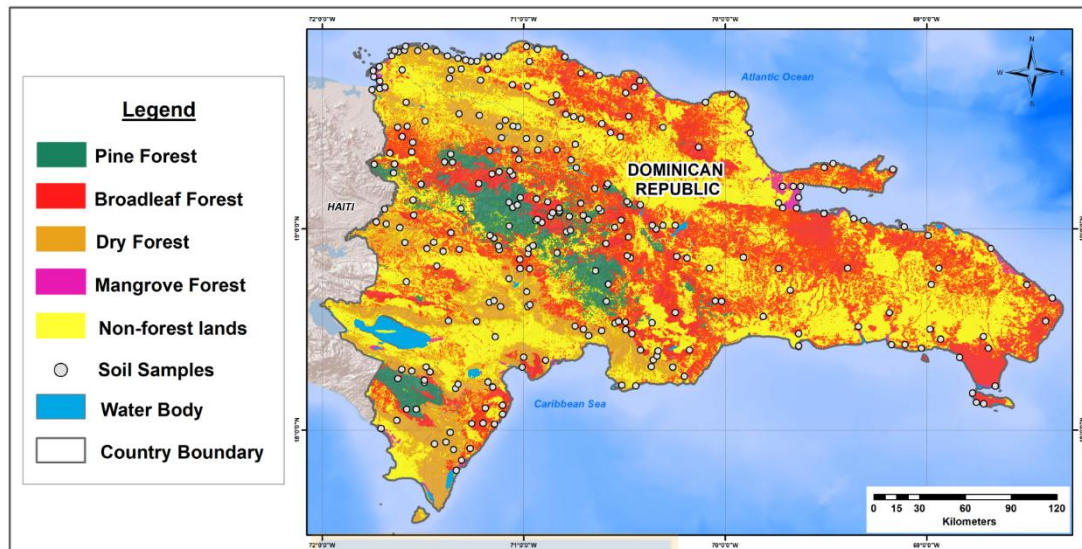
Figure 3.3 Map of soil sample location and forest types of the Dominican Republic.
Fuente: Elaboración propia

3.2.2.2 Soil organic carbon database

In this study, 268 soil samples from the National Forest Inventory (NFI) were used. The NFI was collected in 2018 by the MARN of the Dominican Republic with the support of the REDD/CCAD-GIZ program and the World Bank's Forest Carbon Partnership Facility (FCPF) (available in: https://www.sica.int/documentos/inventario-forestal-nacional-de-republica-dominicana_1_126744.html) (MARN, 2021). The sampling design adopted by the NFI corresponds to stratified sampling for each forest stratum. For building the NFI, the methodology proposed by the REDD/CCAD-GIZ program was used (Feliz et al., 2019; MARN, 2021).

Data used were taken from 268 plots located in the different forest types (**Figure 2**) (**Table 1**). Each soil sample was collected from the NFI at a depth of 0 to 15 cm and their geographic coordinates were recorded with a global positioning system device (GPS). In the NFI, SOC data were reported with an extrapolation from depth of 0 to 15 cm. The samples were numbered, bagged, and brought back to the Laboratory of Soils and Water of the Dominican Institute of Agricultural and Forestry Research Agricultural Technology Center (CENTA) and the Dominican Agribusiness Laboratory (LAD). After drying, the samples were weighed and passed through a 2-mm sieve. The determination of SOC content (Mg C ha^{-1}) is based on the Walkley & Black chromic acid wet oxidation method (Walkley & Black, 1934).

Bulk density (BD) was determined on subsamples dried at 105 °C as described by (Dane Topp G. Clarke. Campbell Gaylon S., 2002). Results were reported as g cm-3 on an oven-dry basis and SOC was reported as g (100 g)$^{-1}$. Soil organic carbon stock (SOCS) was computed as the product of three variables, organic carbon content (C), bulk density (BD), and thickness (D). SOCS was calculated according to equation 1:

$$SOC_{stock} = C \times BD \times D \times \left(1 - \frac{gravel[\%]}{100}\right) \qquad (1)$$

Where C is the concentration of soil carbon (g C (100g)-1); BD is bulk density (g cm-3), D is the thickness of the layer (cm), gravel [%] is the percentage of gravel in the soil sample.

Table 3.1 shows the descriptive statistics of SOC (0 – 15 cm depth) samples collected from the NFI. SOC contents ranged from 15.95 to 282.38, with a mean value of 110.35 and a median of 101.34 Mg C ha^{-1}. The coefficient of skewness is -0.46 Mg C ha^{-1}. The sampling point´s standard deviation (SD) is 63.78 Mg C ha^{-1} and is lower than the mean value.

Table 3.1 Descriptive statistics of soil organic carbon (SOC) stocks (Mg C ha^{-1}) (0–15 cm depth) collected from the National Forest Inventory (NFI).

Description	n	SOC (Mg C ha^{-1})					Skewness	Kurtosis
		Mean	Median	Min	Max	SD		
Dry forest	52	126.47	124.97	18.88	282.38	64.02	0.49	-0.12
Pine forest	43	69.68	57.64	19.78	187.81	41.53	1.03	0.34
Broadleaf forest	129	105.57	94.64	15.95	274.49	61.93	0.65	-0.30
Mangrove forest	44	145.06	140.53	27.92	261.87	63.43	-0.07	-0.88
Total forest	*268*	*110.35*	*101.34*	*15.95*	*282.38*	*63.78*	*0.57*	*-0.46*

Note: Min: minimum; Max: maximum; SD: standard deviation

Fuente: Elaboración propia

Once the forest mask and environmental variables was defined, a method to estimate SOC content in forest soils with a geospatial dataset was developed,

they were combined with the soil samples collected in the field, and a regression model was applied for each of the three models; the covariates were divided using the RF algorithm in the GEE platform; finally, the models were evaluated and the spatial distributions and SOC stock map was built (**Figure 3.4**).
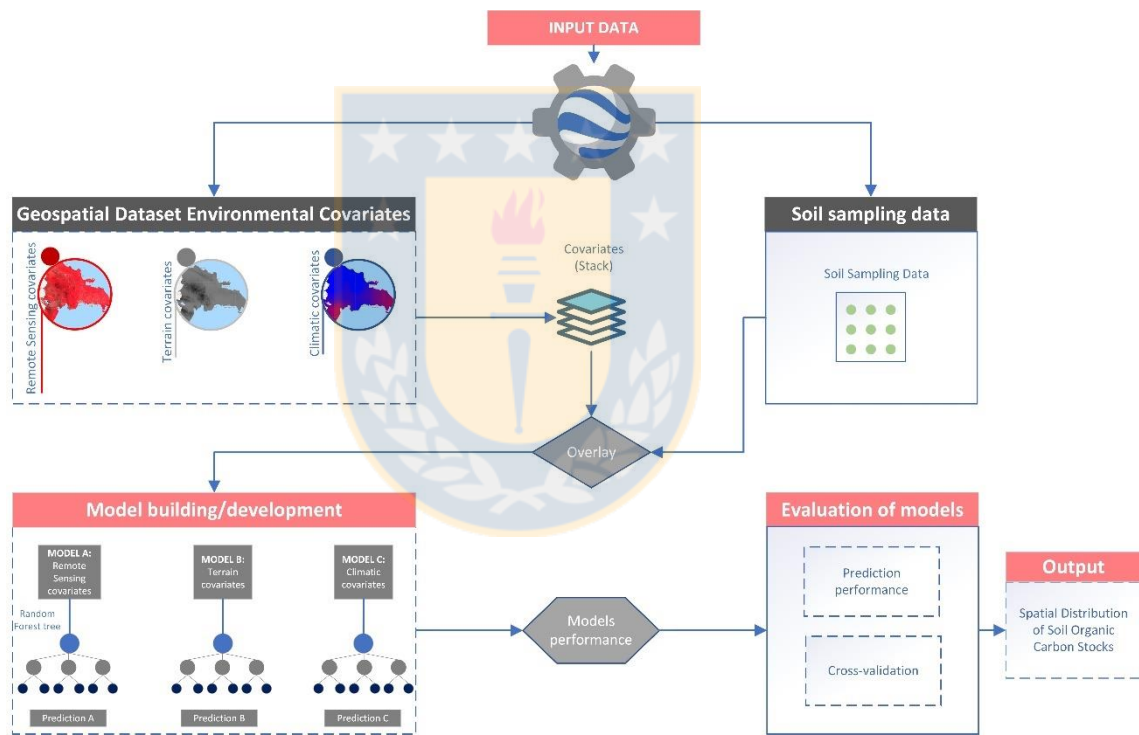
Figure 3.4 Flowchart: Dataset of environmental covariates obtained from remote sensing and combined with soil samples to define different predictive models using machine learning for mapping the spatial distribution of SOC.
Fuente: Elaboración propia

3.2.2.3 Environmental predictors

For the digital mapping of SOC, we selected 20 accessible and commonly used predictive environmental dataset covariates, which represent key factors for the spatial distribution and formation of SOC content such as: vegetation, soil, topography, and climate (McBratney et al., 2003). These covariates represent factors of soil formation according to (Jenny, 1941). Further spectral vegetation indices (SVIs) were calculated using Landsat-8 images (**Table 2**). From the combination of these dataset covariates with data soil samples, three models with different combinations of predictive variables were built, using the RF algorithm for the digital mapping of SOC. The models were as follows:

- **Model A:** Multispectral remote sensing variables + topographic variables + climatic variables.

- **Model B:** Topographic and climatic variables.

- **Model C:** Multispectral remote sensing variables.

For each of the models, the RF algorithm was applied, and its accuracy was evaluated. The relative importance of the variables in the model was also assessed. All the datasets were obtained and processed in the GEE cloud-based platform.

3.2.2.3.1 Multispectral imagery

We used Landsat 8 collection 1 Tier 1 Operational Land Imager (OLI) surface reflectance data with 16 days and 30 m resolutions, available in the GEE platform, using the Landsat Digital Number (DN) values, representing scaled, calibrated at-sensor radiance; GEE collection snippet: *ee.ImageCollection("LANDSAT/LO08/C01/T1_RT").*

All Landsat-8 OLI surface reflectance data from the year 2018 ± 0.5 (available in the GEE platform) were used in this study: a total of 92 images from path 123 and row 32; 89 images from path 125; and row 34 and 94 images from path 125 and row 33. Landsat surface reflectance data were atmospherically corrected using the Landsat Surface Reflectance Corrected (LaSRC) (OLI) algorithms (Masek et al., 2013); We used methods provided by Earth Engine for filtering image collections using the code "imageCollection.filterDate()" and we built a composite mosaic multiband and multi-date formed by combining spatially overlapping images into a single image based on a function of multiple spectral and temporal aggregation ranges; compositing in GEE refers to the process of combining spatially overlapping images into a single image based on an aggregation function, and mosaicking refers to the process of spatially assembling image datasets to produce a spatially continuous image (GEE Developers, 2021). (For more details see GEE codes developed in this study available in Appendix A

or visit https://developers.google.com/earth-engine/guides/ic_composite_mosaic)

The CFmask algorithm was used to mask cloud and shadow produced, as well as a per-pixel saturation mask (Zhu & Woodcock, 2012). In addition, the empirical Earth rotation model (ERM) was used as a basis to perform a terrain illumination correction algorithm (Tan et al., 2013), which allowed conducting the topographic correction for each image. For reflectance images, we used the medoid method (Flood, 2013).

