



COLEGIO DE POSTGRADUADOS

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN EN CIENCIAS AGRÍCOLAS

CAMPUS MONTECILLO

POSTGRADO EN CIENCIAS FORESTALES

**MODELACIÓN ESPACIO-TEMPORAL DEL
CONTENIDO DE CARBONO EN EL
MANTILLO DE UN BOSQUE BAJO MANEJO
FORESTAL**

ZAIRA ROSARIO PÉREZ VÁZQUEZ

T E S I S

**PRESENTADA COMO REQUISITO PARCIAL
PARA OBTENER EL GRADO DE:**

MAESTRA EN CIENCIAS

MONTECILLO, TEXCOCO, ESTADO DE MÉXICO

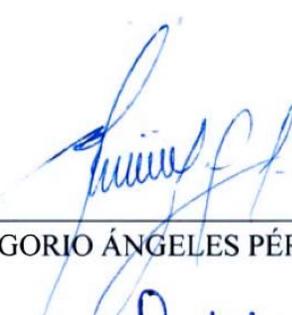
2020

La presente tesis titulada “**Modelación espacio-temporal del contenido de carbono en el mantillo de un bosque bajo manejo forestal**”, realizada por la alumna **Zaira Rosario Pérez Vázquez**, bajo la dirección del Consejo Particular indicado, ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

**MAESTRA EN CIENCIAS
CIENCIAS FORESTALES**

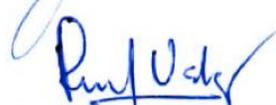
CONSEJO PARTICULAR

CONSEJERO



DR. GREGORIO ÁNGELES PÉREZ

ASESOR



DR. JOSÉ RENE VALDEZ LAZALDE

ASESOR



DR. BRUNO MANUEL CHÁVEZ VERGARA

Montecillo, Texcoco, Estado de México, febrero de 2020

Este estudio fue financiado por el Programa de Paisajes Sostenibles de la Agencia para el Desarrollo Internacional de los Estados Unidos de América, a través de la Oficina de Programas Internacionales del Servicio Forestal de la USDA, proyecto “Reducing Greenhouse Gas Emissions and Improving Forest Management in Mexico (Acuerdo No. 12-IJ-11242306-033).

MODELACIÓN ESPACIO-TEMPORAL DEL CONTENIDO DE CARBONO EN EL MANTILLO DE UN BOSQUE BAJO MANEJO FORESTAL

Zaira Rosario Pérez Vázquez, M. C.
Colegio de Postgraduados, 2020

RESUMEN

El piso forestal es un importante almacén de carbono (C) en los ecosistemas forestales y los procesos que ocurren en éste influyen directamente en la productividad forestal y la fertilidad de los suelos. La cantidad de carbono almacenado en el piso forestal es altamente variable en espacio y tiempo, como resultado de la interacción entre diferentes factores ambientales y perturbaciones naturales y/o humanas. El desarrollo de modelos para estimar el contenido de C en el piso forestal a nivel de paisaje es esencial para un manejo forestal sostenible y para crear estrategias de mitigación contra el cambio climático. No obstante, esta información es poco conocida en México y las evaluaciones precisas de C en este almacén aun representan un importante desafío metodológico a diversas escalas. El objetivo de este estudio fue modelar la distribución espacial y temporal del contenido de carbono (CC) en el piso forestal de un bosque templado bajo manejo para la producción maderable en el Sitio de Monitoreo Intensivo de Carbono (SMIC-Atopixco), en Zacualtipán, Hidalgo, México. Primero, se compararon tres técnicas de modelación espacial (interpolación, regresión y Machine Learning) para mapear CC y se seleccionó el método con mejor desempeño predictivo. Segundo, se evaluó la relación entre la distribución espacial del CC en el piso forestal y cada variable predictora. El piso forestal se clasificó en dos capas (hojarasca y capa de fermentación) y las estimaciones fueron realizadas para 2013 y 2018. Las variables empleadas como predictoras representan la estructura espacial, topográfica y del dosel presente en el área de estudio. El algoritmo Random Forest fue el método con mejor desempeño ($R^2 > 0.86$) para modelar CC en el piso forestal. Para ambos años, los resultados indicaron que el CC incrementó con el grado de descomposición del material orgánico, con valores totales promedio de CC de $12.19 \pm 5.48 \text{ Mg ha}^{-1}$ para 2013 y de $5.76 \pm 2.34 \text{ Mg ha}^{-1}$ para 2018. La variabilidad espacial del CC en el piso forestal fue determinada por la edad del rodal, el área basal, la altura dominante, la cobertura arbórea y la elevación. Este estudio proporciona un esquema de mapeo rentable en apoyo a las actividades de inventario y monitoreo de C en ecosistemas forestales.

Palabras clave: almacenes de carbono; geoestadística; piso forestal; Random Forest, variabilidad espacio-temporal.

SPATIO-TEMPORAL MODELLING OF FOREST FLOOR CARBON CONTENT IN A MANAGED FOREST

Zaira Rosario Pérez Vázquez, M. C.
Colegio de Postgraduados, 2020

ABSTRACT

The forest floor is an important carbon (C) stock in forest ecosystems and the processes that occur there, directly influence forest productivity and soil fertility. Forest floor carbon stocks are highly variable in space and time, as a result of the interaction among different environmental factors and natural and/or human disturbances. The development of models to estimate forest floor carbon content at landscape level is essential for sustainable forest management and for creating climate change mitigation strategies. However, this information is poorly understood in Mexico and precise assessments of C in this stock still represent an important methodological challenge at various scales. The aim of this study was to model the spatial and temporal distribution of forest floor carbon content (CC) in a temperate forest under management for timber production at the Intensive Carbon Monitoring Site (SMIC-Atopixco), in Zacualtipán, Hidalgo, México. First, three spatial modeling techniques (interpolation, regression and Machine Learning) were compared to map CC and the method with the best predictive performance was selected. Second, the relationship between each predictor variable and the spatial distribution of forest floor CC was evaluated. The forest floor was classified into two layers (Ho and fermentation layer) and CC estimates were made in 2013 and 2018. The variables used as predictors were the spatial and topographic derived variables, and canopy structure of the study site. The Random Forest algorithm was the best performing method to model forest floor CC. For both years, the results indicated that the CC increased with the decomposition degree of organic material, with average total CC values of $12.19 \pm 5.48 \text{ Mg ha}^{-1}$ for 2013, and $5.76 \pm 2.34 \text{ Mg ha}^{-1}$ for 2018. The spatial variability of forest floor CC was determined by stand age, basal area, dominant height, tree canopy cover and elevation. This study provides a cost-effective mapping scheme in support of C inventory and monitoring activities in forest ecosystems.

Keywords: carbon stocks; geostatistics; forest floor; Random Forest; spatiotemporal variability.

AGRADECIMIENTOS

Al **Consejo Nacional de Ciencia y Tecnología (CONACyT)**, por el financiamiento otorgado para la realización de mis estudios.

Al **Colegio de Posgraduados**, por la enseñanza académica de excelencia y el apoyo para la obtención del grado.

Al **Dr. Gregorio Ángeles Pérez**, por su gran contribución en mi formación profesional y por la confianza brindada en la dirección de esta investigación. Su excelencia como profesor, sus acertadas sugerencias y todas sus enseñanzas son motivo de admiración y representan un ejemplo a seguir, personal y profesionalmente.

Al **Dr. Bruno Chávez Vergara** y al **Dr. Rene Valdez Lazalde**, por formar parte de mi Consejo Particular, sus valiosos aportes, sugerencias y recomendaciones en el desarrollo de esta investigación permitieron que mis estudios de Maestría sean concluidos con éxito.

A la **Dra. Martha Elva Ramírez Guzmán**, por su disponibilidad de tiempo y sus atinadas sugerencias y aportaciones en el análisis estadístico de este estudio.

A los **Ejidos: Atopixco, La Mojonera, El Reparo, Santo Domingo, Tzincoatlán, Zahuastipán y pequeñas propiedades**. Por permitir hacer uso de sus bosques para el desarrollo de esta investigación y por su valioso apoyo en el levantamiento de datos en campo.

A **Benjamín Méndez López**, por ser un gran amigo y por su gran apoyo en el muestreo en campo.

Al **personal académico y administrativo del Posgrado en Ciencias Forestales**, cada uno de ustedes me brindó su amistad, apoyo y enseñanzas desde que ingrese a esta institución. Ha sido una dicha para mi poder conocerlos y formar parte de este posgrado.

A los **compañeros y amigos** con los que coincidí en esta etapa y compartí momentos agradables e inolvidables. En especial, a **Iván Velázquez**, por esos momentos en los que siempre me apoyaste y motivaste a seguir adelante a pesar de todo, gracias por recordarme siempre mi fuerza interior.

El agradecimiento es la memoria del corazón - Lao Tse

DEDICATORIA

Para mi amada familia.

A mi *mamá*, mi *papá* y mi *hermana*. Porque creyeron en mí y siempre me han apoyado para cumplir mis propósitos. *Mamá*, tu siempre me has enseñado que la constancia y la perseverancia mueven montañas, eres el pilar de mi vida. *Papá*, gracias por luchar siempre por nuestro bienestar, tú me has enseñado a ver la vida siempre de forma positiva y eres un ejemplo a seguir. Tu fortaleza es una inspiración para mí. *Vania*, porque siempre estás conmigo y además de ser mi hermana, eres mi confidente y amiga. Al *Dr. Francisco Ruíz García* (†), porque él desde mi infancia, me inspiró a hacer “ciencia”, su gran apoyo y sabiduría médica en esos momentos siempre permanecerá en nuestros corazones. A *Lizzie*, mi razón de ser y lo más valioso que tengo. A *negri*, porque hay ladridos que nos cambian el alma.

Un camino de millas comienza con un solo paso.

CONTENIDO

RESUMEN	iv
ABSTRACT	v
LISTA DE CUADROS	xi
LISTA DE FIGURAS	xii
INTRODUCCIÓN GENERAL	1
OBJETIVOS	3
HIPÓTESIS	4
REVISIÓN DE LITERATURA	5
Antecedentes	5
Capas del piso forestal	5
Funciones del piso forestal.....	6
Producción y descomposición de hojarasca.....	7
Métodos de modelación espacial	8
CAPÍTULO I. ENFOQUES ESPACIALES PARA LA MODELACIÓN DE CARBONO EN EL PISO FORESTAL DE BOSQUES BAJO MANEJO	10
1.1 RESUMEN	10
1.2 ABSTRACT.....	11
1.3 INTRODUCCIÓN	12
1.4 MATERIALES Y MÉTODOS	13
1.4.1 Área de estudio	13
1.4.2 Diseño de muestreo.....	15
1.4.3 Colecta y análisis de muestras	15
1.4.4 Variables predictoras	16
1.4.5 Análisis estadístico.....	17
1.5 RESULTADOS.....	21
1.5.1 Entrenamiento de los modelos	23
1.5.2 Desempeño de los modelos.....	25
1.5.3 Mapas de CC a partir de las técnicas de predicción.....	28

1.6 DISCUSIÓN	31
1.6.1 Selección del mejor método.....	31
1.6.2 Fuentes de incertidumbre	33
1.6.3 Implicaciones para la estimación del CC.....	34
1.7 CONCLUSIONES	35
CAPÍTULO II. PATRONES ESPACIO-TEMPORALES DEL CONTENIDO DE	
CARBONO EN EL PISO FORESTAL DE UN BOSQUE TEMPLADO MANEJADO	36
2.1. RESUMEN	36
2.2. ABSTRACT.....	37
2.3. INTRODUCCIÓN	38
2.4. MATERIALES Y MÉTODOS	40
2.4.1 Área de estudio	40
2.4.2 Diseño de muestreo.....	41
2.4.3 Colecta de muestras	42
2.4.4 Análisis de laboratorio	42
2.4.5 Variables predictoras	43
2.4.6 Análisis estadístico.....	45
2.5. RESULTADOS.....	46
2.5.1 Contenido de carbono en el piso forestal.....	46
2.5.2 Desempeño del modelo e importancia de variables.....	49
2.5.3 Variación espacial del CC en el piso forestal	50
2.5.4 Interacciones entre variables y el CC.....	52
2.6. DISCUSIÓN	53
2.6.1 Dinámica de carbono en el piso forestal	53
2.6.2 Variaciones temporales del CC en el piso forestal	54
2.6.3 Efectos de la estructura del dosel sobre el CC	55
2.6.4 Efectos de la topografía sobre el CC.....	57
2.6.5 Enfoque utilizado para estimar CC en el piso forestal.....	57
2.7. CONCLUSIONES	58
DISCUSIÓN GENERAL	60
CONCLUSIONES GENERALES.....	62

LITERATURA CITADA	63
ANEXOS	76

LISTA DE CUADROS

Cuadro 1. Descripción de las variables ambientales utilizadas para predecir el contenido de carbono en el piso forestal.	16
Cuadro 2. Estadística descriptiva del contenido de carbono (Mg ha^{-1}), estratificado por año y capa del piso forestal.	22
Cuadro 3. Coeficientes de correlación de Spearman (r^2) entre el CC por capa del piso forestal y cada una de las variables predictoras, para 2013 y 2018.	22
Cuadro 4. Parámetros de los modelos de semivariograma con mejor ajuste, estratificado por año y capa del piso forestal.	23
Cuadro 5. Parámetros y significancia de los predictores en los modelos GLM y GAM, para predecir CC en las capas del piso forestal, en 2013 y 2018.	24
Cuadro 6. Parámetros de ajuste de los modelos GLM, GAM y RF para predecir CC en el piso forestal.	25
Cuadro 7. Estadísticas de validación cruzada de las tres técnicas empleados para modelar el CC en el piso forestal.	26
Cuadro 8. Descripción de las variables predictoras utilizadas para modelar el contenido de carbono en el piso forestal.	43
Cuadro 9. Estadística descriptiva del contenido de carbono y las propiedades del piso forestal, estratificado por año y capa de del piso forestal.	47
ANEXOS:	
Cuadro S1. Matriz de correlaciones de Spearman entre variables predictoras, para 2013.	79
Cuadro S2. Matriz de correlaciones de Spearman entre variables predictoras, para 2018.	79
Cuadro S3. Parámetros de calibración y desempeño del modelo Random Forest.	86

LISTA DE FIGURAS

- Figura 1.** Perfil representativo de un suelo forestal, con la presencia de la capa orgánica y sus capas de fermentación. La densidad de puntos refleja la concentración de materia orgánica en el suelo. Obtenido de a) Chapin *et al.* (2002), b) Woodall *et al.* (2012)..... 6
- Figura 2.** Triángulo de descomposición, la conceptualización clásica de los factores dominantes que regulan las tasas de descomposición de la hojarasca en los ecosistemas terrestres. Obtenido de Bradford *et al.* (2016). 8
- Figura 3.** Componentes del Sitio de Monitoreo Intensivo de Carbono (SMIC-Atopixco): a) ubicación de las unidades de muestreo y anualidades; b) ubicación geográfica del área de estudio; c) diseño del conglomerado tipo Inventario Nacional Forestal y de Suelos (INFyS) y esquema de muestreo para la colecta de información. 14
- Figura 4.** Valores observados vs predichos del conjunto de datos de validación cruzada (k=10) del modelo “Kriging ordinario”, para predecir CC en las capas del piso forestal: Ho (hojarasca), F (fermentación) y Tot (Totales), en 2013 y 2018. La línea negra es de regresión y la línea punteada es la relación 1:1 26
- Figura 5.** Valores observados vs predichos del conjunto de datos de validación cruzada (k=10) de los modelos GAM, GLM y RF, para predecir CC en el piso forestal, en 2013 y 2018. La línea negra es de regresión y la línea punteada es la relación 1:1..... 27
- Figura 6.** Predicciones espaciales para cada modelo espacial del CC en el piso forestal. Año 2013. Capas del piso forestal: Ho: hojarasca, F: fermentación, Tot: Suma de Ho y F. Modelos espaciales: Kriging Ordinario (KO), Modelo Lineal Generalizado (GLM), Modelo Aditivo Generalizado (GAM) y RF (Random Forest). 29
- Figura 7.** Predicciones espaciales para cada modelo espacial del CC en el piso forestal. Año 2018. Capas del piso forestal: Ho: hojarasca, F: fermentación, Tot: Suma de Ho y F. Modelos espaciales: Kriging Ordinario (KO), Modelo Lineal Generalizado (GLM), Modelo Aditivo Generalizado (GAM). 30
- Figura 8.** Componentes del Sitio de Monitoreo Intensivo de Carbono (SMIC-Atopixco): a) ubicación de las unidades de muestreo y anualidades; b) ubicación geográfica

del área de estudio; c) diseño del conglomerado tipo Inventario Nacional Forestal y de Suelos (INFyS) y esquema de muestreo para la colecta de información.	40
Figura 9. Contenido de Carbono para cada capa del piso forestal para 2013 y 2018.	48
Figura 10. Relación C: N por capa del piso forestal, para 2013 y 2018.	48
Figura 11. Importancia relativa de variables predictoras (%) de Random Forest para predicciones del contenido de carbono en el piso forestal.	49
Figura 12. Patrones espaciales del contenido de carbono por capa de descomposición en el piso forestal y su coeficiente de variación a 1 m de resolución, en el Sitio de Monitoreo Intensivo de Carbono-Atopixco en el año 2013.	50
Figura 13. Patrones espaciales del contenido de carbono por capa de descomposición en el piso forestal y su coeficiente de variación a 1 m de resolución, en el Sitio de Monitoreo Intensivo de Carbono-Atopixco en el año 2018.	51
Figura 14. Interacción entre el contenido de carbono (CC) por capa del piso forestal (Ho: hojarasca, F: fermentación; Tot: Suma de Ho y F) y: a) área basal ($m^2 ha^{-1}$), b) altura dominante (m), c) edad del rodal (años) y d) elevación (msnm).	53

ANEXOS

Figura S1. Histogramas de frecuencia relativa del CC por capa del piso forestal, para el año 2013 y 2018. La línea roja muestra la función de densidad de probabilidad. Ho: hojarasca, F: fermentación, Tot: suma de Ho y F.	76
Figura S2. Histogramas de frecuencia relativa (eje Y) de cada variable predictora (eje X), para el conjunto de datos de 2013. La línea roja muestra la distribución teórica normal, la línea azul muestra la función de densidad de probabilidad. Nomenclatura: X, Y: Coordenadas UTM X, UTM Y (m); ELE: elevación (m); SLO: pendiente (grados); ASP: exposición (grados); SA: edad del rodal (años); BA: área basal ($m^2 ha^{-1}$); DH: altura dominante (m); S: riqueza de especies; H: índice de Shannon-Wiener; TCO: cobertura arbórea (%).	77
Figura S3. Histogramas de frecuencia relativa (eje Y) de cada variable predictora (eje X), para el conjunto de datos de 2018. La línea roja muestra la distribución teórica normal, la línea azul muestra la función de densidad de probabilidad. Nomenclatura: X, Y: Coordenadas UTM X, UTM Y (m); ELE: elevación (m); SLO: pendiente (grados); ASP: exposición (grados); SA: edad del rodal (años);	

BA: área basal ($\text{m}^2 \text{ ha}^{-1}$); DH: altura dominante (m); S: riqueza de especies; H: índice de Shannon-Wiener; TCO: cobertura arbórea (%).	78
Figura S4. Semivariograma empírico y ajuste de modelo teórico por capa de descomposición en el piso forestal, para 2013 y 2018. Ho: hojarasca, F: fermentación, Tot: suma de Ho y F.	80
Figura S5. Gráficos de verificación de supuestos del modelo GLM, para el año 2013.	81
Figura S6. Gráficos de verificación de supuestos del modelo GLM, para el año 2018.	82
Figura S7. Gráficos de verificación de supuestos del modelo GAM, para el año 2013.	83
Figura S8. Gráficos de verificación de supuestos del modelo GAM, para el año 2018.	84
Figura S9. Diagramas de caja de residuales de cada modelo de predicción espacial del CC en el piso forestal. GAM: modelo aditivo generalizado, GLM: modelo lineal generalizado y RF: Random Forest.	85
Figura S10. Valores observados vs predichos del conjunto de datos de validación cruzada ($k=10$) del modelo "Random Forest" para: a) 2013 y b) 2018. La línea negra es de regresión y la línea punteada es la relación 1:1.	86
Figura S11. Variación temporal del contenido de carbono por capa de descomposición del piso forestal, bajo enfoque de cronosecuencia. El rodal de 85 años corresponde a un bosque de referencia que no ha sido cosechado.	87

INTRODUCCIÓN GENERAL

En los suelos forestales se depositan residuos provenientes de distintos componentes de la vegetación (hojas, material leñoso fino, inflorescencias y frutos), que en conjunto se denominan hojarasca (Anaya *et al.*, 2007; Colín-Vargas *et al.*, 2018). La deposición y acumulación de estos residuos sobre el suelo mineral dan origen al piso forestal, el cual se compone de hojarasca recientemente caída y materia orgánica parcialmente descompuesta, la cual tiende a estratificarse (Yanai *et al.*, 2003).

El Panel Intergubernamental sobre Cambio Climático (IPCC) reconoce al piso forestal como el tercer almacén de carbono (C) más importante de los ecosistemas terrestres (IPCC, 2006). Globalmente, el C en este almacén representa aproximadamente entre el 5 y 7% (43 ± 3 Pg de C) de todas las reservas de C de los ecosistemas forestales (Pan *et al.*, 2011; Domke *et al.*, 2016). El piso forestal es un componente clave en los ciclos biogeoquímicos. Su capacidad para almacenar C desempeña un papel fundamental en el funcionamiento de los ecosistemas terrestres, por influir en la productividad forestal y en la fertilidad del suelo (Schlesinger y Benhardt, 2013). La mejor comprensión posible de la dinámica de C en este almacén permitirá proponer acciones para un manejo forestal sostenible y desarrollar estrategias de mitigación contra el cambio climático.

El contenido de carbono (CC) en el piso forestal depende del balance entre la producción de hojarasca y su posterior descomposición, los cuales a su vez, están determinados por una serie de factores ambientales y perturbaciones humanas que al interactuar simultáneamente, ocasionan una compleja heterogeneidad espacial y temporal (Colín Vargas *et al.*, 2018). En México, se han realizado estudios que cuantifican, bajo diversos enfoques, los almacenes de carbono en los distintos componentes del ecosistema, incluido el piso forestal (López-Hernández *et al.*, 2013; Galicia *et al.*, 2015; Pando-Moreno *et al.*, 2018); sin embargo, en pocas investigaciones se ha explorado el CC en este almacén, desde una perspectiva espacial (Martínez-Yrizar *et al.*, 1999; Anaya Merchant, 2007).

En la última década, a través del manejo de Sistemas de Información Geográfica (SIG) se han desarrollado diversos enfoques para obtener distribuciones espaciales y patrones del contenido de carbono en el suelo, y en particular, el Mapeo Digital del Suelo (DSM, por sus siglas en inglés) ha sido uno de los enfoques más aplicados para modelar carbono en las distintas capas del suelo

(Hartemink, *et al.*, 2008; Van Zijl, 2019). En el DSM se ponen de manifiesto las relaciones entre la variable de interés y las variables predictoras, mediante el uso de diferentes modelos estadísticos (Dharumarajan *et al.*, 2017). Las técnicas de DSM más reportadas en la literatura científica incluyen la interpolación espacial, modelos de regresión y, más recientemente, métodos de aprendizaje automático (Machine Learning, en inglés). Estas técnicas han demostrado ser un medio potencial para superar las limitaciones de los enfoques tradicionales, al considerar la posición geográfica de los puntos medidos y hacer predicciones en áreas no muestreadas (Beguin *et al.*, 2017).

Bajo este contexto, el objetivo general del presente estudio fue modelar la distribución espacial y temporal del contenido de carbono en el piso forestal de un bosque templado gestionado para la producción maderable en el Sitio de Monitoreo Intensivo de Carbono (SMIC-Atopixco), en Zacualtipán, Hidalgo, México (Ángeles-Pérez *et al.*, 2015). Primero, se evaluó el desempeño de tres técnicas DSM para modelar la distribución espacial del CC, a partir de un conjunto de datos levantados en 2013 y 2018. Segundo, se evaluó la relación entre cada variable predictora y la distribución espacial del CC en el piso forestal. Este estudio es de los primeros análisis realizados en México que aplica el enfoque DSM, en combinación con datos derivados de fuentes múltiples, para cuantificar la variabilidad espacial y temporal del CC en el piso forestal de un bosque bajo manejo maderable. Esta investigación proporciona una alternativa metodológica viable para modelaciones futuras de carbono en ecosistemas forestales.

OBJETIVOS

1. Objetivo general

Modelar la distribución espacial y temporal del contenido de carbono en el piso forestal de un bosque templado bajo manejo para la producción maderable, en el Sitio de Monitoreo Intensivo de Carbono (SMIC-Atopixco), en Zacualtipán, Hidalgo, México.

2. Objetivos particulares

Capítulo I. Enfoques espaciales para la modelación de carbono en el piso forestal de bosques bajo manejo

- Comparar el desempeño de tres técnicas, en el marco del DSM (interpolación espacial, modelos de regresión y Machine Learning) para modelar la distribución espacial del contenido de carbono en el piso forestal.

Capítulo II. Patrones espacio-temporales del contenido de carbono en el piso forestal de un bosque templado manejado

- Determinar la variabilidad espacial del contenido de carbono en el piso forestal mediante el algoritmo Random Forest.
- Analizar la relación de las variables topográficas y de estructura del dosel sobre la distribución espacial del contenido de carbono en el piso forestal.
- Comparar la variación en los patrones espaciales del contenido de carbono por capa de descomposición del piso forestal, para 2013 y 2018.

HIPÓTESIS

Capítulo I. Enfoques espaciales para la modelación de carbono en el piso forestal de bosques bajo manejo

- El uso de datos armonizados, provenientes de fuentes múltiples, mejora la capacidad predictiva de los modelos espaciales.
- El desempeño de los modelos está en función de las variables predictoras utilizadas para representar las relaciones espaciales del contenido de carbono en el piso forestal.

Capítulo II. Patrones espacio-temporales del contenido de carbono en el piso forestal de un bosque templado manejado

- Entre las capas que conforman el piso forestal, el contenido de carbono es diferente ya que están afectados por su grado de descomposición.
- El contenido de carbono en el piso forestal es diferente entre 2013 y 2018, debido a cambios en la estructura del dosel, topográfica y espacial en el sistema forestal estudiado.
- La distribución espacial del contenido de carbono en el piso forestal está determinada por la estructura del dosel, topográfica y espacial en el sistema forestal estudiado.

REVISIÓN DE LITERATURA

Antecedentes

El Sitio de Monitoreo Intensivo del Carbono (SMIC-Atopixco), localizado en Zacualtipán, Hidalgo, México, forma parte la red de sitios denominada “Red MEX-SMIC”, la cual ha sido establecida en algunos ecosistemas forestales de México, considerados relevantes por su contribución a la emisión y remoción del CO₂ atmosférico. La finalidad de esta red es contribuir, al desarrollo e implementación del sistema de Monitoreo, Reporte y Verificación (MRV), indispensable en una estrategia REDD+ en México. Los SMIC proporcionan información detallada de reservorios y flujos de carbono, desarrollan metodologías para la colecta de información en campo y generan modelos validados de la dinámica de carbono forestal (Birdsey *et al.*, 2013; Ángeles-Pérez *et al.*, 2015).

Dentro del SMIC-Atopixco se han realizado diversas investigaciones encaminadas a mejorar el sistema de manejo forestal practicado en la región (Ángeles-Pérez, 1995; Santiago-García *et al.*, 2012; Ramírez Martínez, 2018), evaluaciones de la composición y diversidad de especies (Villegas-Macedo, 2019). modelos de dinámica de acumulación de biomasa y carbono en rodales (Aguirre-Salado *et al.*, 2009; Figueroa-Navarro *et al.*, 2010; Ortiz-Reyes *et al.*, 2015, Soriano Luna *et al.*, 2018, entre otros). En particular, para el piso forestal se han realizado algunas evaluaciones sobre almacenamiento de carbono (Cadena Morales, 2006; Zaragoza-Castañeda, 2016), cuantificación de la producción de hojarasca y descomposición (Martínez Castillo, 2014; Ángeles-Pérez *et al.*, 2015) y determinación de emisiones de CO₂ (Barrales Brito, 2013).

Capas del piso forestal

El piso forestal puede dividirse en diferentes capas, en función del grado de descomposición del material orgánico presente (Flores-Garnica *et al.*, 2018; Liu *et al.*, 2018). En los bosques templados, frecuentemente pueden distinguirse dos o tres capas, dependiendo del tipo de vegetación y de las condiciones ambientales predominantes. Las clasificaciones más comunes son las propuestas por Fisher y Binkley (2000) y por Chapin *et al.* (2002), basadas en el grado de descomposición del sustrato (**Figura 1**):

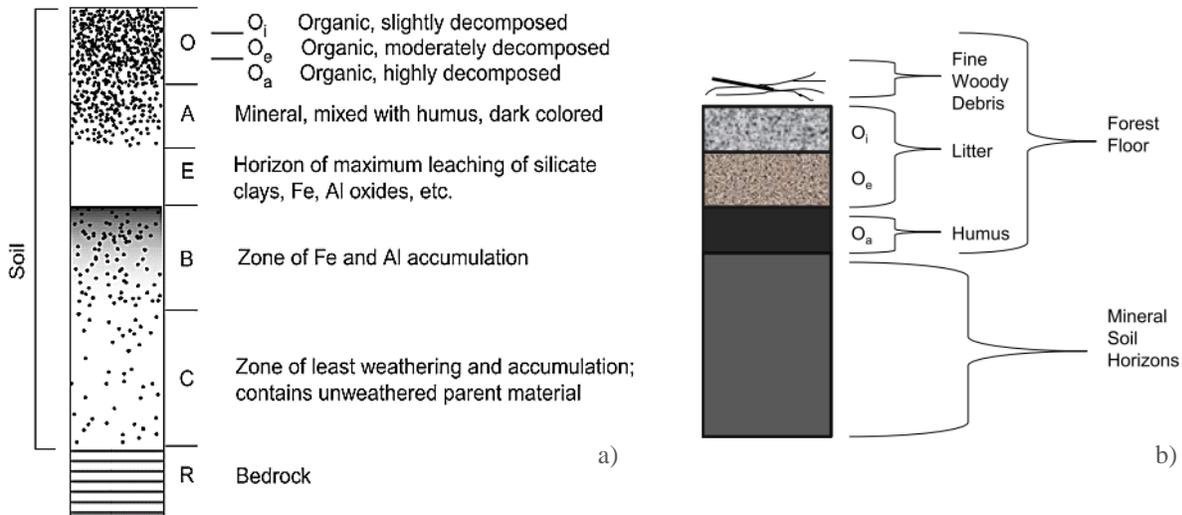


Figura 1. Perfil representativo de un suelo forestal, con la presencia de la capa orgánica y sus capas de fermentación. La densidad de puntos refleja la concentración de materia orgánica en el suelo. Obtenido de a) Chapin *et al.* (2002), b) Woodall *et al.* (2012).

El contenido de carbono en el piso forestal es altamente variable y sensible a pequeñas modificaciones en su entorno, y se encuentra ligado estrechamente a factores como el clima, la vegetación, la topografía y el material parental (Ponge *et al.* 2013). Las actividades silvícolas y los incendios forestales a menudo son las principales perturbaciones que alteran el almacenamiento de carbono en el piso forestal (León *et al.*, 2011; Cameron *et al.*, 2015). Por lo cual, un leve impacto en este almacén podría llegar a afectar las propiedades del suelo y a la vegetación de los bosques (Bauhus *et al.*, 2004; Cameron *et al.*, 2015).

Funciones del piso forestal

Además de ser un almacén de carbono importante, el piso forestal regula la mayoría de los procesos funcionales en los ecosistemas forestales (Moreno-Valdez *et al.*, 2018). Esta capa orgánica protege al suelo contra altas temperaturas (Dang *et al.*, 2018; Kreye *et al.*, 2018), amortigua los procesos de erosión (López-Escobar *et al.*, 2017), favorece la infiltración (Barksdale *et al.*, 2014) e influye directamente en la fertilidad del suelo, por modificar sus propiedades fisicoquímicas (Liu *et al.*, 2018).

Dado que el piso forestal modifica las condiciones microambientales del suelo, también ejerce influencia en las interacciones bióticas (depredación de semillas, herbivoría, ataque de patógenos, interfiere en la competencia planta-planta, etc.) y, por lo tanto, puede determinar el reclutamiento de plantas a través de efectos directos e indirectos múltiples (Dupuy y Chazdon, 2008; Gessner *et al.*, 2010; Bansal *et al.*, 2013). Algunos autores han reconocido que el piso forestal puede afectar la estructura y la organización de las comunidades vegetales, por su efecto en la germinación de las semillas, la sobrevivencia de plántulas y el desarrollo de retoños vegetativos (Bartuszevige *et al.*, 2007; Dupuy y Chazdon, 2008).

Producción y descomposición de hojarasca

La acumulación de carbono en el piso forestal depende del balance entre la producción de hojarasca y su descomposición posterior (Berg, 2018). La producción de hojarasca representa la mayor vía de flujo de material orgánico y nutrientes al piso forestal (Berg, 2018; Bueis *et al.*, 2018). A escala local, la cantidad y la concentración de nutrientes en la hojarasca producida está condicionada por características y procesos que interactúan simultáneamente, como la disponibilidad de nutrimentos en el suelo, la translocación de nutrientes durante el proceso de senescencia, la distribución y competencia interna por los recursos, las condiciones climáticas, entre otros (Bubb *et al.*, 1998; Blanco *et al.*, 2006).

