

ALMACENES Y DINÁMICA DEL CARBONO ORGÁNICO DE LOS SUELOS DE MÉXICO: INCERTIDUMBRE Y ESTADO ESTACIONARIO

STOCKS AND DYNAMIC OF ORGANIC CARBON IN SOILS OF MEXICO: UNCERTAINTY AND STATIONARY STATE

Fernando Paz-Pellat^{1†}, Alma Velázquez-Rodríguez², Aurelio Báez-Pérez³, Fernando Ayala-Niño⁴, Bruno Chávez-Vergara⁴, Gilberto Vela-Correa⁵, Helena Cotler-Ávalos⁶, Sara Covaleda-Ocón⁷, José Cueto-Wong³, Ben de Jong⁸, Fernando de León-González⁵, Jorge Etchevers-Barra⁹, Mariela Fuentes-Ponce⁵, Felipe García-Oliva⁴, Lucila González-Molina³, Armando Guerrero-Peña⁹, Mario Guevara-Santamaría⁴, Jorge Herrera-Silveira¹⁰, Claudia Hidalgo-Moreno⁹, Antoine Libert-Amico¹¹, Jorge Mendoza-Vega⁸, Blanca Prado-Pano⁴, Julio César Rodríguez¹², Víctor Salas-Aguilar¹³, Zulia Sánchez-Mejía¹⁴, Vinisa Saynes-Santillán⁹, Claudia Teutli-Hernández⁴, Rodrigo Vargas¹⁵, Enrico Yépez-González¹⁴

¹ Programa Mexicano del Carbono y Colegio de Postgraduados, Estado de México, México.

² Universidad Autónoma del Estado de México, Toluca, Estado de México, México.

³ Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, México.

⁴ Universidad Nacional Autónoma de México, México.

⁵ Universidad Autónoma Metropolitana, Xochimilco, Ciudad de México.

⁶ Centro GEO, Ciudad de México, México.

⁷ Kibeltik Clima y Medio Ambiente A.C., San Cristóbal de las Casas, Chiapas, México.

⁸ El Colegio de la Frontera Sur, Campeche, Campeche, México.

⁹ Colegio de Postgraduados, México.

¹⁰ CINVESTAV-IPN, Mérida, Yucatán, México.

¹¹ Programa Mexicano del Carbono, Texcoco, Estado de México.

¹² Universidad Autónoma de Sonora, Hermosillo, Sonora, México.

¹³ Universidad Autónoma de Ciudad Juárez, Chihuahua, México.

¹⁴ Instituto Tecnológico de Sonora, Ciudad Obregón, Sonora, México.

¹⁵ University of Delaware, Newark, Delaware, USA.

† Autor para correspondencia: ferpazpel@gmail.com

RESUMEN

El carbono orgánico de los suelos (COS) es uno de los almacenes más importantes de los ecosistemas terrestres, por lo que la estimación de este almacén y sus flujos (dinámica) en México permite la generación de escenarios de intervención para analizar el impacto de políticas públicas, además de dar soporte a la implementación de mercados del carbono o similares para incentivar esquemas de manejo mejorado los suelos para incrementar sus almacenes o reducir emisiones de gases efecto invernadero. En este trabajo se analizan los métodos de estimación de almacenes en el país y se propone un esquema generalizado para el uso de datos multi-fuente, bajo el principio de conservación de la incertidumbre, con el uso del método del Kriging indicador Bayesiano o de geoestadística multi-punto, donde los métodos de aprendizaje de máquina son incorporados en el esquema. Adicionalmente, se revisa la hipótesis de estado estacionario utilizada en los modelos de la dinámica del COS y se propone el uso de la saturación potencial del COS para el desarrollo de métricas para definir la lejanía o cercanía al estado estacionario para poder usar modelos de la dinámica del COS.

Palabras clave: estimación de almacenes de carbono; Kriging indicador Bayesiano; aprendizaje de máquina; modelos de dinámica; saturación potencial del carbono orgánico de los suelos.

ABSTRACT

Soil organic carbon (SOC) is one of the most important stocks of terrestrial ecosystems, so the estimation of this stock and its flows (dynamics) in Mexico allows the generation of intervention scenarios to analyze the impact of public policies, in addition to supporting the implementation of carbon markets or similar instruments to encourage improved soil management schemes to increase carbon stocks or reduce greenhouse gas emissions. In this paper, the methods for estimating carbon stocks in the country are analyzed and a generalized scheme for the use of multi-source data is proposed, under the principle of conservation of uncertainty, with the use of the Bayesian indicator Kriging or geostatistics multi-point method, where machine learning methods are incorporated into the scheme. Additionally, the steady state hypothesis used in the SOC dynamics models is reviewed and the use of the concept of SOC potential saturation is proposed for the development of metrics to define the distance or proximity to the steady state in order to be able to use models of the SOC dynamics.

Keywords: carbon stock estimation; Bayesian indicator Kriging; machine learning; dynamics models; soil organic carbon potential saturation.