Following radiometric, geometric, and atmospheric corrections, digital numbers for the blue (B1), green (B2), red (B3), near-infrared (B4), and shortwave IR-2 bands (B6) were extracted. Several spectral indices were then calculated: The Bare Soil Index (BSI), the Normalized Difference Vegetation Index (NDVI), the Soil-Adjusted Vegetation Index (SAVI), the Index-Based built-up Index (IBI), the Enhanced Vegetation Index (EVI) and the Green Normalized Difference Vegetation Index (GNDVI) with a spatial resolution of 30 m. (**Table 3.2**) and (**Figure 3.5**).

Table 3.2 Predictive covariates derived from Landsat 8 OLI.

Dataset (Covariates)	Abbr.	Formula	References
Remote Sensing-derived covariates			
Normalized Difference Vegetation Index	NDVI	$NDVI = \dfrac{(NIR - \text{Red})}{(NIR + Red)}$	(Sobrino & Raissouni, 2000)
Enhanced Vegetation Index	EVI	$EVI = 2.5 * \dfrac{(NIR + Red)}{(NIR + 6 * Red - 7.5 * Blue + L)}$	(Huete et al., 1997)
Soil-Adjusted Vegetation Index	SAVI	$SAVI = \dfrac{(NIR - Red)}{(NIR + Red + L)} * (1 + L)$	(Huete, 1988)
Index-Based built-up Index	IBI	$IBI = \dfrac{NDBI - (SAVI + MNDWI)/2}{NDBI + (SAVI + MNDWI)/2}$	(Xu, 2008)
Bare Soil Index	BSI	$BSI = \dfrac{(SWIR1 + Red) - (NIR + Blue)}{(SWIR1 + Red) + (NIR + Blue)}$	(Piyoosh & Ghosh, 2018)
Green Normalized Difference Vegetation Index	GNDVI	$GNDVI = \dfrac{(NIR - Green)}{(NIR + Green)}$	(Alba et al., 2017)
Near-infrared reflectance of vegetation	NIRv	$NIRv = (NDVI_{median\,monthly} - 0.08)\; x\; NIR_{median\,monthly}$	(Badgley et al., 2017)
Band 2 Blue	B_{BLUE}	Wavelength of 0.450–0.515 µm	
Band 3 Green	B_{GREEN}	Wavelength of 0.525–0.600 µm	
Band 4 Red	B_{RED}	Wavelength of 0.630–0.680 µm	
Band 5 Near Infrared	NIR	Wavelength of 0.845–0.885 µm	
Band 6 Shortwave Infrared-1	SWIR1	Wavelength of 1.560–1.660 µm	
Band 7 Shortwave Infrared-2	SWIR2	Wavelength of 2.100–2.300 µm	

Fuente: Elaboración propia
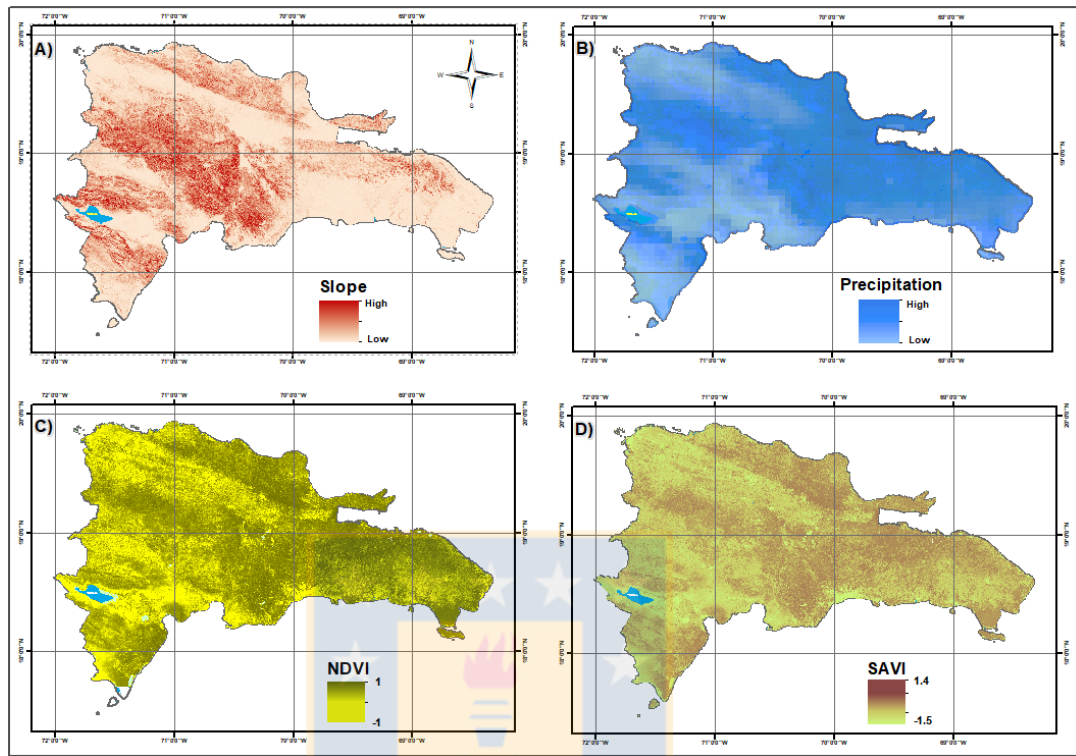
Figure 3.5 Example of the geospatial dataset used and computed on the Google Earth Engine platform. A) slope; B) precipitation; C) Normalized Difference Vegetation Index (NDVI); D) Soil-Adjusted Vegetation Index (SAVI).
Fuente: Elaboración propia

3.2.2.3.2 Climatic variables

We used the climatic datasets available in the GEE platform closest to the date soil sampling, such as Moderate Resolution Imaging Spectroradiometer (MODIS) MOD11A1 V6 product, which provides daily land surface temperature (LST) and emissivity Daily Global 1 km. *GEE collection snippet:*

114

ee.ImageCollection("MODIS/006/MOD11A1"). The temperature value is derived from the MOD11_L2 swath product (Wan, 2014).

Relative humidity data (2 m above ground) were obtained from the Global Forecast System (GFS). This is a weather forecast model produced by the National Centers for Environmental Prediction (NCEP) and National Aeronautics and Space Administration (NASA). The GFS is a coupled model, composed of an atmosphere model, an ocean model, a land/soil model, and a sea ice model, which work together to provide an accurate picture of weather conditions (Saha et al., n.d.). *GEE Collection snippet: ee.ImageCollection ("NOAA/GFS0P25").*

For the precipitation data, we used the Climate Hazards Group InfraRed Precipitation with Stations (CHIRPS) dataset, which builds on previous approaches to "smart" interpolation techniques and high resolution, long period of record precipitation estimates based on infrared Cold Cloud Duration (CCD) observations (Funk et al., 2015). *GEE collection snippet: ee.ImageCollection("UCSB-CHG/CHIRPS/DAILY").*

3.2.2.3.3 Topographic variables

Terrain analysis is crucial for modeling environmental systems. Specifically, the topography is considered as a variable that can largely explain SOC changes.

In fact, models that take topographic attributes into account can provide better estimates of SOC stocks (McBratney et al., 2003).

We used the digital elevation model (DEM) derived from NASA's Shuttle Radar Topography Mission (SRTM DEM) (Farr et al., 2007). We calculated the topographic slope, aspect and elevation from this SRTM V3 product (SRTM Plus) at a resolution of 1 arc-second (approximately 30 m). GEE collection Snippet: *ee.Image("USGS/SRTMGL1_003").* The number of terrain and climate-based covariates used within each dataset is shown in **Table 3.3**, and an example of its geographical representation is shown in **Figure 3.5.**

Table 3.3 Terrain and climate-based covariates

Dataset (Covariates)	Abbreviation	Definition
Terrain-based covariates		
1. Elevation	Elev	Height above sea level (m)
2. Slope	Slo	Average gradient above flow path
3. Aspect	Asp	The compass direction of the maximum rate of change
4. Topographic Wetness index	TWI	Combined local upslope contributing area and slope
Climatic-based covariates		
1. Temperature	Temp	It is derived from the daily temperature values
2. Precipitation	Prec	It is derived from the daily precipitation values
3. Relative Humidity	RH	Water vapor in the air, compared to how much it could hold a specific temperature

Fuente: Elaboración propia

3.2.3. Data Processing

3.2.3.1 Google Earth Engine (GEE) platform

The model for mapping SOC stocks in forests developed in the present study was built in the GEE cloud-based computing platform. GEE is a platform designed for scientific analysis at the petabyte (PB) scale and has an extensive public data catalog for earth observation (Gorelick et al., 2017). One way to use this platform is using an online tool called The Code Editor, which lets the user access the platform using a scripting language (JavaScript).

GEE has hosted historical images of the Earth for more than forty years. The images collected daily are made available to the public for data mining on a global scale. GEE allows processing massive data of a raster format for large areas and with high volumes of information. In our case, topographic, climatic and vegetational variables were analyzed (**Table 2 and Table 3**) with high performance and minimum user involvement in the processing. The algorithm used in GEE to estimate SOC contents was the Random Forest (RF). The codes developed in this study using the GEE cloud-based computing platform are available in **Appendix A.**

3.2.4 Random Forest (RF) modelling

In this study, the RF algorithm was selected to predict the SOC stocks in the forest ecosystem of the Dominican Republic. RF is one of the most popular and most powerful supervised ML algorithms that can perform both regression and classification tasks. As the name suggests, this algorithm builds a set of regression trees. Each of the trees predicts the result in each pixel, while the final prediction is obtained averaging these values (Breiman, 2001). We used the RF to estimate the relative importance of the predictive variables.

For the prediction accuracy, the 268 SOC samples were randomly divided into 2 sets: 70% of the total samples were used as model training data (n = 188), and the remaining 30% for model validation and accuracy assessment (n = 80). RF modelling was performed using the GEE cloud computing platform applying the following line of code: *ee.Classifier.smileRandomForest.setOutputMode('REGRESSION').* The principal parameters of the algorithm were: number of decision trees= 100 and default values to min leaf population (1), variables per split (square root of the number of variables), bag fraction (0.5), max nodes (defaults to no limit), seed (0) and set output mode= regression.

3.2.5. Model validation

To evaluate the performance of the SOC model, five indexes were calculated using the following formulas: coefficient of determination (R^2), Lin's concordance correlation coefficient (LCCC) (Lin, 1989), root-mean-square error (RMSE), mean absolute percentage error (MAPE), and mean absolute deviation (MAD).

$$R^2 = \frac{\left(\sum_{i=1}^{n}(y_i - \bar{y}_i)(f_i - \bar{f}_i)\right)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2 \sum_{i=1}^{n}(f_i - \bar{f}_i)^2} \tag{2}$$

$$LCCC = \frac{2r\partial_x\partial_y}{\partial_y^2 + \partial_f^2 + (\bar{y} + \bar{f})^2} \tag{3}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - f_i)^2} \tag{4}$$

$$MAPE = \frac{\sum\left|\frac{y_i - f_i}{y_i}\right|}{n} \times 100 \tag{5}$$

$$MAD = \frac{\sum|y_i - f_i|}{n} \tag{6}$$

where $n(i = 1,2,...,n)$ is the number of samples used for the ML model, y_i is the value observed (Mg C ha^{-1}), $\bar{y}_i$ is the corresponding mean value, f_i is the predicted value (Mg C ha^{-1}), $\bar{f}_i$ is mean value. ∂_x and ∂_y are the variances of the predicted and measured values; r is the correlation coefficient between the predicted value and the measured value.