La descomposición de la hojarasca involucra un conjunto complejo de procesos que incluyen agentes físicos, químicos y biológicos para reducir la hojarasca a sus componentes químicos elementales (Aerts, 1997; Berg y McLaugherty, 2008). Es un proceso biogeoquímico clave para el funcionamiento de los ecosistemas, al ser el principal mecanismo por el cual se liberan nutrientes en formas disponibles para las plantas, dióxido de carbono a la atmósfera y da lugar a la materia orgánica del suelo (Aber y Melillo, 1980; Zheng *et al.*, 2017).

Las tasas de descomposición de la hojarasca son reguladas en función de su escala espacial, de macro a micro por: el clima, la composición química de la materia orgánica y la comunidad de organismos descomponedores (**Figura 2**). Además de estos factores, se ha demostrado que tanto la producción como la descomposición de la hojarasca están sujetos a la influencia de otras variables

bióticas y abióticas, además de perturbaciones externas al ecosistema (Berg y Staaf, 1980; Berg *et al.*, 1993; Bradford *et al.*, 2016).

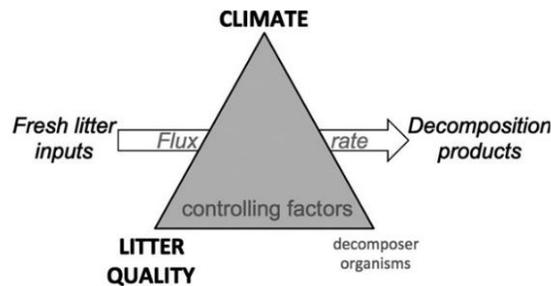


Figura 2. Triángulo de descomposición, la conceptualización clásica de los factores dominantes que regulan las tasas de descomposición de la hojarasca en los ecosistemas terrestres. Obtenido de Bradford *et al.* (2016).

Métodos de modelación espacial

En las últimas dos décadas, se ha desarrollado un considerable número de modelos espaciales para cuantificar y mapear almacenes de carbono en los distintos componentes del ecosistema (Zhang y McGrath, 2004; Camera *et al.*, 2017). La introducción de tecnologías digitales como sensores remotos y sistemas de información geográfica, la velocidad alta de procesamiento computacional, los métodos cuantitativos para describir patrones y procesos del suelo, etc., han proporcionado nuevas oportunidades para predecir y cuantificar estos almacenes (Grunwald, 2009; Beguin *et al.*, 2017).

En particular, el Mapeo Digital del Suelo (DSM) es un enfoque desarrollado recientemente, que involucra el uso de información georreferenciada y modelos estadísticos para mapear predicciones e incertidumbres relacionadas a las propiedades del suelo, además explora las relaciones espaciales y las interacciones entre una variable de interés y sus covariables ambientales (Adhikari *et al.*, 2014; Beguin *et al.*, 2017). Un componente clave de cualquier técnica DSM es el método utilizado para definir la relación espacial entre las observaciones de suelo y las variables ambientales (Brungard *et al.*, 2015). Los métodos más comunes de DSM incluyen los siguientes modelos: modelos geoestadísticos, regresión lineal múltiple (MLR), modelo lineal generalizado (GLM), modelo aditivo generalizado (GAM), Random Forest (RF), entre otros (Zhang *et al.*, 2017).

Los métodos geoestadísticos como Kriging Ordinario, consideran la autocorrelación espacial y la heterogeneidad inherente del atributo de interés; sin embargo, este método pasa por alto la influencia de otras variables ambientales (Liu *et al.*, 2015). Los modelos lineales múltiples explican relaciones estocásticas entre variables de respuesta y algunas variables utilizadas como predictoras, con resultados frecuentemente satisfactorios (Meersmans *et al.*, 2008; Malone *et al.*, 2017). Sin embargo, su aplicación suele complicarse cuando los datos de entrada violan algunos supuestos estadísticos (Antúnez *et al.*, 2017). Para tratar con estas dificultades, se han incluido a las técnicas DSM, otro tipo de modelos más flexibles, como el Modelo Lineal Generalizado (GLM) y el Modelo Aditivo Generalizado (GAM) (Mckenzie y Ryan, 1999; Soriano-Luna *et al.*, 2018). Los modelos GLM y GAM son una extensión matemática de los modelos de regresión lineal, estos modelos consideran las relaciones no lineales, la heterogeneidad de varianzas en los datos de entrada y la regresión no paramétrica (Antúnez *et al.*, 2017; Carvalho Gomes *et al.*, 2019).

Los métodos de Machine Learning (ML) incluyen un amplio conjunto de modelos utilizados para describir patrones y hacer predicciones (Brungard *et al.*, 2015). Aunque ML es más frecuentemente aplicado a grandes bases de datos, es una herramienta atractiva para hacer predicciones espaciales sobre propiedades del suelo, incluido contenidos de carbono, y para determinar sus relaciones con diversas variables ambientales (Beguin *et al.*, 2017; Cao *et al.*, 2019). Uno de los algoritmos de ML más aplicados para el mapeo de carbono es Random Forest (Grimm *et al.*, 2008; Beguin *et al.*, 2017; Cao *et al.*, 2019). La implementación de este algoritmo ha demostrado tener ventajas en comparación con los métodos mencionados anteriormente (Zhang *et al.*, 2017; Chen *et al.*, 2019). Grimm *et al.* (2008) resume sus ventajas principales: a) su capacidad para modelar relaciones no lineales de gran dimensión; b) manejo simultáneo de predictores continuos y categóricos; c) robustez contra sobreajuste; d) mediciones de la tasa de error e importancia de variables; e) requiere solo tres parámetros de entrada definidos por el usuario; f) acepta multicolinealidad entre variables y g) no requiere forzosamente una distribución normal.

CAPÍTULO I. ENFOQUES ESPACIALES PARA LA MODELACIÓN DE CARBONO EN EL PISO FORESTAL DE BOSQUES BAJO MANEJO

1.1 RESUMEN

En ecosistemas forestales, el piso forestal es un importante almacén de carbono y nutrientes que influyen la productividad forestal y la fertilidad de los suelos. Cambios pequeños en este almacén pueden alterar el funcionamiento del ecosistema, por lo cual es fundamental desarrollar metodologías que mejoren las estimaciones de carbono a nivel de paisaje. No obstante, obtener estimaciones con mayor precisión es un desafío a causa de su gran variabilidad y dependencia espacial. En este estudio se compararon tres técnicas de modelación espacial (interpolación, regresión y Machine Learning). para mapear los contenidos de carbono (CC) en el piso forestal de un bosque templado. Los métodos incluyen Kriging Ordinario (KO), Modelo Lineal Generalizado (GLM), Modelo Aditivo Generalizado (GAM) y Random Forest (RF). Las estimaciones del CC fueron realizadas para 2013 y 2018. Las variables empleadas como predictoras representan la estructura espacial, del dosel y topográfica presentes en el área de estudio. Todos los modelos fueron evaluados mediante validación cruzada ($k=10$) y se determinó el error medio absoluto (MAE), error cuadrático medio (RMSE) y el coeficiente de determinación (R^2). Los resultados mostraron que el desempeño de los métodos fue, en orden decreciente, RF, GAM, GLM y KO. El método KO reflejó el grado de dependencia espacial del CC, pero las estimaciones espaciales fueron poco realistas ($R^2 < 0.35$). GAM y GLM mostraron buen desempeño ($R^2 > 0.70$), pero mayor sobreestimación del CC. RF obtuvo el mejor ajuste ($R^2 > 0.86$) para modelar CC en ambos años evaluados. Se concluye que el modelo RF es un método prometedor y con gran potencial para mejorar las estimaciones de carbono en el piso forestal a escala de paisaje. **Palabras clave:** almacenes de carbono; geoestadística; hojarasca; Modelo Aditivo Generalizado; Modelo Lineal Generalizado; Random Forest.

SPATIAL APPROACHES FOR CARBON MODELING IN THE FOREST FLOOR OF FORESTS UNDER MANAGEMENT

1.2 ABSTRACT

In forest ecosystems, the forest floor is an important carbon stock and nutrients that influences forest productivity and soil fertility. Small changes in this stock can alter the ecosystem functioning, thus it is essential to develop methodologies that improve C estimates at the landscape scale. However, obtaining estimates with high precision is still a challenge due to its large variability and spatial dependence. In this study, three spatial modeling techniques (interpolation, regression and Machine Learning) were compared to map forest floor carbon content (CC) in a temperate forest. The methods include Ordinary Kriging (KO), Generalized Linear Model (GLM), Generalized Additive Model (GAM) and Random Forest (RF). CC estimates were made for 2013 and 2018. The variables used as predictors represent the spatial, canopy and topographic structure present in the study site. All models were evaluated using ten-fold cross validation ($k=10$) and the mean absolute error (MAE), root mean square error (RMSE) and the coefficient of determination (R^2) were determined. The results showed that performance of the methods was, in decreasing order, RF, GAM, GLM and KO. The KO method reflected the degree of spatial dependence of CC but the spatial estimates were unrealistic ($R^2 < 0.35$). GAM and GLM showed good performance ($R^2 > 0.70$), but higher overestimation of CC. RF obtained the best fit ($R^2 > 0.86$) to model CC in both evaluated years. It is concluded that RF model is a promising method with high potential to improve estimates of carbon in the forest floor at landscape scale. **Keywords:** carbon stocks; geostatistics; forest floor litter; Generalized Additive Model; Generalized Lineal Model; Random Forest.

1.3 INTRODUCCIÓN

El piso forestal es la capa superficial del suelo formada por residuos orgánicos con diferentes grados de descomposición (Martínez-Yrizar y Sarukhan, 1993). Esta capa orgánica puede almacenar grandes cantidades de carbono (C) y regular la mayoría de los procesos funcionales de los ecosistemas, por influir directamente en la productividad forestal y en la fertilidad de los suelos (Schlesinger y Benhardt, 2013). El Panel Intergubernamental sobre Cambio Climático (2006) reconoce al piso forestal como el tercer almacén de C más importante de los ecosistemas terrestres, representando a nivel global ~5% (43 ± 3 Pg de C) de todas las reservas de C de los ecosistemas forestales (Pan *et al.*, 2011). Dada su importancia, es crítico estimar y monitorear con precisión los almacenes de carbono en el piso forestal (Yanai *et al.*, 2003; Woodall *et al.*, 2012).

En México, las estimaciones de los almacenes de carbono del piso forestal se basan generalmente en mediciones de campo (CONAFOR, 2012; López-Escobar *et al.*, 2017), y a la fecha, son pocos los estudios que exploran su variabilidad espacial (Martínez-Yrizar *et al.*, 1999; Anaya Merchant, 2007), ya que este tipo de evaluaciones son a menudo un desafío metodológico, debido a la elevada heterogeneidad del paisaje y a las condiciones particulares de cada micrositio. Tales variaciones están determinadas por factores biofísicos y son altamente sensibles a cambios en el ecosistema, como resultado de prácticas silvícolas, incendios forestales, cambio climático, etc., que al interactuar entre sí, forman complejos patrones espaciales y temporales (Grunwald, 2009; Zhang *et al.*, 2014).

La introducción de tecnologías digitales, como los sensores remotos y los sistemas de información geográfica, así como una mejora en la velocidad del procesamiento computacional, el manejo de datos espaciales y el desarrollo de nuevos métodos cuantitativos, han proporcionado nuevas oportunidades para describir patrones, propiedades y procesos del suelo. El Mapeo Digital del Suelo (DSM) es uno de los principales enfoques que combinan las tecnologías anteriores para cuantificar y modelar la distribución espacial de los almacenes de carbono en el suelo, desde escalas locales hasta globales (Grunwald, 2009). El enfoque DSM utiliza diversas técnicas y métodos numéricos para calibrar y validar modelos estadísticos de una variable regionalizada, a partir de covariables ambientales asociadas (Adhikari *et al.*, 2014; Beguin *et al.*, 2017). En DSM, se suele incluir métodos geoestadísticos (Zhang y McGrath, 2004; Bhunia *et al.*, 2018; Yao *et al.*,

2019), modelos de regresión (Meersmans *et al.*, 2008; Black *et al.*, 2014), y más recientemente, métodos de Machine Learning (Chen *et al.*, 2019; Zhang *et al.*, 2019).

Estos métodos han demostrado ser una alternativa efectiva y poco costosa para estimar carbono en los distintos componentes del ecosistema, en comparación con enfoques de estimación tradicionales. Sin embargo, debido a la complejidad del atributo de interés, así como a la calidad y disponibilidad de información, algunos de estos métodos no son capaces de mapear con precisión las existencias de C y en ocasiones, seleccionar un método óptimo de predicción no siempre es prometedor. Por lo cual, una adecuada selección del método con menor incertidumbre proporcionará la base para evaluar el potencial de almacenamiento de C en el piso forestal, además contribuirá al desarrollo de estrategias para un manejo forestal sostenible.

El objetivo de este estudio fue comparar el desempeño de tres técnicas, en el marco del DSM (interpolación espacial, modelos de regresión y Machine Learning) para modelar la distribución espacial del CC en el piso forestal. Estas técnicas incluyen cuatro métodos que representan las rutinas estadísticas más utilizadas en DSM: Kriging Ordinario (KO), Modelo Lineal Generalizado (GLM), Modelo Aditivo Generalizado (GAM) y Random Forest (RF). Los modelos fueron creados a partir de conjunto de datos obtenidos en 2013 y se probó su aplicabilidad con datos obtenidos en 2018. Por lo que este estudio ofrece una oportunidad para desarrollar nuevas técnicas de modelación estadística que permitan mejorar el mapeo predictivo de los almacenes de carbono en el piso forestal.

1.4 MATERIALES Y MÉTODOS

1.4.1 Área de estudio

Este estudio se desarrolló en el Sitio de Monitoreo Intensivo del Carbono (SMIC-Atopixco), localizado en el Municipio de Zacualtipán, Hidalgo, México. El SMIC-Atopixco forma parte de la Red Mexicana de Sitios de Monitoreo Intensivo de Carbono (MEX-SMIC) (Ángeles-Pérez *et al.*, 2015) y se encuentra delimitado dentro del paisaje forestal por un polígono de 9 km² (900 ha), dividido en nueve cuadros de 1 km² (Figura 3).

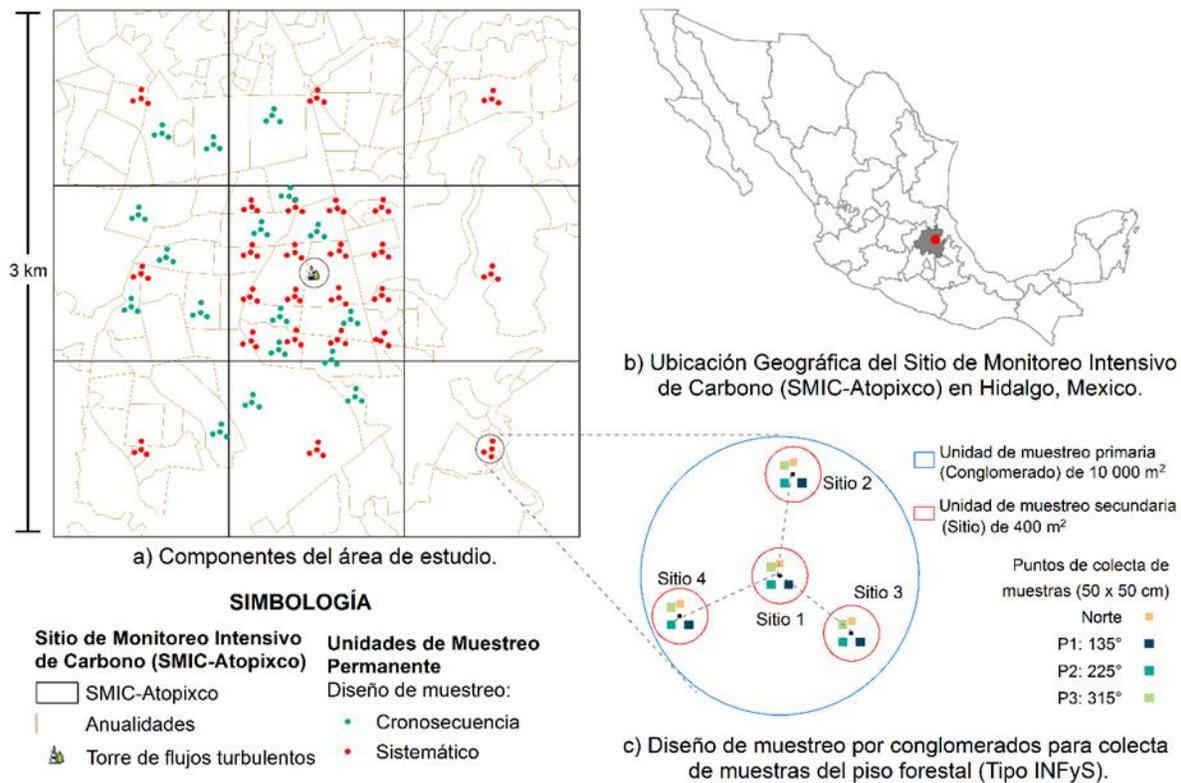


Figura 3. Componentes del Sitio de Monitoreo Intensivo de Carbono (SMIC-Atopixco): a) ubicación de las unidades de muestreo y anualidades; b) ubicación geográfica del área de estudio; c) diseño del conglomerado tipo Inventario Nacional Forestal y de Suelos (INFyS) y esquema de muestreo para la colecta de información.

Se ubica entre las coordenadas 20°40'17'' y 20°34'51'' latitud norte y entre las coordenadas 98°40'07'' y 98°34'22'' longitud oeste. La topografía es variable, con elevaciones entre 1836 y 2097 msnm y pendientes entre 0 y 40°. Los suelos son de origen volcánico, profundos y ricos en materia orgánica; en las partes bajas se encuentran suelos Feozem háplico y en las partes altas con mayor pendiente, se encuentra el Regosol calcárico (Santiago-García *et al.*, 2013). El clima predominante es C(fm)w'b(e)g, templado húmedo con una marcada estación de lluvias entre junio y octubre (García, 1981). La temperatura media anual es de 13.5° y la precipitación total anual es de 2,050 mm y presenta aproximadamente 174 días con neblinas (Godínez-Ibarra *et al.*, 2007). Dentro del SMIC-Atopixco se ha llevado a cabo el aprovechamiento de sus recursos forestales maderables desde principios de la década de los 80 en el siglo pasado. El manejo del bosque se ha encaminado a desarrollar bosques coetáneos dominados por *Pinus patula*, que a la fecha presentan una cronosecuencia cercana a 40 años (Ángeles-Pérez *et al.*, 2015).

1.4.2 Diseño de muestreo

El SMIC-Atopixco sigue un esquema de muestreo propuesto por Hollinger (2008). En el cuadro central se ubica una torre de flujos de torbellinos y tiene establecidas 16 unidades de muestreo permanente (de aquí en adelante, referidas como conglomerados), similares a los utilizados en el Inventario Nacional Forestal y de Suelos (INFyS) (CONAFOR, 2012). En el centro de cada uno de los ocho km² periféricos se localiza un conglomerado, lo que da un total de 24 conglomerados dispuestos en forma sistemática. Adicionalmente, se establecieron 16 conglomerados distribuidos bajo un enfoque de cronosecuencia (establecidos en rodales con diferente edad que no fueron representados dentro del esquema sistemático). De esta forma, el diseño de muestreo comprende un total de 40 conglomerados (**Figura 3**).

1.4.3 Colecta y análisis de muestras

Para los propósitos de este estudio, el piso forestal fue caracterizado en dos capas que representan las dos etapas de descomposición del material orgánico: 1) la capa Ho, identificada como la materia orgánica recién caída, en donde se reconocen todos sus componentes; y 2) la capa de fermentación (F), en el que la materia orgánica está descompuesta a tal grado que no se reconocen estructuras completas (CONAFOR, 2012). Con fines de comparación, se incluyó la capa total del piso forestal (Tot), obtenida por la suma de la capa Ho y F. La colecta de muestras se realizó en los 160 sitios de muestreo y fue llevada a cabo en dos periodos. El primer periodo fue dirigido por Ángeles-Pérez *et al.* (2015), de noviembre 2013 a enero 2014. El segundo periodo se realizó de noviembre 2018 a enero 2019. Por cada sitio, se determinaron tres parcelas de muestreo, dispuestos a cinco metros del centro del sitio en azimut 135° (P₁), 225° (P₂) y 315° (P₃), obteniendo así, un total de 480 muestras por cada capa del piso forestal. Las muestras se colectaron con base en el protocolo seguido por Ángeles-Pérez *et al.* (2015).

En laboratorio, se obtuvo el peso seco de una submuestra representativa por conglomerado (10 g) y se determinó la concentración de carbono por digestión seca a 900°C, en el determinador automático de C (TOC SSM 5050A Shimadzu). Con los valores de concentración se calculó para cada muestra, el contenido de carbono (CC) mediante la Ecuación 1 (López-Merlín *et al.*, 2016):

$$CC = PS_m * Conc (C) * 0.01 \quad (1)$$

Donde CC = Contenido de C ($g\ m^{-2}$), Conc (C) = concentración de C (%) y PS_m = peso seco de la muestra (g). Finalmente, el CC fue convertido a $Mg\ ha^{-1}$.

1.4.4 Variables predictoras

Se utilizó la información de doce variables ambientales predictoras obtenidas para los años 2013 y 2018. Estas variables se clasificaron en tres categorías (**Cuadro 1**): topográficas (elevación, pendiente, exposición), de estructura espacial (coordenadas geográficas X y Y) y de estructura de dosel (edad del rodal, altura dominante, área basal, apertura del dosel, riqueza de especies, índice de Shannon-Wiener, cobertura arbórea). La información fue proyectada en la misma referencia geográfica (WGS84 UTM Zona 14), remuestreada a una misma resolución espacial (1 m) y transformada a formato raster mediante el software ArcGIS 10.5 (Environmental System Research Institute, ESRI Inc., Redlands, CA).

Cuadro 1. Descripción de las variables ambientales utilizadas para predecir el contenido de carbono en el piso forestal.

Categoría	Variable predictora	Resolución espacial	Fuente	Referencia	Unidades
Topografía	Elevación	1 m	LiDAR	Soriano-Luna <i>et al.</i> , 2018	msnm
	Pendiente	1 m	LiDAR	-	grados
	Exposición	1 m	LiDAR	-	grados
Estructura espacial	Coordenadas X	1 m	LiDAR	-	m
	Coordenadas Y	1 m	LiDAR	-	m
Estructura del dosel	Edad del rodal	1 m	Archivo vectorial	Soriano-Luna <i>et al.</i> , 2015	años
	Altura dominante	5 m	Inventario/LiDAR	Soriano-Luna <i>et al.</i> , 2018	m
	Área basal	sitio*	Inventario	-	$m^2\ ha^{-1}$
	Fracción de apertura del dosel	Sitio	Fotografías hemisféricas	-	%
	Riqueza de especies	Sitio	Inventario	Villegas-Macedo, 2019	árboles sitio ⁻¹
	Índice de Shannon-Wiener	1 m	Archivo raster	Villegas-Macedo, 2019	-
	Cobertura arbórea	5 m	LiDAR	Soriano-Luna <i>et al.</i> , 2018	%

*Sitio: unidad de muestreo secundaria con una superficie de $400\ m^2$.

Las variables topográficas y de estructura espacial se derivaron de un Modelo de Elevación Digital (DEM) con resolución espacial de 1 m. El DEM fue generado para toda el área de estudio por Soriano-Luna *et al.* (2018), a partir de datos derivados de métricas LiDAR (Light Detection and Ranging).

La información utilizada para caracterizar la estructura del dosel fue extraída de tres fuentes: métricas LiDAR, inventarios dasométricos y fotografías hemisféricas. Los datos de altura dominante, área basal, apertura del dosel y cobertura arbórea se obtuvieron a partir de métricas LiDAR (Soriano-Luna *et al.*, 2018) y de dos inventarios dasométricos, los cuales se realizaron en 2013 y 2018, correspondientes al mismo periodo de colecta de muestras. Este conjunto de datos fue obtenido en campo mediante la metodología establecida por el INFyS (CONAFOR, 2012). La edad del rodal se definió como el tiempo transcurrido a partir del año en que se aplicó la corta de repoblación (Ángeles-Pérez *et al.*, 2015); la información de esta variable fue extraída de polígonos en formato vectorial delimitados por Soriano Luna *et al.* (2015). Los datos de riqueza de especies (S) y la estimación espacialmente explícita del Índice de Shannon-Wiener fue obtenida por Villegas-Macedo (2019) y convertida a la misma resolución espacial.

1.4.5 Análisis estadístico

1.4.5.1 Análisis exploratorio

El análisis estadístico fue realizado en R v.3.5.1. (R Development Core Team, 2018). Se realizó un análisis preliminar de calidad de la información para identificar datos atípicos o faltantes. Posteriormente, se obtuvieron las distribuciones de frecuencia y se comprobó la normalidad de los datos mediante la prueba de Kolmogorov-Smirnov (Brown y Forsythe, 1974). El CC en cada capa del piso forestal fue descrito mediante la estadística de resumen tradicional (media, mediana, mínimo, máximo, desviación estándar, coeficiente de variación, curtosis y asimetría). Se utilizó el coeficiente de correlación de Spearman para detectar problemas de multicolinealidad entre variables predictoras, seleccionando para el entrenamiento de los modelos, aquellas variables que presentaron una correlación menor al 80% ($r < 0.8$).

1.4.5.2 Métodos de predicción espacial

Se comparó el desempeño predictivo de cuatro métodos DSM, pertenecientes a tres técnicas diferentes, para estimar la distribución espacial del CC por capa de descomposición del piso forestal, en 2013 y 2018. Los métodos evaluados para cada técnica fueron: 1) interpolación espacial (Kriging Ordinario), 2) regresión espacial (Modelo Lineal Generalizado y Aditivo Generalizado), y 3) Machine Learning (Random Forest). La elección de estos métodos responde a tres aspectos: a) incluyen algoritmos que han demostrado ser eficientes en el modelado predictivo de carbono en el suelo, b) abarcan la mayoría de las rutinas estadísticas más aplicadas y representativas de las técnicas DSM, y c) pueden manejar estructuras de dependencia espacial (Scull *et al.*, 2003).

1.4.5.2.1 Método de interpolación espacial

Kriging Ordinario (KO) es un método geoestadístico univariado que estima el valor de un atributo continuo z a cualquier ubicación no muestreada x_0 (Goovaerts, 1999; Webster y Oliver, 2007). Para ello, requiere el cumplimiento de dos supuestos: normalidad y estacionariedad espacial de los datos. Una vez verificados estos supuestos, y previo a realizar estimaciones en áreas no muestreadas, se realizó un análisis estructural de la variable de interés, en el que se obtuvo el semivariograma empírico del CC de cada una de las capas del piso forestal y posteriormente se ajustó a uno de los modelos teóricos disponibles (Bivand *et al.*, 2008). Los semivariogramas obtenidos miden la fuerza de la correlación espacial del CC en función de la distancia, utilizando la Ecuación 2:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \quad (2)$$

Donde $\gamma(h)$ es la semivarianza; h es la distancia de retraso; $N(h)$ es el número de pares de ubicaciones separadas por una distancia de retraso h ; $Z(x_i)$ y $Z(x_i + h)$ son los valores medidos de la variable a una ubicación x_i y $(x_i + h)$ (Berndt y Haberlandt, 2018; Bhunia *et al.*, 2018). El ajuste del semivariograma está en función de tres parámetros: rango, meseta y nugget. El rango expresa la distancia a la cual la autocorrelación se desvanece, la meseta muestra la máxima variabilidad en ausencia de dependencia espacial y el nugget define la variabilidad intrínseca de

los datos. La relación nugget/meseta caracteriza la dependencia espacial de los valores, considerando los siguientes umbrales: a) $< 25\%$ indica fuerte dependencia espacial, b) entre 25 y 75% indica moderada dependencia espacial y c) $> 75\%$ los valores exhiben dependencia espacial débil (Cambardella *et al.*, 1994). Los paquetes utilizados para la generación de semivariogramas y la interpolación kriging fueron *geoR* (Ribeiro y Diggle, 2006), *gstat* (Pebesma, 2004) y *automap* (Hiemstra, 2013).

1.4.5.2.2 Métodos de regresión espacial

El Modelo Lineal Generalizado (GLM) cuenta con un componente aleatorio que identifica a la variable de respuesta y su distribución de probabilidad, un componente sistemático que especifica las variables explicativas utilizadas en la función predictora lineal y una función de enlace del valor esperado de Y , $E(Y)$, como una combinación lineal de las variables predictoras (Wood y Augustin, 2002; Antúnez *et al.*, 2017; Wood, 2017). El Modelo Aditivo Generalizado (GAM) es un método que incorpora la regresión semi paramétrica y es capaz de modelar relaciones no lineales entre las variables (Antúnez *et al.*, 2017). GAM está construido por la suma de funciones suavizadas de las variables predictoras, denominadas “*splines*”. Estas funciones permiten identificar el tipo de efecto y la relación no lineal entre variables.

Tanto GLM como GAM permiten la inclusión de una distribución teórica definida en su estructura inicial. En este caso, debido a la distribución de la variable de respuesta, se asumió una distribución gaussiana para todos los análisis. Los modelos de regresión fueron generados para 2013 y 2018, mediante los paquetes de R: “*stats*” (Bolar, 2019) para GLM y “*mgcv*” (Wood y Wood, 2019) para GAM. Se definió al CC como variable dependiente y se agregó (como variable factorial) la capa del piso forestal perteneciente, a fin de obtener las estimaciones para cada una de las capas evaluadas. Las variables predictoras fueron seleccionadas posteriormente al análisis exploratorio. Para minimizar el efecto de autocorrelación espacial y sobreajuste, se incluyó la dependencia espacial dentro de la estructura del algoritmo, como una función de las coordenadas geográficas y se determinó, mediante el Índice de Moran, la autocorrelación espacial de los residuales con el paquete “*spdep*” (Bivand, 2019). Por último, para controlar la complejidad de los modelos, se realizó la comprobación del cumplimiento de supuestos en residuales y se determinó el Criterio de Información de Akaike (AIC) y el Criterio de Información Bayesiano (BIC).

1.4.5.2.3 Método de Machine Learning

Random Forest (RF) es una de los métodos más populares de Machine Learning (ML), es un modelo no paramétrico de clasificación y regresión que permite modelar variables tanto continuas como categóricas (Breiman, 2001; Biau 2012; Biau y Scornet, 2016). En comparación con los métodos anteriores, RF no hace suposiciones con respecto a la distribución de los datos de entrada e incorpora aleatoriedad en sus predicciones, a través de un muestreo bootstrap reiterativo y la selección aleatoria de variables al generar cada árbol de decisión (Liaw y Wiener, 2002; Zeraatpisheh *et al.*, 2017).

Se generaron dos modelos RF para el conjunto de datos de 2013 y 2018 mediante los paquetes “*randomForest*” (Liaw y Wiener, 2002) y “*ModelMap*” (Freeman *et al.*, 2018). Se agregó como variable factorial, la capa del piso forestal perteneciente, a fin de obtener las estimaciones para cada una de las capas evaluadas y se utilizaron los mismos predictores de GAM y GLM. Para el entrenamiento de cada modelo se desarrolló un análisis de sensibilidad extensivo sobre dos parámetros contenidos en el algoritmo de RF, hasta obtener el mejor desempeño predictivo. Los parámetros fueron el máximo número de árboles (*n*tree) y el número de predictores para cada árbol (*m*try). Para estabilizar el error de cada modelo se definió un valor de *n*tree igual a 2000, mientras que *m*try consideró al menos un tercio del número de variables predictoras seleccionadas.

1.4.5.3 Validación y desempeño de los modelos

El desempeño del modelo en la predicción fue evaluado a través del método de validación cruzada (k=10). La validación cruzada proporciona una estructura para crear varias divisiones aleatorias de entrenamiento/prueba en el conjunto de datos, de esta forma, se garantiza que cada conjunto de datos se encuentre en el conjunto de prueba al menos una vez (Zeraatpisheh *et al.*, 2019). Por lo cual, del conjunto total de datos, 90% fueron aleatoriamente seleccionados para entrenamiento y 10% para la validación del modelo; este proceso fue repetido 100 veces. Los paquetes de R utilizados para cada modelo fueron: “*gstat*” (Pebesma, 2004) para KO; “*boot*” (Canty y Ripley, 2019) para GLM; “*gamclass*” (Maindonald, 2018) para GAM y “*ModelMap*” (Freeman *et al.*, 2018) para RF.

Los criterios utilizados para evaluar el desempeño de estos modelos fueron: a) el porcentaje de varianza explicada, obtenida con el paquete “*modEVA*” (Barbosa, 2020); b) el coeficiente de determinación (R^2) (Ecuación 3); c) el error medio absoluto (MAE) (Ecuación 4) y d) el error cuadrático medio (RMSE) (Ecuación 5)

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (3)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (4)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (5)$$

Donde, \hat{y}_i es el i ésimo valor obtenido de la predicción, y_i es el i ésimo valor observado, \bar{y}_i es el valor medio de y_i , y n es el número de valores predichos u observados con $i = 1, 2, \dots, n$. Este procedimiento mejora la precisión en la estimación de la incertidumbre del modelo y previene que los análisis presenten problemas de sobreajuste (Beguin *et al.*, 2017).