INTRODUCCIÓN

El carbono orgánico de los suelos (COS) muestra un gran potencial para el secuestro de carbono (FAO, 2017), motivo por el cual se han planteado diferentes iniciativas internacionales para incrementar los almacenes y reducir las emisiones de gases de efecto invernadero (GEI), tales como: iniciativa 4x1000 (Lal, 2016; MAAF, 2015; Minasny *et al.*, 2017), agricultura climáticamente inteligente (Banco Mundial-CIAT-CATIE, 2014) y RecSoil (FAO, 2018, 2019, 2020a y b), entre otras, orientadas hacia los suelos agrícolas. Aunque existen diversos ejercicios sobre el potencial de secuestro de COS con diferentes prácticas de manejo o cambios de uso del suelo (*e.g.* The World Bank, 2012a y b), la dinámica del COS, particularmente los almacenes subterráneos del carbono, presentan grandes retos para su entendimiento (Vargas *et al.*, 2013). La necesidad de modelar la dinámica del COS permite generar escenarios para evaluar acciones y programas de políticas públicas y de implementación de opciones de intervención con prácticas mejoradas (Lal, 2009).

El uso de modelos de la dinámica del COS, tales como el Century (Parton *et al.*, 1988) y el RothC (Coleman y Jenkinson, 1996), después de su parametrización, son herramientas que permiten fundamentar estrategias de implementación de acciones de intervención para el secuestro de carbono o reducción de emisiones de CO₂. En lo particular, el modelo RothC requiere de información fácilmente disponible, por lo que se ha usado ampliamente en las

iniciativas internacionales (*e.g.* FAO, 2020a) y para el caso de México en experimentos de corto plazo (<20 años) con resultados medianamente satisfactorios al producir subestimaciones y sobreestimaciones (González-Molina *et al.*, 2010, 2018, 2020). Adicionalmente, algunos modelos simples de la dinámica del COS (IPCC, 2006; Ogle *et al.*, 2005) se han utilizado en los inventarios nacionales GEI (de Jong *et al.*, 2006, 2009, 2010), generando fuertes diferencias en las emisiones del almacén suelo, con relación a los resultados oficiales a nivel de país (CONAFOR-SEMARNAT, 2014, 2015, 2020a y b) lo que plantea la necesidad de un análisis fundamentado en la información y el conocimiento, sin sesgos y en forma transparente y documentada. Particularmente la cuantificación de la incertidumbre en las estimaciones debe considerarse como prioritaria, además de la interoperabilidad en la información (Vargas *et al.*, 2017).

La fusión de datos-modelos (Wang *et al.*, 2009) es una estrategia que permite la incorporación de información y conocimiento, estableciendo guías para su uso conjunto. La información e incertidumbre, pueden analizarse bajo dos enfoques complementarios. En la teoría generalizada de la información (Klir, 2006) o GIT (*Generalized Information Theory*), la incertidumbre se conceptualiza como una deficiencia de información (incompleta, imprecisa, fragmentada, no confiable, vaga, contradictoria) y plantea la expansión de la teoría clásica de probabilidad al relajar los requerimientos de aditividad y de límites claros entre clases, para integrar diversas teorías (*e.g.* probabilidades imprecisas,

conjuntos borrosos, capacidades, etc.). Pese al potencial de la teoría generalizada para caracterizar la incertidumbre y sus elementos, este trabajo se restringe a la teoría de la probabilidad, con algunas extensiones bajo el Teorema de Bayes y otras conceptualizaciones, dado su uso predominante en la estimación de emisiones GEI (IPCC, 2006). Por otro lado, en la teoría generalizada de la incertidumbre (Zadeh, 2006) o GTU (*Generalized Theory of Uncertainty*), se conceptualiza a la incertidumbre como un atributo de la información. La información es vista en términos de restricciones generalizadas en los valores que toman las variables. Cualquier restricción que se introduzca (*e.g.* usar conocimiento) reduce la incertidumbre y se considera una fuente de información. Este marco teórico permite la integración de conocimiento para reducir la incertidumbre de la información.

El contexto de este trabajo se basa en el Primer Reporte del Estado del Ciclo del Carbono en México (Paz-Pellat *et al.*, 2019a), realizado por el Programa Mexicano del Carbono (PMC). La parte de los suelos se incluyó en el reporte (Paz-Pellat *et al.*, 2019b) para poder sintetizarlo, en términos de estimaciones del COS, a escala de país. Como parte de la preparación hacia el Segundo Reporte, en este trabajo se analizan los elementos constitutivos para el desarrollo de una estrategia hacia el objetivo planteado de generar escenarios, científicamente sólidos, para evaluar diferentes opciones de manejo, además de los efectos del cambio climático y perturbaciones antropogénicas (Covaleda *et al.*, 2016, 2018, 2019).