The prediction algorithm with the lowest MAD, MAPE and RMSE, and highest R^2 and LCCC values are determined as the best model for SOC prediction.

3.3 Results

3.3.1 Exploratory data analysis

3.3.1.1 Geospatial environmental predictive datasets

The summary statistics for each of the 20 environmental covariates at each sampling site used in the present study are shown in **Table 3.4**. To describe the environmental covariates, they were divided into three groups that fit different models in our study: (i) imagery remote sensing data set, (ii) terrain-based covariate data set, and (iii) climatic-based covariates.

Table 3.4 Descriptive statistics of remote sensing-derived environmental variables at sample sites.

Dataset (Covariates)	Unit	Mean	Median	Min	Max	SD	Skewness	Kurtosis
Remote Sensing imagery								
NDVI	digital number	0.79	0.80	0.43	0.93	0.08	-1.02	1.22
EVI	digital number	2.36	2.35	0.90	3.56	0.44	-0.20	0.09
SAVI	digital number	1.18	1.21	0.65	1.40	0.13	-1.02	1.22
IBI	digital number	-0.27	-0.28	-0.54	0.02	0.10	0.00	-0.06
BSI	digital number	-0.29	-0.30	-0.58	0.07	0.12	0.28	-0.19
GNDVI	digital number	0.72	0.73	0.49	0.85	0.07	-0.76	0.40
NIRv	digital number	2,195.10	2,131.65	788.13	4,209.52	623.93	0.28	-0.29
B_{BLUE}	digital number	303.68	301.00	125.00	600.00	81.57	0.55	0.65
B_{GREEN}	digital number	497.71	488.00	245.00	1,007.00	114.20	0.63	1.22
B_{RED}	digital number	351.11	327.00	115.00	1,048.00	131.60	1.52	4.17
NIR	digital number	3,057.91	2,987.00	1,445.00	5,235.00	624.17	0.28	0.08
SWIR1	digital number	1,510.44	1,537.00	583.00	2,698.00	404.05	0.11	0.16
SWIR2	digital number	648.53	609.00	178.00	1,606.00	256.88	0.93	1.45
Terrain-based covariates								
Elevation	m	438.60	279.75	-1.17	2,359.85	499.42	1.65	2.43
Slope	degree	10.51	8.90	0.00	45.32	9.38	0.84	0.14
Aspect	degree	187.18	185.99	0.09	359.61	116.74	-0.07	-1.29
TWI	digital number	9.35	8.99	0.00	19.19	2.56	0.37	3.29
Climatic-based covariates								
Temperature means	degree Celsius	26.66	27.85	0.00	33.88	5.79	-3.31	12.76
Precipitation means	mm/day	3.70	3.58	0.00	7.41	1.45	0.37	-0.62
Relative humidity Means	%	77.15	76.36	62.47	86.69	5.66	0.06	-1.06

Note: NDVI, Normalized Difference Vegetation Index; EVI, Enhanced Vegetation Index; SAVI, Soil Adjust Vegetation Index; IBI, Index-Based built-up Index; BSI, Bare Soil Index; GNDVI, Green Normalized Difference Vegetation Index;

NIRv, Near-infrared reflectance of vegetation; B_{BLUE}, Landsat 8 blue band reflectance; B_{GREEN}, Landsat 8 green band reflectance; B_{RED}, Landsat 8 red band reflectance; NIR, Landsat 8 band 5 near-infrared; SWIR1, Landsat 8 band 6 Shortwave infrared-1; SWIR2, Landsat 8 band 7 Shortwave infrared-2; TWI, Topographic Wetness Index.

Fuente: Elaboración propia

3.3.1.2 Soil organic carbon models

SOC content in forest-covered soils of the Dominican Republic were estimated using three different models that grouped a series of geospatial datasets. **Table 3.5** shows the descriptive statistics of soil organic carbon (values in Mg C ha^{-1}) for each model. Model A recorded the highest mean value of SOC (110.35 Mg C ha^{-1}). The analysis of variance applied showed no significant differences ($p < 0.05$) between the three models evaluated. The use of climatic and topographic covariates helped improve the model, but not significantly. Using only multispectral imaging can produce good results in digital mapping of SOC at a depth of 0-15 cm.

Table 3.5 Descriptive statistics of predicted soil organic carbon (Mg C ha^{-1}) models.

Model	Mean	Median	Min	Max	SD	SE	Lower	Upper	Skewness	Kurtosis
Model A: Multispectral remote sensing variables + topographic variables + climatic variables	110.35	107.74	46.46	191.33	36.10	2.21	105.78	114.92	0.31	-0.85
Model B: Topographic and climatic variables	110.87	108.86	49.54	200.33	33.53	2.05	106.47	115.27	0.30	-0.59
Model C: Multispectral remote sensing variables	110.45	108.04	51.76	198.28	35.96	2.20	106.23	114.67	0.36	-0.73

Note: Min: minimum; Max: maximum; SD: standard deviation; SE: standard error; Lower and Upper: the lower and upper limits of the mean at 95% probability.

Fuente: Elaboración propia

3.3.2. Spatial model performance

Comparisons of the performance for the three dataset models by cross-validation is shown in **Table 3.6, Figure 3.6 and Figure 3.7.** We found that *Model A* is the model with the best performance, explaining 83% of the spatial variation of SOC. In general, the more predictors, the better the model. This model groups 20 predictive variables in a dataset: multispectral remote sensing variables + topographic variables + climatic variables. Model B (topographic and climatic variables) yielded an R^2 = 0.77 and an RMSE = 38.57 Mg C ha^{-1}. Interestingly,

Model C (only multispectral remote sensing-derived variables) yielded an R^2 = 0.79 and an RMSE = 35.69 Mg C ha^{-1}. These results are consistent with the significant correlations between covariates and SOC (see **Figure 3.8**).

Table 3.6 Comparison and evaluation of predicted model performance by cross-validation.

Model	Mg C ha^{-1}	R^2	RMSE (Mg C ha^{-1})	MAD	MAPE	LCCC
Model A: Multispectral remote sensing variables + topographic variables + climatic variables	110.35	0.83	35.02	27.90	40.18	0.78
Model B: Topographic and climatic variables	110.87	0.77	38.57	30.00	44.13	0.72
Model C: Multispectral remote sensing variables	110.45	0.79	35.69	28.76	41.00	0.76

Note: R^2: coefficient of determination; RMSE: root mean square error; MAD: mean absolute deviation; MAPE: mean absolute percentage error and, LCCC: Lin's concordance correlation coefficient.

Fuente: Elaboración propia
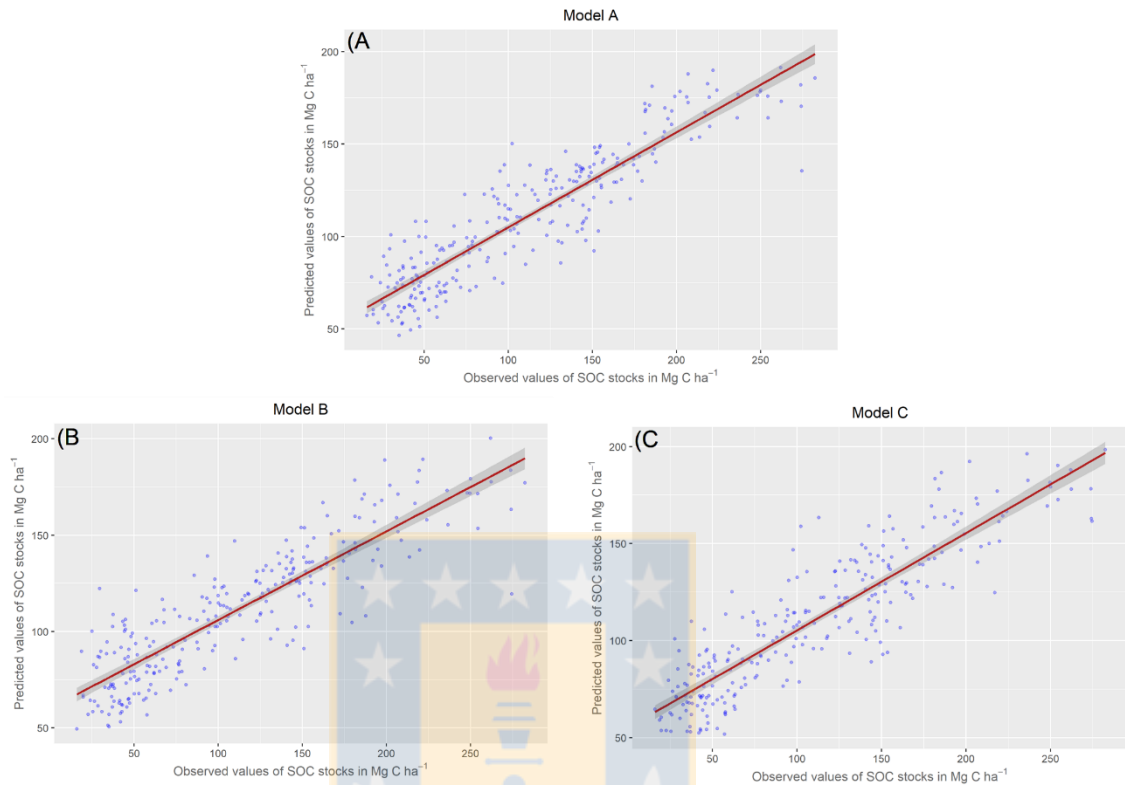
Figure 3.6 Scatter plot of observed vs. predicted values (Mg C ha⁻¹) for SOC content (0–15 cm) using the Random Forest Algorithm. A) The predicted data are derived from Model A: Multispectral remote sensing variables + topographic variables + climatic variables; B) Model B: Topographic and climatic variables and C) Model C: Multispectral remote sensing variables.
Fuente: Elaboración propia