1.4.5.4 Mapas de distribución espacial

Para cada capa y año evaluado, se generaron mapas de distribución espacial del CC, utilizando los modelos ajustados. Los paquetes utilizados fueron: “*gstat*” (Pebesma, 2004) para KO, “*raster*” (Hijmans, 2019) para GLM y GAM, “*ModelMap*” (Freeman *et al.*, 2018) para RF.

1.5 RESULTADOS

El carbono en el piso forestal fue mayor en la capa F, lo cual sugiere un efecto del grado de descomposición del material orgánico. Se observó una disminución del CC en las dos las capas del piso forestal de 2013 a 2018. El valor medio total del CC fue de $12.19 \pm 5.48 \text{ Mg ha}^{-1}$ para 2013 y de $5.76 \pm 2.34 \text{ Mg ha}^{-1}$ para 2018. La estimación del CC para 2013 y 2018 reveló una moderada variabilidad (CV), sin embargo, la variabilidad de las estimaciones en cada capa se

mantuvo. Para la capa Ho, se obtuvo un CV de 36.36% y 36.85% en 2013 y 2018, respectivamente; para la capa F, el CV fue 52.53% para 2013 y 50.45% para 2018. En todas las capas y en ambos años evaluados, el CC presentó una distribución normal, por lo que no se requirió transformaciones adicionales para la generación de los modelos (**Cuadro 2; Anexo, Figura S1**).

Cuadro 2. Estadística descriptiva del contenido de carbono (Mg ha^{-1}), estratificado por año y capa del piso forestal.

Año	Capa	Media	Mediana	Min	Max	DS	CV	Kur	Skew	KS _p
2013	Ho	2.03	1.95	0.00	4.72	0.74	36.36	4.75	0.58	0.34*
	F	10.16	9.37	0.00	26.33	5.34	52.53	3.27	0.54	0.61*
	Tot	12.19	11.47	0.77	27.75	5.48	44.98	3.18	0.48	0.54*
2018	Ho	1.85	1.78	0.28	5.23	0.68	36.85	6.63	1.07	0.33*
	F	3.91	3.83	0.00	8.81	1.97	50.45	2.55	0.08	0.99*
	Tot	5.76	5.65	0.44	11.80	2.34	40.54	2.65	-0.06	0.99*

Ho: hojarasca, F: fermentación, Tot: suma de Ho y F; Min: mínimo, Max: máximo, DS: desviación estándar; CV: coeficiente de variación (%); Kur: curtosis; Skew: asimetría y KS_p: Prueba de Kolmogórov-Smirnov (*nivel de significancia $\alpha > 0.05$).

Las variables que presentaron una relación lineal significativa con el CC fueron el área basal, la edad del rodal y la altura dominante (**Cuadro 3**). El grado de correlación entre variables predictoras y sus distribuciones de frecuencia se muestra en el anexo, en las **Figura S2, Figura S3; y Cuadro S1, Cuadro S2**. Las únicas variables con $r^2 \geq 0.8$ en ambos años, fueron la Riqueza de especies (S) y el Índice de Shannon (H). Dado que H fue estimado a partir de S (Villegas-Macedo, 2019), sólo se seleccionó a la variable S para los análisis posteriores.

Cuadro 3. Coeficientes de correlación de Spearman (r^2) entre el CC por capa del piso foresta y cada una de las variables predictoras, para 2013 y 2018.

Año	Capa	X	Y	ELE	SLO	ASP	SA	BA	DH	S	H	TCO	CO
2013	Ho	0.01	0.11	0.13	-0.06	0.01	^b 0.21	^b 0.24	^a 0.19	^a 0.19	^a 0.20	0.13	^a -0.17
	F	0.07	0.12	0.12	^a 0.01	^a 0.04	0.33	0.37	0.35	0.16	0.11	0.20	-0.01
	Tot	0.06	0.14	0.15	0.00	0.05	^b 0.36	^b 0.39	^b 0.36	^a 0.20	0.15	^b 0.23	-0.03
2018	Ho	0.11	^a 0.16	0.05	-0.09	0.06	^b 0.42	^b 0.36	^b 0.38	^b 0.23	^b 0.29	0.10	0.10
	F	0.12	0.00	-0.02	-0.09	0.05	^b 0.24	^b 0.32	^b 0.25	^a 0.20	^a 0.17	^a 0.19	^a -0.19
	Tot	0.14	0.04	-0.02	-0.09	0.04	^b 0.29	^b 0.35	^b 0.29	^b 0.23	^b 0.21	0.11	-0.11

X, Y: Coordenadas UTM X, UTM Y; ELE: elevación; SLO: pendiente; ASP: exposición; SA: edad del rodal; BA: área basal; DH: altura dominante; S: riqueza de especies; H: índice de Shannon-Wiener; TCO: cobertura arbórea; CO: Apertura del dosel. Nivel de significancia: (a < 0.05), (b < 0.01).

1.5.1 Entrenamiento de los modelos

1.5.1.1 Kriging Ordinario

El análisis de semivariogramas no mostró anisotropía, por lo que la variable se ajustó a semivariogramas omnidireccionales (**Cuadro 4**). Los modelos teóricos Matern, Stein y Periódico presentaron el mejor ajuste (**Anexo, Figura S4**). Los datos mostraron una variación considerable a pequeña escala, principalmente para 2018, con excepción de las capas F y Tot en 2013. Este comportamiento espacial sugiere que el diseño y la intensidad del muestreo no fue suficientemente intensivo para revelar patrones espaciales del CC bajo este método.

Cuadro 4. Parámetros de los modelos de semivariograma con mejor ajuste, estratificado por año y capa del piso forestal.

Año	Capa	Ajuste de modelo teórico	Nugget	Meseta	Rango (m)	Kappa	Nugget/meseta
2013	Ho	Matern	0.39	0.56	137	1.9	0.69
	F	Stein	0.00	29.0	50	1.1	0.00
	Tot	Stein	0.00	31.0	53	0.5	0.00
2018	Ho	Periódico	0.37	0.41	58	0.0	0.90
	F	Stein	1.5	3.8	116	0.2	0.39
	Tot	Stein	1.5	5.4	91	0.2	0.27

Ho: hojarasca, F: fermentación, Tot: suma de Ho y F.

Por el contrario, la relación nugget/meseta mostró diferencias en el grado de dependencia espacial del CC en las tres capas del piso forestal. Las capas F y Tot en 2013, mostraron una fuerte dependencia espacial, con valores inferiores a 0.25. Para 2018, la capa Ho mostró una débil dependencia espacial, con un valor de 0.90. El resto de las capas exhibieron una moderada dependencia espacial.

1.5.1.2 Modelos GLM, GAM y RF

Se detectó que evaluar el CC por capas del piso forestal contribuyó significativamente en el ajuste de ambos modelos por año (**Cuadro 5**). Para GLM, las variables con mayor contribución en las estimaciones de 2013 fueron la elevación y el área basal ($\alpha < 0.001$), la altura dominante y la

cobertura arbórea ($\alpha < 0.01$) y en menor medida, la estructura espacial ($\alpha < 0.05$). La significancia de las variables en GLM se redujo en 2018, pero se mantuvieron las mismas variables. Para 2013, GAM indicó que las variables con mayor efecto en la predicción de CC fueron la elevación, la edad del rodal y la cobertura arbórea ($\alpha < 0.001$), con grados de libertad estimados (EDF) ligeramente superiores a 8.0. En 2018, se observó un mayor efecto de la estructura espacial sobre el CC y se apreciaron incrementos en la significancia del resto de las variables, reflejando además un mayor grado de relación no lineal.

Cuadro 5. Parámetros y significancia de los predictores en los modelos GLM y GAM, para predecir CC en las capas del piso forestal, en 2013 y 2018.

Ajuste del modelo	(X_i)	GLM				GAM			
		2013		2018		2013		2018	
		β_i	α	β_i	α	EDF	α	EDF	α
Capa	Ho	2645.0	0.0323 ^a	-504.2	0.3765	2.048	0.0001 ^c	3.915	0.0001 ^c
	F	2653.0	0.0317 ^a	-502.1	0.3748	10.149	0.0001 ^c	1.851	0.0001 ^c
	Tot	2655.0	0.0316 ^a	-500.3	0.3785	12.177	0.0001 ^c	5.765	0.0001 ^c
	X	0.0012	0.0266 ^a	-0.0004	0.1086	1.000	0.5407	9.000	0.0001 ^c
	Y	-0.0015	0.0121 ^a	0.0003	0.2479	2.186	0.2298	8.470	0.0005 ^c
	ELE	0.0454	0.0008 ^c	-0.0123	0.0374 ^a	8.426	0.0001 ^c	8.308	0.0001 ^c
	SLO	0.0495	0.1397	-0.0320	0.0177 ^a	1.821	0.0699	6.320	0.1293
	ASP	0.0001	0.9560	-0.0006	0.4951	1.000	0.1640	7.775	0.0903
	SA	0.0160	0.2730	-0.0039	0.5189	8.548	0.0001 ^c	7.789	0.0001 ^c
	BA	0.1422	0.0001 ^c	0.0394	0.0251 ^a	3.162	0.0015 ^b	8.483	0.0001 ^c
	DH	0.1324	0.0035 ^b	-0.0312	0.0234 ^a	7.455	0.0575	1.000	0.9246
	S	-0.0478	0.6370	-0.0021	0.9569	8.152	0.0439 ^a	8.463	0.0009 ^c
	TCO	-0.0387	0.0082 ^b	0.0252	0.0001 ^c	8.636	0.0001 ^c	5.733	0.1048
	CO	0.0175	0.6868	0.0382	0.0826	5.910	0.0093 ^b	4.690	0.0002 ^c

β_i : Parámetro desconocido de cada predictor; α : nivel de significancia de los predictores ^a($\alpha \leq 0.05$), ^b($\alpha \leq 0.01$) y ^c($\alpha \leq 0.001$); EDF: grados de libertad estimados para GAM; X, Y: Coordenadas UTM X, UTM Y; ELE: elevación; SLO: pendiente; ASP: exposición; SA: edad del rodal; BA: área basal; DH: altura dominante; S: riqueza de especies; D: índice de Shannon-Wiener; TCO: cobertura arbórea, CO: apertura del dosel.

GAM y GLM presentaron AIC y BIC ligeramente similares entre años, con un mejor ajuste para 2013 en comparación con 2018. No obstante, entre modelos, los valores de AIC y BIC son contradictorios (**Cuadro 6**). A pesar de que AIC y BIC miden la bondad del ajuste a partir de la máxima verosimilitud del modelo, BIC penaliza más la complejidad del modelo que AIC, lo que sugiere que el modelo GLM de 2018 fue el modelo con mejor ajuste.

Cuadro 6. Parámetros de ajuste de los modelos GLM, GAM y RF para predecir CC en el piso forestal.

Año	GLM				GAM				RF	
	AIC	BIC	I. Moran	α	AIC	BIC	I Moran	α	Ntree	Mtry
2013	2590.5	2652.2	-0.05	0.6604	2498.2	2746.3	-0.31	1.00	2000	9
2018	1773.7	1835.4	0.14	0.0028	1631.4	1960.8	-0.24	1.00	2000	7

AIC: Criterio de Información de Akaike; BIC: Criterio de Información Bayesiano. I Moran: Índice de autocorrelación espacial de Moran ($\alpha > 0.05$).

Los residuales para GAM y GLM idealmente deben asumir una distribución idéntica e independiente, sin embargo, se obtuvieron valores más cercanos a cero para GLM en ambos años. Para GAM, los residuales presentaron una ligera autocorrelación espacial. La significancia de autocorrelación espacial para ambos modelos fue alta ($\alpha > 0.05$), con excepción de GLM en 2018 (consistente con el valor más bajo de BIC), lo que puede indicar un ligero sobreajuste en el resto de los modelos.

Respecto al entrenamiento de RF, en ambos años el error fue estabilizado mediante la inclusión de 2000 árboles (*n tree*), mientras que el análisis de sensibilidad del modelo permitió definir *mtry* en nueve y siete variables por cada nodo, para 2013 y 2018, respectivamente (**Cuadro 6**).

1.5.2 Desempeño de los modelos

Aunque hay marcadas diferencias en el desempeño de los modelos predictivos, la superioridad de RF, GAM y GLM sobre KO en la predicción del CC en el piso forestal fue muy notable, con valores de R^2 superiores a 0.70 (**Cuadro 7**).

En vista de que KO es un método univariado que fue generado de forma individual para cada una de las capas del piso forestal, su principal limitante es que los valores de MAE y RMSE no pueden ser comparados con el resto de los modelos. Sin embargo, los valores de R^2 dan una idea del bajo desempeño que presentó KO para modelar CC en el piso forestal, con valores de 0.3 para 2013 y entre 0.7 a 0.3 para 2018. Los valores observados vs predichos revelaron una sobreestimación del CC en todos los valores antes de intersección, mientras que los valores localizados después de la intersección mostraron una subestimación del CC (**Figura 4**).

Cuadro 7. Estadísticas de validación cruzada de las tres técnicas empleados para modelar el CC en el piso forestal.

Año	Técnica	Método de estimación	Bondad de ajuste		Error (Mg ha ⁻¹)	
			R ²	Varianza explicada (%)	MAE	RMSE
2013	Interpolación espacial	KO (Capa Ho)	0.34	ND	0.51	0.69
		KO (Capa F)	0.34	ND	3.97	5.01
		KO (Capa Tot)	0.31	ND	4.15	5.21
	Regresión espacial	GLM	0.75	84.1	2.96	4.21
		GAM	0.77	89.4	3.20	4.02
Machine Learning	RF	0.90	86.2	1.81	2.62	
2018	Interpolación espacial	KO (Capa Ho)	0.07	ND	0.53	0.72
		KO (Capa F)	0.29	ND	1.56	1.88
		KO (Capa Tot)	0.31	ND	1.80	2.21
	Regresión espacial	GLM	0.72	86.6	1.25	1.73
		GAM	0.76	92.7	1.25	1.58
Machine Learning	RF	0.86	69.1	0.85	1.51	

R²: coeficiente de determinación, MAE: error medio absoluto, RMSE: error cuadrático medio. ND: No definido para interpolación con Kriging Ordinario (KO); GLM: Modelo Lineal generalizado; GAM: Modelo aditivo Generalizado; RF: Modelo Random Forest.

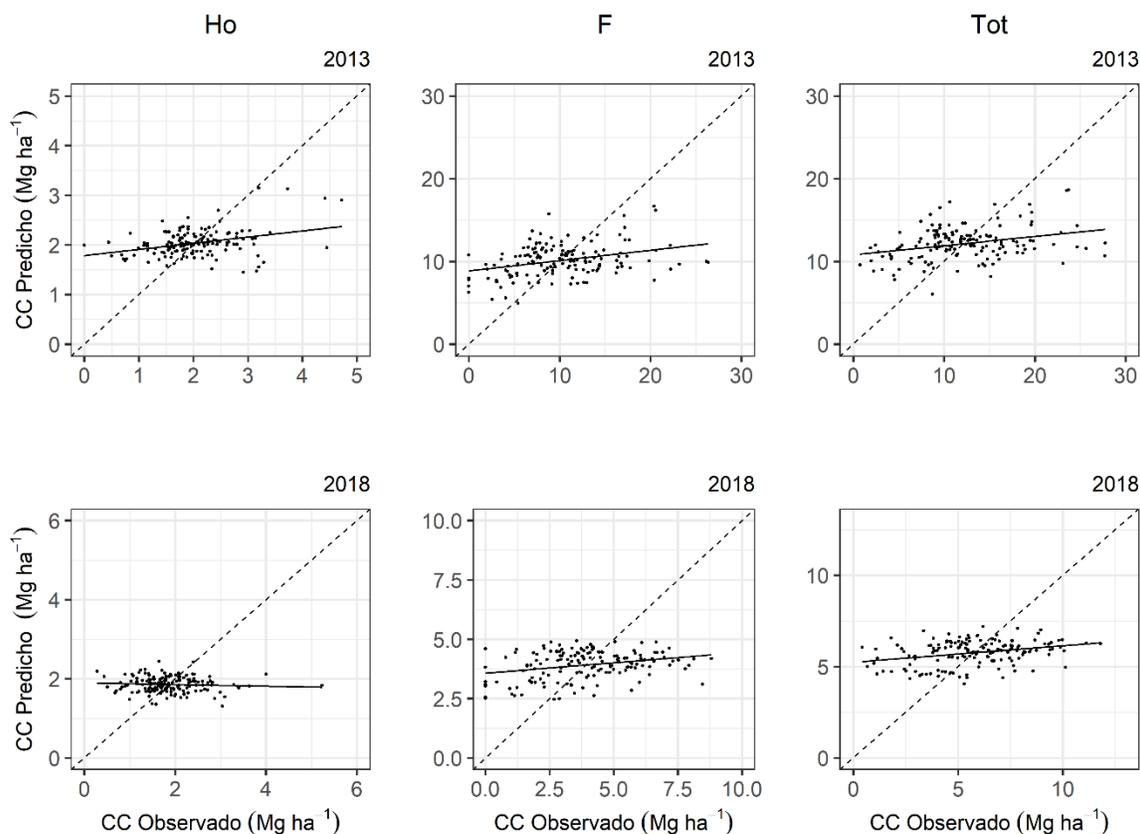


Figura 4. Valores observados vs predichos del conjunto de datos de validación cruzada (k=10) del modelo “Kriging ordinario”, para predecir CC en las capas del piso forestal: Ho (hojarasca), F (fermentación) y Tot (Totales), en 2013 y 2018. La línea negra es de regresión y la línea punteada es la relación 1:1

Al comparar las técnicas de regresión y ML, los resultados indicaron que RF fue el método con mejor desempeño para modelar la distribución del CC en el piso forestal, con valores de R^2 de 0.90 para 2013 y 0.86 para 2018 (Cuadro 7, Figura 5). Respecto a la varianza explicada, las variables predictoras utilizadas explicaron para RF el 86.2 y 69.1% en 2013 y 2018, respectivamente. Por su parte, GAM y GLM mostraron valores más altos para este criterio, sin embargo, considerando la ligera dependencia espacial en los residuales, puede existir un ligero sobreajuste. Por tanto, la relación espacial entre el CC en el piso forestal contiene información valiosa que no fue totalmente capturada por RF, especialmente para 2018; sin embargo, RF presentó valores más bajos de MAE y RMSE, en comparación con GLM y GAM. Lo anterior indicó un mejor ajuste y menor incertidumbre para esta RF. Ambos modelos de regresión presentaron una mayor sobreestimación en comparación con RF.

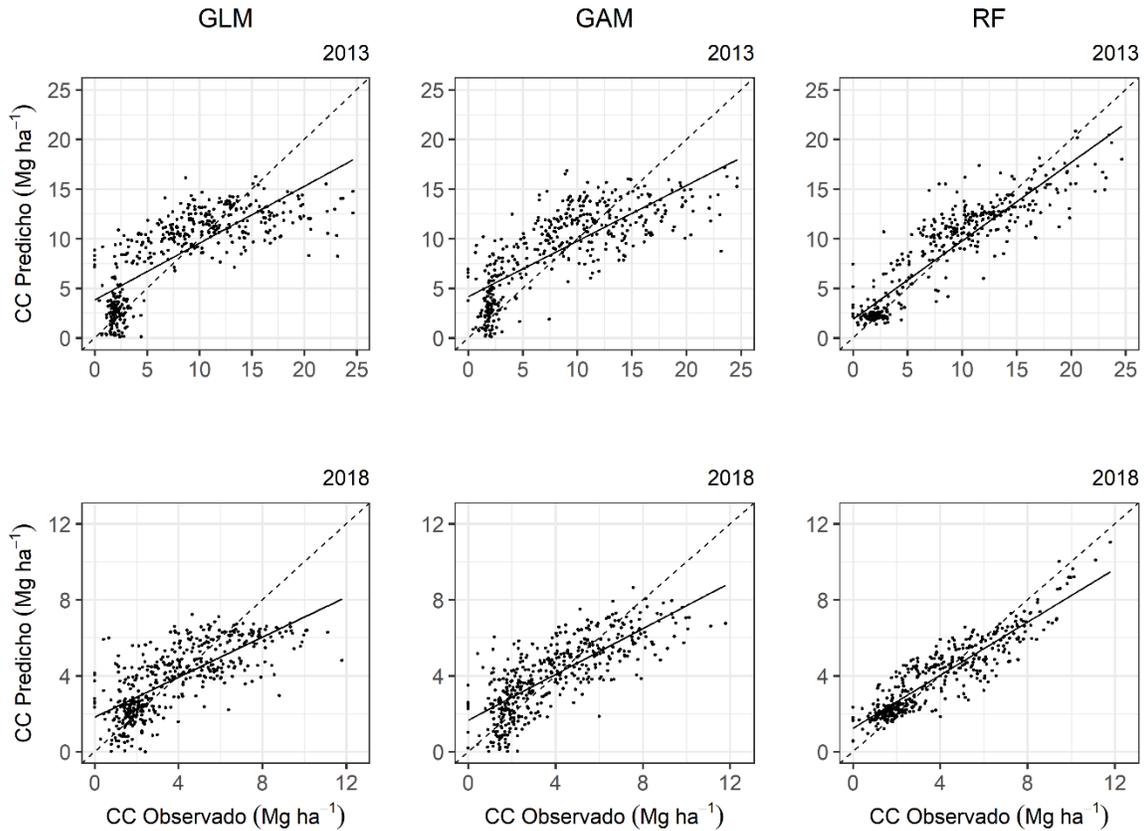


Figura 5. Valores observados vs predichos del conjunto de datos de validación cruzada ($k=10$) de los modelos GAM, GLM y RF, para predecir CC en el piso forestal, en 2013 y 2018. La línea negra es de regresión y la línea punteada es la relación 1:1.

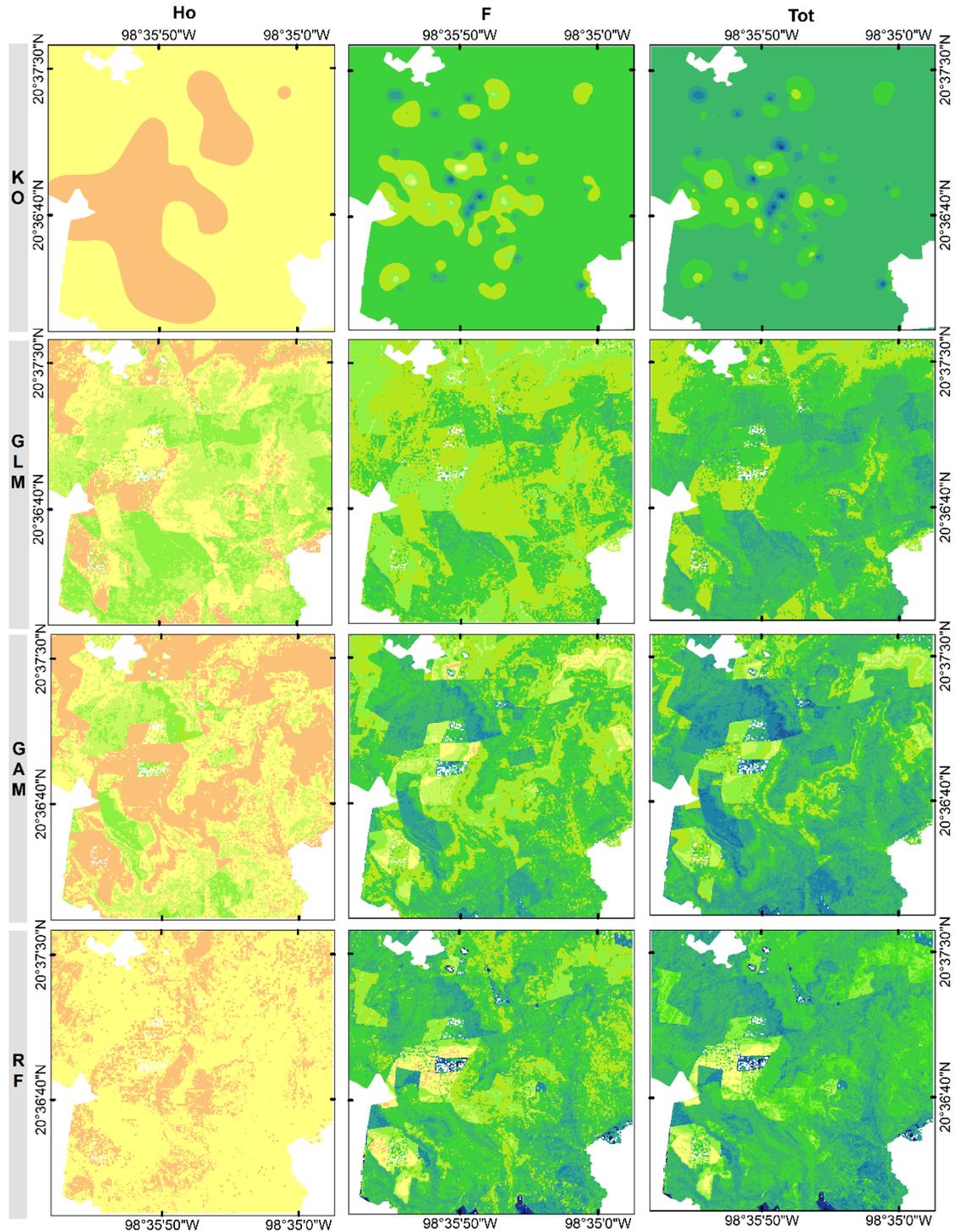
1.5.3 Mapas de CC a partir de las técnicas de predicción

Los mapas de distribución espacial del CC en cada una de las capas del piso forestal se muestran en la **Figura 6** para el año 2013, y en **Figura 7** para el año 2018. Debido a que el CC varía a distancias muy cortas, la interpolación espacial fue generada a partir de una malla fina de píxeles (1 m de resolución) que cubre toda el área de estudio (3 x 3 km). Para facilitar la comparación en las predicciones entre años, capas y modelos, las predicciones espaciales fueron reclasificadas en intervalos geométricos.

El CC exhibió una alta variación espacial dentro del SMIC-Atopixco. Para ambos años, se observó una diferencia notoria entre los patrones de distribución espacial del CC en las capas Ho y F. La interpolación con KO exhibió patrones circulares alrededor de los puntos de muestreo y presentó clústeres en áreas con baja densidad de muestreo, por lo que se consideró una representación poco realista de la distribución espacial del CC en las diferentes capas del piso forestal.

Las predicciones espaciales de los modelos regresión GAM y GLM fueron notablemente diferentes entre ellos. Para 2013, ambos modelos sobreestimaron el CC con valores superiores al máximo valor observado en el análisis exploratorio. GLM mostró una mayor sobrestimación para la capa Ho y una subestimación para la capa F, lo que fue acorde con los gráficos de validación cruzada. GAM sobrestimó en menor medida el CC en la capa Ho, mientras que los patrones espaciales modelados para la capa F fueron similares a los obtenidos por RF para la misma capa.

Las predicciones en 2018 también exhibieron problemas de sub y sobreestimación del CC en los modelos de regresión, sobre todo en la región noreste de la zona de estudio, sin embargo, RF logró reducir este tipo de clústeres espaciales. GLM mostró una menor sobreestimación, contrario a 2013. Por lo que se asume que GAM presentó un mejor desempeño para modelar el CC en el año 2013 y GLM en el año 2018. Las predicciones espaciales de RF, para ambos años fueron muy cercanas a los valores mínimos y máximos observados en el análisis exploratorio y fueron consistentes con las estadísticas de validación, por lo cual, se consideró como la representación espacial más aproximada a la realidad.



2013 - Valores estimados del Contenido de Carbono (CC) en el piso forestal (Mg/ha)

0.0 - 2.0	4.0 - 6.0	8.0 - 10.0	12.0 - 14.0	16.0 - 18.0	20.0 - 22.0
2.0 - 4.0	6.0 - 8.0	10.0 - 12.0	14.0 - 16.0	18.0 - 20.0	22.0 - 24.0

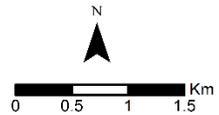
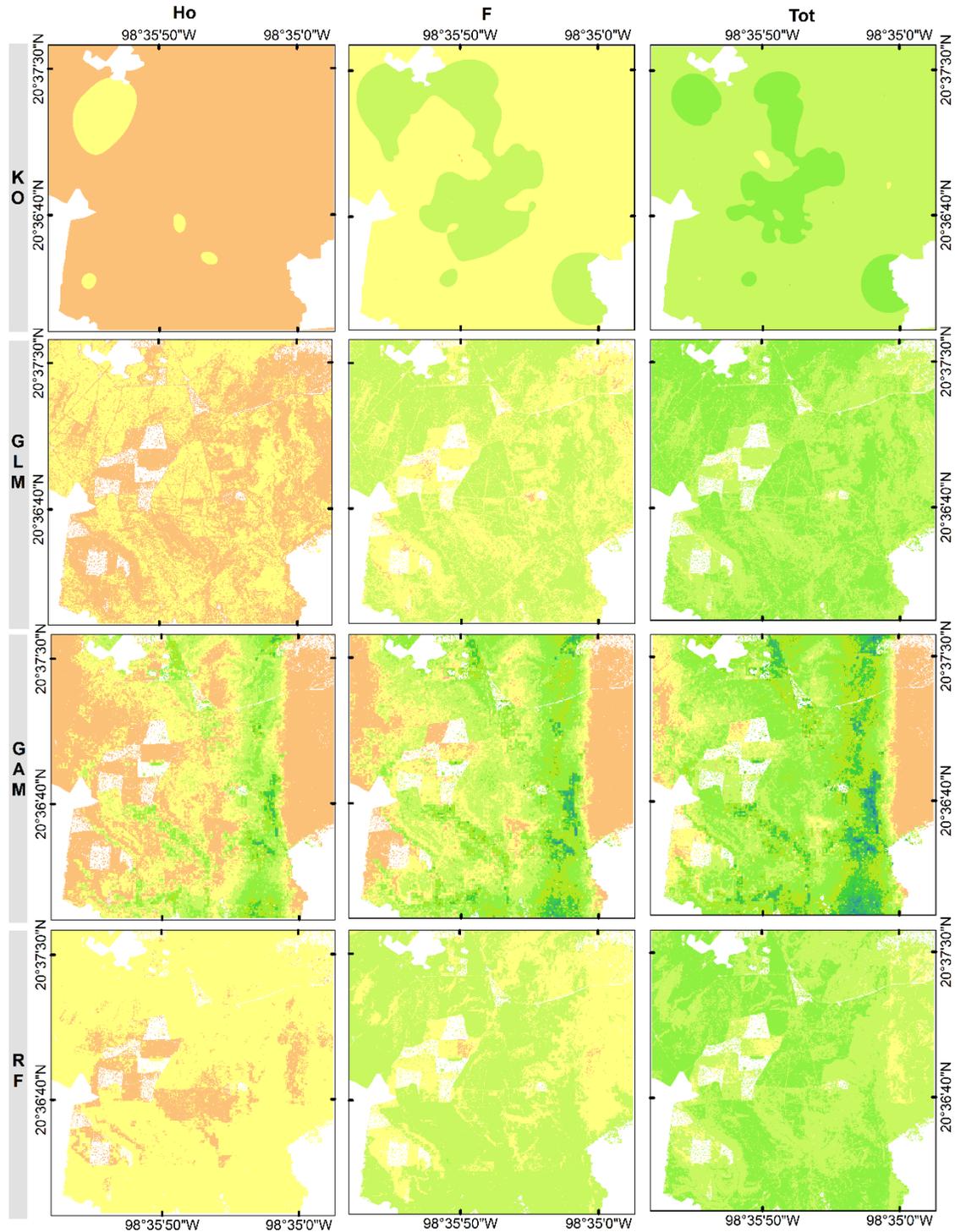


Figura 6. Predicciones espaciales para cada modelo espacial del CC en el piso forestal. Año 2013. Capas del piso forestal: Ho: hojarasca, F: fermentación, Tot: Suma de Ho y F. Modelos espaciales: Kriging Ordinario (KO), Modelo Lineal Generalizado (GLM), Modelo Aditivo Generalizado (GAM) y RF (Random Forest).



2018 - Valores estimados del Contenido de Carbono (CC) en el piso forestal (Mg/ha)

0.0 - 2.0	4.0 - 6.0	8.0 - 10.0	12.0 - 14.0	16.0 - 18.0	20.0 - 22.0
2.0 - 4.0	6.0 - 8.0	10.0 - 12.0	14.0 - 16.0	18.0 - 20.0	22.0 - 24.0

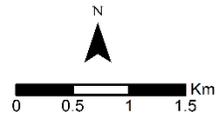


Figura 7. Predicciones espaciales para cada modelo espacial del CC en el piso forestal. Año 2018. Capas del piso forestal: Ho: hojarasca, F: fermentación, Tot: Suma de Ho y F. Modelos espaciales: Kriging Ordinario (KO), Modelo Lineal Generalizado (GLM), Modelo Aditivo Generalizado (GAM).