Información, estimaciones y modelos: alcances y limitaciones

Información del COS y otras propiedades

En lo general, la mayoría de la información disponible en México respecto al carbono del suelo, se refiere al carbono orgánico total y la textura del suelo (Etchevers, 2013). En mediciones relativamente recientes, realizadas con auto analizadores, no se hace una diferenciación entre el carbono orgánico e inorgánico y solo se mide el total, por lo que se introduce una fuente de incertidumbre no evaluada, que requiere considerarse en futuras estimaciones del COS.

La información disponible de mediciones de COS está asociada a diferentes tipos de muestreo (soporte

geométrico en términos geoestadísticos). Por ejemplo, los muestreos del INEGI (2004) se basan en perfiles de suelo, con un soporte geométrico entre 5 y 10cm a diferentes profundidades (estratos). CONAFOR (2009) utiliza transectos en una de sus unidades de muestreo y presentan un soporte geométrico de superficie de alrededor de 15 m x 15 m, dos transectos perpendiculares de 15m que se interceptan en el centro, cuando fueron analizadas en laboratorio las muestras colectadas en espaciamientos de dos metros en los transectos o corresponde a un soporte geométrico entre 5 y 10 cm, cuando solo se colecto en el punto central del sitio de muestreo. Algo similar ocurre con los datos de SEMARNAT-CP (2002), con soporte geométrico de 20x20x20cm. Por otro lado, FAO-SAGARPA (2012) colectaron muestras de suelo usando dos líneas de muestreo en zigzag, en una parcela agrícola (dimensiones no conocidas). Los diferentes soportes geométricos no pueden compararse entre sí y requieren el uso de técnicas multiescalar para que sean interoperables (Paz-Pellat *et al.*, 2010). En las bases de datos más antiguas del COS, la localización geográfica de los muestreos tiene alta incertidumbre. Por ejemplo, se estima que el error de los muestreos de perfiles del INEGI (2004) es de ± 250 m, para muestras de antes de 1998 y de ± 10 m para el periodo posterior, reduciéndose el error en muestreos más actuales a ± 5 m. Para los datos de CONAFOR, el error de posición geográfica varía entre ± 10 m (primer inventario) y ± 5 m para los posteriores, los errores están asociados a los cambios en la tecnología de los GPS. Para el caso de los datos de SEMARNAT-CP (2002) se estima que tienen un error geográfico de ± 30 m. Estas incertidumbres deben considerarse en las estimaciones del COS. De los análisis realizados por Paz-Pellat *et al.* (2010), el error asociado a diferentes soportes geométricos de muestreo es muy importante y debe considerarse en las estimaciones, de otra manera presentarán un sesgo y alta incertidumbre.

Una de las limitaciones de las bases del COS en México, tales como la del INEGI (2004), es que las campañas de muestreo corresponden a diferentes años (Figura 1). Aunque hay bases de datos (CONAFOR y SEMARNAT-CP) con periodos cortos de muestreo para coberturas nacionales. Se reconoce que esta limitación es común para las bases de COS que reportan los países en las bases de datos internacionales (Batjes *et al.*, 2017).



Figura 1. Años de las campañas de muestreo realizadas en México para la integración de la base de datos del COS por el INEGI (2004).

Las variaciones temporales del COS en tiempos cortos son difíciles de estimar (Post *et al.*, 2001), por lo que las campañas nacionales de muestreo entre 2 y 4-5 años son confiables para el uso conjunto de las mediciones, pero debe tenerse precaución para cambios de uso del suelo contrastantes (*e.g.* de bosque a cultivo), ya que los cambios de corto plazo pueden ser importantes (IPCC, 2006). Asimismo, la incertidumbre asociada a las mediciones y su extrapolación a nivel nacional puede ser mucho más grande que la diferencia aparente entre los resultados de las distintas campañas de muestreo (Guevara *et al.*, 2020a).

Desde la perspectiva de la suficiencia de información, casi todas las mediciones del COS en México están incompletas (falta de datos de densidad aparente y fragmentos gruesos), especialmente en lo que se refiere al momento de pasar de datos de concentración (g C kg^{-1} suelo) a datos de densidad (t C ha^{-1}), lo que puede conducir a sobreestimaciones importantes del COS (Poeplau *et al.*, 2017) y hace indispensable el desarrollo de técnicas precisas y no sesgadas para estimar los datos faltantes.

Para la implementación de modelos de la dinámica del COS, el PMC ha generado bases de datos a nivel nacional, resolución espacial de 250 m (Salas-Aguilar y Paz-Pellat, 2017; Salas-Aguilar y Paz-Pellat, 2018; entre otros). Asimismo, se encuentran disponibles públicamente bases de datos de mediciones del COS

y otras propiedades (Paz-Pellat y Ortiz-Solorio, 2019; Paz-Pellat y Velázquez-Rodríguez, 2018; entre otros).