We iterated the model A, B, and C and calculated the average standard deviation (SDs) to analyze the uncertainty of each model in predicting topsoil SOC (**Figure 3.7**). We found the highest SD in Model C for all forest types and Model A with the lower uncertainty compared to models B and C.
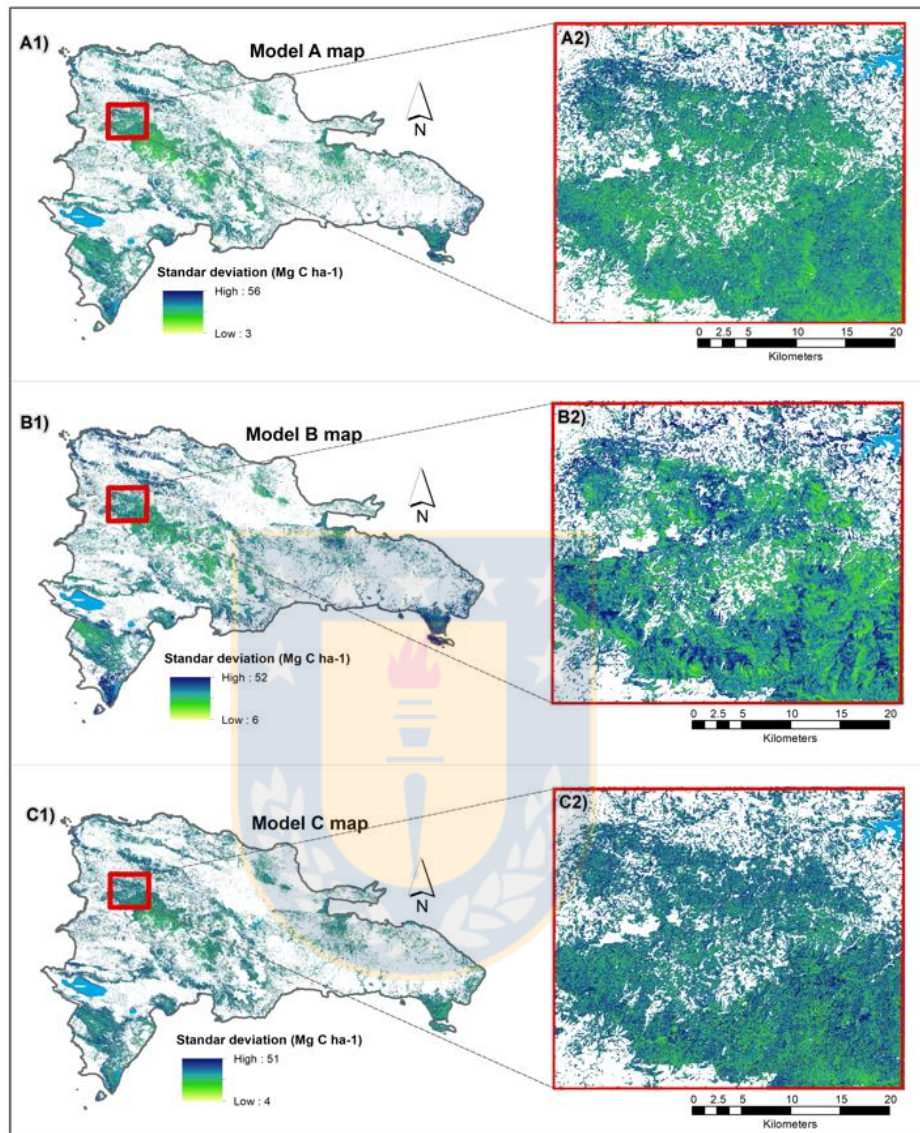
Figure 3.7 Standard deviation (SD) of SOC stock at 0–15 cm predicted from the random forest (RF) model
Fuente: Elaboración propia

The relationships between SOC content and each predictive covariate is of great importance to know information about the contribution of each covariate in the model. Pearson`s correlation analysis between SOC and predictive

covariates was derived as shown in **Figure 3.8**. SOC stock was positively correlated with temperature, TWI, B_{blue}, B_{red} and B_{green} band, but negatively correlated with precipitation, relative humidity, elevation, and slope. Interestingly, the correlations with the images were the most significant. Finally, we found that there was multicollinearity between the vegetation indexes derived from remote sensing and SOC.
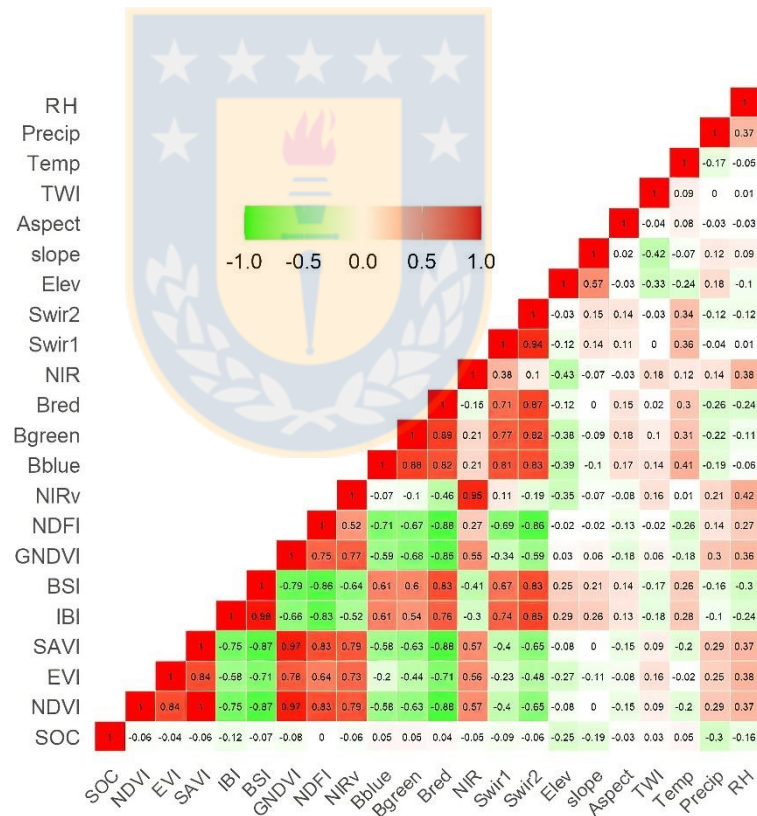
Figure 3.8 Pearson`s correlation coefficient analysis between observed SOC stocks and all environmental variables based on 268 sample sites.
Fuente: Elaboración propia

3.3.3 Covariates relative importance

The average relative importance of each covariate derived from the geospatial dataset to estimate SOC was calculated. To facilitate the analysis for each model, we combined the relative importance of all environmental covariates to 100% **(Figure 3.9**). For the dataset grouped in Model A, the 3 most important covariates were slope, temperature and NDVI (34% of the total relative importance). The vegetation indices were ranked at different levels. In Model B, which groups climatic and topographic covariates, the covariates of elevation and precipitation recorded the highest relative importance (48% of the total relative importance).

For model C, which only groups covariates derived from Landsat 8 satellite images, the Index-Based built-up Index (IBI) was the covariate with the highest relative importance in the model. For the IBI index, three thematic indices were used: the Modified Normalized Difference Water Index (MNDWI), the Soil Adjusted Vegetation Index (SAVI), and the Normalized Difference Built-up Index (NDBI) (Xu, 2008). **Table 3.2** provides further details of the indices used and **Figure 3.10** provides further details of the indices per each type of forest.
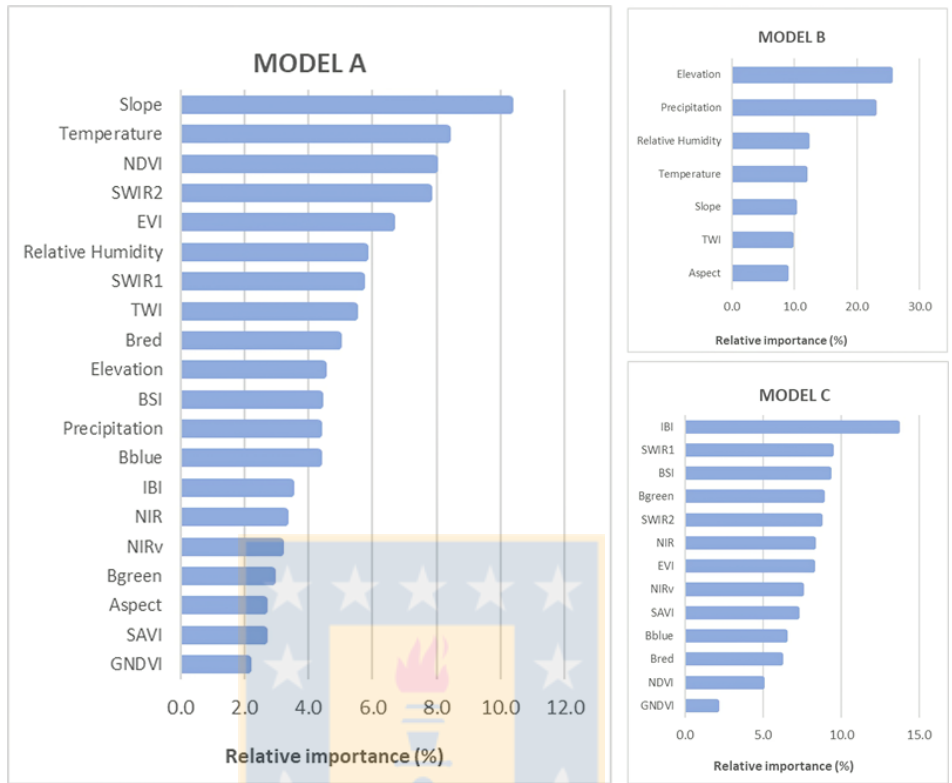
Figure 3.9 Relative importance in the Random Forest (RF) models trained for different geospatial datasets.
Fuente: Elaboración propia
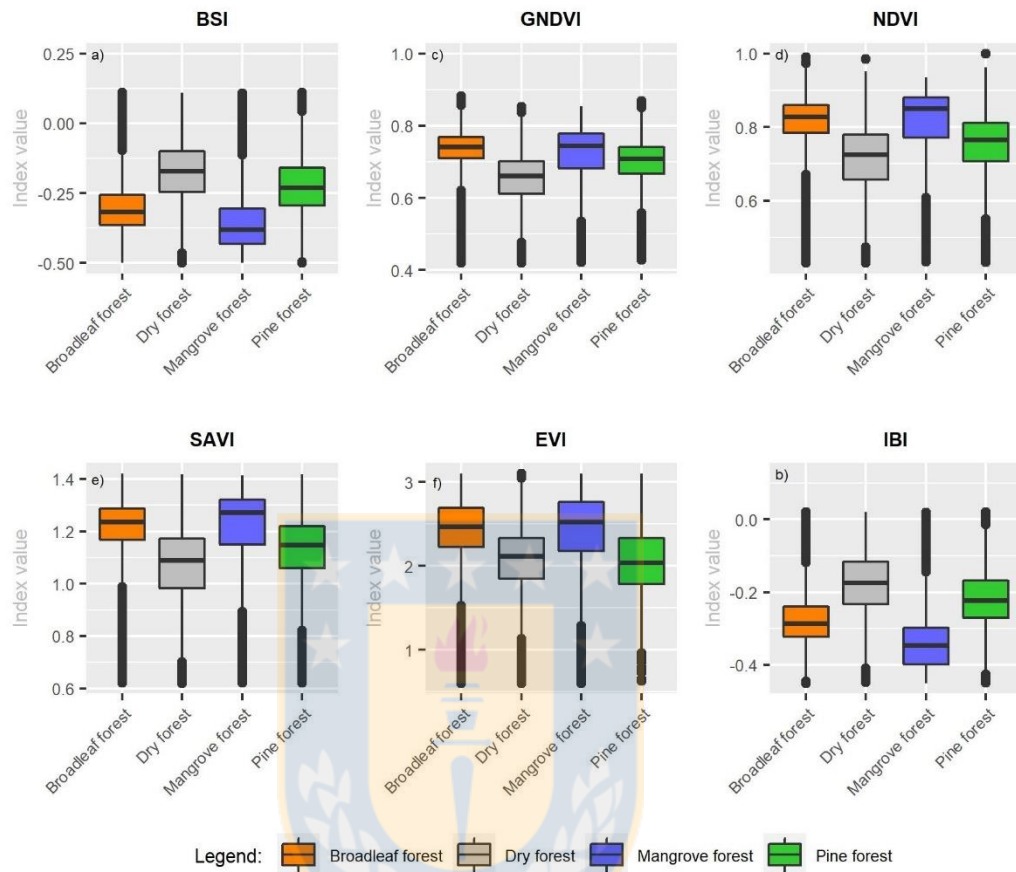
Figure 3.10 Distribution of the spectral indices for each type of forest derived from Landsat 8.
Fuente: Elaboración propia

3.3.4 SOC stock spatial distribution

A spatially explicit SOC map was created using the GEE cloud computing platform. The results obtained with the three spatial distribution models to predict SOC content in forest-covered areas of the Dominican Republic are shown in

Figure 3.11. Non-forest areas were excluded from the analysis and are shown without color on the map.