1.6 DISCUSIÓN

1.6.1 Selección del mejor método

La selección de un modelo apropiado es un aspecto central en la estimación de almacenes de carbono en los ecosistemas forestales (Zhang *et al.*, 2014). El empleo de diferentes técnicas como la interpolación espacial, modelos de regresión y ML permitió identificar y comparar los requerimientos estadísticos, el desempeño y las principales limitantes de cada modelo para estimar la distribución espacial del CC en el piso forestal. Aunado a lo anterior, el emplear dos conjuntos de datos de diferentes años reflejó una aproximación del potencial de aplicación de cada modelo para estimaciones futuras. Los resultados obtenidos indicaron que la interpolación espacial con KO subestima en gran medida el CC en cada una de las capas del piso forestal y consecuentemente, no representa correctamente la distribución espacial del CC. Por su parte, los modelos de regresión presentaron un desempeño aceptable, aunque ligeramente sesgado por el incumplimiento del supuesto estadístico de dependencia, que puede incurrir en un sobreajuste y/o sobreestimaciones en el CC.

La utilización de KO permitió explorar el comportamiento espacial del CC dentro del área de estudio mediante el análisis estructural y, a pesar de haber obtenido valores de R^2 relativamente bajos ($R^2 < 0.35$), estos fueron similares a los reportados en otros estudios sobre estimaciones de reservas de carbono en las capas orgánicas del suelo (Dai *et al.*, 2018; Yao *et al.*, 2019). KO es un método univariado cuyo principal supuesto es que la distribución espacial del CC es una superficie espacialmente continua y sus predicciones en puntos no muestreados estarán únicamente en función de la distancia de separación, sin embargo, el CC en el piso forestal es muy dinámico y está influenciado por otras variables ambientales y perturbaciones humanas, lo que genera discontinuidades espaciales que no logran ser explicadas por KO.

Contrastando con evaluaciones como la de Jeyanni *et al.* (2016), cuyo análisis estructural del CC fue similar a esta investigación, pero con un diseño de muestreo sistemático por transectos a 10 m de distancia, obtuvieron desempeños superiores en KO para modelar CC en mantillo y capas de fermentación. Los autores sugieren que la determinación del rango efectivo puede contribuir al momento de definir estrategias de muestreo; además indican para bosques templados, distancias de muestreo efectivas entre 4 a 12.5 m en estimaciones de CC en la capa Ho y entre 20

a 80 m para la capa F. Estos valores proporcionados son similares a los obtenidos en el ajuste de semivariogramas de este estudio. En este sentido, el bajo desempeño de KO se atribuye a tres aspectos: a) las características intrínsecas del modelo, b) la alta variabilidad del CC a pequeñas distancias y c) a la densidad y distribución espacial de los puntos de muestreo.

Por otra parte, GLM y GAM obtuvieron valores de R^2 entre 0.72 a 0.77, lo que sugiere un desempeño aceptable en ambos modelos, sin embargo, las predicciones espaciales mostraron sobreestimaciones substanciales en los almacenes de cada capa. En general, estos modelos de regresión no presentan discrepancias substanciales entre ellos (Antúnez *et al.*, 2017), sin embargo, la mayor explicación de varianza obtenida para GAM se relaciona con la inclusión de relaciones no lineales mediante funciones suavizadas, mismas que fueron sugeridas por los bajos coeficientes de correlación lineal obtenidos en el análisis exploratorio y la significancia de variables obtenidas en el entrenamiento de los modelos.

Con respecto a la técnica de ML, el alto ajuste obtenido por RF para ambos años ($R^2 > 0.85$) sugiere que esta técnica tiene un gran potencial para mejorar las estimaciones de carbono en el piso forestal, tanto a escalas locales como regionales. Esto se atribuye a tres ventajas sobresalientes de RF con respecto al resto de los modelos evaluados: a) es un método robusto con mayor flexibilidad en los supuestos estadísticos clásicos, b) mejor resistencia al sobreajuste, c) no requiere la preselección de variables, por efectos de multicolinealidad (Breiman, 2001; Beguin *et al.*, 2017). El ajuste del RF fue similar al reportado por Wiesmeier *et al.* (2014), quienes reportaron el 52 % de la varianza explicada por RF para el CC en el suelo, en una escala regional. Por su parte, Deng *et al.* (2018) reportaron, para esa misma región, un alto desempeño de RF en predecir reservas de carbono en los primeros 30 cm del perfil del suelo, con un R^2 de 0.76, en comparación con otras técnicas DSM.

La disminución en el porcentaje de varianza explicada entre 2013 y 2018, puede atribuirse a la influencia de otros factores ambientales que afectan la variabilidad del CC y que no han sido investigados en este estudio. Pese a esta disminución en la explicación de la varianza, los valores obtenidos son mucho mayores en comparación con otros estudios que se han realizado a escalas regionales y nacionales. Cao *et al.* (2019) realizaron una evaluación similar dentro de E.U., obteniendo un 15% de varianza explicada, R^2 de 0.20 y un RMSE de 9.23 Mg ha⁻¹ en predicciones espaciales de carbono en el piso forestal a escala nacional. Todos estos contrastes pueden indicar

que RF se desempeña mejor en escalas mucho más pequeñas, cuyo rango de variabilidad de los predictores utilizados sea menor y pueda ser totalmente caracterizado por el modelo. Lo que permite afirmar que RF es un algoritmo de ML adecuado para realizar estimaciones precisas de carbono en el piso forestal, a diferentes escalas espaciales.

1.6.2 Fuentes de incertidumbre

El desempeño de los modelos aplicados puede ser altamente variable y como lo sugieren Zeraatpisheh *et al.* (2017, 2019), la precisión en las predicciones estará en función de cuatro aspectos clave: a) la complejidad inherente del atributo de interés; b) las características de los datos de entrada; c) el esquema de muestreo y d) la selección adecuada de los análisis estadísticos y algoritmos requeridos para la predicción espacial.

Una de las principales limitantes para la aplicación de los métodos DSM es la disponibilidad de datos ($n = 40$, con 160 sitios de muestreo), la cual puede conducir a estimaciones imprecisas con cierto grado de sub o sobrestimación, a pesar del substancial esfuerzo para asegurar datos de alta calidad en las mediciones de campo. De acuerdo con lo señalado por Carvalho Gomes *et al.* (2019), una disponibilidad limitada de mediciones en campo y baja representatividad de puntos de muestreo en la zona de interés representan una fuente importante de error en DSM, que pueden conducir a un bajo rendimiento del modelo aplicado, principalmente para KO.

No obstante, lo anterior se compensa mediante la combinación de datos auxiliares derivados de distintas fuentes. En este caso, la utilización de información LiDAR con alta resolución redujo el esfuerzo de muestro, mejoró la capacidad predictiva de los modelos multivariados y consecuentemente, permitió estimaciones más precisas del CC desde el nivel de parcela a una superficie continua dentro de toda el área de estudio. Estos resultados son consistentes con algunas evaluaciones similares, por ejemplo, Zhang *et al.* (2014) compararon los tres enfoques para predecir biomasa arbórea (AGB) mediante información armonizada de fuentes múltiples, demostrando que la combinación de información de inventarios y LiDAR provee un esquema costo-efectivo de mapeo que mejora en gran medida el desempeño predictivo de RF a escalas regionales. Por su parte, Beguin *et al.* (2017) y Chen *et al.* (2019) compararon el desempeño de las mismas técnicas DSM para predecir la variabilidad espacial CC en las capas orgánicas del

suelo, sugiriendo que RF tiene un mejor desempeño con respecto a los métodos de interpolación, por incorporar información armonizada de variables auxiliares.

El incumplimiento de supuestos en los modelos de regresión también representa una importante fuente de error. El CC es una variable espacialmente dependiente y puede tener algún grado de redundancia con la información contenida en las covariables ambientales, por lo que frecuentemente violan el supuesto de independencia en residuales. Cuando la independencia es asumida incorrectamente, las predicciones e inferencia estadística puede ser inexacta o mal interpretada (Dalthorp, 2004; Wood, 2017). En este caso, la inclusión de la estructura espacial dentro del algoritmo de GAM y GLM permitió reducir significativamente el efecto de la dependencia espacial, que fue corroborado con los gráficos de verificación del modelo (**Figura S8**).

1.6.3 Implicaciones para la estimación del CC

Estos resultados tienen una serie de implicaciones importantes para futuras modelaciones con estas técnicas DSM. Bajo las condiciones de este estudio, la utilización de semivariogramas permitió identificar el grado de dependencia espacial, la distancia a la cual el CC varía dentro del área de estudio y las diferencias entre capas del piso forestal. Por consiguiente, el conocimiento del comportamiento espacial del CC sugirió incluir la estructura espacial y la especificación del CC por capa del piso forestal como variable factorial en los métodos de regresión y RF, lo que permitió estimaciones más robustas para estas técnicas. Algunas evaluaciones espaciales que han clasificado el piso forestal en capas, de acuerdo al grado de descomposición, han enfatizado en este aspecto, al obtener mejores evaluaciones, en términos de precisión, de las reservas de carbono en este componente (Penne *et al.*, 2010; Jeyanny *et al.*, 2016).

Una segunda implicación parte de la selección de covariables utilizadas en la predicción del CC. El desempeño de los modelos depende del tipo de variables predictoras utilizadas para representar las relaciones espaciales dentro del paisaje. En este caso, las covariables fueron seleccionadas a priori en función de evaluaciones previas dentro del área de estudio (Ángeles-Pérez *et al.*, 2015; Soriano-Luna *et al.*, 2018; Villegas-Macedo, 2019). Por otra parte, en este estudio se demostró que la combinación de datos provenientes de inventarios dasométricos e

información derivada de sensores remotos, así como la resolución a la cual se desplegó toda la información, permiten una mejora en el entrenamiento y validación de los modelos GAM, GLM y RF. Consecuentemente, mejoran su capacidad predictiva, lo cual es consistente con los resultados obtenidos por otros autores (Domke *et al.*, 2016; Cao *et al.*, 2019).

1.7 CONCLUSIONES

Entre las técnicas DSM comparadas, el método de interpolación espacial (KO) tuvo la menor precisión, atribuida a la estrategia de muestreo, la complejidad intrínseca del CC y las características inherentes del modelo. Debido a las relaciones complejas entre el CC y las variables predictoras, los modelos GAM y RF presentaron los mejores desempeños, sin embargo, GAM y GLM mostraron un ligero sobreajuste por incumplimiento de supuestos estadísticos. El alto desempeño de RF se atribuye su mayor flexibilidad en el cumplimiento de supuestos estadísticos clásicos.

La distribución espacial del contenido de carbono, obtenida con kriging ordinario fue poco realista; los mapas derivados de GAM y GLM mostraron sobreestimaciones superiores al máximo valor observado en campo. RF presentó los mapas más aproximados a la realidad, con valores muy cercanos a los mínimos y máximos obtenidos en campo.

En comparación con estimaciones tradicionales de carbono en el piso forestal, la metodología empleada en este estudio proporciona tres ventajas para su aplicación: 1) permite combinar información derivada de fuentes múltiples, la cual mejora el desempeño predictivo de los modelos evaluados; 2) reduce el esfuerzo de muestreo; y 3) puede aplicarse a escalas regionales o nacionales.

CAPÍTULO II. PATRONES ESPACIO-TEMPORALES DEL CONTENIDO DE CARBONO EN EL PISO FORESTAL DE UN BOSQUE TEMPLADO MANEJADO

2.1.RESUMEN

El carbono en el piso forestal es altamente variable en espacio y tiempo, como resultado de la interacción entre diferentes factores ambientales. Comprender la dinámica espacial del carbono en el piso forestal es esencial para un manejo forestal sostenible y para crear estrategias de mitigación contra el cambio climático. No obstante, esta información es aún poco conocida en ecosistemas forestales mexicanos. En este estudio se evaluó el efecto de factores biofísicos sobre la distribución espacial del contenido de carbono (CC) en el piso forestal de un bosque templado bajo manejo para la producción maderable. El piso forestal se clasificó en dos capas (Ho y capa de fermentación) y las evaluaciones fueron realizadas para 2013 y 2018. Las estimaciones del CC se obtuvieron mediante el algoritmo Random Forest (RF). Para ambos años, el CC incrementó en función del grado de descomposición, con estimaciones totales promedio de CC de 12.19 ± 5.48 Mg ha⁻¹ para 2013 y de 5.76 ± 2.34 Mg ha⁻¹ para 2018. La validación cruzada (k=10) mostró el alto desempeño de RF para predecir la distribución espacial del CC ($R^2 > 0.85$). La edad del rodal explicó en mayor medida la variabilidad espacial del CC en el piso forestal. La relación C: N sugirió un incremento en las tasas de descomposición entre 2013 y 2018. La importancia de variables indicó que el área basal, la altura dominante, la cobertura arbórea y la elevación desempeñaron un papel relevante en la estimación del CC. Se concluye que la aplicación de RF ofrece la oportunidad de predecir y caracterizar los patrones, relaciones e interacciones espaciales del carbono en el piso forestal, al ser un método novedoso de predicción espacial. **Palabras clave:** almacenes de carbono; hojarasca; variabilidad espacio-temporal; manejo forestal; random forest.

SPATIO-TEMPORAL PATTERNS OF FOREST FLOOR CARBON CONTENT IN A TEMPERATE MANAGED FORESTS

2.2.ABSTRACT

The carbon in the forest floor is highly variable in space and time, as a result of the interaction among different environmental factors. Understanding the spatial dynamics of forest floor carbon is essential for sustainable forest management and for creating climate change mitigation strategies. However, this information is still poorly understood in Mexican forest ecosystems. In this study, the effect of biophysical factors on the spatial distribution of forest floor carbon content (CC) of a temperate forest under management for timber production was evaluated. The forest floor was classified into two layers (Ho and fermentation layer) and the evaluations were carried out in 2013 and 2018. CC estimates were obtained using the Random Forest (RF) algorithm. For both years, the CC increased according to the degree of decomposition, with average total CC values of $12.19 \pm 5.48 \text{ Mg ha}^{-1}$ for 2013, and $5.76 \pm 2.34 \text{ Mg ha}^{-1}$ for 2018. Ten-fold cross validation ($k=10$) showed the high RF performance to predict the spatial distribution of CC ($R^2 > 0.85$). Stand age determined the spatial variability of forest floor CC. The C:N ratio suggested an increase in decomposition rates between 2013 and 2018. The variable importance indicated that basal area, dominant height, tree canopy cover and elevation played an important role in estimating CC. It is concluded that an RF application offers the opportunity to predict and characterize the spatial patterns, relationships and interactions of forest floor carbon content, as it is a novel method of spatial prediction. **Keywords:** carbon stocks; forest floor litter; spatiotemporal variability; forest management; random forest.

2.3.INTRODUCCIÓN

Los principales almacenes de carbono (C) en los ecosistemas forestales son el suelo, la vegetación y el piso forestal (Schlesinger y Benhardt, 2013). El piso forestal es la capa orgánica del suelo que resulta de la diferencia entre la caída de hojarasca y su tasa de descomposición (Berg, 2018). A nivel global, este almacén representa 5% (43 Pg) de las reservas de C en ecosistemas forestales (Pan *et al.*, 2011) y es donde ocurren los procesos que regulan la circulación de materia orgánica, energía y nutrimentos entre las plantas y el suelo. Por tal motivo, es un componente dinámico fundamental en el mantenimiento de la productividad forestal y la fertilidad del suelo (Colín-Vargas *et al.*, 2018; Liu *et al.*, 2018; Lucena *et al.*, 2019).

La variación espacial y temporal es un aspecto primordial en la evaluación de la dinámica de C en el piso forestal (Guo *et al.*, 2004; Dai *et al.*, 2018). Cambios pequeños en los almacenes de C pueden tener un impacto importante sobre diversos procesos ecosistémicos (van Zijl, 2019; Wang *et al.*, 2019). Entender la variabilidad espacio-temporal del C en el piso forestal permitirá una mejor comprensión del funcionamiento del ecosistema; además, representa una herramienta potencial para predecir variaciones bajo escenarios futuros de cambio climático, definir estrategias de mitigación y desarrollar principios de manejo forestal (Birdsey *et al.*, 2013; Funes *et al.*, 2019).

De manera general, la variabilidad espacio-temporal del carbono puede ser inducida por fluctuaciones en la dinámica interna del ecosistema o por la influencia de perturbaciones naturales y humanas (Anaya *et al.*, 2007). Este tipo de interacciones genera patrones en la matriz de paisaje que pueden ser diferenciados por factores bióticos y abióticos, los cuales influyen y/o están correlacionados de forma simultánea con procesos ecológicos y/o biogeoquímicos (Fu *et al.*, 2014; Cao *et al.*, 2019). La combinación de estos factores crea condiciones locales únicas que conducen a incrementos o disminuciones en los almacenes de C del piso forestal (Carvalho Gomes *et al.*, 2019).

En México, las estimaciones de carbono en el piso forestal, por diferentes métodos siguen sin ser concluyentes. La mayoría están basadas en métodos convencionales, los cuales generalmente reportan valores promedio de acumulación de carbono, combinaciones lineales entre algunos factores ambientales o diferencias bajo condiciones particulares (Galicia *et al.*, 2015; López-Hernández *et al.*, 2018). Aunque estos métodos son muy valiosos, pueden arrojar

estimaciones con un alto grado de incertidumbre, debido a aspectos como: 1) interacciones complejas entre las variables ambientales y los almacenes de carbono, 2) representación limitada de puntos de muestreo, y 3) dificultad de elaborar mapas de variación espacial (Yao *et al.*, 2019).

Recientemente han surgido diferentes enfoques para modelar espacialmente los almacenes de carbono en el suelo y cuantificar las relaciones con diversas variables ambientales (FAO, 2018; Bhunia *et al.*, 2018). En particular, el algoritmo Random Forest (RF) es uno de los métodos de Machine Learning más utilizados en el Mapeo Digital del Suelo (MDS) (Nussbaum *et al.*, 2018). Estudios recientes indican que RF supera a otros modelos espaciales por representar relaciones no lineales complejas (Leo Breiman, 2001; Hounkpatin *et al.*, 2018; Chen *et al.*, 2019) y por su flexibilidad en el cumplimiento de supuestos estadísticos clásicos (Carvalho Gomes *et al.*, 2019; Zeraatpisheh *et al.*, 2019). No obstante, son pocas las investigaciones que, mediante RF, han integrado interacciones entre variables ambientales para modelar las variaciones espaciales y temporales del carbono en el piso forestal (Zhu *et al.*, 2017; Cao *et al.*, 2019; Zhang *et al.*, 2019).

Por lo anterior, RF ofrece la oportunidad de predecir la variabilidad espacial del carbono en el piso forestal y modelar las posibles interacciones entre los almacenes de carbono y covariables ambientales, permitiendo la exploración y comprensión de algunos procesos biogeoquímicos y ecológicos involucrados. Este estudio, realizado dentro del Sitio de Monitoreo Intensivo de Carbono (SMIC-Atopixco) en Zacualtipán, Hidalgo, es de los primeros análisis realizados en México que utiliza el algoritmo Random Forest, a partir de datos derivados de fuentes múltiples, para cuantificar la variabilidad espacial y temporal del CC en las capas que componen el piso forestal de un bosque templado bajo manejo para producción maderable. En este sentido, los objetivos fueron: 1) determinar la variabilidad espacial del contenido de carbono (CC) en el piso forestal mediante el modelo RF, 2) analizar la relación de las variables topográficas y de estructura del dosel sobre la distribución espacial del CC y 3) comparar la variación en los patrones espaciales del CC por capa del piso forestal, para 2013 y 2018.

2.4.MATERIALES Y MÉTODOS

2.4.1 Área de estudio

El Sitio de Monitoreo Intensivo de Carbono (SMIC-Atopixco) se ubica entre las coordenadas 20°40'17", 20°34'51" N y 98°40'07", 98°34'22" W en el Municipio de Zacualtipán, Hidalgo, México (**Figura 8**). EL SMIC-Atopixco cuenta con una superficie de 900 ha y forma parte de la Red MEX-SMIC (Ángeles-Pérez *et al.*, 2015).

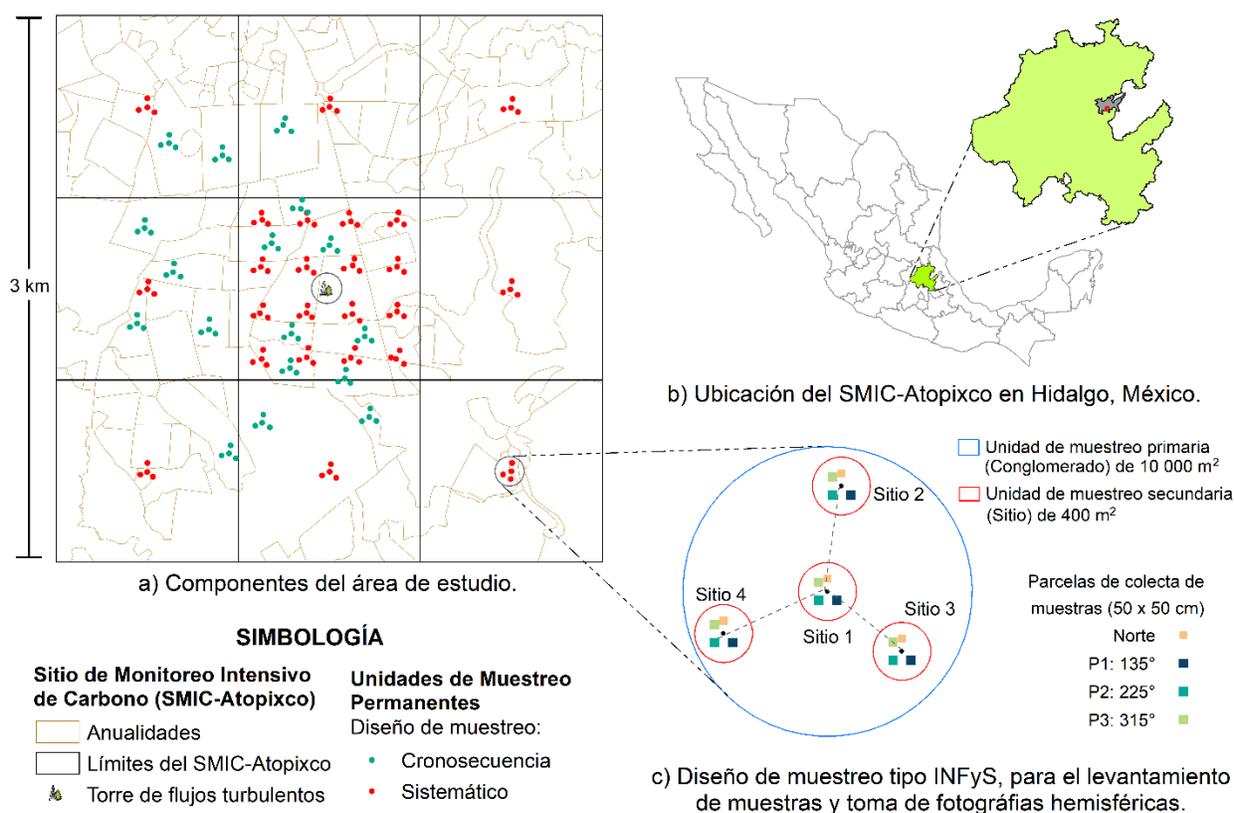


Figura 8. Componentes del Sitio de Monitoreo Intensivo de Carbono (SMIC-Atopixco): a) ubicación de las unidades de muestreo y anualidades; b) ubicación geográfica del área de estudio; c) diseño del conglomerado tipo Inventario Nacional Forestal y de Suelos (INFyS) y esquema de muestreo para la colecta de información.

El clima predominante es C(fm)w'b(e)g, templado húmedo con una marcada estación de lluvias entre junio y octubre (García, 1981). La temperatura media anual es de 13.5°C, la precipitación total anual es de 2,050 mm y presenta aproximadamente 174 días con neblinas (Godínez-Ibarra *et al.*, 2007). Fisiográficamente, el área de estudio pertenece a la Sierra Madre

Oriental, la altitud varía de 1836 a 2097 msnm y presenta desde zonas elevadas con pendientes suaves hasta terrenos con pendientes pronunciadas. Los suelos son de origen volcánico, generalmente profundos y ricos en materia orgánica. Es común encontrar luvisoles órticos y crómicos, así como cambisoles éutricos (Soriano Luna *et al.*, 2015).

El SMIC-Atopixco se encuentra en la región forestal Zacualtipán-Molango, destacada por su alta productividad y contribución contra el cambio climático. Las comunidades rurales que conforman el SMIC-Atopixco desde 1980, cuentan con un programa de manejo forestal para producción maderable, basado en los principios del Método de Desarrollo Silvícola (SARH, 1984). Bajo este método, el aclareo es una práctica silvícola recurrente dentro del SMIC-Atopixco, para producir madera con fines comerciales y para controlar la densidad y la composición arbórea (Soriano-Luna *et al.*, 2018). En las últimas décadas, el manejo del bosque en el área de estudio se ha encaminado a desarrollar bosques coetáneos dominados por *Pinus patula*. Actualmente se cuenta con una cronosecuencia de intervenciones cercana a 40 años, lo cual genera un paisaje heterogéneo derivado de esta práctica. Además, existen rodales que no han sido intervenidos desde hace aproximadamente 85 años y que representan el 30% de la superficie total forestal del área de estudio (Ángeles-Pérez *et al.*, 2015; Soriano-Luna *et al.*, 2018).

2.4.2 Diseño de muestreo

El SMIC-Atopixco sigue un esquema de muestreo propuesto por Hollinger (2008). Se encuentra delimitado dentro del paisaje forestal por un polígono de nueve km² (900 ha), dividido en 9 cuadros de 1 km² (**Figura 8a**). El cuadro central cuenta con una torre de flujos turbulentos y tiene establecidos 24 unidades de muestreo permanente con arreglo sistemático (de aquí en adelante, referidos como conglomerados): 16 en cuadrante central y uno en cada centro de los ocho km² periféricos. Adicionalmente, se establecieron 16 conglomerados bajo un enfoque de cronosecuencia (en rodales con diferente edad que no fueron representados dentro del esquema sistemático). De esta forma, el muestreo comprende un total de 40 conglomerados.

Cada conglomerado tiene una superficie de 1 ha (56.42 m de radio) y está constituido por cuatro unidades de muestreo secundarias (sitios) de 400 m², dispuestos en forma de Y invertida, similar al propuesto para el Inventario Nacional Forestal y de Suelos (INFyS) (CONAFOR, 2012).

2.4.3 Colecta de muestras

El piso forestal fue caracterizado en dos capas que representan las dos etapas de descomposición del material orgánico (de aquí en adelante, referidas como capas de descomposición): 1) la capa de mantillo (Ho), identificada como la materia orgánica recién caída, en donde se identifican todos sus componentes; y 2) capa de fermentación (F), en el que la materia orgánica está descompuesta a tal grado que no se reconocen estructuras completas (CONAFOR, 2012). Con fines de comparación, se incluyó la capa total del piso forestal (Tot), obtenida por la suma de la capa Ho y F.

La primera colecta de muestras fue dirigida por Ángeles-Pérez *et al.* (2015), de noviembre 2013 a enero 2014. El segundo periodo de muestreo fue realizado de noviembre 2018 a enero 2019. Para cada periodo se colectaron 480 muestras para cada capa del piso forestal: 40 conglomerados, 4 sitios por conglomerado, 3 muestras por sitio dispuestos a 5 m de su centro, en azimut 135° (P₁), 225° (P₂) y 315° (P₃) (**Figura 8c**).

Para su levantamiento, se colocó un marco de PVC de 50 x 50 cm y se dividió en cuatro cuadrantes, se midió la profundidad de la capa Ho en el centro de cada cuadrante y se obtuvo su profundidad promedio (mm). Se removió cuidadosamente componentes ajenos al piso forestal (piedras, insectos, material leñoso superior a 1 cm de diámetro, etc.) y se registró el peso fresco (g), de toda la muestra y de una submuestra homogénea y representativa del sitio para su análisis en laboratorio. El procedimiento fue repetido para la capa F.

2.4.4 Análisis de laboratorio

Las submuestras se secaron en una estufa de circulación forzada a 70°C , hasta peso constante. Se registró el peso seco de la submuestra y se determinó la fracción de peso seco, utilizada para estimar la biomasa acumulada por cada capa del piso forestal. Posteriormente, las muestras para cada sitio se mezclaron y molieron para obtener una muestra compuesta de 10 gramos, representativa del conglomerado. Las muestras compuestas fueron analizadas en el Laboratorio de Fertilidad de Suelos y Química Ambiental del Colegio de Postgraduados, para determinar su concentración de C y N.

La concentración de carbono se obtuvo por digestión seca a 900°C, en el determinador automático de C (TOC SSM 5050A Shimadzu). La concentración de nitrógeno se determinó por digestión sulfúrica y fue cuantificada por el método Kjeldahl. Con los valores de concentración se calculó para cada muestra, la relación C:N, el contenido de carbono (CC) y el contenido de nitrógeno (CN) (Ecuación 6, 7) (López-Merlín *et al.*, 2016):

$$CC = PS_m * Conc (C) * 0.01 \quad (6)$$

$$CN = PS_m * Conc (N) * 0.01 \quad (7)$$

Donde CC = Contenido de C (g m⁻²), CN = Contenido de N (g m⁻²), Conc (C) = concentración de C (%), Conc (N) = concentración de N (%) y PS_m = peso seco de la muestra (g). Finalmente, el CC y el CN fueron convertidos a Mg ha⁻¹.

2.4.5 Variables predictoras

Se utilizó la información de doce variables ambientales predictoras obtenidas para los años 2013 y 2018. Estas variables fueron clasificadas en: topográficas, de estructura espacial y de estructura de dosel (**Cuadro 8**).

Cuadro 8. Descripción de las variables predictoras utilizadas para modelar el contenido de carbono en el piso forestal.

Categoría	Variable predictora	Resolución espacial	Fuente	Referencia	Unidades
Topografía	Elevación	1 m	LiDAR	Soriano-Luna <i>et al.</i> , 2018	msnm
	Pendiente	1 m	LiDAR	-	grados
	Exposición	1 m	LiDAR	-	grados
Estructura espacial	Coordenadas X	1 m	LiDAR	-	m
	Coordenadas Y	1 m	LiDAR	-	m
Estructura del dosel	Edad del rodal	1 m	Archivo vectorial	Soriano-Luna <i>et al.</i> , 2015	Años
	Altura dominante	5 m	Inventario/LiDAR	Soriano-Luna <i>et al.</i> , 2018	m
	Área basal	Sitio	Inventario	-	m ² ha ⁻¹
	Fracción de apertura del dosel	Sitio	Fotografías hemisféricas	-	%
	Riqueza de especies	Sitio	Inventario	Villegas-Macedo, 2019	árboles sitio ⁻¹
	Índice de Shannon-Wiener	1 m	Archivo raster	Villegas-Macedo, 2019	-
	Cobertura arbórea	5 m	LiDAR	Soriano-Luna <i>et al.</i> , 2018	%

*Sitio: unidad de muestreo secundaria con una superficie de 400 m²;

La información fue transformada a formato raster, proyectada en la misma referencia geográfica (WGS84 UTM Zona 14) y remuestreada a la misma resolución espacial (1 m) mediante el software ArcGIS 10.5. (Environmental System Research Institute, ESRI Inc., Redlands, CA).

a) *Topografía y estructura espacial*: Los datos de elevación, pendiente y exposición fueron derivados de un modelo de elevación digital (MED), con resolución espacial de 1 m. El MED fue generado por Soriano-Luna *et al.* (2018) para toda el área de estudio. La estructura espacial fue representada mediante las coordenadas geográficas X, Y.

b) *Estructura del dosel*: Las variables utilizadas para caracterizar la estructura del dosel fueron: edad del rodal, altura dominante, área basal, riqueza de especies, Índice de Shannon-Wiener, cobertura arbórea y fracción de apertura del dosel. La información fue extraída a partir de métricas LiDAR (Soriano-Luna *et al.*, 2018) y de dos inventarios dasométricos realizados en 2013 y 2018, correspondientes al mismo periodo de colecta de muestras.

Los datos de apertura del dosel se obtuvieron mediante fotografías hemisféricas, las cuales fueron capturadas con cámara digital Nikon Coolpix 4500, en combinación con un lente circular “FishEye” (Frazer *et al.*, 2001). La clasificación binaria de píxeles entre el dosel y el cielo se obtuvo mediante un algoritmo de umbral automático basado en la detección de bordes, propuesto por Nobis y Hunziker (2005). Las imágenes fueron procesadas en los programas de cómputo SideLook (SL) (Nobis, 2005) y GapLight Analyzer (GLA) (Frazer, 1999).

Los datos de riqueza de especies (S) y la estimación espacialmente explícita del Índice de Shannon-Wiener fue obtenida por Villegas-Macedo (2019) y convertidos a la misma resolución espacial. El área basal por sitio (AB_s) fue estimada mediante la Ecuación 8 (Cruz-Leyva *et al.*, 2010):

$$AB_s = \sum_{i=1}^n \left(\frac{\pi}{40000} \right) * DN^2 \quad (8)$$

donde AB_s = área basal por sitio ($m^2 ha^{-1}$), DN = diámetro normal (cm) y n = número de árboles por sitio. La edad del rodal se definió como el tiempo transcurrido a partir del año en que se aplicó la corta de repoblación, la información de esta variable fue extraída de polígonos en formato vectorial delimitados por Soriano Luna *et al.* (2015).