Estimaciones nacionales del COS

Paz-Pellat *et al.* (2016) presentan una síntesis de las diferentes estimaciones nacionales del COS realizadas en México hasta el 2016. Una actualización de los métodos utilizados sería: a) métodos de imputación de COS a clases de uso del suelo y vegetación o USyV (de Jong *et al.*, 2006, 2009, 2010; Paz-Pellat *et al.*, 2019b) y su distribución a profundidad (Paz y Etchevers, 2016); b) métodos geoestadísticos (Cruz-Cárdenas *et al.*, 2014; Segura *et al.*, 2004) usando la técnica Kriging y la base de datos de SEMARNAT-CP y, c) métodos de aprendizaje de máquina o ML (*Machine Learning*), los cuales representan las estimaciones más actuales (Guevara *et al.*, 2020c; Guevara y Vargas, 2021).

En el caso de los mapas nacionales generados por el proceso de imputación del COS (media y varianza) a clase de USyV (*sensu* INEGI), la Figura 2 muestra los resultados para la profundidad de 0 a 30 y 0 a 100cm, con el uso de la base de datos del INEGI actualizada hasta la fecha de generación de los mapas (2007). Desafortunadamente, por falta de bases de datos independientes e interoperables, no se generaron métricas de incertidumbre y se desconoce su precisión.

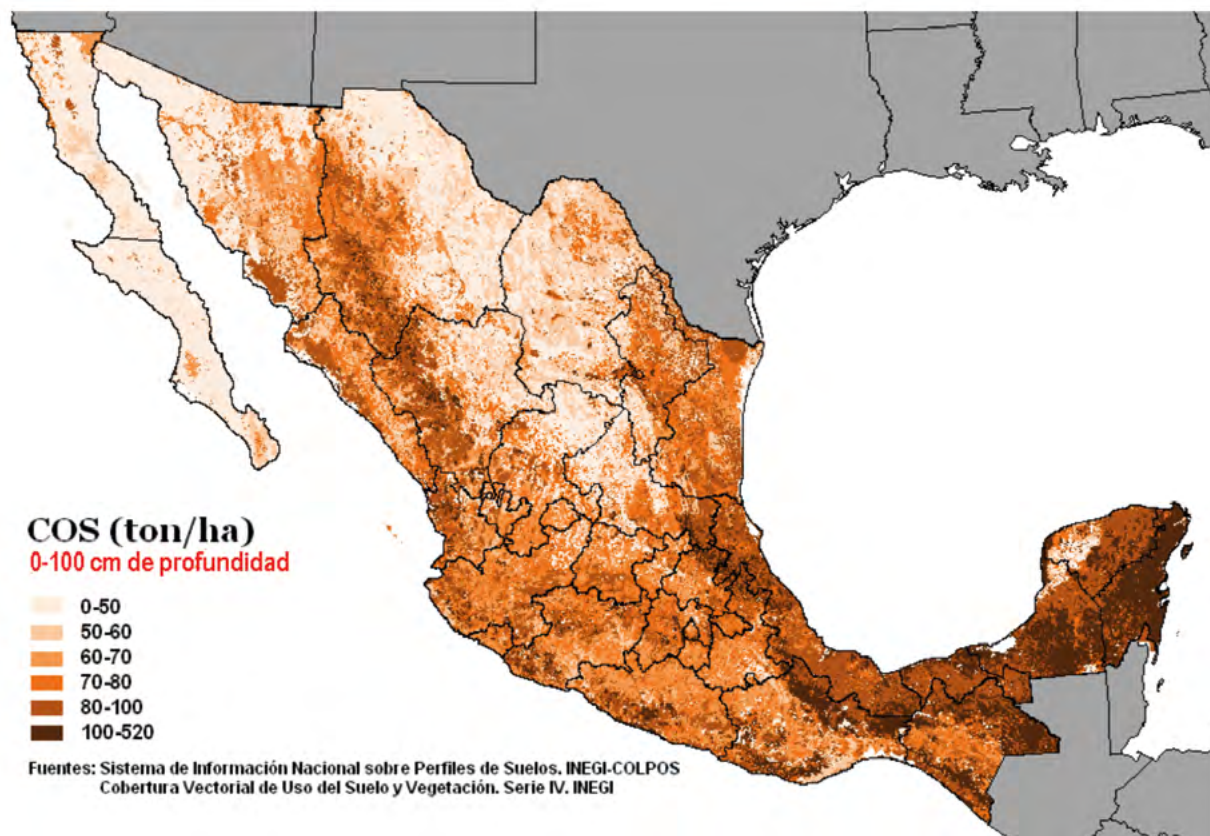
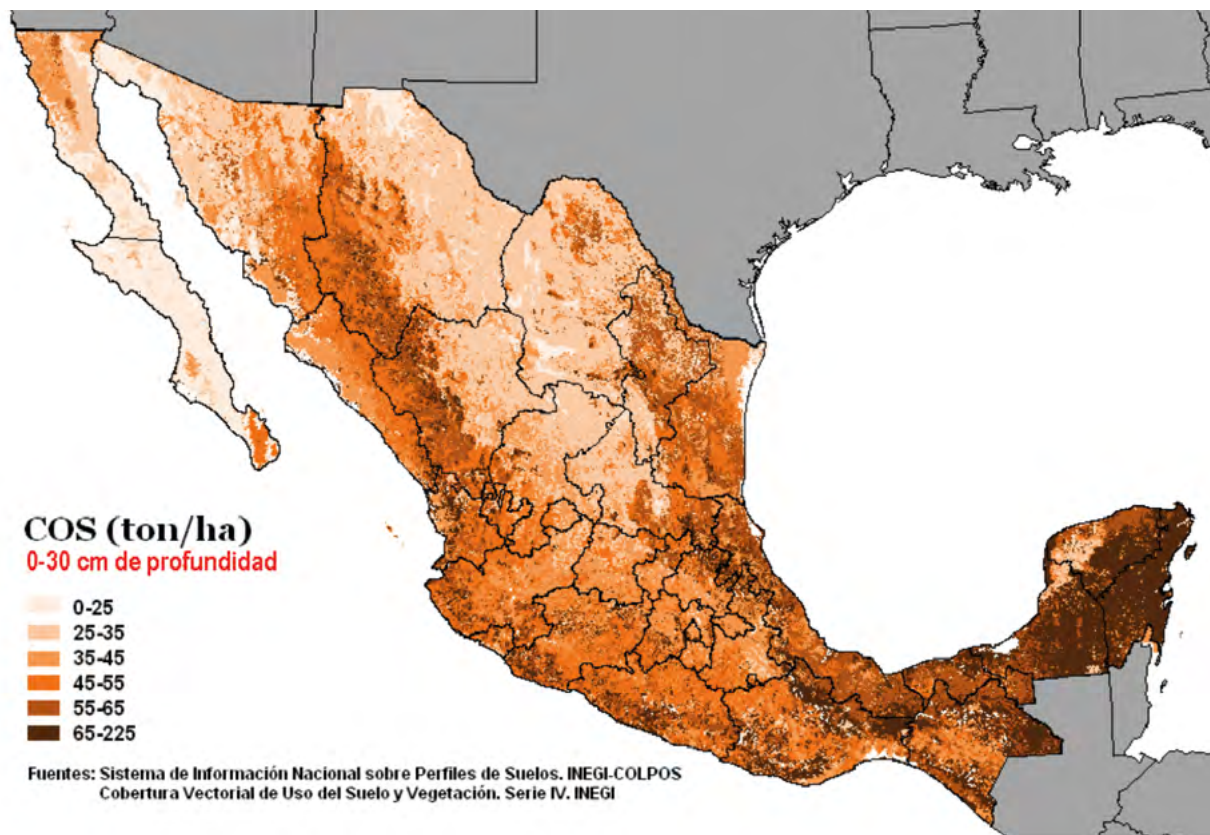