Analyzing the best model obtained in our study (Model A), identified that the spatial patterns of SOC are closely related to the type of forest. The highest SOC (0-15 cm depth) contents are found in the mangrove forests that are located in the coastal areas of the country, with average estimates of 131.87 Mg C ha^{-1}, and maximum and minimum values of 193.09 Mg C ha^{-1} and 63.91 Mg C ha^{-1}, respectively (**Table 3.7**). The lowest SOC content is found in the soils covered by pine forests, especially those located in the elevated and steep slopes, in the central and southern region of the country. These soils have a SOC mean value of 89.06 Mg C ha^{-1} and a minimum value of 44.76 Mg C ha^{-1}. Most of these soils are dominated by degraded forests with low productivity and dry shrub vegetation. Their low SOC content is attributed to steeper slopes, which make soils more susceptible to erosion and greater water discharge.

We found that in the soils covered by forests in the Dominican Republic, a total of 144,051,831 Mg C is stored in the topsoil (0-15 cm depth), with 52.1% corresponding to soils covered by broadleaf forests, 31.2% covered by dry forests, 14.3% covered by pine forests, and 2.4% covered by mangrove forests. **Table 3.7** shows more details of the results of SOC obtained in the three models for forest type.
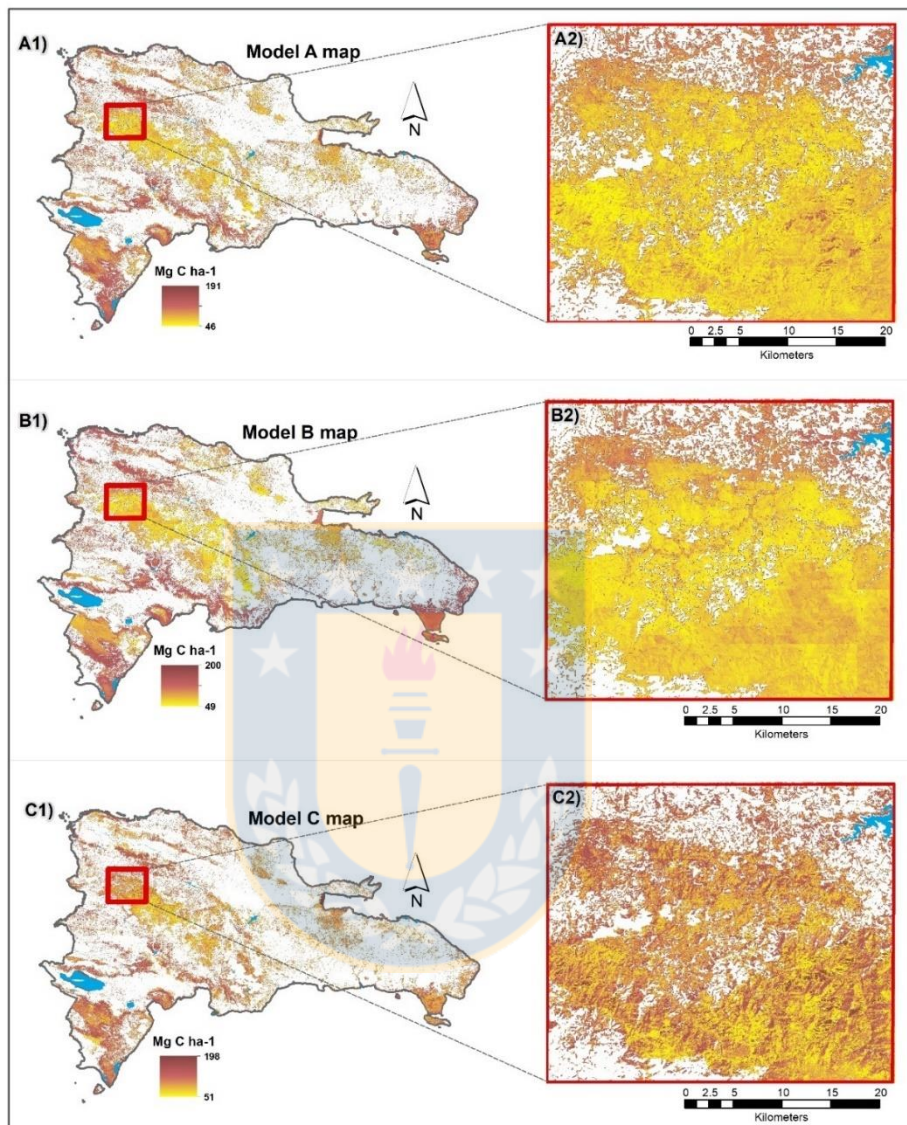
Figure 3.11 Distribution maps of soil organic carbon (Mg C ha^{-1}) derived with the Random Forest Algorithm. Maps are shown with a masking layer of non-forest land. A1) Model A included all predictive covariates (Multispectral remote sensing variables + topographic variables + climatic variables); A2) zoomed-in image to SOC map derived from Model A. B1). Model B included topographic and climatic variables; B2) zoomed-in image to SOC map derived from Model B. C1) Model C included only multispectral remote sensing variables, C2) zoomed-in image to SOC map derived from Model C.

Fuente: Elaboración propia

Table 3.7 Descriptive statistics of soil organic carbon (Mg C ha^{-1}) by forest type using different models

Model A: Multispectral remote sensing variables + topographic variables + climatic variables						
Forest type	**Mean**	**Max**	**Min**	**SD**	**Total (Mg C)**	**% of SOC**
Mangrove forest	131.87	193.09	63.95	25.88	3,451,749	2.4%
Dry forest	120.77	185.90	54.32	19.36	44,909,164	31.2%
Broadleaf forest	100.13	180.39	45.00	22.86	75,065,227	52.1%
Pine forest	89.06	167.88	44.76	19.29	20,625,691	14.3%
Total					*144,051,831*	*100.0%*

Model B: Topographic variables + climatic variables						
Forest type	**Mean**	**Max**	**Min**	**SD**	**Total (Mg C)**	**% of SOC**
Mangrove forest	129.42	201.51	64.75	18.04	3,387,413	2.3%
Dry forest	126.95	193.30	67.07	15.28	47,208,022	32.6%
Broadleaf forest	99.98	191.38	45.61	23.86	74,947,851	51.7%
Pine forest	83.73	171.13	45.63	14.62	19,392,685	13.4%
Total					*144,935,971*	*100.0%*

Model C: Multispectral remote sensing variables						
Forest type	**Mean**	**Max**	**Min**	**SD**	**Total (Mg C)**	**% of SOC**
Mangrove forest	130.86	199.48	48.58	33.62	3,277,858	2.4%
Dry forest	111.82	199.54	44.72	24.68	39,552,568	29.3%
Broadleaf forest	102.26	196.56	41.58	24.84	70,913,089	52.6%
Pine forest	94.72	189.15	41.99	22.87	21,150,587	15.7%
Total					*134,894,102*	*100.0%*

Note: Min: minimum; Max: maximum; SD: standard deviation.

Fuente: Elaboración propia

3.4. Discussion

3.4.1 Method for measuring and monitoring SOC stocks: a contribution to regional and global initiatives

SOC stocks have acquired great relevance due to the role they play in climate regulation and as an important indicator of soil quality. International organizations such as The United Nations Framework Convention on Climate Change (UNFCCC), the United Nations Convention to Combat Desertification (UNCCD) and the Convention on Biodiversity (CBD) have widely recognized the importance of SOC in the international framework of climate change mitigation. In this sense, there have been emerging regional initiatives aiming at the sustained production of soil information, such as the Soil Information System for Latin America and the Caribbean (SISLAC), or global initiatives such as ISRIC - World Soil Information, legally registered as the International Soil Reference and Information Center, which has the mission to serve the international community as custodian of global soil information.

Similarly, the Group of Earth Observation (GEO) has established a Global Soil Information System (GLOSIS) as part of the Global Earth Observation System of Systems (GEOSS). All of these initiatives have encouraged countries to establish national systems for monitoring and measuring SOC. Therefore, there

is a need to develop accurate, replicable and low-cost methods to quantify and monitor SOC stock changes.

In the present study, we performed a digital mapping of SOC stocks (in the area under study), using a combination of freely accessible geospatial datasets provided in the GEE cloud computing platform, and field data using the RF algorithm to predict the distribution of SOC in forest soils of the Dominican Republic. To our knowledge, this is the first attempt to map the SOC stocks using this type of technique, in the country and the tropical region of Central America and the Caribbean. The results obtained are encouraging because the three models used had a good performance, even with variables derived only from Landsat 8 OLI images (Model C).

Compared with other methods used to map SOC stocks in tropical regions (Guevara et al., 2018; Ramesh et al., 2015; Rossi et al., 2009; Vasques et al., 2016), our method eliminated excessive soil sampling, which can result in a high cost, particularly in those territories covered by forests where access is very difficult.

3.4.2 Importance of variables in the SOC prediction model

The CLORPT model (CL: Climate; O: Organism, vegetation; R: Relief; P: Parent material; and T: time) (Jenny, 1941) and the SCORPAN model (S: property or soil

class; C: climate; O: organisms; R: topography; P: parent material; A: age or time factor; and N: space, spatial position) (McBratney et al., 2003) are conceptual models commonly used for digital soil mapping since they relate environmental covariates to soil properties. However, these relationships between covariates and properties differ depending on the geographical area of the soil being analyzed. A review conducted by (McBratney et al., 2003) indicated that the key environmental covariates to infer soil properties were relief (80% of studies), followed by soil class (S) (35%), vegetation (O) and parent material (P) (both 25%), spatial position (N) (20%) and climate (C) (5%).

Several studies have indicated that topographic factors such as elevation and slope have a higher correlation with SOC changes (J.L. Boettinger et al., 2010; Hinge et al., 2018). The NDVI allowed us to understand the importance of the amount of biomass and vegetation cover to predict SOC stocks. Other studies have also reported that SOC can be estimated only by the presence of vegetation (Yang et al., 2008; Zhao & Shi, 2010). Therefore, the NDVI can be used as an approximation to determine SOC. The BSI allowed us to understand the importance of analyzing bare soil, especially in highly fragmented secondary forests such as the forests of the Dominican Republic and many tropical forests. This type of forest structure indicated that vegetation and bare soil combine to generate a forest with a low canopy density (mainly fragmented forests due to human intervention). Our results show that the IBI can significantly enhance the

model to predict SOC in fragmented forests with a low canopy density effectively suppressing background noise caused for bare soil. Xu, (2008) found that the IBI possesses a positive correlation with land surface temperature and negative correlations with the NDVI.

In terms of the climatic covariates used in our study, temperature was the second most important variable, which explains the SOC changes in Model A. In this sense, previous research has indicated that temperature is a direct predictor of SOC since it has a major influence on determining the type of vegetation, its growth and the microbial decomposition of organic matter (M. Wang et al., 2014).