2.4.6 Análisis estadístico

El análisis estadístico fue realizado en cuatro etapas: a) análisis descriptivo de los datos, b) generación y validación del modelo de predicción del CC en el piso forestal, c) análisis de la distribución espacio temporal y d) obtención de la importancia de las variables. Todos los análisis fueron realizados en R 3.5.1. (R Development Core Team, 2018).

a) Análisis descriptivo: De forma preliminar, se realizó un análisis de calidad de los datos y el análisis descriptivo de las propiedades del piso forestal (contenido de carbono (CC), concentración de carbono (Conc. C), profundidad promedio de la capa (Prof.), biomasa y relación C: N). Se calculó la media aritmética, la desviación estándar (SD), mínimo (min), máximo (max), coeficiente de variación (CV), coeficiente de curtosis (kur) y coeficiente de asimetría (skew). La normalidad de cada variable fue comprobada mediante la prueba de Kolmogórov-Smirnov (KSp) ($\alpha > 0.05$) (Brown y Forsythe, 1974). La diferencia entre capas de descomposición y años evaluados fue probada estadísticamente con un análisis de varianza de dos factores (Two-way ANOVA), tanto para CC, como para la relación C: N.

b) Modelo de predicción espacial: Para cada año evaluado, la distribución espacial del CC en cada una de las capas del piso forestal fue modelada mediante el algoritmo Random Forest (RF) (Breiman, 2001; Biau, 2010; Biau y Scornet, 2016). Se utilizó el paquete de R “*RandomForest*” (Liaw y Wiener, 2002) para generar un modelo por año evaluado. El mejor desempeño predictivo fue obtenido al evaluar y optimizar dos parámetros de ajuste (*ntree* y *mtry*), mediante el entrenamiento de cada modelo (**Anexo, Cuadro S3**). Asimismo, se agregó como variable factorial, la capa del piso forestal a fin de obtener las estimaciones para cada una de las capas evaluadas. Las predicciones espaciales de RF y sus respectivos coeficientes de variación se obtuvieron mediante el paquete “*ModelMap*” (Freeman *et al.*, 2018).

Se estimó la importancia de las variables predictoras (Vimp), entendida como el aumento del error de predicción del modelo por cada variable ausente, cada predictor fue permutado aleatoriamente mientras el resto de los predictores permanecían inalterables (Breiman, 2002; Liaw y Wiener, 2002).

c) Validación del modelo: el desempeño del modelo en la predicción fue evaluado a través del método de validación cruzada (k=10). La validación cruzada proporciona una estructura para

crear varias divisiones aleatorias de entrenamiento/prueba en el conjunto de datos, para garantizar que cada conjunto de datos se encuentre en el conjunto de prueba al menos una vez (Zeraatpisheh *et al.*, 2019). Por lo cual, del conjunto total de datos, 90% fueron aleatoriamente seleccionados para entrenamiento y 10% para validación del modelo. Los criterios de error utilizados para evaluar el desempeño del modelo fueron: a) el coeficiente de determinación (R^2) (Ecuación 9) y b) el error cuadrático medio (RMSE) (Ecuación 10).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (9)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (10)$$

Donde, \hat{y}_i es el *i*ésimo valor obtenido de la predicción, y_i es el *i*ésimo valor observado, \bar{y}_i es el valor medio de y_i , y n es el número de valores predichos u observados con $i = 1, 2, \dots, n$. La estrategia de validación para RF fue repetida tres veces para obtener un R^2 y RMSE promedio.

2.5.RESULTADOS

2.5.1 Contenido de carbono en el piso forestal

La estadística descriptiva de las propiedades del piso forestal (**Cuadro 9**) y las gráficas de cajas y bigotes (**Figura 9**) mostraron simetría en la distribución de los datos. El análisis de varianza de dos vías reveló que existen diferencias significativas ($\alpha < 0.05$) del CC por capa del piso forestal ($F = 293.9, \alpha < 2e^{-16}$) y año evaluado ($F = 305.08, \alpha < 2e^{-16}$), consistente con los coeficientes de variación obtenidos.

El valor medio total del CC en el piso forestal fue de $12.19 \pm 5.48 \text{ Mg ha}^{-1}$ para 2013, y de $5.76 \pm 2.34 \text{ Mg ha}^{-1}$ para 2018. Para ambos años, el carbono en el piso forestal fue mayor en la capa F, lo cual sugiere un efecto del grado de descomposición del material orgánico. La estimación del CC para 2013 y 2018 reveló una moderada variabilidad; sin embargo, la variabilidad (CV) de las estimaciones en cada capa se mantuvo. Para la capa Ho, se obtuvo un CV de 36.36% y 36.85%

en 2013 y 2018 respectivamente; para la capa F el CV fue 52.53% en 2013 y 50.45% en 2018. En todas las capas y en ambos años evaluados, el CC presentó una distribución normal.

Cuadro 9. Estadística descriptiva del contenido de carbono y las propiedades del piso forestal, estratificado por año y capa de del piso forestal.

Variable	Año	Capa	Media	Mediana	Min	Max	DS	CV	Kur	Skew	KS _p
CC (Mg ha ⁻¹)	2013	Ho	2.03	1.95	0.00	4.72	0.74	36.36	4.75	0.58	0.34*
		F	10.16	9.37	0.00	26.33	5.34	52.53	3.27	0.54	0.61*
		Tot	12.19	11.47	0.77	27.75	5.48	44.98	3.18	0.48	0.54*
	2018	Ho	1.85	1.78	0.28	5.23	0.68	36.85	6.63	1.07	0.33*
		F	3.91	3.83	0.00	8.81	1.97	50.45	2.55	0.08	0.99*
		Tot	5.76	5.65	0.44	11.80	2.34	40.54	2.65	-0.06	0.99*
Conc. C (%)	2013	Ho	45.71	45.90	0.00	47.70	1.07	2.34	5.56	-1.08	0.01
		F	47.19	47.00	0.00	49.00	0.75	1.58	2.52	0.35	0.03
		Tot	46.45	46.45	44.40	47.50	0.62	1.33	5.40	-1.30	0.01
	2018	Ho	51.43	51.00	50.00	58.00	1.14	2.21	22.49	3.65	0.001
		F	47.17	50.00	0.00	52.00	10.5	22.33	18.89	-4.18	0.002
		Tot	50.46	50.50	48.00	54.50	1.01	2.01	6.69	0.77	0.003
Biomasa (Mg ha ⁻¹)	2013	Ho	4.43	4.25	0.00	10.20	1.60	36.06	4.56	0.52	0.34*
		F	21.51	19.81	0.00	56.88	11.29	52.46	3.37	0.55	0.53*
		Tot	25.95	24.41	1.69	59.83	11.61	44.73	3.23	0.48	0.51*
	2018	Ho	3.60	3.46	0.56	10.06	1.31	36.42	6.57	1.03	0.42*
		F	7.92	7.77	0.00	17.98	3.99	50.41	2.58	0.09	0.97*
		Tot	11.52	11.28	0.75	23.96	4.69	40.72	2.69	-0.04	0.96*
Prof. (m)	2013	Ho	0.04	0.04	0.00	0.09	0.02	45.92	2.80	0.05	0.88*
		F	0.04	0.04	0.00	0.09	0.02	44.96	2.87	0.03	0.90*
		Tot	0.08	0.08	0.00	0.18	0.03	45.26	2.87	0.04	0.88*
	2018	Ho	0.03	0.03	0.01	0.06	0.01	32.82	3.31	-0.01	0.55*
		F	0.03	0.03	0.00	0.06	0.01	45.99	2.80	-0.19	0.72*
		Tot	0.06	0.06	0.01	0.12	0.02	34.34	3.11	-0.35	0.59*
Relación C: N	2013	Ho	37.35	37.32	0.00	47.39	5.57	14.92	15.27	-1.95	0.04
		F	54.98	56.14	0.00	86.85	13.82	25.13	9.38	-1.69	0.01
		Tot	51.19	51.07	34.32	79.57	7.57	14.79	4.40	0.66	0.44*
	2018	Ho	45.69	46.36	36.43	65.00	5.55	12.14	5.10	1.00	0.001
		F	34.52	35.38	0.00	46.36	8.36	24.21	13.90	-3.09	0.001
		Tot	39.22	38.19	31.62	65.00	4.33	11.05	13.29	2.52	0.001

CC: Contenido de carbono (Mg ha⁻¹); Conc. C: Concentración de carbono (%); Prof.: profundidad promedio (mm); Ho: hojarasca, F: fermentación, Tot: suma de Ho y F; Min: mínimo, Max: máximo, DS: desviación estándar; CV: coeficiente de variación (%); Kur: curtosis; Skew: asimetría y KS_p: Prueba de Kolmogórov-Smirnov (*nivel de significancia p > 0.05).

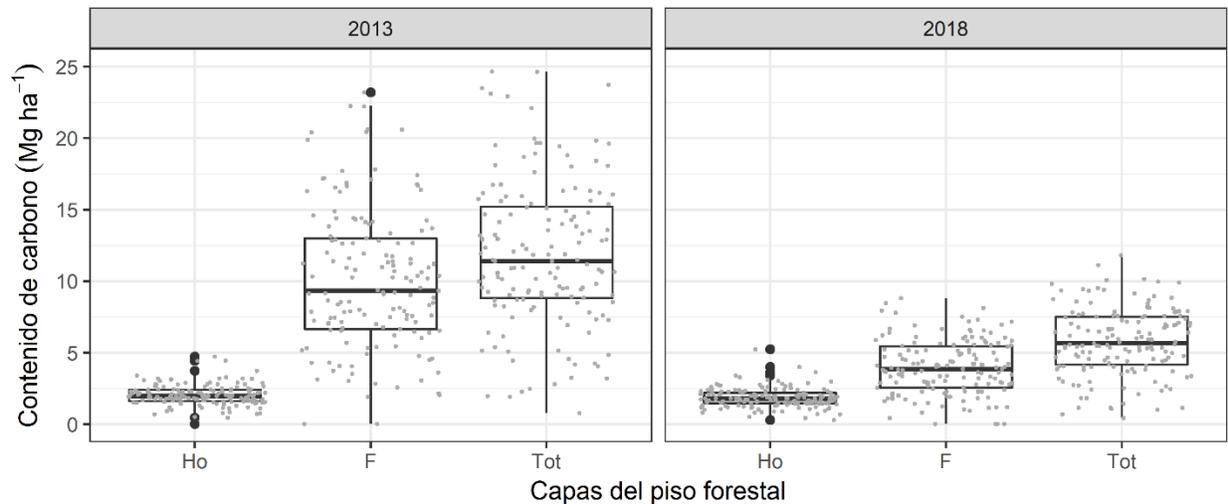


Figura 9. Contenido de Carbono para cada capa del piso forestal para 2013 y 2018.

La relación C: N mostró un comportamiento inverso entre 2013 y 2018 (**Figura 10**). El análisis de varianza de dos vías permitió detectar diferencias significativas entre las capas del piso forestal. ($F = 11.83, \alpha < 8.5e^{-06}$) y año evaluado ($F = 141.89, \alpha < 2e^{-16}$) para este parámetro. En 2013, se obtuvo una mayor relación C: N en la capa F con respecto a la capa Ho, mientras que en 2018 el comportamiento se invirtió. Estas diferencias pueden sugerir un cambio en la velocidad de descomposición en las capas del piso forestal, entre 2013 y 2018.

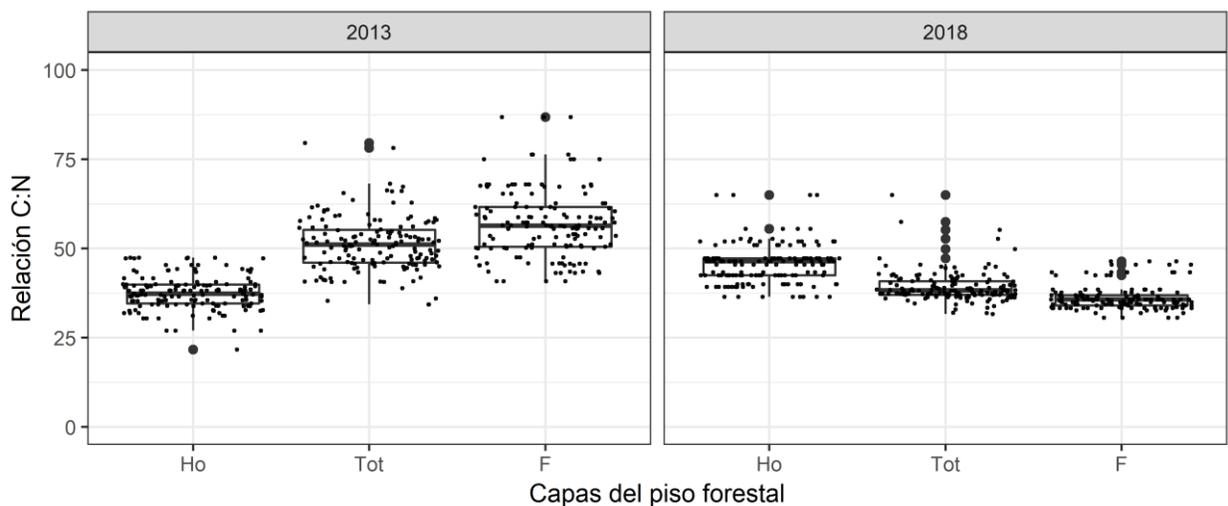


Figura 10. Relación C: N por capa del piso forestal, para 2013 y 2018.

2.5.2 Desempeño del modelo e importancia de variables

En la validación del modelo, RF explicó el 85.81% de la varianza con un RMSE de 2.62 Mg ha⁻¹ para 2013. En 2018, RF obtuvo 69% de la varianza explicada y un RMSE de 1.51 Mg ha⁻¹. Estos resultados indicaron que el ajuste fue significativo y la precisión de RF fue alta para ambos años evaluados, con un R² superior a 0.85 (Anexo, Cuadro S3, Figura S10).

La importancia de variables (Vimp) caracteriza la influencia de las diferentes variables predictoras sobre la variable de respuesta (Figura 11). Los valores obtenidos de Vimp fueron iguales para 2013 y 2018. Los resultados indicaron que evaluar el CC por capas del piso forestal contribuyó significativamente en el ajuste de RF, por lo que tuvo el efecto más importante en las estimaciones, con un 42.93% del total de la importancia relativa. La segunda variable más importante para estimar el CC fue el área basal, con un 13.14%. La cobertura arbórea, la altura dominante, la edad del rodal y la elevación también desempeñaron un papel importante en la distribución espacial de CC, aportando 6.93, 5.85, 5.31 y 4.78%, respectivamente. El resto de los factores predictores mostraron una aportación a la predicción del modelo menor al 2.5%. m² ha⁻¹

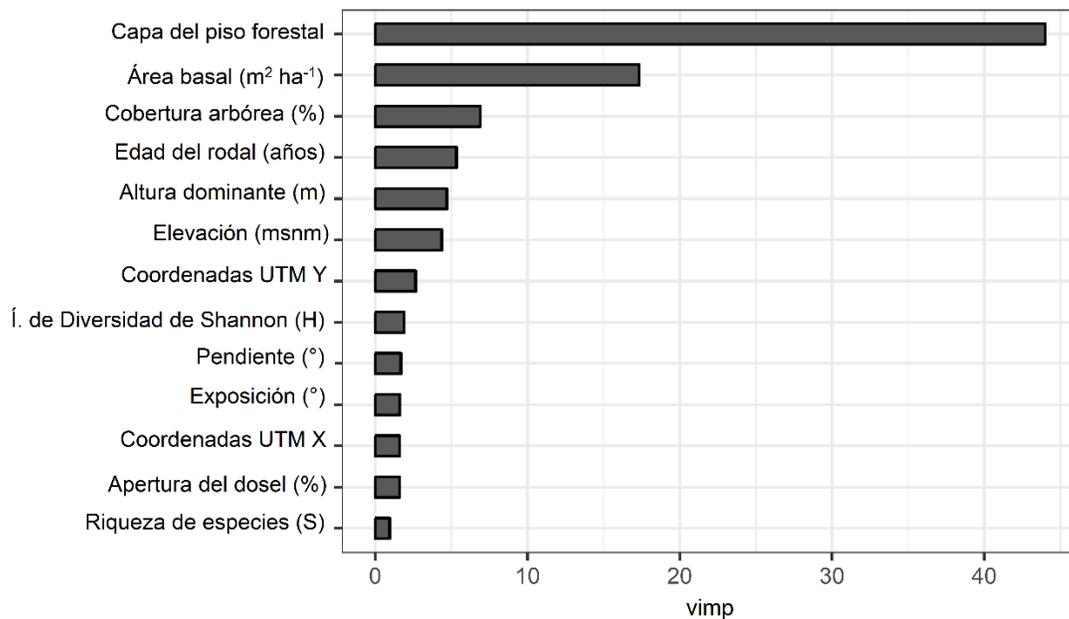


Figura 11. Importancia relativa de variables predictoras (%) de Random Forest para predicciones del contenido de carbono en el piso forestal.

2.5.3 Variación espacial del CC en el piso forestal

Los mapas de distribución espacial del CC en cada una de las capas de descomposición para 2013 y 2018, así como sus respectivos coeficientes de variación (CV) fueron desplegados a una resolución de 1 m y reclasificados en intervalos geométricos, para facilitar las comparaciones. El CC exhibió una alta variación espacial dentro del SMIC-Atopixco y fue determinada principalmente por la edad del rodal. Para ambos años, se observó una diferencia notoria entre los patrones de distribución espacial del CC en ambas capas del piso forestal (Ho y F).

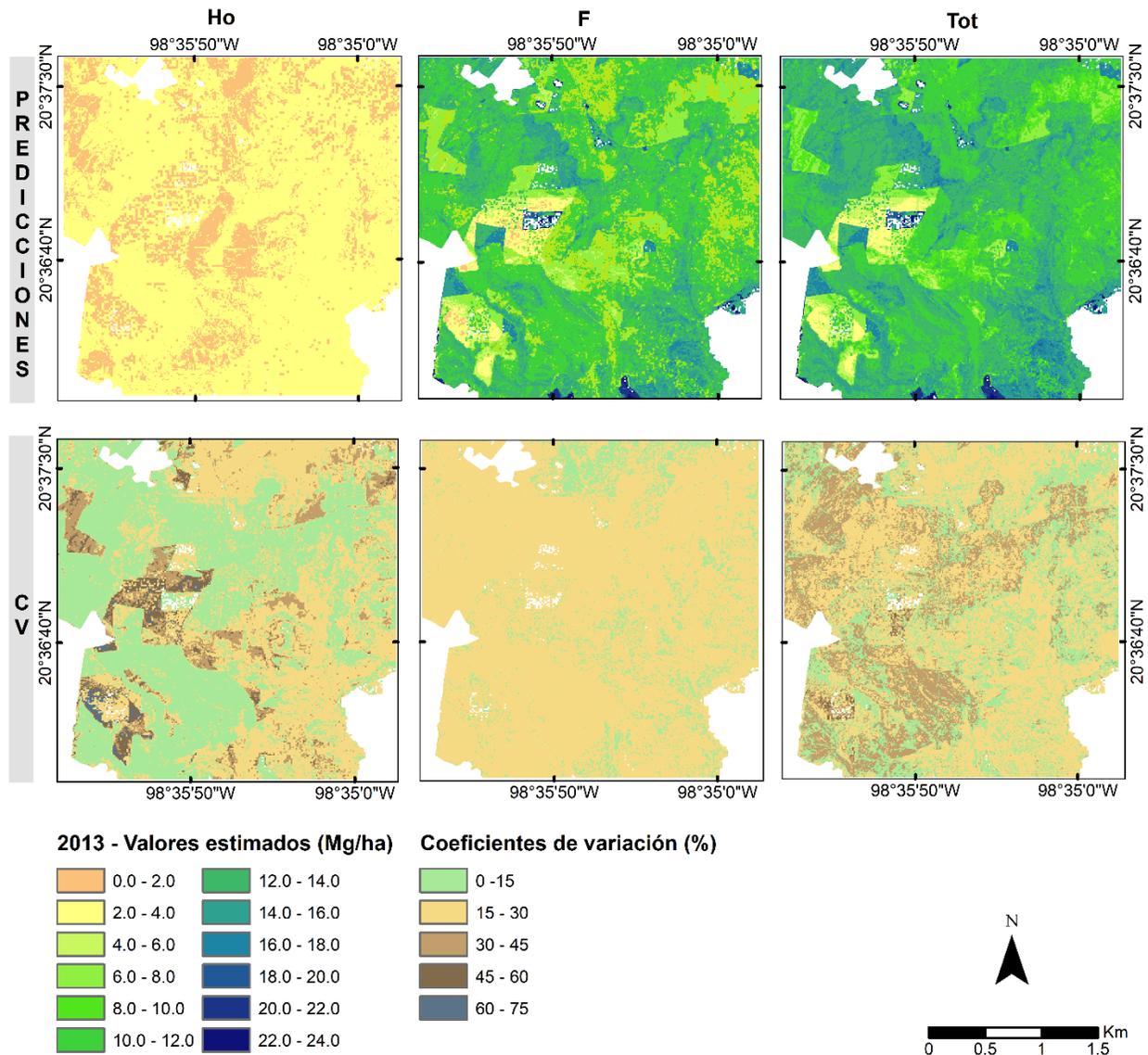


Figura 12. Patrones espaciales del contenido de carbono por capa de descomposición en el piso forestal y su coeficiente de variación a 1 m de resolución, en el Sitio de Monitoreo Intensivo de Carbono-Atopixco en el año 2013.

En 2013, el CC en la capa Ho osciló mayoritariamente entre 2 a 4 Mg ha⁻¹. En la capa F, el CC predominante varió entre 6 a 12 Mg ha⁻¹. Esta capa presentó mayor heterogeneidad espacial en comparación con la capa Ho (**Figura 12**). Los rodales con edades inferiores a 5 años y superiores a 20 años presentaron los mayores CC. En contraste, las regiones con menor CC se localizaron en rodales con edades entre 5 a 10 años. Respecto al coeficiente de variación, en la capa Ho predominaron valores entre 0 y 15%, los mayores valores de error fueron ubicados en la zona suroeste del SMIC-Atopixco. La capa F mostró CV más homogéneos, con valores predominantes entre 15 a 30%.

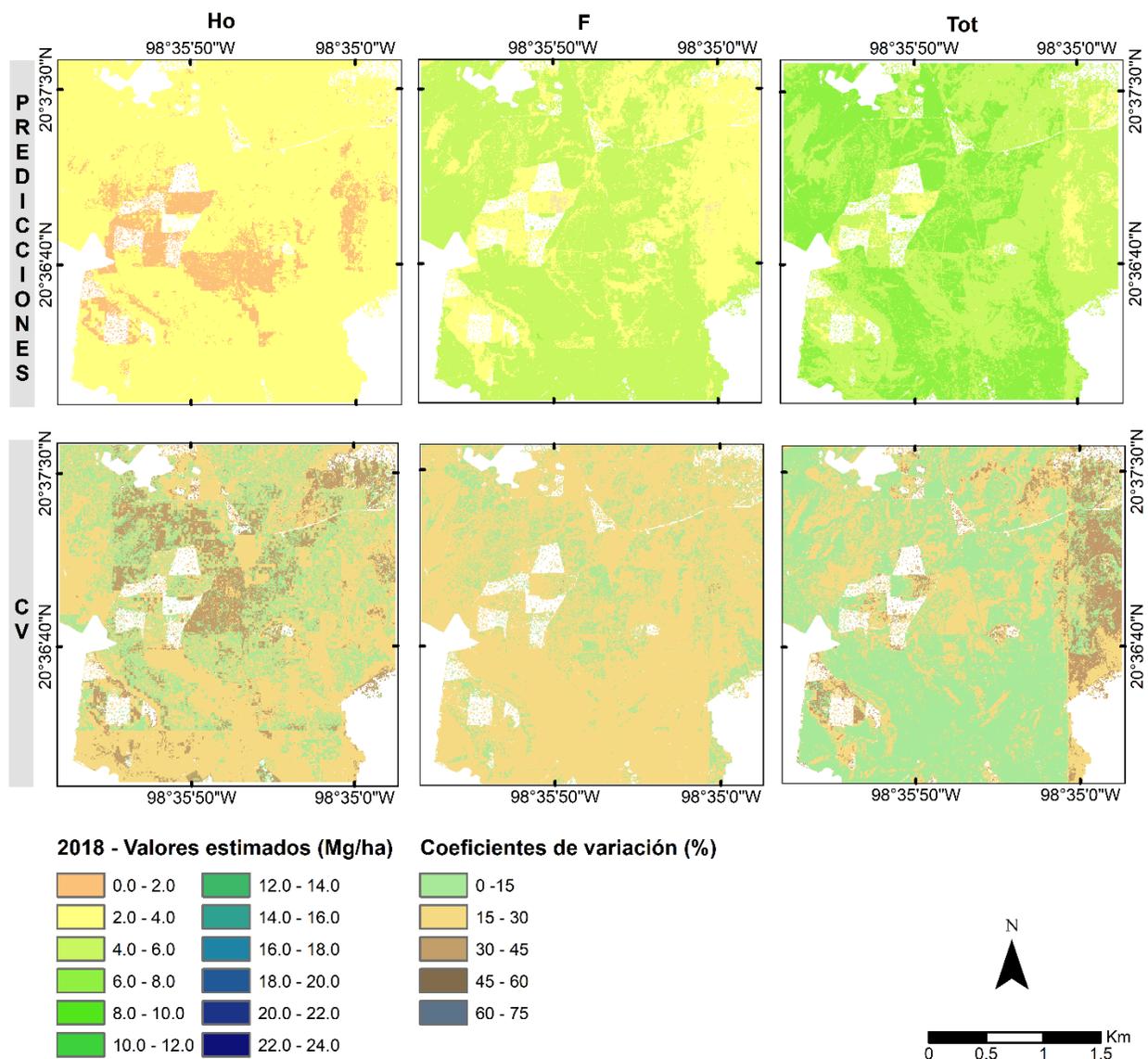


Figura 13. Patrones espaciales del contenido de carbono por capa de descomposición en el piso forestal y su coeficiente de variación a 1 m de resolución, en el Sitio de Monitoreo Intensivo de Carbono-Atopixco en el año 2018.

En 2018, el CC de la capa Ho también mostró una variabilidad espacial similar a 2013, con el mismo rango de valores predominantes. La capa F presentó una mayor homogeneidad espacial del CC, con valores entre 4 a 6 Mg ha⁻¹ (Figura 13). En cuanto al CV, tanto para la capa Ho como para la capa F, predominaron valores entre 15 a 30% y entre 0 a 15% para la capa Tot.

2.5.4 Interacciones entre variables y el CC

La capa de fermentación muestra las curvas de cambio entre cada uno de los predictores y el CC (Figura 14). En vista de que la capa Ho presentó CC mucho más bajo que la capa F, las tendencias no pueden ser gráficamente representadas; sin embargo, su variabilidad es absorbida y representada por la capa Tot.

Las áreas basales entre 30 a 50 m² ha⁻¹ presentaron los mayores almacenes de carbono y a partir de valores inferiores a 25 m² ha⁻¹, el CC disminuyó rápidamente hasta alcanzar valores cercanos a 5 Mg ha⁻¹ (Figura 14a). La altura dominante (Figura 14b) y la edad del rodal (Figura 14c) son variables estrechamente ligadas entre sí, en este caso, la altura dominante mostró una ligera tendencia a incrementar el CC a partir de una altura superior a 25 m, mientras que la interacción entre el CC y la edad del rodal fue mucho más marcada.

La pendiente de la curva mostró la velocidad en que el CC cambió a medida que incrementó la edad del rodal. Los rodales entre 0 y 5 años presentaron un alto CC en el piso forestal, esta tendencia disminuyó rápidamente hasta alcanzar un periodo de estabilización entre 10 a 20 años. A partir de ese periodo, el CC incrementó ligeramente a medida que el bosque alcanzaba un mayor estado sucesional. Con respecto a la elevación (Figura 14d), las regiones con menor CC fueron localizadas entre 2000 a 2050 msnm; valores inferiores y superiores a este rango presentaron mayor CC en el piso forestal.

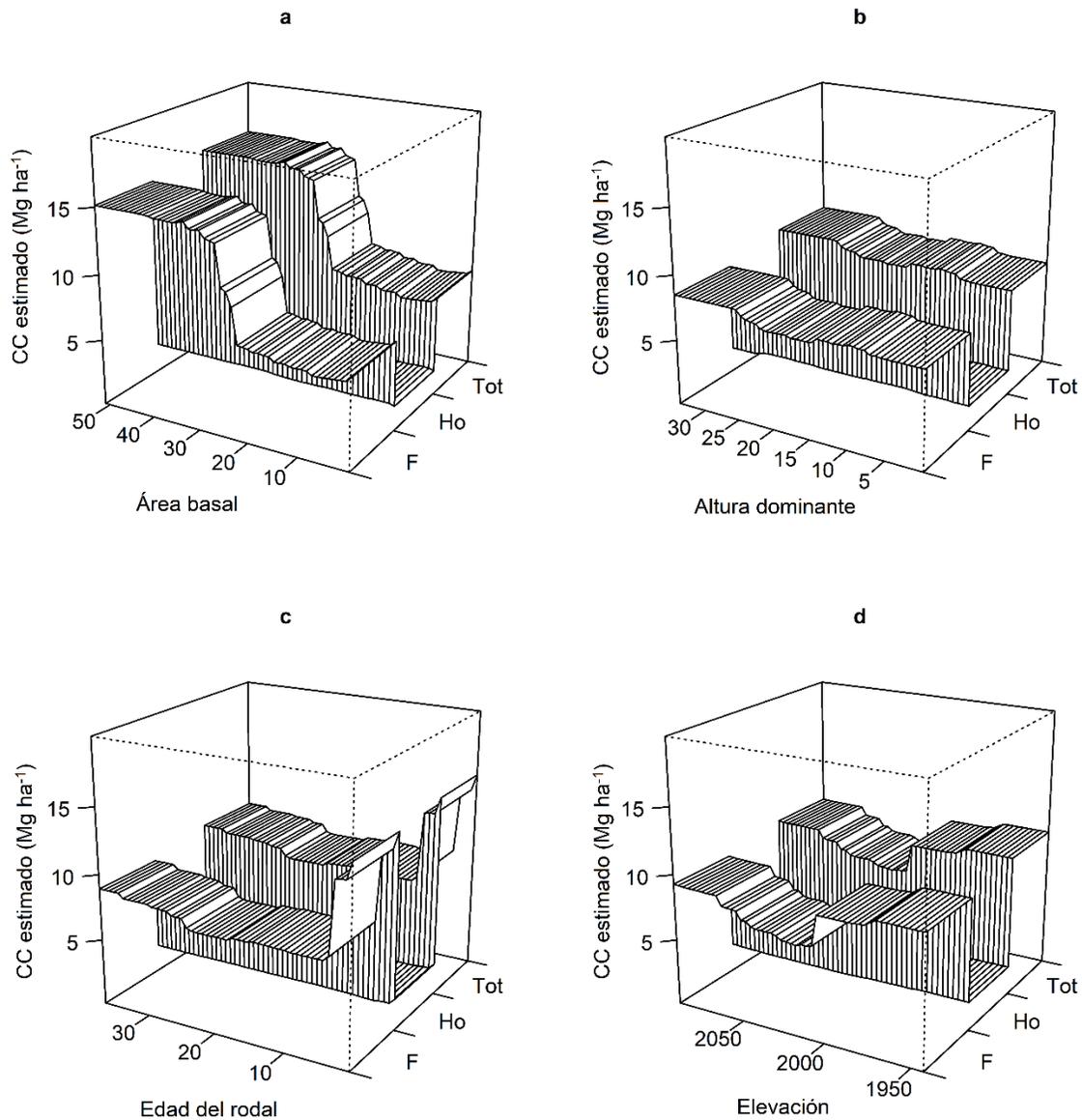


Figura 14. Interacción entre el contenido de carbono (CC) por capa del piso forestal (Ho: hojarasca, F: fermentación; Tot: Suma de Ho y F) y: a) área basal ($m^2 ha^{-1}$), b) altura dominante (m), c) edad del rodal (años) y d) elevación (msnm).

2.6.DISCUSIÓN

2.6.1 Dinámica de carbono en el piso forestal

El CC exhibió notables diferencias en los patrones espaciales, debido a la dependencia y variabilidad espacial, así como a su dependencia con variables biofísicas y perturbaciones

humanas. Las estimaciones promedio del contenido de carbono en la capa Total (que representa la suma de las capas Ho y F), fueron significativamente diferentes para los años 2013 y 2018, con $12.19 \pm 5.48 \text{ Mg ha}^{-1}$ y $5.76 \pm 2.34 \text{ Mg ha}^{-1}$, respectivamente.