Figura 2. Distribución espacial del COS a la profundidad de 0 a 30 cm y 0 a 100 cm para el año base 2007.

El uso de clases de USyV para imputar los datos medidos del COS se basa en que los ingresos de carbono (mantillo y raíces) condicionan los valores del almacén, aunque esto es dependiente del estado de equilibrio (estado estacionario) de la dinámica de la vegetación.

El uso de técnicas de Kriging ordinario y sus variantes de la teoría geoestadística (Chiles y Delfiner, 2012; Goovaerts, 1997), tiene el inconveniente de suavizar las estimaciones, ya que tiene un dominio

de atracción Gaussiano que simetriza y generalmente sobreestima los valores, se considera que un valor grande tiene una fuerte influencia en los resultados, dado que el Kriging es un estimador global, debido al uso de funciones de covarianza o variogramas a escala global. La Figura 3 muestra un mapa de la distribución del COS en México, en donde se observa que los valores altos del COS influyen en que las estimaciones circundantes adopten también valores altos.

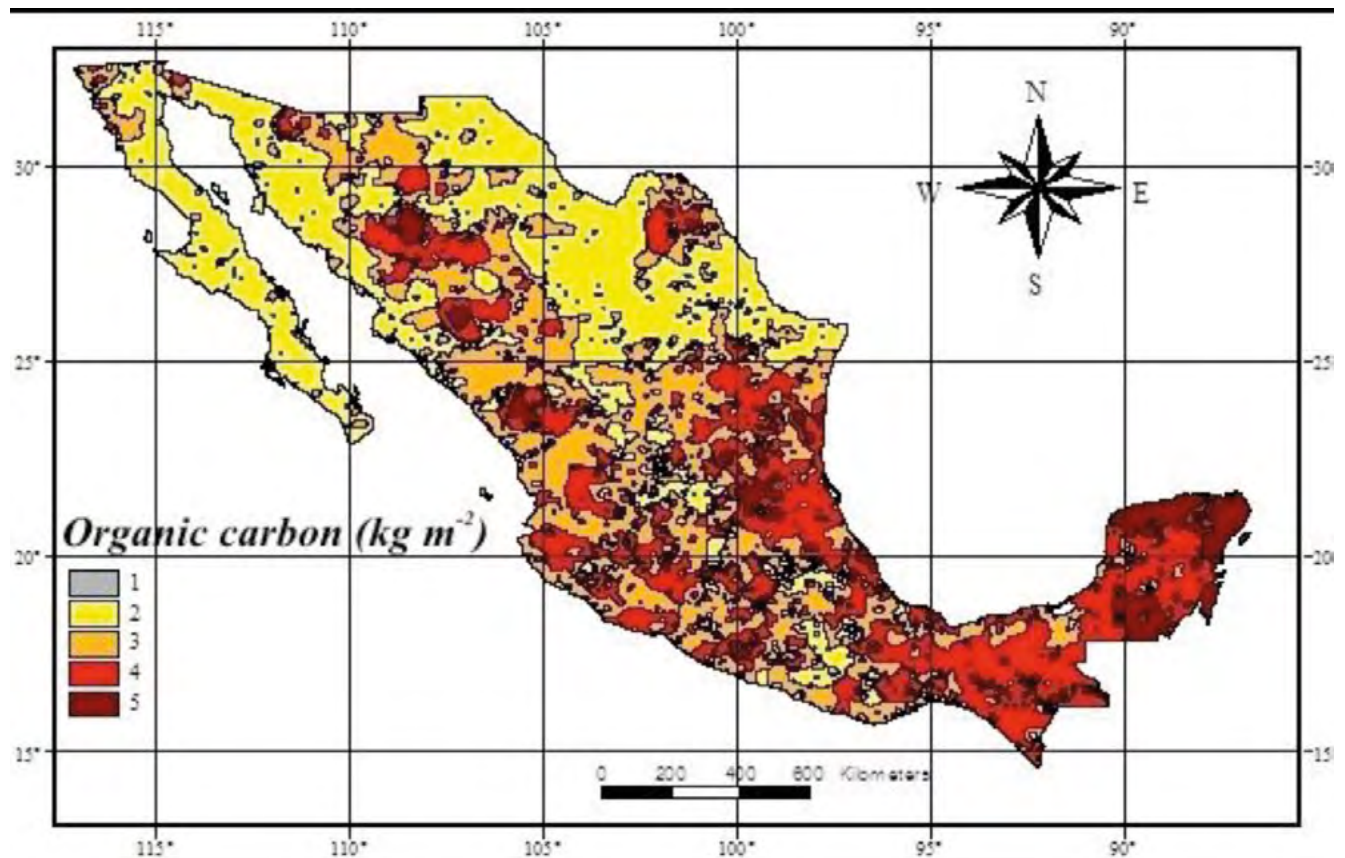


Figura 3. Mapa de la distribución, 0 a 20 cm de profundidad, del COS en México. Fuente: Etchevers *et al.* (2006).

El uso de métodos de ML en la generación de mapas digitales de suelos o MDS, es relativamente reciente con numerosas aplicaciones (Guevara *et al.*, 2018; Ramcharan *et al.*, 2018; Wadoux *et al.*, 2020). En México se han aplicado estos métodos para realizar estimaciones (con incertidumbres asociadas) del COS, con estimaciones indirectas incluidas de la densidad aparente y fragmentos gruesos, a nivel nacional a la profundidad de 0 a 30 cm (Guevara *et al.*, 2020c) y 0 a 100 cm (Guevara y Vargas, 2021). Los mapas

generados están disponibles públicamente (Guevara *et al.*, 2020b; Guevara y Vargas, 2020a), incluidas las métricas de incertidumbre. La FAO (2018) discute la implementación de este tipo de métodos, considerando el uso de software para tal fin.

En la Figura 4 se muestran los resultados de las estimaciones para México y Estados Unidos, mediante el uso de métodos de ML (Guevara *et al.*, 2020c), lo que pone en perspectiva su incertidumbre.