3.4.3 Comparative analysis of other SOC measurements and mapping initiatives in the region

There have been different local initiatives for the measurement of SOC in the Central American Region in recent years. These have focused on collecting soil C data as part of a multipurpose methodology of local forest inventories, including the 5 pools of carbon defined by the IPCC. However, a wall-to-wall mapping of SOC has not been generated yet. Our study is the first report in which soil C data is used in combination with ML techniques and open-access dataset and available in the GEE cloud computing platform for geographically explicit mapping of the SOC.

ML techniques are widely used in digital SOC mapping as they combine complex and non-linear relationships between different soil attributes and predictive environmental covariates (Drake et al., 2006). Although various prediction algorithms have different capabilities, size of the training sample affects more than the selection of models to improve the prediction accuracy of SOC (Somarathna et al., 2017).

By using the RF algorithm, Model A yielded a SOC mean value (0 -15 cm) of 110.35 Mg C ha^{-1} (**Table 3.5)**. This value is higher than the mean value of 81.04 Mg C ha^{-1} reported in the Global Soil Organic Carbon Map (GSOCmap V1.5) prepared by the Food and Agricultural Organization (FAO). The global soil carbon map consists of national SOC maps, developed as 1 km soil grids, at a depth of 0-30 cm (FAO & ITPS, 2020); this is the main and most recent SOC mapping initiative existing in the region with which we can compare the results obtained herein. Another comparison with actual SOC mapping, shows that the SOC mean value obtained in the present study is lower than that of 128.80 Mg C ha^{-1} reported in SoilGrids - global gridded soil information; this is a system for digital soil mapping based on a global compilation of soil profile data (WoSIS) and environmental layers, which uses state-of-the-art ML methods to map SOC contents at a depth of 0-15 cm and 250 m resolution (Hengl et al., 2017).

We believe that, due to the source of the data, depth of the soil sample, spatial resolution, scale of the map and the technique used to predict SOC contents, our results are slightly different from the results reported by these two global initiatives for the Dominican Republic. **Figure 3.12** shows a comparative analysis of the three maps described.
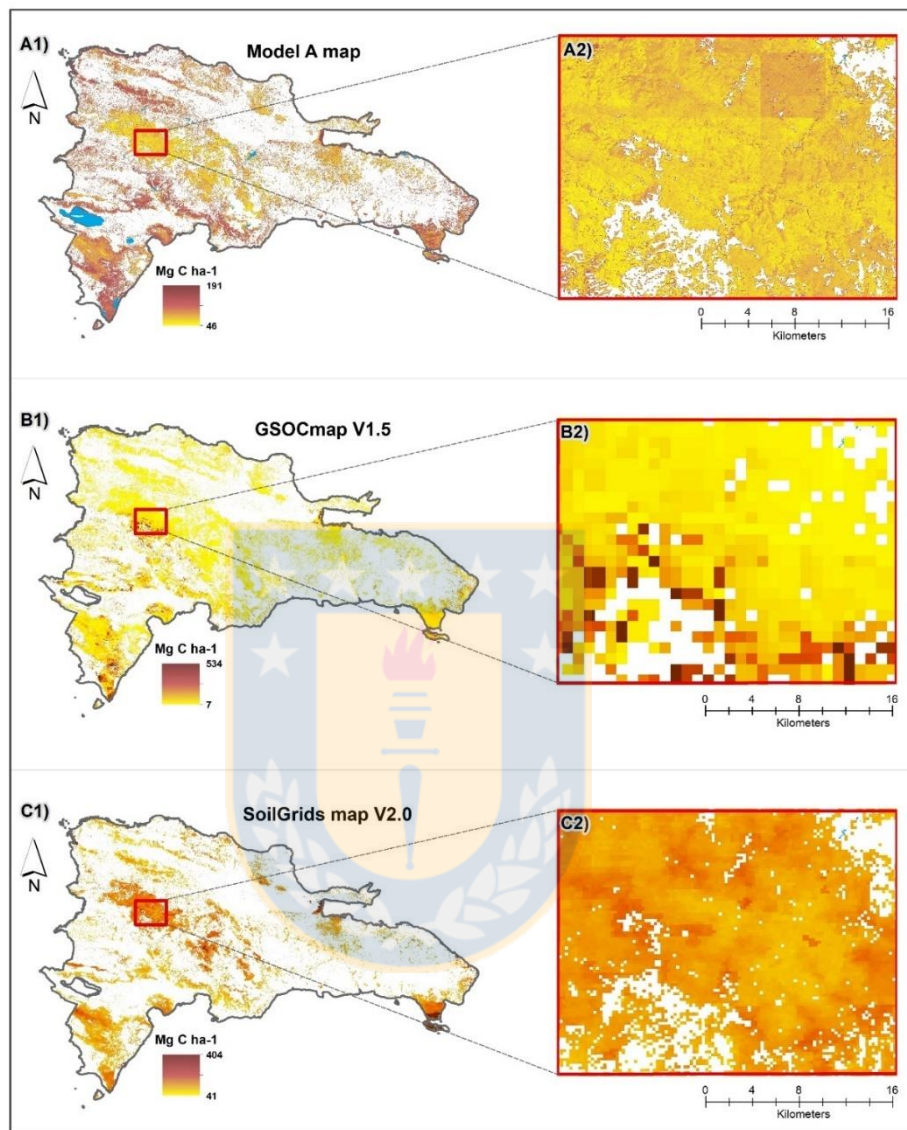
Figure 3.12 Comparative map of soil organic carbon (Mg C ha^{-1}) obtained with Model A versus Global Soil Organic Carbon Map (GSOCmap V1.5) and SoilGrid map 2.0. A1). Model A included all predictive covariates (Multispectral remote sensing variables + topographic variables + climatic variables); A2) zoomed-in image to SOC map derived from Model A. B1). Global Soil Organic Carbon Map (GSOCmap V1.5); B2) zoomed-in image to GSOCmap V1.5. C1). SoilGrids map V2.0); C2) zoomed-in image to SoilGrids map V2.0.
Fuente: Elaboración propia

Other reports on SOC in Central America are found in El Salvador and Costa Rica, where soil C was measured at 0-20 cm and 0-30 cm depths, respectively. Both measurements were part of the report of the National Multipurpose Forest Inventory, which was built to quantify the SOC stocks of those countries in order to report Forest Reference Emission Levels (FREL/REDD+) to the UNFCCC. However, geographically explicit maps of SOC stocks were not developed. In El Salvador, the mean SOC value in soils covered by forests was 137.45 Mg C ha^{-1} at 0-20 cm depth (MARN & García, 2018), and Costa Rica, the mean value was 108.81 Mg C ha^{-1} at 0-30 cm depth (Emanuelli et al., 2015); both reports are close to those obtained in our study (110.35 Mg C ha^{-1}). As mentioned above, these are the most recent reports for the region, and they provide evidence of the potential that our methodology has as it can be replicated in other countries in the future, and thus contribute to SOC mapping at the regional and global levels.

3.5. Conclusions

The present study developed and applied a methodology for SOC mapping in forest lands, using geospatial datasets available in the GEE platform. This approach opens new possibilities for applying ML techniques that will allow countries to develop robust, transparent, consistent, and replicable systems for measuring and monitoring C in soils. In our study, we found a better coefficient of determination in Model A (topographic, climatic, and Landsat 8 OLI imagery

datasets), and determined that the SOC content is mostly related to slope, temperature, and NDVI.

This study shows that freely accessible multispectral optical imagery available in the GEE platform, such as Landsat 8 OLI, can by itself estimate SOC with adequate accuracy in the tropical forests of the Dominican Republic. Adding climatic and topographic covariates improved the model, but not significantly. The results obtained allow indicating that multispectral images are a good tool for SOC digital mapping at 0-15 cm depth.

ML helps simplify model adjustments. This is a great advantage because ML allows mapping SOC content using many predictive variables of climatic, topographic, or vegetation type with minimum human interaction; however, using soil samples distributed and stratified by forest type was crucial to improve model prediction of SOC content with an R^2 of 83%.

We found that the GEE platform has excellent potential in "wall-to-wall" SOC mapping in forest lands. Further research is required on the use of these tools for SOC mapping in different land uses (e.g., agricultural and livestock soils), different ecosystems beyond the tropics, and at further depths. We can conclude that the methodology developed may encourage new research that favors the fulfillment of the Pillar 4 Implementation Plan towards a Global Soil Information

System within the framework of the Global Soil Alliance, and especially the support of Indicator 15.3.1 of the Objectives of Sustainable Development.

Future research is needed to evaluate; i) new spatial datasets available such as SoilGrid database, Sentinel-1 Synthetic Aperture Radar (SAR) data, and other satellite images like MODIS, Sentinel - 2 MultiSpectral Instrument (MSI), or Planet. ii) other algorithms such as deep learning (neural network), iii) perform hyperparameter estimation and optimization of the RF model to ensure maximum model accuracy, iv) increase the number of SOC stock samples to improve model performance, and v) assessing the land use and land cover change effect on soil organic carbon from landscape to national scale are some outlooks for future studies.

Credit authorship contribution statement:

Efraín Duarte: Conceptualization, Methodology, Software, Formal analysis, Writing - original draft, Data curation. **Erick Zagal:** Supervision, Reviewing and Editing. **Juan A. Barrera:** Conceptualization. **Francis Dube:** Reviewing and Editing. **Fabio Casco:** Conceptualization, Methodology, Software. **Alexander J Hernández:** Conceptualization.

Appendix A

The codes developed in this study using the GEE cloud-based computing platform are available at the following link: https://github.com/EDuarteCode/SOC_Code.git

Capítulo III: Glosario de abreviaturas, siglas y acrónimos

BSI	Bare Soil Index
CCD	Cold Cloud Duration
CENTA	Agricultural and Forestry Research Agricultural Technology Center
CHIRPS	Climate Hazards Group InfraRed Precipitation with Stations
DEM	Digital Elevation Models
DN	Digital Number
DSM	Digital Soil Mapping
ECVs	Essential Climate Variables
ERM	Earth rotation model
EVI	Enhanced Vegetation Index
FAO	United Nations Food and Agriculture Organization
FCPF	Forest Carbon Partnership Facility
FRA	Global Forest Resource Assessment
FREL/FRL	Reference Emission Levels/Forest Reference Levels
GEE	Google Earth Engine
GEO	Group of Earth Observation
GCOS	Global Climate Observing System
GFS	Global Forecast System
GSOCmap	Global Soil Organic Carbon Map
GNDVI	Green Normalized Difference Vegetation Index
GSP	Global Soil Partnership
IBI	Index-Based built-up Index
IPCC	Intergovernmental Panel on Climate Change
ITPS	Intergovernmental Technical Panel on Soils
LAD	Dominican Agribusiness Laboratory
LaSRC	Landsat Surface Reflectance Corrected
LCCC	Lin's concordance correlation coefficient

LST	Land Surface Temperature
MAD	Mean Absolute Deviation
MAPE	Mean Absolute Percentage Error
MARN	Ministry of the Environment and Natural Resources
m.a.s.l	Meters Above Sea Level
MODIS	Moderate Resolution Imaging Spectroradiometer
Mg C	Megagramme Carbon
MNDWI	Modified Normalized Difference Water Index
NDBI	Normalized Difference Built-up Index
ML	Machine Learning
MLR	Multiple Linear Regression
MRV	Monitoring, Reporting, and Verification
NASA	National Aeronautics and Space Administration
NCEP	National Centers for Environmental Prediction
NDVI	Normalized Difference Vegetation Index
NFI	National Forest Inventory
OLI	Operational Land Imager
PB	Petabyte
Pg C	Petagram of carbon
REDD+	Reducing Emissions from Deforestation and Forest Degradation
RF	Random Forest
RMSE	Root-Mean-Square Error
RS	Remote Sensing
SAVI	Soil Adjust Vegetation Index
SOC	Soil Organic Carbon
SOCS	Soil organic carbon stock
SISLAC	System for Latin America and the Caribbean
SD	Standard Deviation
SVIs	spectral vegetation indices