En los bosques templados se ha estimado que el piso forestal almacena entre 3.6 y 11.5 Mg ha^{-1} de carbono, sin embargo, con el aprovechamiento de estos bosques se llega a remover hasta el 50% de biomasa (Meier y Leuschner, 2010). Los resultados obtenidos son comparables a estas cifras, así como a los reportados por Galicia *et al.* (2015), quienes obtuvieron valores entre 8.7 y 17.9 Mg ha^{-1} de carbono en el piso forestal de una cronosecuencia de bosques templados sujetos a manejo forestal, localizados en la Sierra Norte de Oaxaca. En comparación, Acosta Mireles *et al.* (2008), reportaron valores promedio de CC de 17.91 Mg ha^{-1} en el piso forestal de rodales dominados por *Pinus patula*, en bosques templados sujetos a manejo forestal ubicados en Tlaxco, Tlaxcala.

2.6.2 Variaciones temporales del CC en el piso forestal

La producción de hojarasca y la posterior descomposición del mantillo son procesos ecológicos complejos que están influenciados por factores ambientales (Xu *et al.*, 2012; Djukic *et al.*, 2018). El CC en el piso forestal está sujeto al balance de estos procesos, los cuales a su vez, muestran variaciones temporales que son regulados por las características del dosel, las condiciones climáticas y las alteraciones derivadas de prácticas silvícolas (Yanai *et al.*, 2003). En este caso, a pesar de que las estimaciones del CC para ambos años fueron realizadas durante el mismo periodo estacional, se registró una disminución en el valor medio total del CC de 2013 a 2018. Asimismo, la relación C: N también presentó variaciones contrastantes entre años.

La relación C: N se ha considerado un índice de la calidad de los residuos vegetales, que permite en cierto grado predecir el proceso de descomposición, en función de la disponibilidad de N (Wu *et al.*, 2014; Djukic *et al.*, 2018). De manera general, la materia orgánica con una alta relación C:N tiende a descomponerse de forma más lenta, al tener condiciones limitantes de compuestos nitrogenados (Averill y Waring, 2018) o estar asociadas a compuestos de lenta degradación, como lignina y taninos (Rinkes *et al.*, 2014).

Para 2013, la menor relación C: N registrada en la capa Ho sugiere potencialmente una mayor tasa de descomposición, en comparación con la capa F, que incorpora de forma más lenta el carbono a otros almacenes y consecuentemente, almacena un mayor CC, lo cual es un patrón general en los sistemas forestales (Urbanov *et al.*, 2015; Berg, 2018). Para 2018, el comportamiento contrario observado en la relación C: N en ambas capas, puede estar influido por una elevada tasa de descomposición con respecto a 2013 (Cotrufo *et al.*, 2015). Este tipo de tendencias también fue observada por Blanco *et al.* (2011), quienes muestran la sensibilidad de la composición química a intervenciones silvícolas, sobre todo, en términos de inmovilización de nutrientes y cambios en la velocidad de descomposición del material orgánico.

No obstante, para probar esta hipótesis se requieren nuevas estimaciones de flujos de carbono y dinámica de nutrientes que puedan contrastarse con estas evaluaciones. Asimismo, es importante ejercer precaución al comparar la pérdida de C y la relación C:N, pues las relaciones pueden ser altamente variables y estarán también sujetas a aspectos como la composición química del material foliar, la fase de descomposición, el grado de recalcitrancia, las condiciones locales del área de estudio, así como las interacciones entre nutrientes y microorganismos (Meier y Leuschner, 2010).

2.6.3 Efectos de la estructura del dosel sobre el CC

Los almacenes de carbono en el piso forestal de sistemas sin manejo son principalmente influenciados por factores climáticos y topográficos (He *et al.*, 2009; Blanco *et al.*, 2011). Sin embargo, en sistemas sujetos a manejo forestal, el carbono en las capas orgánicas del suelo es, a escalas locales, altamente sensible a modificaciones en la estructura del dosel (Nave *et al.*, 2010). Bajo el esquema de este estudio, RF definió al área basal, la cobertura arbórea, la altura dominante y la edad del rodal como los predictores bióticos más importantes en la modelación espacial del CC en el piso forestal. Estas variables, a su vez, incluyen indirectamente, el efecto del manejo forestal.

Las prácticas silvícolas modifican la estructura del dosel y pueden alterar significativamente el almacenamiento del carbono en el piso forestal por diferentes mecanismos (Bravo-Oviedo *et al.*, 2015; Lado-Monserrat *et al.*, 2016). Las actividades silvícolas realizadas en el SMIC-Atopixco frecuentemente involucran reducciones en la densidad arbórea, con el objetivo de disminuir la

competencia y mejorar el crecimiento de los árboles remanentes (Ángeles-Pérez *et al.*, 2015; Chávez-Aguilar *et al.*, 2016; Soriano-Luna *et al.*, 2018). Este tipo de alteraciones genera cambios en las tasas de producción de hojarasca y modifica las condiciones microambientales del suelo, lo que repercute significativamente en la acumulación de carbono en el piso forestal.

Las interacciones entre el área basal y la altura dominante se encuentran estrechamente relacionados con la cobertura arbórea y la edad del rodal (Barlow *et al.*, 2007; Penne *et al.*, 2010; Rupšys y Petruskas, 2019). La cobertura arbórea regula la cantidad de luz penetrante y controla algunos procesos ecosistémicos fotodependientes (Aguirre-Salado *et al.*, 2011). Autores como Bueis *et al.* (2018) y Blanco *et al.* (2011) reportan un efecto significativo del área basal y la cobertura arbórea sobre el proceso de descomposición, como una respuesta a la alteración de las condiciones microambientales piso forestal. Los cambios en estas variables permiten observar ciertos efectos (p. e. cambios en la calidad química del sustrato, alteraciones en la temperatura y humedad del suelo, modificaciones en la incidencia de luz, mayores entradas de residuos vegetales al sistema) que pueden alterar significativamente las tasas de descomposición del material orgánico acumulado en el piso forestal (Giunta y Shaw, 2018).

Con respecto a la edad del rodal, un meta análisis realizado por Nave *et al.* (2010) para bosques templados con intervenciones silvícolas, se observó que el carbono en el piso forestal tiende a disminuir drásticamente en los cinco años posteriores a un aprovechamiento, lo cual resulta similar a las tendencias obtenidas en este estudio y que pueden apreciarse bajo una cronosecuencia. En este caso, la edad del rodal y el CC en el piso forestal mostraron un patrón con tendencia en forma de “U” (**Figura 14c; Anexo, Figura S11**). Esta tendencia permite suponer que después de un aprovechamiento, una gran parte de los residuos de la biomasa aérea son depositados en el piso forestal, y a medida que el proceso de descomposición incorpora el carbono en otros almacenes, el CC tiende a disminuir (Zhu *et al.*, 2017). La rápida disminución de carbono en los primeros cinco años puede fundamentarse en dos aspectos: 1) modificaciones en las condiciones micro climáticas generadas por la apertura del dosel, las cuales pueden acelerar el proceso de descomposición, y 2) la producción de hojarasca y otros residuos vegetales es relativamente menor en rodales jóvenes, pues se encuentran en una etapa activa de crecimiento con mayor absorción de carbono y nutrientes del suelo (Grier *et al.*, 1989; He *et al.*, 2009).

2.6.4 Efectos de la topografía sobre el CC

La topografía representa el resultado de la historia geológica y climática de una región (Grier *et al.*, 1989). Los gradientes de latitud y altitud han sido ampliamente utilizados para dilucidar algunos mecanismos involucrados en la acumulación de carbono en el piso forestal (Griffiths *et al.*, 2010; Jeyanny *et al.*, 2016). En este estudio, la elevación fue la única variable topográfica con importancia para explicar la variabilidad espacial del CC en el piso forestal. Asimismo, no se encontraron interacciones notorias entre el CC y la pendiente, la exposición o la estructura espacial, lo que se atribuye a la menor variabilidad de estos predictores dentro del área de estudio.

Es probable que la interacción entre el CC y la elevación no sea estrictamente directa, y en su lugar, refleje el efecto de otras variables que no fueron analizadas en este estudio, como la temperatura, el viento y la precipitación (Powers y Schlesinger, 2002; Xu *et al.*, 2018). En el SMIC-Atopixco, las zonas con mayor altitud tienden a ser ligeramente más frías y húmedas, lo que podría explicar una disminución en la velocidad de descomposición del material orgánico, y consecuentemente, reflejar una mayor acumulación de carbono en el piso forestal. En altitudes bajas, los mayores almacenes de carbono pueden estar influenciados también por la posición geográfica en la que se encuentran los rodales coetáneos de menor edad, lo que sugiere que la interacción de la elevación con el CC se ve oscurecida por el efecto de la edad del rodal y otras variables bióticas. No obstante, se recomiendan mayores evaluaciones dentro del área de estudio que permitan comprobar estas hipótesis.

2.6.5 Enfoque utilizado para estimar CC en el piso forestal

En esta investigación, la estimación del CC en el piso forestal está basada en datos armonizados, provenientes de colecta de muestras e inventarios forestales realizados en parcelas permanentes de muestreo, así como en información derivada de LiDAR que cubre la totalidad del área de estudio. Las evaluaciones fueron realizadas en rodales bajo un enfoque de cronosecuencia, donde el tiempo se sustituye por el espacio (He *et al.*, 2009; Ángeles-Pérez *et al.*, 2015). De esta forma, en combinación con la medición en dos puntos en el tiempo (2013 y 2018), se obtuvieron tendencias y estimaciones del potencial de acumulación de carbono en este componente del paisaje forestal (Chávez-Aguilar *et al.*, 2016; Casiano-Domínguez *et al.*, 2018).

A menudo es difícil extrapolar los resultados de evaluaciones realizadas desde una escala de parcela a paisaje debido a la heterogeneidad del piso forestal, la vegetación y las condiciones abióticas. En México no se han realizado evaluaciones que consideren todos estos aspectos, y a nivel internacional, son pocos los estudios que han empleado protocolos similares para estimar el almacenamiento de carbono y su variabilidad espacial en el piso forestal (Penne *et al.*, 2010; Fu *et al.*, 2014; Domke *et al.*, 2016; Jeyanny *et al.*, 2016; Paul *et al.*, 2018; Cao *et al.*, 2019). Sin embargo, la efectividad de este enfoque depende de la fuerza de las correlaciones entre las distribuciones de CC en el piso forestal y las variables predictoras utilizadas, y sobre todo, si esas relaciones son constantes a lo largo del paisaje (Powers y Schlesinger, 2002).

El buen desempeño predictivo de RF muestra que este algoritmo puede estimar el CC en el piso forestal en áreas no muestreadas del paisaje y describir patrones espaciales a escala local. Sus principales ventajas de aplicación son: 1) reduce el esfuerzo de obtención de información de campo, al no requerir una alta densidad de puntos de muestreo; 2) modela relaciones no lineales complejas entre la variable de respuesta y los predictores; y 3) su capacidad para sobreponerse a problemas de multicolinealidad e información ambiental limitada (Breiman, 2001; Liaw y Wiener, 2002).

Las estimaciones espaciales del CC pueden constituir una guía importante para el monitoreo del carbono en el suelo y las estimaciones a nivel de ecosistema. Los mapas de distribución espacial permiten detectar zonas que requieren nuevas estrategias de muestreo en áreas con altos niveles de incertidumbre. Por otro lado, este tipo de estimaciones permiten promover estrategias efectivas de manejo forestal encaminadas a conservar áreas con material orgánico significativo, proteger el suelo de la erosión o regular los depósitos de hojarasca para prevenir incendios forestales.

2.7.CONCLUSIONES

Las estimaciones de carbono en las capas del piso forestal indicaron que el CC fue mayor en la capa de fermentación en comparación con la capa de mantillo, como un efecto del grado de descomposición del material orgánico. Las diferencias observadas en la relación C: N sugieren un cambio en la velocidad de descomposición en las capas del piso forestal, entre 2013 y 2018.

El algoritmo Random Forest mostró que la elevada variación espacial del CC dentro del SMIC-Atopixco fue determinada principalmente por la edad del rodal, el área basal, la altura arbórea dominante y la elevación. La precisión del modelo de predicción RF fue alta para ambos años evaluados, con un R^2 superior a 0.85. Evaluar el CC por capas del piso forestal tuvo el efecto más importante en las estimaciones y contribuyó significativamente en el ajuste de RF. Asimismo, la armonización de las mediciones de campo con datos de sensores remotos mejora la capacidad predictiva del modelo. Lo anterior indica que el modelo de regresión RF puede ser utilizado para modelar almacenes de carbono en este componente del ecosistema. Finalmente, este estudio proporciona un método estadístico viable y novedoso para futuras modelaciones de la dinámica del carbono en el piso forestal en sistemas bajo aprovechamiento.

DISCUSIÓN GENERAL

La distribución espacial del CC fue modelada bajo el enfoque del Mapeo Digital de Suelo (DSM), comparando el desempeño predictivo de tres técnicas que representan las rutinas de modelación más aplicadas para estudios de carbono en el suelo (interpolación espacial, regresión estadística y Machine Learning). El desempeño de cada uno de los modelos es altamente variable y diversos autores han sugerido que está en función del diseño de la investigación, de las variables de entrada consideradas, la escala y resolución espacial de la información, así como de la complejidad inherente del atributo de interés (Zeraatpisheh *et al.*, 2017; Wang *et al.*, 2018). Por lo cual, definir una “mejor” técnica de modelado es aún un desafío para DSM. Para la zona de estudio, Random Forest (RF) superó al resto de los modelos, expresando altos valores de R^2 y errores de predicción más pequeños.

Los resultados obtenidos son similares a otros estudios que han utilizado RF para modelar carbono en este componente (Wiesmeier *et al.*, 2014; Deng *et al.*, 2018), y tal como era esperado en el planteamiento de las hipótesis, el uso de datos provenientes de fuentes múltiples, mejoró la capacidad predictiva de los modelos. La superioridad de RF se atribuye a una menor susceptibilidad al sobreajuste (por incorporar aleatoriedad en sus predicciones mediante muestreos bootstrap reiterativos) y a una mejor capacidad para manejar las relaciones complejas existentes entre el CC y las variables predictoras empleadas (Breiman, 2001; Cao *et al.* 2019; Zeraatpisheh *et al.* 2019).

Por otra parte, las relaciones y posibles interacciones del CC en el piso forestal con diversos factores bióticos y abióticos han sido poco explorados y aun es controversial cómo estos factores están ligados al almacenamiento de carbono (Cao *et al.*, 2019). Diversos estudios han encontrado relaciones consistentes entre el CC de carbono en el piso forestal y las variables definidas en este estudio, como la diversidad de especies, la topografía, la estructura del dosel, la edad del rodal, entre otros (Barlow *et al.*, 2007; Penne *et al.*, 2010; Domke *et al.*, 2016; Cao *et al.*, 2019).

Estas relaciones pueden ser expresadas mediante funciones lineales, polinomiales, o incluso funciones mucho más complejas; aunado a esto, no todos estos factores son igualmente importantes, desde un punto de vista espacial, y su importancia relativa puede cambiar bajo diferentes condiciones (Powers y Schlesinger, 2002). Para el SMIC-Atopixco, el CC exhibió

patrones espaciales contrastantes entre las distintas capas del piso forestal y entre los años evaluados. El análisis de interacciones espaciales del CC en el piso forestal con las variables predictoras empleadas reveló que la edad del rodal y el área basal, son las variables que tienen una mayor influencia sobre los almacenes de carbono en el piso forestal y su distribución espacial, seguido de la altura dominante, la cobertura arbórea y la elevación. Cabe destacar que, a excepción de la elevación, las variables definidas por RF se encuentran estrechamente vinculadas con el manejo forestal realizado en la zona de estudio.

La estimación del CC en los bosques bajo manejo forestal es necesaria para tener bases sólidas para la evaluación de la sostenibilidad del mismo y para comprender mejor la contribución de este tipo de bosques al ciclo global del carbono. Las prácticas silvícolas realizadas en el SMIC-Atopixco tienen como finalidad el aprovechamiento de sus recursos maderables, lo que modifica la estructura del dosel y puede afectar significativamente el almacenamiento de carbono en el piso forestal por diferentes mecanismos, incluyendo alteraciones en la cantidad y calidad de la hojarasca, la composición de la actividad microbiana del suelo y las condiciones micro climáticas, que a su vez, modifican procesos como la absorción de nutrientes, senescencia, caída y descomposición de hojarasca, entre otros (Bravo-Oviedo *et al.*, 2015; Lado-Monserrat *et al.*, 2016).

Este tipo de efectos son reflejados, de forma indirecta, en los patrones espaciales descritos por RF y son consistentes con algunas evaluaciones realizadas previamente en el área de estudio (Cadena Morales, 2006; Barrales-Brito, 2013; Martínez Castillo, 2014; Ángeles-Pérez, *et al.*, 2015). Por ejemplo, Martínez Castillo (2014) reportó variaciones significativas del C y N en el mantillo en función de la edad del rodal, así como una relación C:N diferente entre rodales con distintas edades, lo cual es similar a la variación temporal observada entre 2013 y 2018, que mostró una disminución en los almacenes de carbono y cambios en la relación C:N, que sugieren un incremento en las tasas de descomposición.

CONCLUSIONES GENERALES

Se observaron diferencias significativas en el contenido de carbono (CC) por capas del piso forestal y por año evaluado. El CC incrementó con el grado de descomposición del material orgánico y se observó una disminución del CC total entre 2013 y 2018. La relación C:N mostró un comportamiento inverso entre 2013 y 2018, lo que sugirió un incremento en las tasas de descomposición entre ese periodo.

El CC exhibió una alta variación espacial dentro del SMIC-Atopixco. Para 2013 y 2018, se observó una diferencia notoria entre los patrones de distribución espacial del CC en las capas Ho y F. Las variables con mayor contribución en las estimaciones del contenido de carbono fueron el área basal, la edad del rodal, la altura dominante, la cobertura arbórea y la elevación. Estas presentaron relaciones no lineales significativas con el contenido de carbono en el piso forestal.

La comparación entre las tres técnicas DSM mostró que la técnica de interpolación espacial tuvo la menor precisión ($R^2 < 0.3$), la técnica de regresión mostró ligeras sub y sobreestimaciones del contenido de carbono ($R^2 > 0.72$), y la técnica de Machine Learning presentó el mejor desempeño ($R^2 > 0.86$). En comparación con los valores de C observados en campo, la distribución espacial del CC, obtenida con kriging ordinario fue poco realista; los mapas derivados de GAM y GLM mostraron sobreestimaciones superiores al máximo valor observado, y RF presentó los mapas más aproximados a la realidad, con valores muy cercanos a los mínimos y máximos obtenidos en campo.

Las interacciones y patrones espaciales descritos muestran el alto grado de sensibilidad de los almacenes de carbono en el piso forestal; la medición en dos puntos en el tiempo (2013 y 2018) permitió observar tendencias y realizar estimaciones del potencial de acumulación de carbono. Asimismo, se detectó que evaluar el CC por capas del piso forestal y utilizar información de fuentes múltiples contribuyó significativamente en el ajuste de los modelos, y consecuentemente, en la precisión de las estimaciones. Las evaluaciones de este tipo pueden constituir una guía importante para el monitoreo del carbono a nivel de ecosistema mediante la aplicación de un método estadístico viable y novedoso para futuras modelaciones de la dinámica del carbono en el piso forestal.

LITERATURA CITADA

- Aber, J.D., Melillo, J.M., 1980. Litter decomposition: measuring relative contributions of organic matter and nitrogen to forest soils. *Canadian Journal of Botany* 58, 416-421.
- Acosta-Mireles, M., Carrillo-Anzures, F., Díaz-Lavariaga, M., 2008. Determination of Total Carbon in Mixed Pine Forests (*Pinus patula* Schl. et Cham). *Terra Latinoamericana* 27 (2), 105-114.
- Adhikari, K., Hartemink, A.E., Minasny, B., Kheir, R.B., Greve, M.B., Greve, M.H., 2014. Digital Mapping of Soil Organic Carbon Contents and Stocks in Denmark. *PLOS One* 9 (8), 1-13.
- Aerts, R., 1997. Climate, leaf litter chemistry and leaf litter decomposition in terrestrial ecosystem: a triangular relationship. *Oikos* 79, 439-449.
- Aguirre-Salado, C.A., Valdez-Lazalde, J.R., Ángeles-Pérez, G., de los Santos-Posadas, H.M., Haapanen, R., Aguirre-Salado, A.I., 2009. Mapeo de carbono arbóreo aéreo en bosques manejados de *Pinus patula* en Hidalgo, México. *Agrociencia* 43, 209-220.
- Aguirre-Salado, C.A., Valdez-Lazalde, J.R., Ángeles-Pérez, G., de los Santos-Posadas, H.M., Aguirre-Salado, A.I., 2011. Mapeo del Índice de Área Foliar y Cobertura Arbórea mediante fotografía hemisférica y datos SPOT 5 HRG: Regresión y K-nn. *Agrociencia* 45 (1), 105-119.
- Anaya Merchant, C.A., 2007. Dinámica de C y N en el mantillo de un bosque tropical caducifolio en Jalisco, México. Tesis de Doctorado. Universidad Nacional Autónoma de México, México. 89 p.
- Anaya Merchant, C.A., García-Oliva, F., Jaramillo, V.J., 2007. Rainfall and labile carbon control litter nitrogen dynamics in a tropical dry forest. *Oecologia* 150 (4), 602-610.
- Ángeles-Pérez, G., 1995. Efecto de la vegetación competidora en el desarrollo inicial de *Pinus patula*. Tesis de Maestría en Ciencias. Colegio de Postgraduados. Montecillo, México. 114 p.
- Ángeles-Pérez, G., Méndez-López, B., Valdez-Lazalde, R., Plascencia-Escalante, F.O., de los Santos-Posadas, H.M., Chávez-Aguilar, G., ... Martínez-López, A., 2015. Estudio de Caso del Sitio de Monitoreo Intensivo del Carbono en Hidalgo. Colegio de Postgraduados. Montecillo, México. 109 p.
- Antúnez, P., Hernández-Díaz, J.C., Wehenkel, C., Clark-Tapia, R., 2017. Generalized Models: An application to Identify Environmental Variables That Significantly Affect the Abundance of Three Tree Species. *Forests* 8 (59), 1-14.
- Averill, C., Waring, B., 2018. Nitrogen limitation of decomposition and decay: How can it occur? *Global Change Biology* 24 (4), 1417-1427.

- Bansal, S., Hallsby, G., Löfvenius, M.O., Nilsson, M.C., 2013. Synergistic, additive and antagonistic impacts of drought and herbivory on *Pinus sylvestris*: Leaf, tissue and whole-plant responses and recovery. *Tree Physiology* 33, 451-463.
- Barbosa, A.M., Brown, J.A., Jiménez-Valverde A., Real, M., 2020. R package “modEva”: Model Evaluation and Analysis. Disponible en línea: <https://cran.r-project.org/web/packages/modEvA/index.html> (Último acceso: enero 28, 2020).
- Barksdale, W.F., Anderson, C.J., Kalin, L., 2014. The influence of watershed run-off on the hydrology, forest floor litter and soil carbon of headwater wetlands. *Ecohydrology* 7, 803-814.
- Barlow, J., Gardner, T.A., Ferreira, L.V., Peres, C.A., 2007. Litter fall and decomposition in primary, secondary and plantation forests in the Brazilian Amazon. *Forest Ecology and Management* 247 (1–3), 91–97.
- Barrales-Brito, 2013. Balance de carbono y tasa de descomposición de mantillo y horizontes de fermentación de suelos forestales. Tesis de Maestría en Ciencias. Colegio de Postgraduados. Montecillo, México. 139 p.
- Bartuszevige, A.M., Hrenko, R.L., Gorchoy, D.L., 2007. Effects of Leaf Litter on Establishment, Growth, and Survival of Invasive Plant Seedlings in a Deciduous Forest. *The American Midland Naturalist* 158, 472-477.
- Bauhus, J., Vor, T., Bartsch, N., Cowling, A., 2004. The effects of gaps and liming on forest floor decomposition and soil C and N dynamics in a *Fagus sylvatica* forest. *Canadian Journal of Forest Research* 34 (3), 509-518.
- Beguín, J., Fuglstad, G.A., Mansuy, N., Paré, D., 2017. Predicting soil properties in the Canadian boreal forest with limited data: Comparison of spatial and non-spatial statistical approaches. *Geoderma* 306, 195-200.
- Berg, B., 2018. Decomposing litter; limit values; humus accumulation, locally and regionally. *Applied Soil Ecology* 123, 494-508.
- Berg, B., McClaugherty, C., 2008. Plant litter. Decomposition, humus formation, carbon sequestration. Second Edition. Springer. 336 p.
- Berg, B., Staaf, H., 1980. Decomposition rate and chemical changes of scots pine needle litter II. Influence of chemical composition. *Ecological Bulletins* 32, 373-390.
- Berndt, C., Haberlandt, U., 2018. Spatial interpolation of climate variables in Northern Germany- Influence of temporal resolution and network density. *Journal of Hydrology: Regional Studies* 15, 184-202.
- Bhunja, G.S., Shit, P.K., Maiti, R., 2018. Comparison of GIS-based interpolation methods for spatial distribution of soil organic carbon (SOC). *Journal of the Saudi Society of Agricultural Sciences* 17, 114-126.

- Biau, G., 2012. Analysis of a Random Forest. *Journal of Machine Learning Research* 13, 1063-1095.
- Biau, G., Scornet, E., 2016. A random forest guided tour. *TEST* 25, 197-227.
- Birdsey, R., Ángeles-Pérez, G., Kurz, W.A., Lister, A., Olguín, M., Pan, Y., ... Johnson, K., 2013. Approaches to monitoring changes in carbon stocks for REDD+. *Carbon Management* 4 (5), 519-537.
- Bivand, R.S., Pebesma, E.J., Gómez-Rubio, V., 2008. *Applied Spatial Data Analysis in R*. Springer. 373 p.
- Bivand, R.S., 2019. R Package “spdep”: Spatial Dependence. Disponible en línea: <https://cran.r-project.org/web/packages/spdep/index.html> (Último acceso: diciembre 21, 2019).
- Black, K., Creamer, R.E., Xenakis, G., Coo, S., 2014. Improving forest soil carbon models using spatial data and geostatistical approaches. *Geoderma* 232-234, 487-499.
- Blanco, J.A., Bosco Imbert, J., Castillo, F.J., 2006. Influence of site characteristics and thinning intensity on litterfall production in two *Pinus sylvestris* L. forests in the western Pyrenees. *Forest Ecology and Management* 237, 342-352.
- Blanco, J.A., Imbert, J.B., Castillo, F.J. 2011. Thinning affects *Pinus sylvestris* needle decomposition rates and chemistry differently depending on site conditions. *Biogeochemistry* 106 (3), 397-414.
- Bolar, K., 2019. Package STAT - Interactive document for Working with Basic Statistical Analysis. Disponible en línea: <https://cran.r-project.org/web/packages/STAT/STAT.pdf> (Último acceso: diciembre 20, 2019).
- Bradford, M.A., Berg, B., Maynard, D.S., Wieder, W.R., Wood, S.A., 2016. Understanding the dominant controls on litter decomposition. *Journal of Ecology* 104, 229-238.
- Bravo-Oviedo, A., Ruiz-Peinado, R., Modrego, P., Alonso, R., Montero, G. 2015. Forest thinning impact on carbon stock and soil condition in Southern European populations of *P. sylvestris* L. *Forest Ecology and Management* 357, 259-267.
- Breiman, L., 2001. Random Forest. *Machine Learning* 45, 5-32.
- Brown, M.B., Forsythe, A.B., 1974. Robust test for the equality of variances. *Journal of American Statistical Association* 69, 364-367.
- Brungard, C.W., Boettinger, J.L., Duniway, M.C., Wills, S.A., Edwards Jr, T.C., 2015. Machine Learning for predicting soil classes in three semi-arid landscapes. *Geoderma* 239-240, 68-83.

- Bubb, K.A., Xu, Z.H., Simpson, J.A., Saffigna, P.G., 1998. Some nutrient dynamics associated with litterfall and litter decomposition in hoop pine plantations of southeast Queensland, Australia. *Forest Ecology and Management* 110, 343-352.
- Bueis, T., Bravo, F., Pando, V., Turrión, M. B. 2018. Local basal area affects needle litterfall, nutrient concentration, and nutrient release during decomposition in *Pinus halepensis* Mill. plantations in Spain. *Annals of Forest Science* 75 (21), 1-12.
- Cadena-Morales, O.I., 2006. Almacenes de carbono en el mantillo en bosques manejados de *Pinus patula* Schiede & Deppe, en el ejido de La Mojonera, Zacualtipán, Hidalgo. Tesis de Licenciatura. Universidad Veracruzana. Córdoba, Veracruz. 53 p.
- Cambardella, C.A., Moorman, T.B., Novak, J.M., Parkin, T.B., Karlen, D.L., Turco, R.F., Konopka, A.E., 1994. Field-Scale Variability of Soil Properties in Central Iowa Soils. *Soil Sciences Society of America Journal* 58, 1501-1511.
- Camera, C., Zomeni, Z., Noller, J.S., Zissimos, A.M., Christoforou, I.C., Bruggeman, A., 2017. A high-resolution map of soil types and physical properties for Cyprus: A digital soil mapping optimization. *Geoderma* 285, 35-49.
- Cameron, E.K., Shaw, C.H., Bayne, E.M., Kurz, W.A., Kull, S.J., 2015. Modelling interacting effects of invasive earthworms and wildfire on forest floor carbon storage in the boreal forest. *Soil biology & biogeochemistry* 88, 189-196.
- Canty, A., Ripley, B., 2019. R Package “boot”: Bootstrap Functions. Disponible en línea: <https://cran.r-project.org/web/packages/boot/index.html> (Último acceso: diciembre 20, 2019).
- Cao, B., Domke, G.M., Russell, M.B., Walters, B.F., 2019. Spatial modeling of litter and soil carbon stocks on forest land in the conterminous United States. *Science of the Total Environment* 654, 94-106.
- Carvahlo Gomes, L., Moniz Faria, R., de Souza, E., Vieira Veloso, G., Schaefer, C.E.G.R., Fernandes Filho, E.I., 2019. Modelling and mapping soil organic carbon stocks in Brazil. *Geoderma* 340, 337-350.
- Casiano-Domínguez, M., Paz-Pellat, F., Rojo-Martínez, M., Covalada-Ocon, S., Aryal, D.R., 2018. El carbono de la biomasa aérea medido en cronosecuencias: primera estimación en México. *Madera y Bosques* 24, 1-22.
- Chapin, F.S., Matson, P.M., Mooney, H.A., 2002. *Principles of terrestrial ecosystem ecology*. Springer-Verlag, New York. 396 p.
- Chávez-Aguilar, G., Ángeles-Pérez, G., Pérez-Suárez, M., López-López, M.A., García-Moya, E., Wayson, C., 2016. Distribución de biomasa aérea en un bosque de *Pinus patula* bajo gestión forestal en Zacualtipán, Hidalgo, México. *Madera y Bosques* 22 (3), 23-36.

- Chen, L., Ren, C., Li, L., Wang., Y., Zhang, B., Wang, Z., Li, L., 2019. A comparative assessment of geostatistical, machine learning, and hybrid approaches for mapping topsoil organic carbon content. *ISPRS International Journal of Geo-Information* 8, 1-18.
- Colín Vargas, C.I., Domínguez Gómez, T.G., González Rodríguez, H., Cantú Silva, I., Guadalupe Colín, J., 2018. Dinámica de nutrientes durante el proceso de degradación de la hojarasca en el Matorral Espinoso Tamaulipeco. *Revista Mexicana de Ciencias Forestales* 9 (49), 87-109.
- Comisión Nacional Forestal (CONAFOR), 2012. Manual y procedimientos para el remuestreo de campo. Re-muestreo 2012. Secretaría de Medio Ambiente y Recursos Naturales. 136 p.
- Cotrufo, M.F., Soong, J.L., Horton, A.J., Campbell, E.E., Haddix, M.L., Wall, D.H., Parton, W.J., 2015. Formation of soil organic matter via biochemical and physical pathways of litter mass loss. *Nature Geoscience* 8, 776-781.
- Cruz-Leyva. I.A., Valdez-Lazalde, J.R., Ángeles-Pérez, G., de los Santos-Posadas, H.M., 2010. Modelación espacial del área basal y volumen de madera en bosques manejados de *Pinus patula* y *P. teocote* en el ejido Atopixco, Hidalgo. *Madera y Bosques* 16 (3), 75-97.
- Dai, W., Fu, W., Jiang, P., Zhao, K., Li, Y., Tao, J., 2018. Spatial pattern of carbon stocks in forest ecosystem of a typical subtropical region of southeastern China. *Forest Ecology and Management* 409, 288-297.
- Dalthorp, D., 2004. The generalized linear model for spatial data: assessing the effects of environmental covariates on population density in the field. *Entomologia Experimentalis et Applicata* 111, 117-131.
- Dang, P., Gao, Y., Liu, J., Yu, S., Zhao, Z., 2018. Effects of thinning intensity on understory vegetation and soil microbial communities of a mature Chinese pine plantation in the Loess Plateau. *Science of the Total Environment* 630, 171-180.
- Deng, X., Chen, X., Ma, W., Ren, Z., Zhang, M., Grieneisen, M.L., ... Lv, X., 2018. Baseline map of organic carbon stock in farmland topsoil in East China. *Agriculture, Ecosystem and Environment* 254, 213-223.
- Dharumarajan, S., Hedge, R., Singh, S.K., 2017. Spatial prediction of major soil properties using Random Forest Techniques – A case of study in semi-arid tropics of South India. *Geoderma Regional* 10, 154-162.
- Djukic, I., Kepfer-Rojas, S., Schmidt, I.K., Larsen, S.K., Beier, C., Berg, B., ... Yahdjian, L., 2018. Early stage litter decomposition across biomes. *Science of The Total Environment* 628-629, 1369-1394.
- Domke, G.M., Perry, C.H., Walters, B.F., Woodall, C.W., Russell, M.B., Smith, J.E., 2016. Estimating litter carbon stocks on forest land in the United States. *Science of the Total Environmental* 557-558, 469-478.