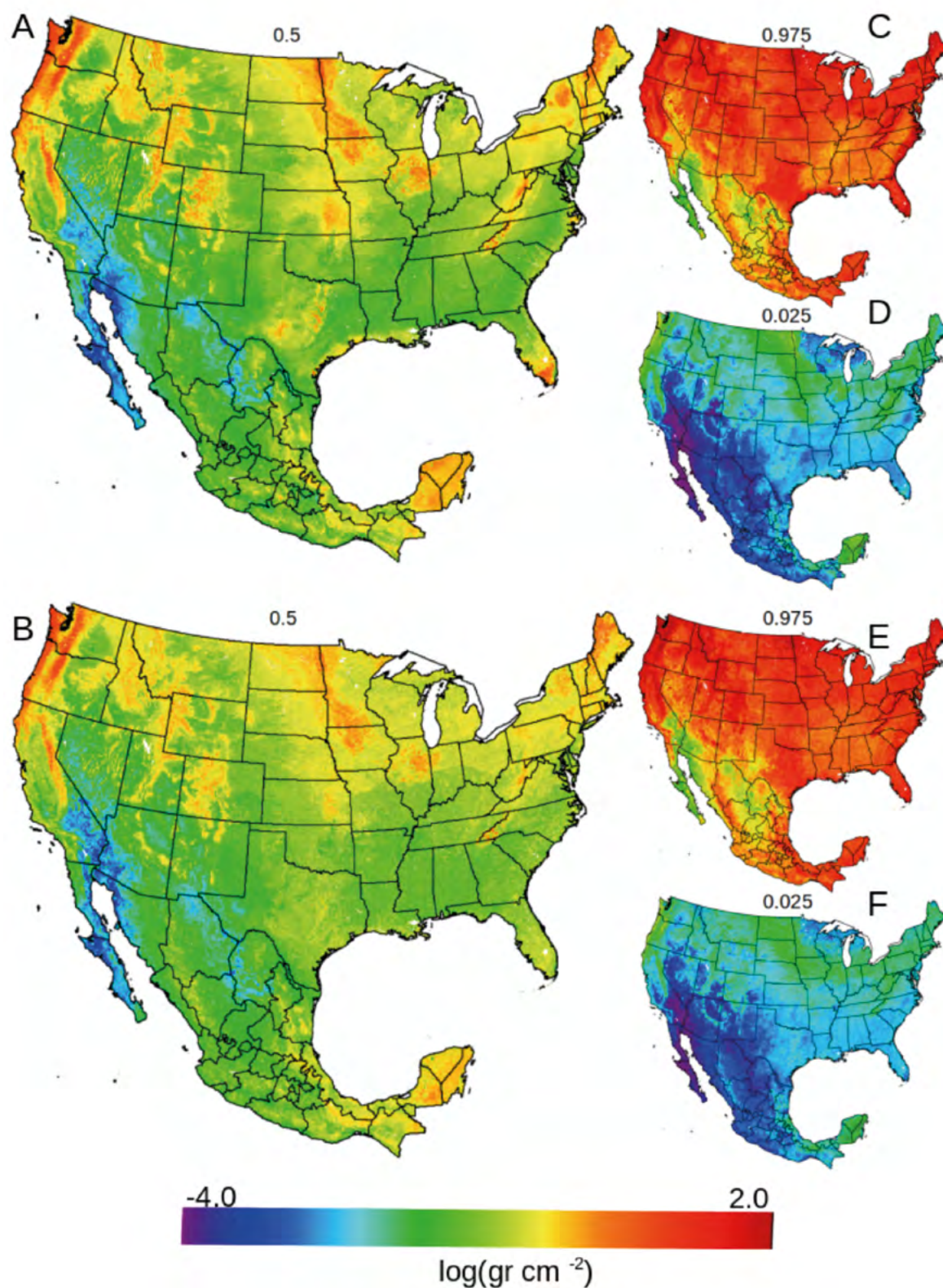


Figura 4. Predicciones espaciales del COS a intervalos de incertidumbre del 95% derivados de regresiones percentiles. (A) conjunto completo de covariados de Soil Grids 250 m (158 covariados) y (B) uso de los mejores covariados correlacionados. Los pequeños mapas (C y E), con intervalos de predicción, superior (0.975) e inferior (0.0265), derivados de la distribución completa condicional del CO a la respuesta de los modelos de la dinámica del COS y su parametrización e hipótesis intrínsecas. Fuente: Guevara *et al.* (2020).

La generación del COS a nivel nacional no es suficiente para el desarrollo de escenarios de intervención, por lo que es necesario el acoplamiento de modelos con los datos generados, de tal forma que el condicionamiento del uso del modelo define los requisitos de información requerida en las estimaciones, lo que generalmente se considera *a posteriori*. Modelos como el RothC se han propuesto para implementación a escala regional (FAO, 2020a), por lo que la discusión de este apartado se centrará principalmente en ese modelo.

El modelo RothC está constituido de cinco almacenes del COS: material de plantas susceptible de descomposición (DPM, por sus siglas en inglés), material de plantas resistente (RPM), biomasa microbiana (BIO), materia orgánica humificada (HUM) y materia orgánica inerte (IOM). Estos almacenes son “virtuales” y de cinéticas de primer orden (lineales). El almacén IOM es inerte al ataque biológico, con un tiempo medio de residencia de 50 000 años. Para inicializar el RothC, así como otros modelos, es necesario conocer los valores de carbono orgánico (CO) de todos los almacenes, más los ingresos o I (mantillo y raíces) de CO al suelo. Sin esta información, el proceso de inicialización de los modelos se hace con simulaciones inversas (hacia el pasado), para estimar I y resto de los almacenes (Dimassi *et al.*, 2018; Foereid *et al.*, 2012; Ludwig *et al.*, 2010) bajo la hipótesis de que el COS se encuentra en estado estacionario (ganancias = pérdidas). La propuesta de invertir el modelo para estimar I, conocido el IOM, se ha utilizado en diferentes aplicaciones o iniciativas (FAO, 2020a; Gottschalk