SRTM   Shuttle Radar Topography Mission

UNFCCC   United Nations Framework Convention on Climate Change

UNCCD   United Nations Convention to Combat Desertification

UNCCD-   Science-Policy Interface of the United Nations Convention to

SPI   Combat Desertification

WMO   World Meteorological Organization

3.6 References

Alba, E., Mello, E. P., Marchesan, J., Silva, E. A., Tramontina, J., & Pereira, R. S. (2017). Spectral characterization of forest plantations with Landsat 8/OLI images for forest planning and management. *Pesquisa AgropecuÃ!`ria Brasileira*, *52*, 1072–1079. http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-204X2017001101072&nrm=iso

Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statist. Surv.*, *4*, 40–79. https://doi.org/10.1214/09-SS054

Baccini, A., Goetz, S. J., Walker, W. S., Laporte, N. T., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P. S. A., Dubayah, R., Friedl, M. A., Samanta, S., & Houghton, R. A. (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change*, *2*(3), 182–185. https://doi.org/10.1038/nclimate1354

Badgley, G., Field, C. B., & Berry, J. A. (2017). Canopy near-infrared reflectance and terrestrial photosynthesis. *Science Advances*, *3*(3), e1602244. https://doi.org/10.1126/sciadv.1602244

Boettinger, J. L., Howell, D. W., Moore, A. C., Hartemink, A. E., & Kienast-Brown, S. (2010). *Digital Soil Mapping*. Springer Netherlands. https://doi.org/10.1007/978-90-481-8863-5

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Cano-Ortiz, A., Musarella, C. M., Piñar, J. C., Spampinato, G., Veloz, A., & Cano, E. (2015). Vegetation of the dry bioclimatic areas in the Dominican Republic. *Plant Biosystems - An International Journal Dealing with All Aspects of Plant Biology*, *149*(3), 451–472. https://doi.org/10.1080/11263504.2015.1040482

Dane Topp G. Clarke. Campbell Gaylon S., J. H. (2002). *Methods of soil analysis. Part 4, Part 4,*. Soil Science Society of America.

Drake, J., Randin, C., & Guisan, A. (2006). Modelling ecological niches with support vector machines. *Journal of Applied Ecology*, *43*(3), 424–432. https://doi.org/https://doi.org/10.1111/j.1365-2664.2006.01141.x

Duarte, E., Barrera, J. A., Dube, F., Casco, F., Hernández, A. J., & Zagal, E. (2020). Monitoring Approach for Tropical Coniferous Forest Degradation Using Remote Sensing and Field Data. In *Remote Sensing* (Vol. 12, Issue 16). https://doi.org/10.3390/rs12162531

Edenhofer, O., Pichs-Madruga, R., Sokona, Y., Seyboth, K., Kadner, S., Zwickel, T., Eickemeier, P., Hansen, G., & Schlömer, S. (2014). *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (F. E. Edenhofer O., Pichs-Madruga R., Sokona Y., Ed.). Cambridge University Press.

Emanuelli, P., Milla, F., Duarte, E., Emanuelli, J., Jiménez, A., & Chavarría, M. (2015). *Inventario Nacional Forestal de Costa Rica 2014-2015*.

Ermida, S. L., Soares, P., Mantas, V., Göttsche, F.-M., & Trigo, I. F. (2020). Google Earth Engine Open-Source Code for Land Surface Temperature Estimation from the Landsat Series. In *Remote Sensing* (Vol. 12, Issue 9). https://doi.org/10.3390/rs12091471

FAO. (2020a). *A protocol for measurement, monitoring, reporting and verification of soil organic carbon in agricultural landscapes – GSOC-MRV Protocol*. https://doi.org/https://doi.org/10.4060/cb0509en

FAO. (2020b). *Forest Resources Assessment, Dominican Republic Report*. http://www.fao.org/3/cb0101es/cb0101es.pdf

FAO, & ITPS. (2020). *Global Soil Organic Carbon Map (GSOCmap) Version 1.5*. https://doi.org/https://doi.org/10.4060/ca7597en

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., & Alsdorf, D. (2007). The Shuttle Radar Topography Mission. *Reviews of Geophysics*, *45*(2). https://doi.org/https://doi.org/10.1029/2005RG000183

Feliz, K., Rodríguez, L., Galán, M., Ovidio, R., Vargas, O., & de Jong, B. (2019). *Dominican Republic Reference Emissions Levels/Forest Reference Levels*. MARN. https://redd.unfccc.int/files/nrfe_-_nrf_rep._dom_revgov2.pdf

Flood, N. (2013). Seasonal Composite Landsat TM/ETM+ Images Using the Medoid (a Multi-Dimensional Median). In *Remote Sensing* (Vol. 5, Issue 12). https://doi.org/10.3390/rs5126481

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, *2*(1), 150066. https://doi.org/10.1038/sdata.2015.66

Gianelle, D., Oechel, W., Miglietta, F., Rodeghiero, M., & Sottocornola, M. (2010). Cataloguing soil carbon stocks. In *Science (New York, N.Y.)* (Vol. 330, Issue 6010, pp. 1476–1477). https://doi.org/10.1126/science.330.6010.1476-c

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/https://doi.org/10.1016/j.rse.2017.06.031

Grinand, C., Maire, G. le, Vieilledent, G., Razakamanarivo, H., Razafimbelo, T., & Bernoux, M. (2017). Estimating temporal changes in soil carbon stocks at ecoregional scale in Madagascar using remote-sensing. *International Journal of Applied Earth Observation and Geoinformation*, *54*, 1–14. https://doi.org/https://doi.org/10.1016/j.jag.2016.09.002

Guevara, M., Olmedo, G. F., Stell, E., Yigini, Y., Aguilar Duarte, Y., Arellano Hernández, C., Arévalo, G. E., Arroyo-Cruz, C. E., Bolivar, A., Bunning, S., Bustamante Cañas, N., Cruz-Gaistardo, C. O., Davila, F., Dell Acqua, M., Encina, A., Figueredo Tacona, H., Fontes, F., Hernández Herrera, J. A., Ibelles Navarro, A. R., … Vargas, R. (2018). No silver bullet for digital soil mapping: country-specific soil organic carbon estimates across Latin America. *SOIL*, *4*(3), 173–193. https://doi.org/10.5194/soil-4-173-2018

Harris, N. L., Brown, S., Hagen, S. C., Saatchi, S. S., Petrova, S., Salas, W., Hansen, M. C., Potapov, P. v, & Lotsch, A. (2012). Baseline map of carbon emissions from deforestation in tropical regions. *Science (New York, N.Y.)*, *336*(6088), 1573–1576. https://doi.org/10.1126/science.1217962

Hengl, T., Mendes de Jesus, J., Heuvelink, G. B. M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S., & Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLOS ONE*, *12*(2), e0169748. https://doi.org/10.1371/journal.pone.0169748

Hinge, G., Surampalli, R. Y., & Goyal, M. K. (2018). Prediction of soil organic carbon stock using digital mapping approach in humid India. *Environmental Earth Sciences*, *77*(5), 172. https://doi.org/10.1007/s12665-018-7374-x

Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, *25*(3), 295–309. https://doi.org/https://doi.org/10.1016/0034-4257(88)90106-X

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*(3), 440–451. https://doi.org/https://doi.org/10.1016/S0034-4257(96)00112-5

Jan, B., & Jeffrey, U. (2018). Achieving the United Nations Sustainable Development Goals: An enabling role for accounting research. *Accounting, Auditing & Accountability Journal*, *31*(1), 2–24. https://doi.org/10.1108/AAAJ-05-2017-2929

Jenny, H. (1941). Factors of soil formation: a system of quantitative pedology. *Macgraw Hill*.

Kennedy, L. M., Horn, S. P., & Orvis, K. H. (2005). Modern pollen spectra from the highlands of the Cordillera Central, Dominican Republic. *Review of Palaeobotany and Palynology*, *137*(1), 51–68. https://doi.org/https://doi.org/10.1016/j.revpalbo.2005.08.007

Köchy, M., Hiederer, R., & Freibauer, A. (2015). Global distribution of soil organic carbon – Part 1: Masses and frequency distributions of SOC stocks for the tropics, permafrost regions, wetlands, and the world. *SOIL*, *1*(1), 351–365. https://doi.org/10.5194/soil-1-351-2015

Lamichhane, S., Kumar, L., & Wilson, B. (2019). Digital soil mapping algorithms and covariates for soil organic carbon mapping and their implications: A review. *Geoderma*, *352*, 395–413. https://doi.org/https://doi.org/10.1016/j.geoderma.2019.05.031

Liang, X., Kankare, V., Hyyppä, J., Wang, Y., Kukko, A., Haggrén, H., Yu, X., Kaartinen, H., Jaakkola, A., Guan, F., Holopainen, M., & Vastaranta, M. (2016). Terrestrial laser scanning in forest inventories. *ISPRS Journal of Photogrammetry and Remote Sensing*, *115*, 63–77. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2016.01.006

Lin, L. I.-K. (1989). A Concordance Correlation Coefficient to Evaluate Reproducibility. *Biometrics*, *45*(1), 255–268. https://doi.org/10.2307/2532051

Mahmoudzadeh, H., Matinfar, H. R., Taghizadeh-Mehrjardi, R., & Kerry, R. (2020). Spatial prediction of soil organic carbon using machine learning techniques in western Iran. *Geoderma Regional*, *21*, e00260. https://doi.org/https://doi.org/10.1016/j.geodrs.2020.e00260

MARN. (2012). *Atlas of the Biodiversity and Natural Resources of the Dominican Republic.* https://biodiversidad-rd.net/atlas-de-la-biodiversidad-y-de-los-recursos-naturales-de-la-republica-dominicana-2012/

MARN. (2019). *Emission Reductions Program Document (ER-PD).* https://www.forestcarbonpartnership.org/country/dominican-republic

MARN. (2021). *National Forest Inventory of the Dominican Republic*. https://www.sica.int/documentos/inventario-forestal-nacional-de-republica-dominicana_1_126744.html

MARN, & García, C. (2018). *Inventario Nacional de Bosques de El Salvador*. https://cidoc.marn.gob.sv/documentos/inventario-nacional-de-bosques/

Martin, P. H., & Fahey, T. J. (2006). Fire history along environmental gradients in the subtropical pine forests of the Cordillera Central, Dominican Republic. *Journal of Tropical Ecology*, *22*(3), 289–302. https://doi.org/DOI: 10.1017/S0266467406003178

Masek, J. G., Vermote, E. F., Saleous, N., Wolfe, R., Hall, F. G., Huemmrich, K. F., Gao, F., Kutler, J., & Lim, T. K. (2013). *LEDAPS Calibration, Reflectance, Atmospheric Correction Preprocessing Code, Version 2*. ORNL Distributed Active Archive Center. https://doi.org/10.3334/ORNLDAAC/1146

McBratney, A. B., Mendonça Santos, M. L., & Minasny, B. (2003). On digital soil mapping. *Geoderma*, *117*(1), 3–52. https://doi.org/https://doi.org/10.1016/S0016-7061(03)00223-4