- Dupuy, J.M., Chazdon, R.L., 2008. Interacting effects of canopy gap, understory vegetation and leaf litter on tree seedling recruitment and composition in tropical secondary forest. *Forest Ecology and Management* 255, 3716-3725.
- Figueroa-Navarro, C.M., Ángeles-Pérez, G., Velázquez-Martínez, A., de los Santos-Posadas, H.M., 2010. Estimación de la biomasa en un bosque bajo manejo de *Pinus patula* Schldl. Et Cham. en Zacualtipán, Hidalgo. *Revista Mexicana de Ciencias Forestales* 1 (1), 105-112.
- Fisher, R.F., Binkley, D., 2000. Ecology and management of forest soils. Third edition. Wiley, 139-160 pp.
- Flores-Garnica, J.G., Wong-González, J.C., Paz-Pellat, F., 2018. Camas de combustibles forestales y carbono en México. *Madera y Bosques* 24 (e2401893), 1-15.
- Food and Agricultural Organization of the United Nations (FAO), 2018. Soil Organic Carbon Mapping Cookbook. 2nd edition. Rome, 203 p.
- Frazer, G.W., Canham, C.D., Lertzman K.P., 1999. Gap Light Analyzer (GLA): Imaging software to extract canopy structure and gap light transmission indices from true-colour fisheye photographs, users manual and program documentation. Copyright © 1999: Simon Fraser University, Burnaby, British Columbia, and the Institute of Ecosystem Studies, Millbrook, New York.
- Frazer, G.W., Fournier, R.A., Trofymow, J.A., Hall, R.J., 2001. A comparison of digital and film fisheye photography for analysis of forest canopy structure and gap light transmission. *Agricultural and Forest Meteorology* 109, 249-263.
- Freeman, E.A., Frescino, T.S., Moisen, G.G., 2018. ModelMap: An R Package for Model Creation and Map Production. Vignette for the ModelMap package. 43 pp. Disponible en línea: <https://cran.r-project.org/web/packages/ModelMap/vignettes/VModelMap.pdf> (Último acceso: noviembre 27, 2019).
- Fu, W.J., Jiang, P.K., Zhou, G.M., Zhao, K.L., 2014. Using Moran's I and GIS to study the spatial pattern of forest litter carbon density in a subtropical region of southeastern China. *Biogeoscience* 11, 2401-2409.
- Funes, I., Savé, R., Rovira, P., Molowny-Horas, R., Alcañiz, J.M., Ascaso, ... Vayreda, J., 2019. Agricultural Soil Organic Carbon Stocks in the north-eastern Iberian Peninsula: Drivers and Spatial Variability. *Science of the Total Environment* 668, 283-294.
- Galicia, L., Saynes, V., Campo, J., 2015. Biomasa aérea, biomasa subterránea y necromasa en una cronosecuencia de bosques templados con aprovechamiento forestal. *Botanical Science* 93, 473-484.
- García, E., 1981. Modificaciones al sistema de clasificación climática de Köppen. Instituto de Geografía. Universidad Nacional Autónoma de México. Serie Libros. 6, 98 p.

- Gessner, M.O., Swan, C.M., Dang, C.K., McKie, B.G., Bardgett, R.D., Wall, D.H., Hättenschwiler, S., 2010. Diversity meets decomposition. *Trends in Ecology and Evolution* 5 (6), 372-380.
- Giunta, A.D., Shaw, J.D., 2018. Environmental, Structural, and Disturbance Influences over Forest Floor Components in Interior Douglas-Fir Forest of the Intermountain West, USA. *Forest* 9, 1-25.
- Godínez-Ibarra, O., Ángeles-Pérez, G., López-Mata, L., García-Moya, E., Valdez-Hernández, J.I., de los Santos-Posadas, H., Trinidad-Santos, T., 2007. Lluvia de semillas y emergencia de plántulas de *Fagus grandifolia* subsp. *mexicana* en La Mojonera, Hidalgo, México. *Revista Mexicana de Biodiversidad* 78, 117-128.
- Goovaerts, P., 1999. Geostatistics in soil science: state-of-the-art and perspectives. *Geoderma* 89, 1-45.
- Grier, C.C., Lee, K.M., Nadkarni, N.M., Klock, G.O., Edgerton, P.J., 1989. Productivity of forests of the United States and its relation to soil and site factors and management practices: a review. In *Forestry sciences*. 51 p.
- Griffiths, R.P., Gray, A.N., Spies, T.A., 2010. Soil Properties in Old-Growth Douglas-Fir Forest Gaps in the Western Cascade Mountains of Oregon. *Northwest Science* 84 (1), 33–45.
- Grimm, R., Behrens, T., Märker, M., Elsenbeer, H., 2008. Soil organic carbon concentrations and stocks on Barro Colorado Island – Digital soil mapping using Random Forest analysis. *Geoderma* 146, 102-113.
- Grunwald, S., 2009. Multi-criteria characterization of recent digital soil mapping and modeling approaches. *Geoderma* 152, 195-207.
- Guo, D., Mou, P., Jones, R.H., Mitchell, R.J., 2004. Spatio-temporal patterns of soil available nutrients following experimental disturbance in a pine forest. *Oecologia* 138, 613-621.
- Hartemink, A.E., McBratney A., Mendonça-Santos, M.E. (Eds.), 2008. *Digital Soil Mapping with Limited Data*. Springer. 445 p.
- He, X., Zhang, P., Lin, Y., Li, A., Tian, X., Zhang, Q. H., 2009. Responses of litter decomposition to temperature along a chronosequence of tropical montane rainforest in a microcosm experiment. *Ecological Research* 24 (4), 781–789.
- Hiemstra, P., 2013. R Package “automap”: Automatic interpolation package. Disponible en línea: <https://cran.r-project.org/web/packages/automap/index.html> (Último acceso: diciembre 20, 2019).
- Hijmans, R.J., 2019. R Package “raster”: Geographic Data Analysis and Modelling. Disponible en línea: <https://cran.r-project.org/web/packages/raster/index.html> (Último acceso: diciembre 21, 2019).

- Hollinger, D., 2008. Defining a Landscape-Scale Monitoring Tier for the North American Carbon Program. En: Field Measurements for Forest Carbon Monitoring. Hoover C. (Ed.), Springer-NY, USA, 3-16 pp.
- Houtpankin, O.K., Op de Hipt, F., Yaovi Bosa, A., Welp, G., Amelung, W., 2018. Soil organic carbon stocks and their determining factors in the Dano catchment (Southwest Burkina Faso). *Catena* 166, 298-309.
- Intergovernmental Panel on Climate Change (IPCC), 2006. IPCC Guidelines for National Greenhouse Gas Inventories. Institute for Global Environmental Strategies, Japan. Disponible en línea: https://www.ipcc-nggip.iges.or.jp/support/Primer_2006GLs.pdf. (Último acceso: noviembre 27, 2019).
- Jeyanny, V., Balasundram, S. K., Ahmad-Husni, M. H., Wan-Rasidah, K., 2016. Spatial variability of forest floor thickness for estimation of refined carbon stocks in a tropical montane forest. *Journal of Tropical Forest Science* 28 (3), 285–297.
- Kreye, J.K., Hiers, J.K., Varner, J.M., Hornsby, B., Drukker, S., O'Brien, J.J., 2018. Effects of solar heating on the moisture dynamics of forest floor litter in humid environments: composition, structure, and position matter. *Canadian Journal of Forest Research* 48, 1331-1342.
- Lado-Monserrat, L., Lidón, A., Bautista, I. 2016. Erratum to: Litterfall, litter decomposition and associated nutrient fluxes in *Pinus halepensis*: influence of tree removal intensity in a Mediterranean forest. *European Journal of Forest Research* 135 (1), 203–214.
- León, J.D., González, M.I., Gallardo, J.F., 2011. Ciclos biogeoquímicos en bosques naturales y plantaciones de coníferas en ecosistemas de alta montaña de Colombia. *Revista de Biología Tropical* 59 (4), 1883-1894.
- Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. *R News*. 2 (3), 18-22.
- Liu, Y., Guo, L., Jiang, Q., Zhang, H., Chen, Y., 2015. Comparing geospatial techniques to predict SOC stocks. *Soil & Tillage Research* 148, 46-58.
- Liu, Y., Zhang, J., Yang, W., Wu, F., Xu, Z., Tan, B., ... Guo, L., 2018. Canopy gaps accelerate soil organic carbon retention by soil microbial biomass in the organic horizon in a subalpine fir forest. *Applied Soil Ecology* 125, 169-176.
- López-Escobar, N.F., Gómez-Guerrero, A., Velázquez-Martínez, A., Fierros-González, M., Castruita-Esparza, L.U., Vera-Castillo, J.A.G., 2017. Reservoirs and nutrient dynamics in two stands of *Pinus montezumae* Lamb. in Tlaxcala, Mexico. *Revista Chapingo Serie Ciencias Forestales y del Ambiente* 24 (1), 115-129.
- López-Hernández, J.M., González-Rodríguez, H., Ramírez-Lozano, R.G., Cantú-Silva, I., Gómez-Meza, M.V., Pando-Moreno, M., Estrada-Castillón, A.E., 2013. Producción de hojarasca y

- retorno potencial de nutrientes en tres sitios del Estado de Nuevo León, México. *Polibotánica* (35), 41-64.
- López-Hernández, J.M., Corral-Rivas, J.J., González-Rodríguez, H., Domínguez-Gómez, T.G., Gómez-Meza, M.V., Cantú-Silva, I., 2018. Depósito y descomposición de hojarasca de *Pinus cooperi* C.E.Blanco en El Salto, Durango, México. México. *Revista Mexicana de Ciencias Forestales* 9 (50), 201-222.
- López-Merlín, D., Maldonado, V., Wayson, C., Carrillo, O., Dupuy Rada, J.M., Ángeles-Pérez, G., Caamal Sosa, J.P., Méndez-López, B., Sánchez-Santos, G., Chávez-Aguilar, G., Johnson, K., Tamayo, M., Puc, S., 2016. Capítulo III. Reservorios del carbono en parcelas permanentes. En: Protocolo para la estimación de la dinámica del carbono forestal en sitios de medición intensiva: un enfoque multi-escala. Fortalecimiento REDD+ y Cooperación Sur-Sur, México, Noruega.
- Lucena, M.S., Rocha Alves, A., Alves Baekke, 2019. Litter and nutrient deposition in semi-arid ecosystem of Brazil (Caatinga) under silvicultural systems. *Madera y Bosques* 25 (1), 1-16.
- Maindonald, J., 2018. Package “gamclass” – Functions and Data for a Course on Modern Regression and Classification. Disponible en línea: <https://cran.r-project.org/web/packages/gamclass/gamclass.pdf> (Último acceso: diciembre 20, 2019).
- Malone, B.P., Jha, S.K., Minasny, B., McBratney, A.B., 2016. Comparing regression-based digital soil mapping and multiple-point geostatistics for the spatial extrapolation of soil data. *Geoderma* 262, 243-253.
- Martínez-Castillo, B.E., 2014. Descomposición de la fracción foliar de *Pinus patula* Schltdl. & Cham. y dinámica de la producción de hojarasca en un bosque manejado, en el Estado de Hidalgo. Tesis de Maestría en Ciencias. Colegio de Postgraduados. Montecillo, México. 113 p.
- Martínez-Yrizar, A., Nuñez, S., Miranda, H., Búrquez, A., 1999. Temporal and Spatial Variation of Litter Production in Sonoran Desert Communities. *Plant Ecology* 145 (1), 37-48.
- Martínez-Yrizar, A., Sarukhan, J., 1993. Cambios estacionales del mantillo en el suelo de un bosque tropical caducifolio y uno subcaducifolio en Chamela, Jalisco, México. *Acta Botanica Mexicana* 21, 1-6.
- McKenzie, N.J., Ryan, P.J., 1999. Spatial prediction of soil properties using environmental correlation. *Geoderma* 89, 67-94.
- Meersmans, J., De Ridder, F., Canters, F., De Baets, S., Van Molle, M., 2008. A multiple regression approach to assess the spatial distribution of Soil Organic Carbon (SOC) at the regional scale (Flanders, Belgium). *Geoderma* 143, 1-1-3.
- Meier, I.C., Leuschner, C. (2010). Variation of soil and biomass carbon pools in beech forests across a precipitation gradient. *Global Change Biology*, 16 (3), 1035–1045.

- Moreno-Valdez, M.E., Domínguez-Gómez, T.G., Alvarado, M.S., Guadalupe Colin, J., Corral Rivas, S., González Rodríguez, H., 2018. Aporte y descomposición de hojarasca en bosques templados de la región de El Salto, Durango. *Revista Mexicana de Ciencias Forestales* 9 (47), 70-93.
- Nave, L.E., Vance, E.D., Swanston, C.W., Curtis, P.S. 2010. Harvest impacts on soil carbon storage in temperate forests. *Forest Ecology and Management* 259 (5), 857–866.
- Nobis, M., 2005. SideLook 1.1 - Imaging software for the analysis of vegetation structure with true-colour photographs. Disponible en línea: <http://www.appleco.ch> (Último acceso: enero 5, 2020).
- Nobis, M., Hunziker, U., 2005. Automatic thresholding for hemispherical canopy-photographs based on edge detection. *Agricultural and Forest Meteorology* 128, 243-250.
- Nussbaum, M., Spiess, K., Baltensweiler, A., Grob, U., Keller, A., Greiner, E., ... Papritz, A., 2018. Evaluation of digital soil mapping approaches with large sets of environmental covariates. *Soil* 4, 1-22.
- Ortiz-Reyes, A.D., Valdez-Lazalde, J.R., de los Santos-Posadas, H.M., Ángeles-Pérez, G., Paz-Pellat, F., Martínez-Trinidad, T., 2015. Inventario y cartografía de variables del bosque con datos derivados de LiDAR: comparación de métodos. *Madera y Bosques* 21 (3), 111-128.
- Pando-Moreno, M., Mendoza-Aguilar, D, Cuellar-Rodríguez, G., Jurado, E., 2018. Descomposición de la hojarasca del Matorral Espinoso Tamaulipeco y de una especie vegetal introducida. *Revista Mexicana de Ciencias Forestales* 9 (50), 174-200.
- Pan, Y., Birdsey, R., 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, 988-993.
- Paul, K.I., England, J.R., Baker, T.G., Cunningham, S.C., Perring, M.P., Polglase, P.J., ... Herrmann, T., 2018. Using measured stocks of biomass and litter carbon to constrain modelled estimates of sequestration of soil organic carbon under contrasting mixed-species environmental plantings. *Science of the Total Environment* 615, 348–359.
- Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. *Computer & Geosciences* 30 (7). 683-691.
- Penne, C., Ahrends, B., Deurer, M., Böttcher, J., 2010. The impact of the canopy structure on the spatial variability in forest floor carbon stocks. *Geoderma* 158, 282-297.
- Ponge, J.F. 2013. Plant – soil feedbacks mediated by humus forms: A review. *Soil Biology and Biogeochemistry* 57, 1048-1060.
- Powers, J.S., Schlesinger, W.H., 2002. Relationships among soil carbon distributions and biophysical factors at nested spatial scales in rain forest of northeastern Costa Rica. *Geoderma* 109, 165-190.

- R Development Core Team, 2018. R: a Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing. Disponible en línea: <http://www.r-project.org/> (Último acceso: diciembre 1, 2019).
- Ramírez Martínez, A., 2018. Manejo forestal con escenarios de Carbono en Zacualtipán, Hidalgo. Tesis de Maestría en Ciencias. Colegio de Postgraduados. Montecillo, México. 82 p.
- Ribeiro, P.J., Diggle, P.J., 2006. geoR: A package for Geostatistical Data Analysis. *R News* 1 (2), 15-18.
- Rinkes, Z.L., DeForest, J.L., Grandy, A.S., Moorhead, D.L., Weintraub, M.N., 2014. Interactions between leaf litter quality, particle size, and microbial community during the earliest stage of decay. *Biogeochemistry* 117, 153–168.
- Rupšys, P., Petrauskas, E., 2019. Evolution of the bivariate tree diameter and height distributions via the stand age: von Bertalanffy bivariate diffusion process approach. *Journal of Forest Research* 24, 16-26.
- Santiago-García, W., de los Santos-Posadas, H.M., Ángeles-Pérez, G., Valdez-Lazalde, J.R., del Valle-Paniagua, D.H., Corral-Rivas, J.J., 2012. Auto-aclareo y guías de densidad para *Pinus patula* mediante el enfoque de regresión de frontera estocástica. *Agrociencia* (47), 75-89.
- Santiago-García, W., de los Santos-Posadas, H.M., Ángeles-Pérez, G., Valdez-Lazalde, J.R., Ramírez-Valverde, G., 2013. Sistema compatible de crecimiento y rendimiento para rodales coetáneos de *Pinus patula*. *Revista Fitotecnia Mexicana* 36 (2), 163-172.
- Schlesinger, W.H., Bernhardt, E.S., 2013. *Biogeochemistry: An Analysis of Global Change*. Third edition. Academic Press. 688 p.
- Scull, P., Franklin, J., Chadwick, O.A., McArthur, D., 2003. Predicting soil mapping: a review. *Progress in Physical Geography* 27, 171-197.
- Secretaría de Agricultura y Recursos Hidráulicos (SARH), 1984. Manual de aplicación del Método de Desarrollo Silvícola. Dirección General de Aprovechamientos Forestales. México, 197 p.
- Soriano-Luna, M.A., Ángeles-Pérez, G., Guevara, M., Birdsey, R., Pan, Y., Vaquera-Huerta, H., ... Vargas, R., 2018. Determinants of Above-Ground Biomass and Its Spatial Variability in a Temperate Forest Managed for Timber Production. *Forests* 9 (490), 1-20.
- Soriano-Luna, M.A., Ángeles-Pérez, G., Martínez-Trinidad, T., Plascencia-Escalante, F.O., Razo-Zarate, R., 2015. Estimación de biomasa aérea por componente estructural en Zacualtipán, Hidalgo, México. *Agrociencia* 49, 423-438.
- Urbanová, M., Šnajdr, J., Baldrían, P., 2015. Composition of fungal and bacterial communities in forest litter and soil is largely determined by dominant trees. *Soil Biology & Biogeochemistry* 84, 53-64.

- Van Zijl, G., 2019. Digital soil mapping approaches to address real world problems in southern Africa. *Geoderma* 337, 1301-1308.
- Villegas-Macedo, A.Y., 2019. Caracterización y mapeo de hábitat en bosque templado bajo manejo maderable. Tesis de Maestría en Ciencias. Colegio de Postgraduados. Montecillo, México. 79 p.
- Wang, X., Li, Y., Gong, X., Niu, Y., Chen, Y., Shi, X., Li, W., 2019. Storage, pattern and driving factors of soil organic carbon in an ecologically fragile zone of northern China. *Geoderma* 343, 155-165.
- Webster, R., Oliver, M.A., 2007. *Geostatistics for Environmental Scientist*. Second Edition. John Wiley & Sons, Ltd, England. 318 p.
- Wiesmeier, M., Barthold, F., Spörlein, P., Geuß, U., Hangen, E., Reischl, A., ... Kögel-Kabner, I., 2014. Estimation of total organic carbon storage and its driving factors in soils of Bavaria (southeast Germany). *Geoderma Regional* 1, 67-78.
- Wood, S., 2017. *Generalized Additive Models. An Introduction with R*. Second Edition. CRC Press. A Chapman & Hall Book. 455 p.
- Wood, S., Augustin, N.H., 2002. GAMs with integrated model selection using penalized regression splines and application to environmental modelling. *Ecological Modelling* 157, 157-177.
- Wood, S., Wood, M.S., 2019. Package “mgcv” – Mixed GAM Computational Vehicle with Automatic Smoothness Estimation. Disponible en línea: <https://cran.r-project.org/web/packages/mgcv/mgcv.pdf> (Último acceso: diciembre 20, 2019).
- Woodall, C.W., Perry, C.H., Westfall, J.A., 2012. An empirical assessment of forest floor carbon stocks components across the United States. *Forest Ecology and Management* 269, 1-9.
- Wu, Q., Wu, F., Yang, W., Zhao, Y., He, W., Tan, B., 2014. Foliar Litter Nitrogen Dynamics as Affected by Forest Gap in the Alpine Forest of Eastern Tibet Plateau. *PlosOne* 9 (5), 1-13.
- Xu, Z.F., Pu, X.Z., Yin, H.J., Zhao, C.Z., Liu, Q., Wu, F.Z., 2012. Warming effects on the early decomposition of three litter types, Eastern Tibetan Plateau, China. *European Journal of Soil Science* 63, 360-367.
- Xu, L., Shi, H., Fang, H., Zhou, G., Xu, X., Zhou, Y., ... Chen, L., 2018. Vegetation carbon stocks driven by canopy density and forest age in subtropical forest ecosystem. *Science of the Total Environment* 631-632, 619-629.
- Yanai, R.D., Stehman, S.V., Arthur, M.A., Prescott, C.E., Friedland A.J., Siccama, T.G., Binkley, D., 2003. Detecting Change in Forest Floor Carbon. *Soil Science Society of America Journal* 67 (5), 1583-1593.

- Yao, X., Yu, K., Deng, Y., Zeng, Q., Lai, Z., Liu, J., 2019. Spatial distribution of soil organic carbon stocks in Masson pine (*Pinus massoniana*) forests in subtropical China. *Catena* 178, 189-198.
- Zaragoza-Castañeda, Z., 2016. Carbono almacenado en mantillo en bosques de Atopixco, Hidalgo. Tesis de Licenciatura. Instituto Tecnológico Superior de Zacapoaxtla. Zacapoaxtla, México. 54 p.
- Zeraatpisheh, M., Ayoubi, S., Jafari, A., Tajik, S., Finke, P., 2017. Comparing the efficiency of digital and conventional soil mapping to predict soil types in a semi-arid region in Iran. *Geomorphology* 285, 186-204.
- Zeraatpisheh, M., Ayoubi, S., Jafari, A., Tajik, S., Finke, P., 2019. Digital mapping of soil properties using multiple machine learning in a semiarid region, central Iran. *Geoderma* 338, 445-452.
- Zhang, C., McGrath, D., 2004. Geostatistical and GIS analyses on soil organic carbon concentrations in grassland of southeastern Ireland from two different periods. *Geoderma* 119, 261-275.
- 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.
- Zhang, J., Huang, S., Hogg, E.H., Lieffers, V., Qin, Y., He, F., 2014. Estimating spatial variation in Alberta forest biomass from a combination of forest inventory and remote sensing data. *Biogeosciences* 11, 2793-2808.
- Zhang, H., Wu, P., Yin, A., Yang, X., Zhang, M., Gao, C., 2017. Prediction of soil organic carbon in an intensively managed reclamation zone of eastern China: A comparison of multiple linear regressions and the random forest model. *Science of the Total Environment* 592, 704-713.
- Zhang, Y., Sui, B., Shen, H., Ouyang, L., 2019. Mapping stocks of soil total nitrogen using remote sensing data: A comparison of random forest models with different predictors. *Computers and Electronics in Agriculture* 160, 23-30.
- Zheng, Z., Mamuti, M., Liu, H., Shu, Y., Hu, S., Wang, X., ... Li, X., 2017. Effects of nutrient additions on litter decomposition regulated by phosphorus-induced changes in litter chemistry in a subtropical forest, China. *Forest Ecology and Management* 400, 123-128.
- Zhu, J., Hu, H., Tao, S., Chi, X., Li, P., Jiang, L., ... Fang, J., 2017. Carbon stocks and changes of dead organic matter in China's forest. *Nature Communications* 8, 1-10.

ANEXOS

CAPITULO 1. ENFOQUES ESPACIALES PARA LA MODELACIÓN DE CARBONO EN EL PISO FORESTAL DE BOSQUES BAJO MANEJO.

1. Análisis exploratorio

- Histogramas de frecuencia del CC por capa del piso forestal, para 2013 y 2018.

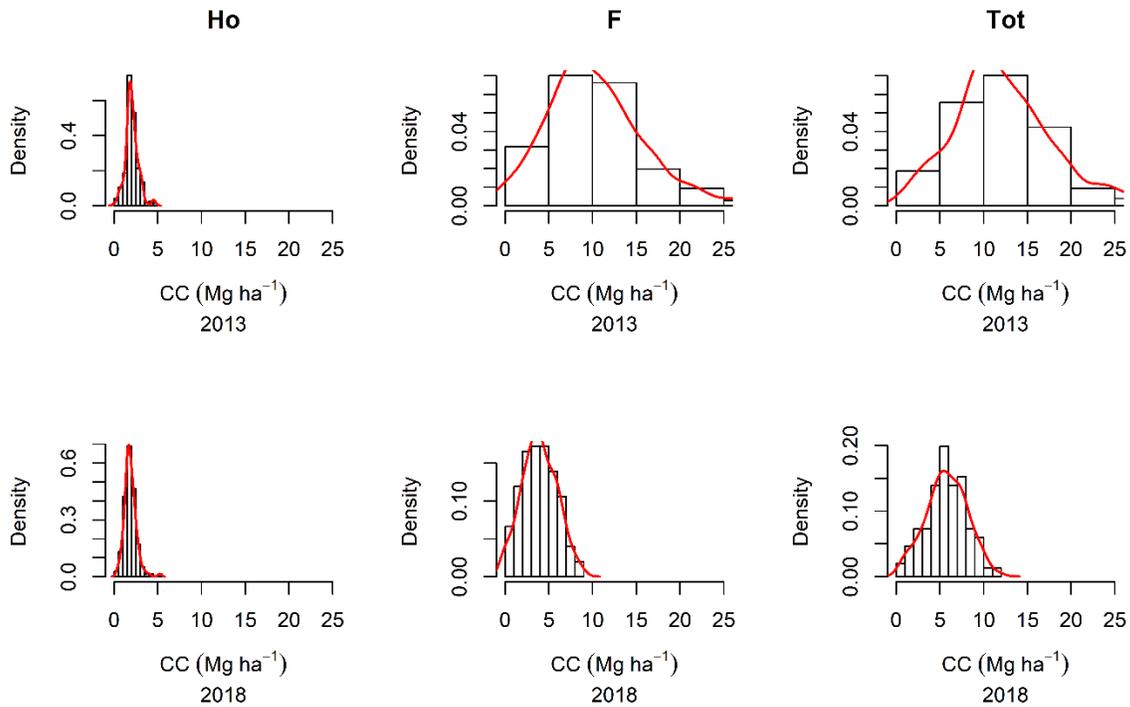


Figura S1. Histogramas de frecuencia relativa del CC por capa del piso forestal, para el año 2013 y 2018. La línea roja muestra la función de densidad de probabilidad. Ho: hojarasca, F: fermentación, Tot: suma de Ho y F.

- Histogramas de frecuencia relativa de cada covariable auxiliar en la predicción, para 2013 y 2018.

-

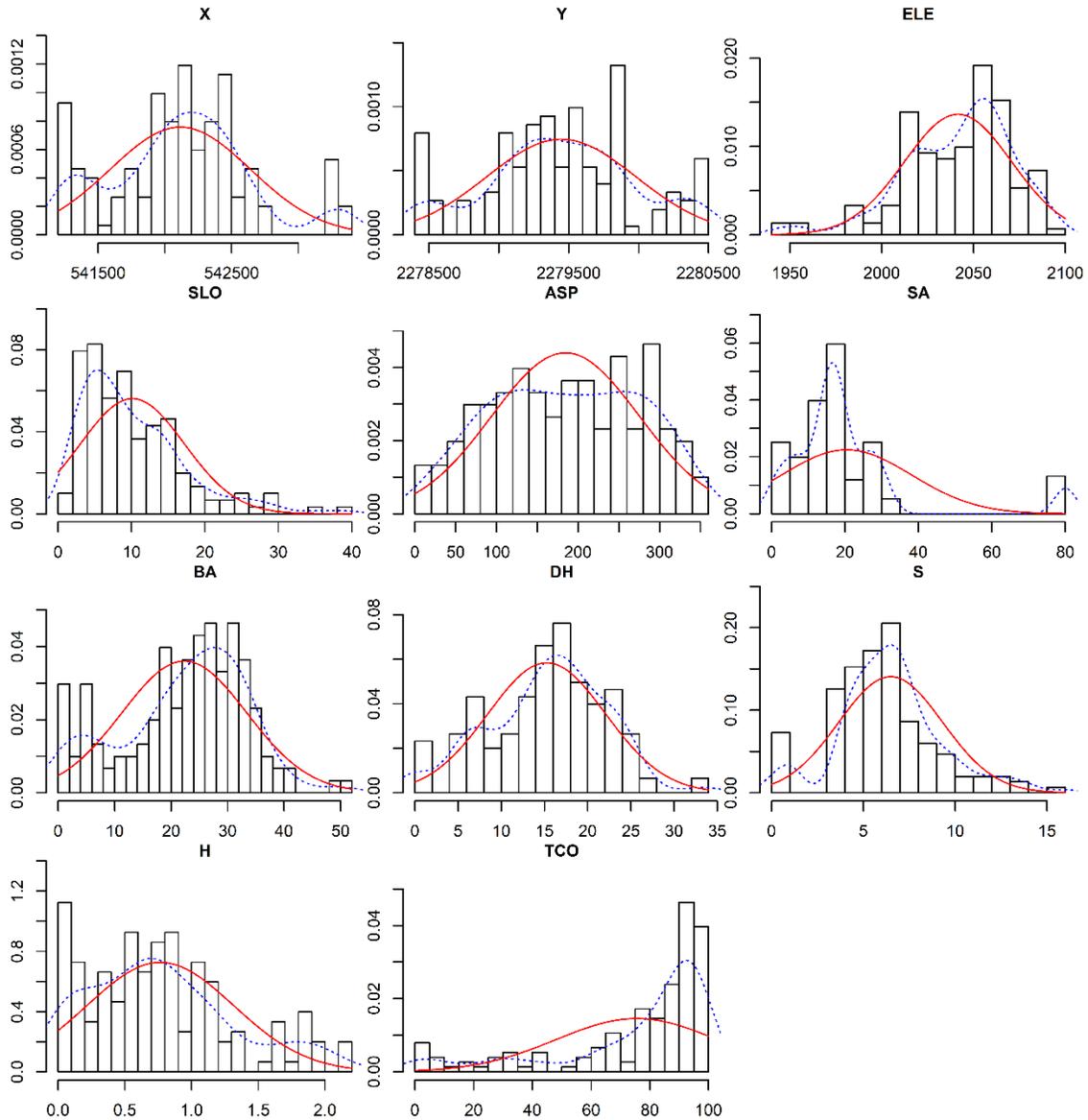


Figura S2. Histogramas de frecuencia relativa (eje Y) de cada variable predictora (eje X), para el conjunto de datos de 2013. La línea roja muestra la distribución teórica normal, la línea azul muestra la función de densidad de probabilidad. Nomenclatura: X, Y: Coordenadas UTM X, UTM Y (m); ELE: elevación (m); SLO: pendiente (grados); ASP: exposición (grados); SA: edad del rodal (años); BA: área basal ($m^2 ha^{-1}$); DH: altura dominante (m); S: riqueza de especies; H: índice de Shannon-Wiener; TCO: cobertura arbórea (%).

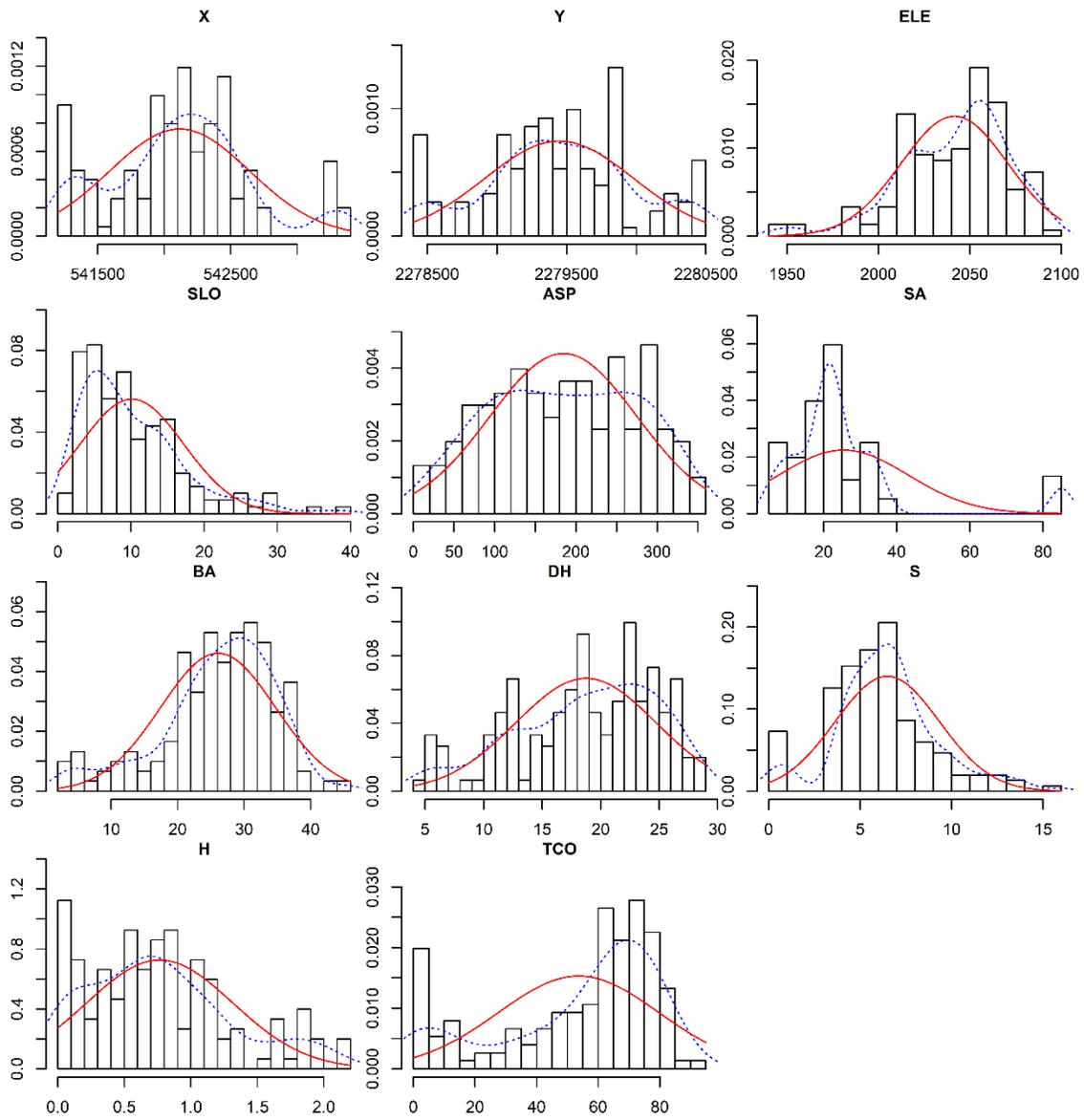


Figura S3. Histogramas de frecuencia relativa (eje Y) de cada variable predictora (eje X), para el conjunto de datos de 2018. La línea roja muestra la distribución teórica normal, la línea azul muestra la función de densidad de probabilidad. Nomenclatura: X, Y: Coordenadas UTM X, UTM Y (m); ELE: elevación (m); SLO: pendiente (grados); ASP: exposición (grados); SA: edad del rodal (años); BA: área basal ($m^2 ha^{-1}$); DH: altura dominante (m); S: riqueza de especies; H: índice de Shannon-Wiener; TCO: cobertura arbórea (%).

- Matrices de correlación de Spearman entre covariables ambientales, para 2013 y 2018.

Cuadro S1. Matriz de correlaciones de Spearman entre variables predictoras, para 2013.

	X	Y	ELE	SLO	ASP	SA	BA	DH	S	H	TCO
X	1.00										
Y	-0.11	1.00									
ELE	^b -0.44	^b 0.68	1.00								
SLO	0.15	^b -0.26	^b -0.51	1.00							
ASP	0.10	0.06	0.05	-0.04	1.00						
SA	0.09	^a 0.21	0.03	-0.08	0.16	1.00					
BA	^b 0.31	^a 0.17	0.00	-0.09	0.14	^b 0.68	1.00				
DH	0.02	^b 0.43	0.14	-0.08	0.10	^b 0.71	^b 0.64	^b 1.00			
S	^a 0.21	-0.08	^a -0.17	0.00	0.15	^b 0.58	^b 0.49	^b 0.44	1.00		
H	0.11	0.05	-0.04	-0.06	0.11	^b 0.63	^b 0.36	^b 0.51	^b 0.85	1.00	
TCO	^b 0.36	0.03	0.00	-0.12	^b 0.28	^b 0.61	^b 0.70	^b 0.51	^b 0.59	^b 0.54	1.00

X, Y: Coordenadas UTM X, UTM Y; ELE: elevación; SLO: pendiente; ASP: exposición; SA: edad del rodal; BA: área basal; DH: altura dominante; S: riqueza de especies; H: índice de Shannon-Wiener; COB: cobertura arbórea. Nivel de significancia: (a < 0.05), (b < 0.01).

Cuadro S2. Matriz de correlaciones de Spearman entre variables predictoras, para 2018.

	X	Y	ELE	SLO	ASP	SA	BA	DH	S	H	TCO
X	1.00										
Y	-0.11	1.00									
ELE	^a -0.44	^b 0.68	1.00								
SLO	0.15	-0.26	^b -0.51	1.00							
ASP	0.10	0.06	0.05	-0.04	1.00						
SA	0.09	0.21	0.03	-0.08	0.16	1.00					
BA	^a 0.21	^b 0.14	0.07	-0.14	-0.01	^b 0.57	1.00				
DH	0.08	0.40	0.12	-0.15	0.16	^b 0.81	^b 0.67	1.00			
S	^a 0.21	^b -0.08	^a -0.17	0.00	0.15	^b 0.58	^b 0.38	^b 0.45	1.00		
H	0.11	0.05	-0.04	-0.06	0.11	^b 0.63	^b 0.29	^b 0.53	^b 0.85	1.00	
TCO	0.08	-0.15	-0.02	-0.04	-0.08	0.11	^b 0.45	^b 0.23	0.11	0.05	1.00

X, Y: Coordenadas UTM X, UTM Y; ELE: elevación; SLO: pendiente; ASP: exposición; SA: edad del rodal; BA: área basal; DH: altura dominante; S: riqueza de especies; D: índice de Shannon-Wiener; COB: cobertura arbórea. Nivel de significancia: (a < 0.05), (b < 0.01).

2. Entrenamiento de los modelos

- Semivariogramas empíricos y ajuste a modelo teórico empleados para la interpolación con Kriging Ordinario

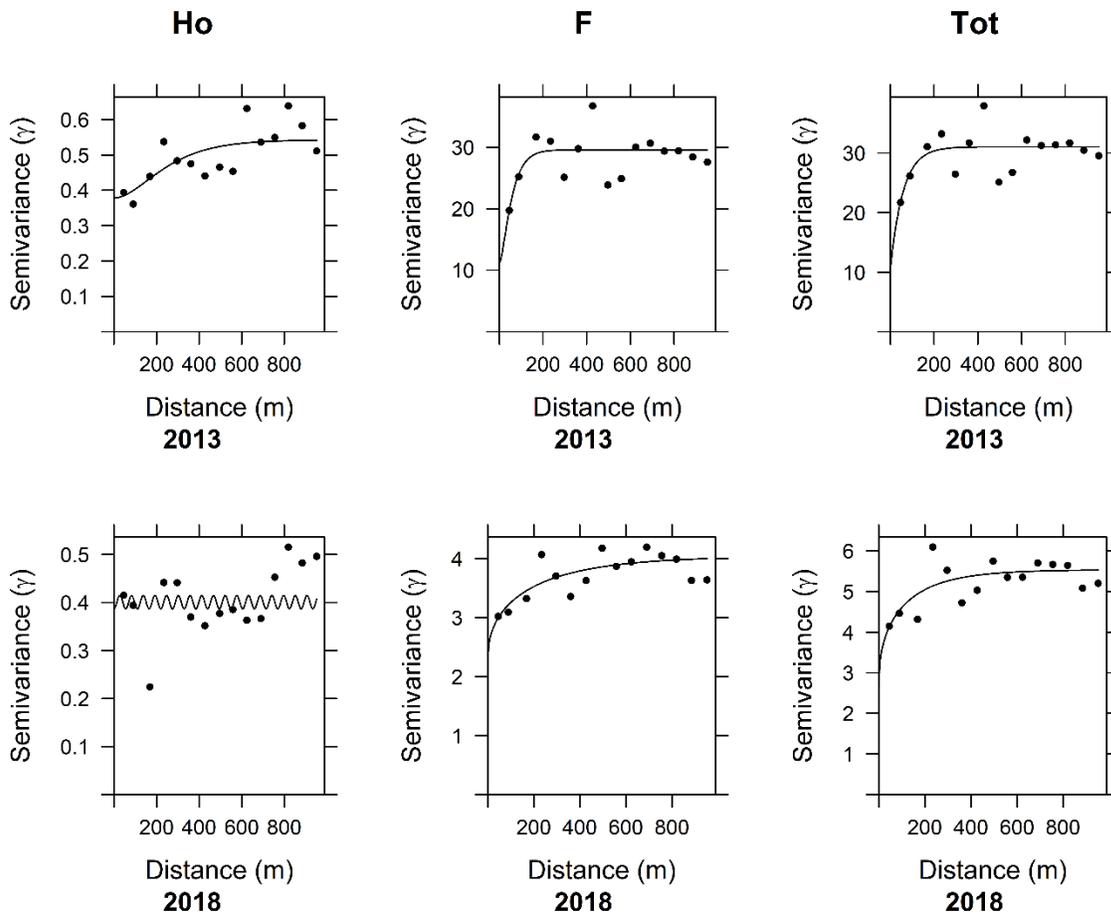


Figura S4. Semivariograma empírico y ajuste de modelo teórico por capa de descomposición en el piso forestal, para 2013 y 2018. Ho: hojarasca, F: fermentación, Tot: suma de Ho y F.

- Gráficos de verificación de supuestos para el modelo de regresión GLM, en 2013 y 2018.

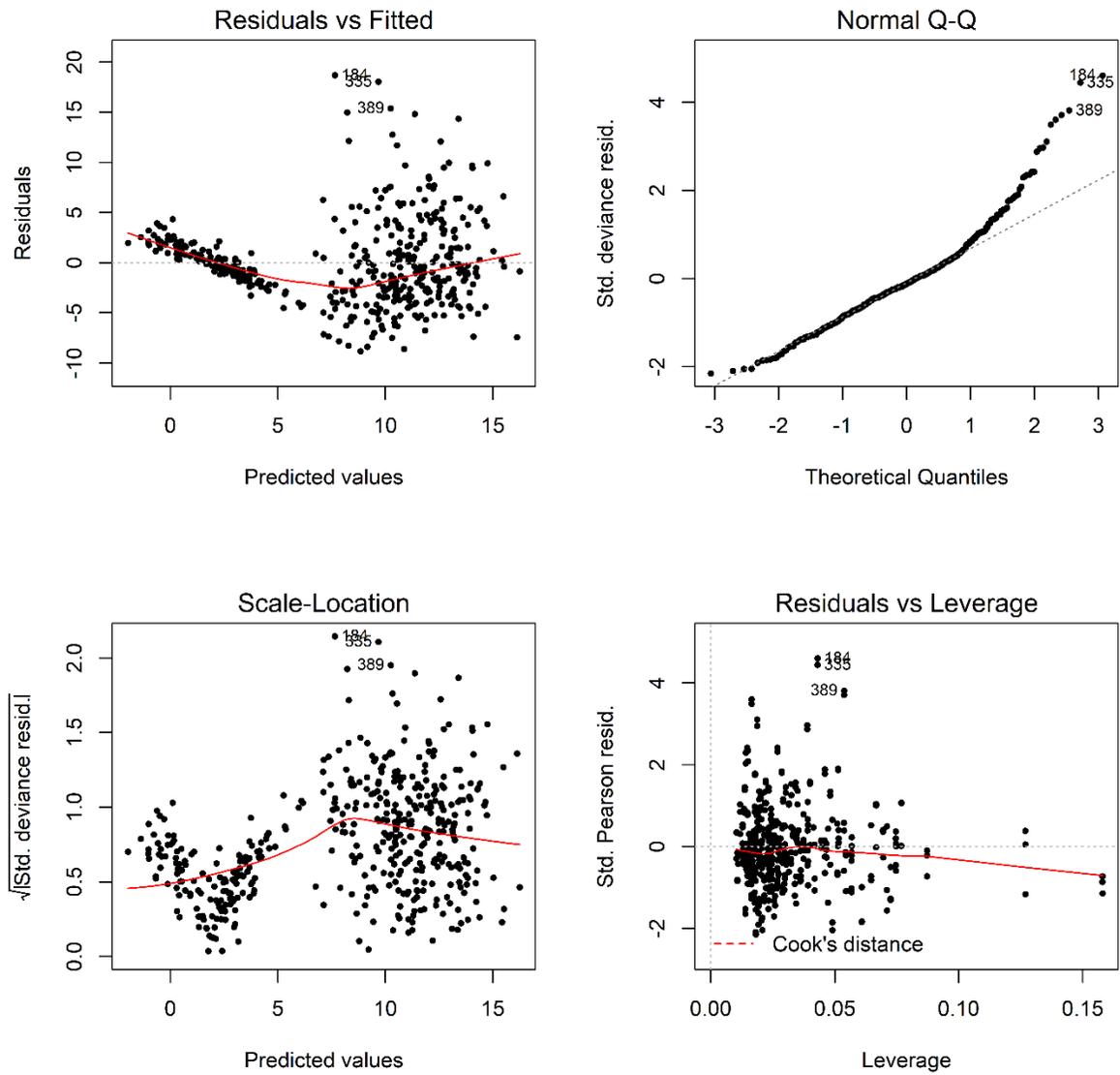


Figura S5. Gráficos de verificación de supuestos del modelo GLM, para el año 2013.

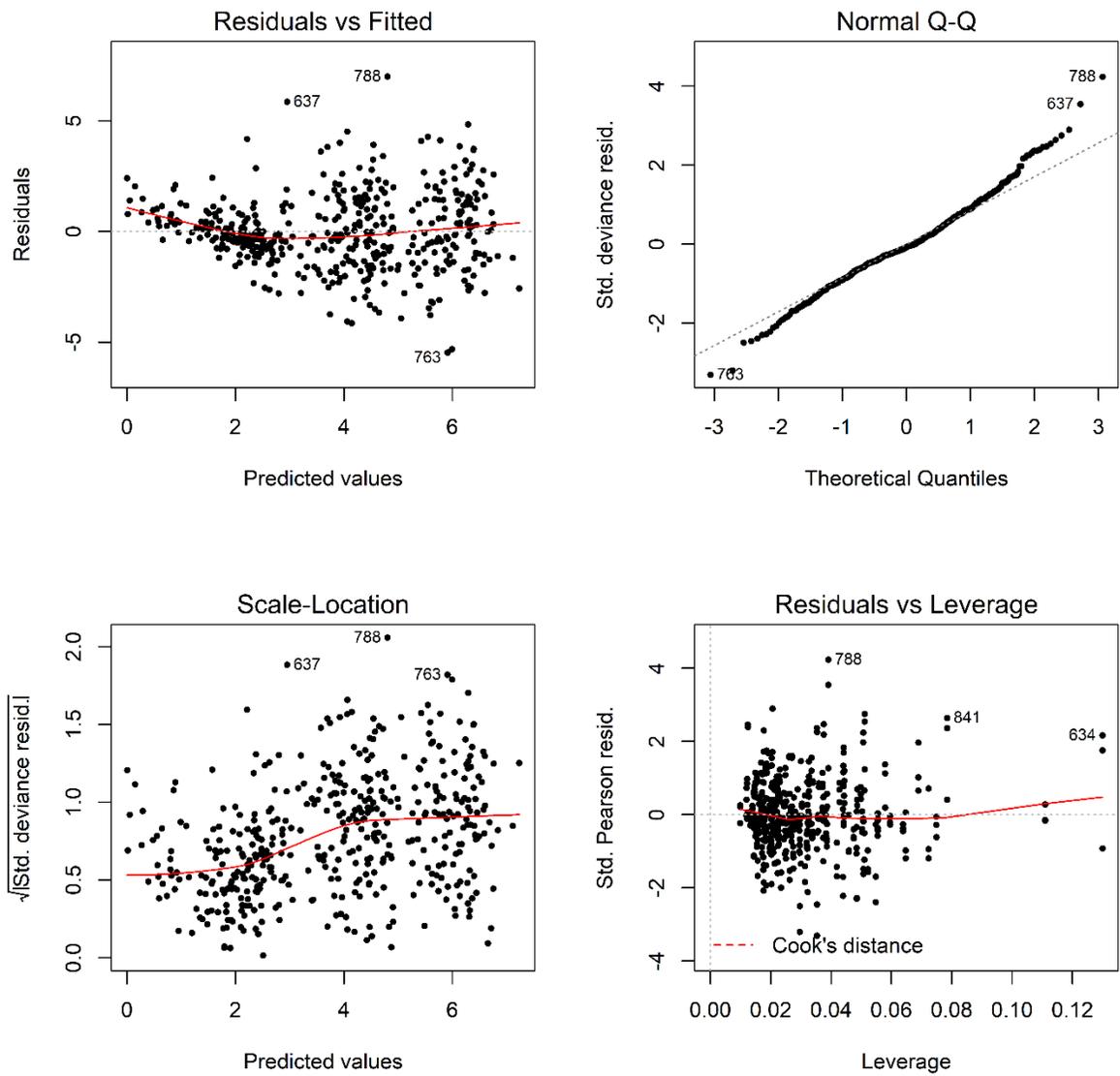


Figura S6. Gráficos de verificación de supuestos del modelo GLM, para el año 2018.

- Gráficos de verificación de supuestos para el modelo de regresión GAM, en 2013 y 2018.

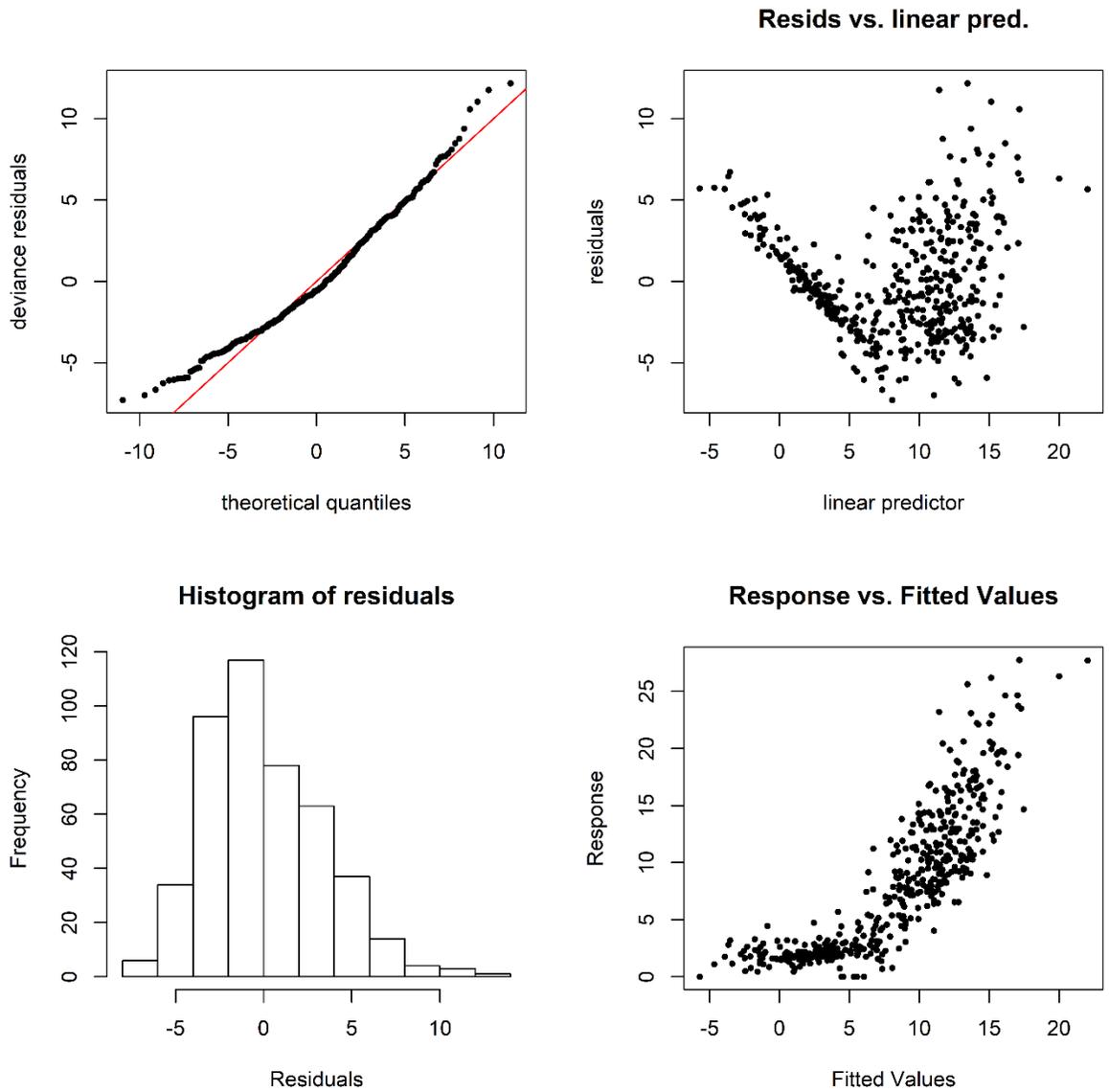


Figura S7. Gráficos de verificación de supuestos del modelo GAM, para el año 2013.

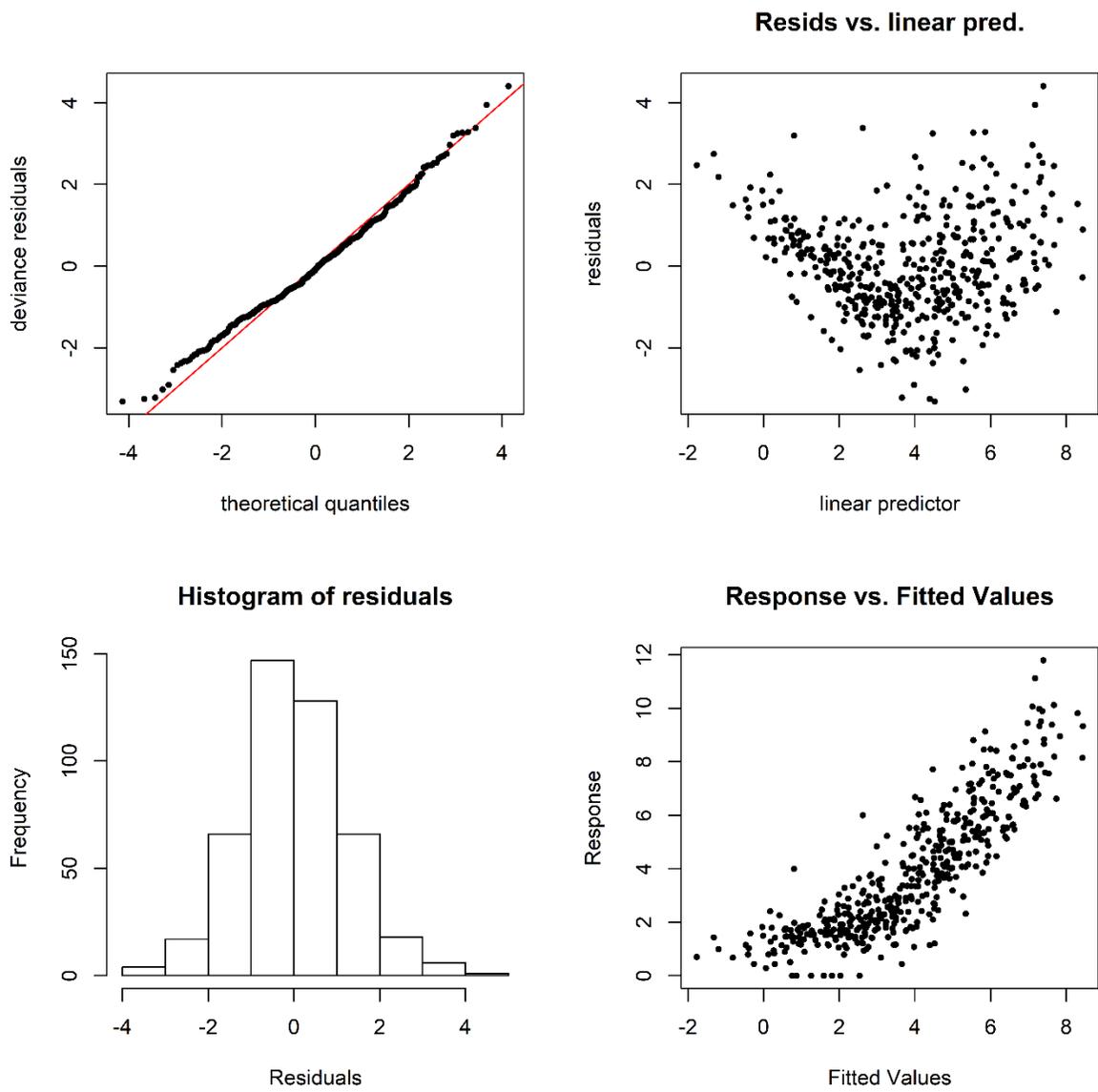


Figura S8. Gráficos de verificación de supuestos del modelo GAM, para el año 2018.

3. Validación de los modelos

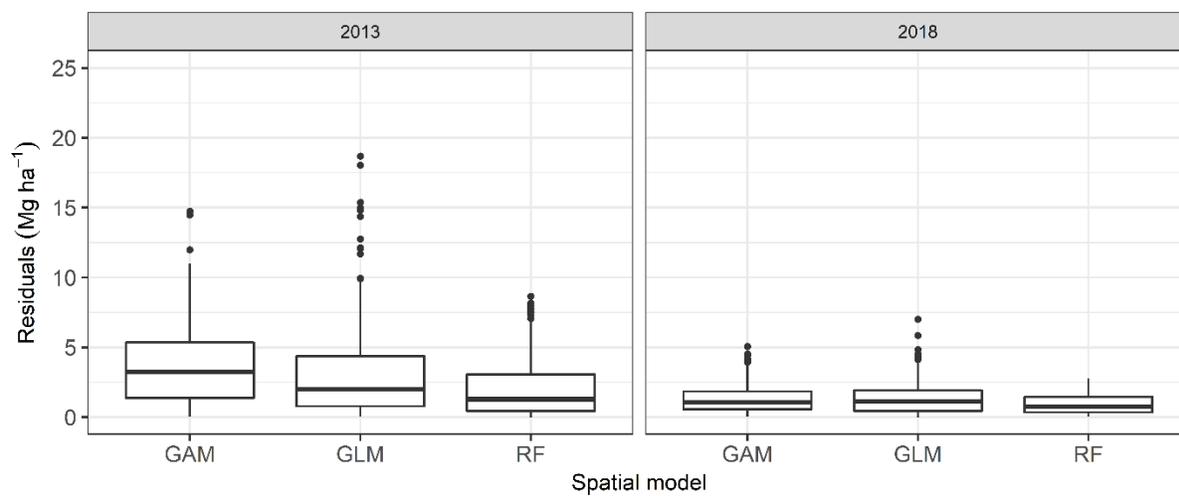


Figura S9. Diagramas de caja de residuales de cada modelo de predicción espacial del CC en el piso forestal. GAM: modelo aditivo generalizado, GLM: modelo lineal generalizado y RF: Random Forest.

CAPITULO 2. PATRONES ESPACIO-TEMPORALES DEL CONTENIDO DE CARBONO EN EL PISO FORESTAL DE UN BOSQUE TEMPLADO MANEJADO

4. Desempeño del modelo RF

Cuadro S3. Parámetros de calibración y desempeño del modelo Random Forest.

Año	Parámetros de entrenamiento		Precisión del modelo		
	Ntree	mtry	Varianza Explicada (%)	RMSE	R ²
2013	2000	9	85.81	2.62	0.90
2018	2000	7	69.07	1.51	0.86

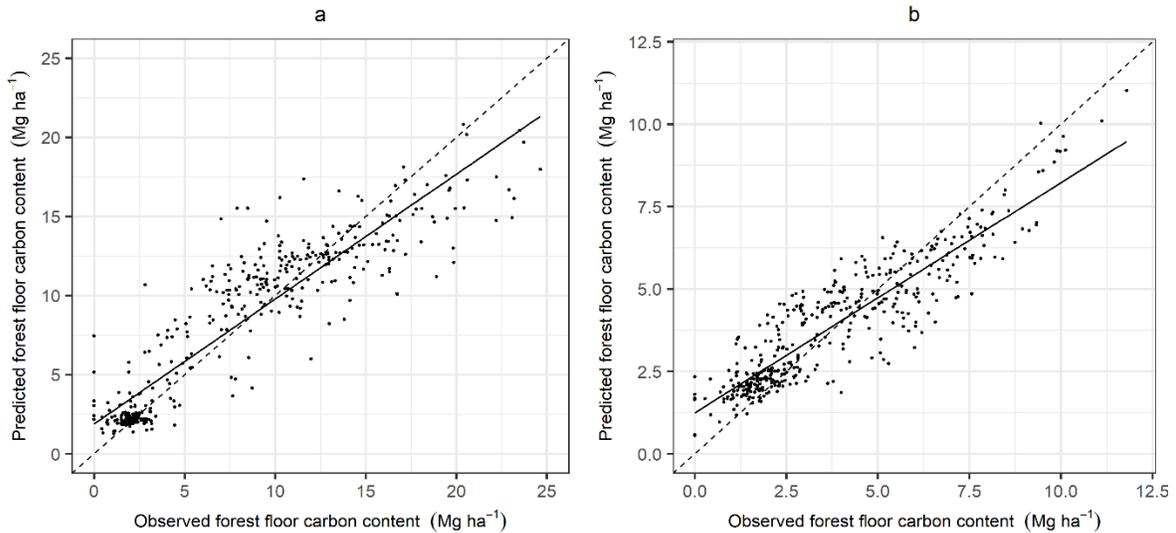


Figura S10. Valores observados vs predichos del conjunto de datos de validación cruzada ($k=10$) del modelo "Random Forest" para: a) 2013 y b) 2018. La línea negra es de regresión y la línea punteada es la relación 1:1.

5. CC en el piso forestal bajo enfoque de cronosecuencia

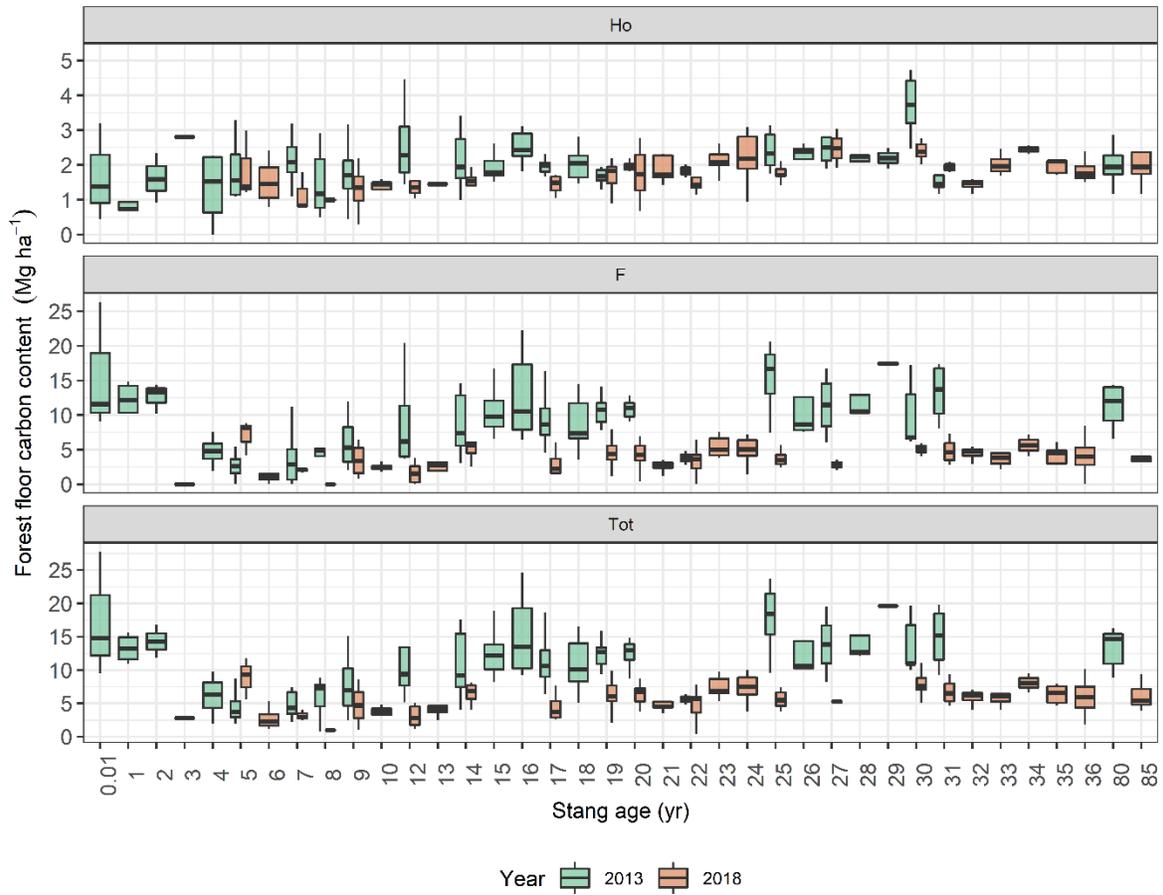


Figura S11. Variación temporal del contenido de carbono por capa de descomposición del piso forestal, bajo enfoque de cronosecuencia. El rodal de 85 años corresponde a un bosque de referencia que no ha sido cosechado.