Padarian, J., Minasny, B., & McBratney, A. B. (2019). Using deep learning for digital soil mapping. *SOIL*, *5*(1), 79–89. https://doi.org/10.5194/soil-5-79-2019

Piyoosh, A. K., & Ghosh, S. K. (2018). Development of a modified bare soil and urban index for Landsat 8 satellite data. *Geocarto International*, *33*(4), 423–442. https://doi.org/10.1080/10106049.2016.1273401

Ramesh, T., Manjaiah, K. M., Mohopatra, K. P., Rajasekar, K., & Ngachan, S. v. (2015). Assessment of soil organic carbon stocks and fractions under different agroforestry systems in subtropical hill agroecosystems of north-east India. *Agroforestry Systems*, *89*(4), 677–690. https://doi.org/10.1007/s10457-015-9804-z

Rossi, J., Govaerts, A., de Vos, B., Verbist, B., Vervoort, A., Poesen, J., Muys, B., & Deckers, J. (2009). Spatial structures of soil organic carbon in tropical forests—A case study of Southeastern Tanzania. *CATENA*, *77*(1), 19–27. https://doi.org/https://doi.org/10.1016/j.catena.2008.12.003

Ruesch, A., Gibbs, H. (2008). *New IPCC Tier-1 Global Biomass Carbon Map for the Year 2000*. https://doi.org/10.15485/1463800

Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T. A., Salas, W., Zutta, B. R., Buermann, W., Lewis, S. L., Hagen, S., Petrova, S., White, L., Silman, M., & Morel, A. (2011). Benchmark map of forest carbon stocks in tropical regions

across three continents. *Proceedings of the National Academy of Sciences*, *108*(24), 9899 LP – 9904. https://doi.org/10.1073/pnas.1019576108

Saha, S., Nadiga, S., Thiaw, C., Wang, J., Wang, W., Zhang, Q., van den Dool, H. M., Pan, H.-L., Moorthi, S., Behringer, D., Stokes, D., Pe?a, M., Lord, S., White, G., Ebisuzaki, W., Peng, P., & Xie, P. (n.d.). The NCEP Climate Forecast System. *Journal of Climate*, *19*(15), 3483–3517. https://doi.org/10.1175/JCLI3812.1

Scharlemann, J. P. W., Tanner, E. V. J., Hiederer, R., & Kapos, V. (2014). Global soil carbon: understanding and managing the largest terrestrial carbon pool. *Carbon Management*, *5*(1), 81–91. https://doi.org/10.4155/cmt.13.77

Scholten, T., Goebes, P., Kühn, P., Seitz, S., Assmann, T., Bauhus, J., Bruelheide, H., Buscot, F., Erfmeier, A., Fischer, M., Härdtle, W., He, J.-S., Ma, K., Niklaus, P. A., Scherer-Lorenzen, M., Schmid, B., Shi, X., Song, Z., von Oheimb, G., … Schmidt, K. (2017). On the combined effect of soil fertility and topography on tree growth in subtropical forest ecosystems—a study from SE China. *Journal of Plant Ecology*, *10*(1), 111–127. https://doi.org/10.1093/jpe/rtw065

Sobrino, J. A., & Raissouni, N. (2000). Toward remote sensing methods for land cover dynamic monitoring: Application to Morocco. *International Journal of Remote Sensing*, *21*(2), 353–366. https://doi.org/10.1080/014311600210876

Somarathna, P. D. S. N., Minasny, B., & Malone, B. P. (2017). More Data or a Better Model? Figuring Out What Matters Most for the Spatial Prediction of Soil Carbon. *Soil Science Society of America Journal*, *81*(6), 1413–1426. https://doi.org/https://doi.org/10.2136/sssaj2016.11.0376

Tan, B., Masek, J. G., Wolfe, R., Gao, F., Huang, C., Vermote, E. F., Sexton, J. O., & Ederer, G. (2013). Improved forest change detection with terrain illumination corrected Landsat images. *Remote Sensing of Environment*, *136*, 469–483. https://doi.org/https://doi.org/10.1016/j.rse.2013.05.013

Tifafi, M., Guenet, B., & Hatté, C. (2018). Large Differences in Global and Regional Total Soil Carbon Stock Estimates Based on SoilGrids, HWSD, and NCSCD: Intercomparison and Evaluation Based on Field Data From USA, England, Wales, and France. *Global Biogeochemical Cycles*, *32*(1), 42–56. https://doi.org/https://doi.org/10.1002/2017GB005678

Trumbore, S. (2009). Radiocarbon and Soil Carbon Dynamics. *Annual Review of Earth and Planetary Sciences*, *37*(1), 47–66. https://doi.org/10.1146/annurev.earth.36.031207.124300

Vasques, G. M., Coelho, M. R., Dart, R. O., Oliveira, R. P., & Teixeira, W. G. (2016). Mapping soil carbon, particle-size fractions, and water retention in tropical dry

forest in Brazil. *Pesquisa AgropecuÃ!`ria Brasileira*, *51*, 1371–1385. http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-204X2016000901371&nrm=iso

Veronesi, F., & Schillaci, C. (2019). Comparison between geostatistical and machine learning models as predictors of topsoil organic carbon with a focus on local uncertainty estimation. *Ecological Indicators*, *101*, 1032–1044. https://doi.org/https://doi.org/10.1016/j.ecolind.2019.02.026

Viscarra Rossel, R. A., Brus, D. J., Lobsey, C., Shi, Z., & McLachlan, G. (2016). Baseline estimates of soil organic carbon by proximal sensing: Comparing design-based, model-assisted and model-based inference. *Geoderma*, *265*, 152–163. https://doi.org/https://doi.org/10.1016/j.geoderma.2015.11.016

Walkley, A., & Black, I. A. (1934). An examination of the degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. *Soil Science*, *37*(1). https://journals.lww.com/soilsci/Fulltext/1934/01000/AN_EXAMINATION_OF_THE_DEGTJAREFF_METHOD_FOR.3.aspx

Wan, Z. (2014). New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. *Remote Sensing of Environment*, *140*, 36–45. https://doi.org/https://doi.org/10.1016/j.rse.2013.08.027

Wang, B., Waters, C., Orgill, S., Cowie, A., Clark, A., Li Liu, D., Simpson, M., McGowen, I., & Sides, T. (2018). Estimating soil organic carbon stocks using different modelling techniques in the semi-arid rangelands of eastern Australia. *Ecological Indicators*, *88*, 425–438. https://doi.org/https://doi.org/10.1016/j.ecolind.2018.01.049

Wang, M., Su, Y., & Yang, X. (2014). Spatial Distribution of Soil Organic Carbon and Its Influencing Factors in Desert Grasslands of the Hexi Corridor, Northwest China. *PLOS ONE*, *9*(4), e94652. https://doi.org/10.1371/journal.pone.0094652

WMO. (2020). *About Essential Climate Variables*. https://gcos.wmo.int/en/essential-climate-variables/about

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/https://doi.org/10.1016/j.rse.2015.11.032

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/https://doi.org/10.1016/j.rse.2019.111383

Xu, H. (2008). A new index for delineating built-up land features in satellite imagery. *International Journal of Remote Sensing*, *29*(14), 4269–4276. https://doi.org/10.1080/01431160802039957

Yang, Y., Fang, J., Tang, Y., Ji, C., Zheng, C., He, J., & Zhu, B. (2008). Storage, patterns and controls of soil organic carbon in the Tibetan grasslands. *Global Change Biology*, *14*(7), 1592–1599. https://doi.org/https://doi.org/10.1111/j.1365-2486.2008.01591.x

Zhang, G., Liu, F., & Song, X. (2017). Recent progress and future prospect of digital soil mapping: A review. *Journal of Integrative Agriculture*, *16*(12), 2871–2885. https://doi.org/https://doi.org/10.1016/S2095-3119(17)61762-3

Zhao, Y.-C., & Shi, X.-Z. (2010). *Spatial Prediction and Uncertainty Assessment of Soil Organic Carbon in Hebei Province, China BT - Digital Soil Mapping: Bridging Research, Environmental Application, and Operation* (J. L. Boettinger, D. W. Howell, A. C. Moore, A. E. Hartemink, & S. Kienast-Brown, Eds.; pp. 227–239). Springer Netherlands. https://doi.org/10.1007/978-90-481-8863-5_19

Zhu, Z., & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, *118*, 83–94. https://doi.org/https://doi.org/10.1016/j.rse.2011.10.028

IV. CONCLUSIONES GENERALES

Mediante la presente investigación se destacan herramientas y enfoques de la teledetección aplicados a la medición y monitoreo de las emisiones de CO_2 causadas por la degradación de los bosques tropicales y se proponen técnicas para la medición del carbono orgánico del suelo; basados en los aspectos anteriores, en esta tesis se concluye:

En la actualidad se identifican muchas formas para definir conceptualmente la degradación de los bosques; esta diversidad de conceptos provoca interpretaciones y desafíos que en la actualidad se debaten en la arena científica. El enfoque para medir y monitorear la degradación del bosque utilizado en este estudio, parte de la aplicación de indicadores cuantitativos como ser la estructura del bosque, la cual se mide con la aplicación de tecnologías de percepción remota espacialmente explícita como ser las imágenes satelitales; el uso de estas imágenes permite medir la magnitud del cambio de un pixel para un periodo de tiempo (1990 – 2018) y lo relaciona con la disminución o aumento del carbono (Mg C ha^{-1}) almacenado en la biomasa viva de los bosques.

La aplicación de las tecnologías de teledetección permite realizar análisis de grandes áreas como son los bosques tropicales que frecuentemente tienen condiciones de acceso muy difíciles; sin embargo, se identificó que en la actualidad la degradación de los bosques tropicales tiene muy poco estudio, en especial los bosques tropicales del continente americano que tienen una alta

amenaza de origen antrópico debido a la tala ilegal, los incendios forestales y las actividades agrícolas; basado en lo anterior, en el presente estudio se propone e implementa una serie de herramientas y tecnologías basadas en datos de libre acceso, con cómputo en la nube y de escala planetaria que permiten medir y monitorear indicadores claves como ser las emisiones de CO_{2eq} causadas por la degradación forestal o las reservas de carbono orgánico del suelo.

La integración de datos satelitales con datos de campo y el uso de técnicas de máquinas de aprendizaje e inteligencia artificial aplicados en este estudio permitieron cuantificar satisfactoriamente el carbono orgánico del suelo en los bosques tropicales de la República Dominicana. Con la combinación de covariables predictivas de carbono relacionadas con la vegetación, el clima y la topografía se calibró un modelo con el cual se estimó el carbono orgánico del suelo para una escala espacial de 30 metros y una escala temporal del año 2018 para la República Dominicana.

El Sexto Informe del Grupo de Trabajo I (AR6) del Panel Intergubernamental sobre Cambio Climático (IPCC) presenta evidencia clara de que el dióxido de carbono (CO_2) es el principal impulsor del cambio climático, incluso cuando otros gases de efecto invernadero (GEI) y contaminantes del aire también afectan el clima; en este sentido, se identifican al suelo y a los bosques como los principales sumideros de carbono de nuestro planeta; basado en lo anterior, a través de la

presente investigación se proponen y aplican innovadores enfoques para la medición y monitoreo del carbono en estos dos grandes reservorios de carbono de manera espacialmente explícita.