et al., 2012; Meersmans *et al.*, 2013). El uso de la hipótesis de estado estacionario del COS inicial sin validarla genera errores importantes en las estimaciones de los modelos (Foereid *et al.*, 2012; Sanderman y Baldock, 2010), en donde la estimación de I es la mayor fuente de incertidumbre (Ludwig *et al.*, 2010). Para estimar I, es necesario conocer el IOM, dado que (COS-IOM) debe usarse para la estimación. Aun cuando existen propuestas para estimar IOM en función de la textura del suelo (Falloon *et al.*, 1998; Körschens *et al.*, 1998; Paz *et al.*, 2016; Rühlmann, 1999), no tienen fundamento, ya que el IOM es independiente de la textura. En las propuestas de Rühlmann (1999) y Paz *et al.* (2016) la textura del suelo define límites en términos generales, por lo que puede utilizarse bajo esta consideración. Una alternativa al uso de la textura, es contemplar estrato secuencias (perfil del COS a diferentes profundidades) para estimar el COI (carbono orgánico inerte) para diferentes tipos y grupos de vegetación en México *sensu* INEGI (Paz y Etchevers, 2016).

La simulación inversa de los modelos del COS para estimar algunos de los parámetros requeridos resulta en el problema de equifinalidad, por lo que es difícil de usar en calibraciones (Taghizadeh-Toosi *et al.*, 2014). La equifinalidad plantea que pueden obtenerse resultados iguales al usar diferentes modelos (diferentes estructuras) o conjunto de parámetros en un mismo modelo (Beven y Freer, 2001; Marschmann *et al.*, 2019). La Figura 5 muestra diferentes condiciones iniciales para obtener un mismo COS inicial (*e.g.* Caruso *et al.*, 2018).

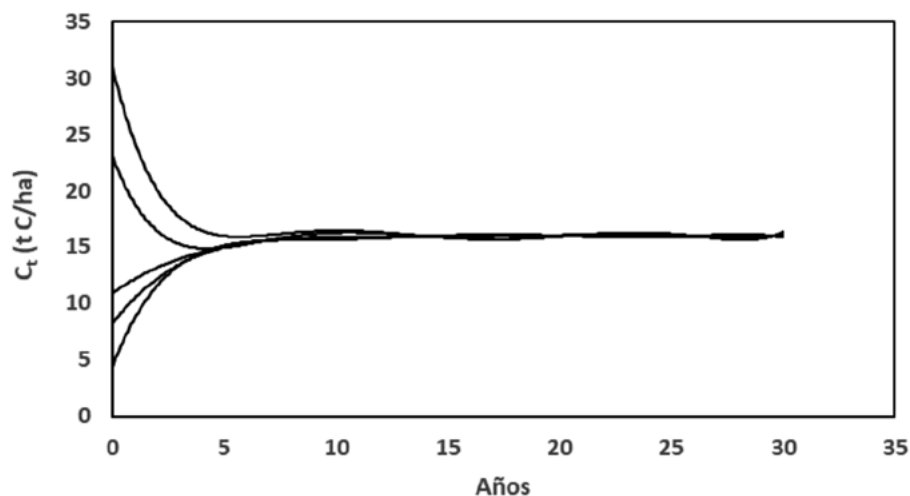


Figura 5. Diferentes condiciones para la convergencia en un valor único del carbono orgánico total (C_t).

El requerimiento de la hipótesis de estado estacionario se utiliza en modelos más simples (IPCC, 2006; Ogle *et al.*, 2005) y, en México, constituye la base de los Modelos de Estados y Transiciones o METs (Covaleda *et al.*, 2016, 2019). Para definir un estado estacionario del COS es necesario, al menos conocer una métrica al respecto o contar con dos mediciones del COS en tiempos diferentes. Un panorama de la dinámica del COS dada su dependencia del I es analizar series temporales de USyV en el pasado, pero en México esta posibilidad está acotada por la escala disponible de los mapas (1:250 000) de tal manera que es necesario generar mapas de USyV al menos a la escala 1:50 000 para este fin (Paz *et al.*, 2020). El valor del COS (C_t) asociado a los cambios de uso del suelo y vegetación,

tiene asociado al menos dos procesos simultáneos (absorción por el nuevo uso y descomposición por el uso anterior), por lo que los patrones temporales (Paz-Pellat y Velázquez-Rodríguez, 2020) requieren de conocer el tiempo en el que ocurrió el cambio, lo que se puede aproximar con el uso de mapas de USyV de mayor detalle.

En la Figura 6 se muestran los resultados de la aplicación del modelo RothC en experimentos de corto y medio plazo (< 20 años) en México, con la hipótesis de estado estacionario y la estimación de IOM con el uso de la textura del suelo (Falloon *et al.*, 1998), en donde se observa que para valores bajos del COS los resultados son buenos, pero no así para valores medios y altos.

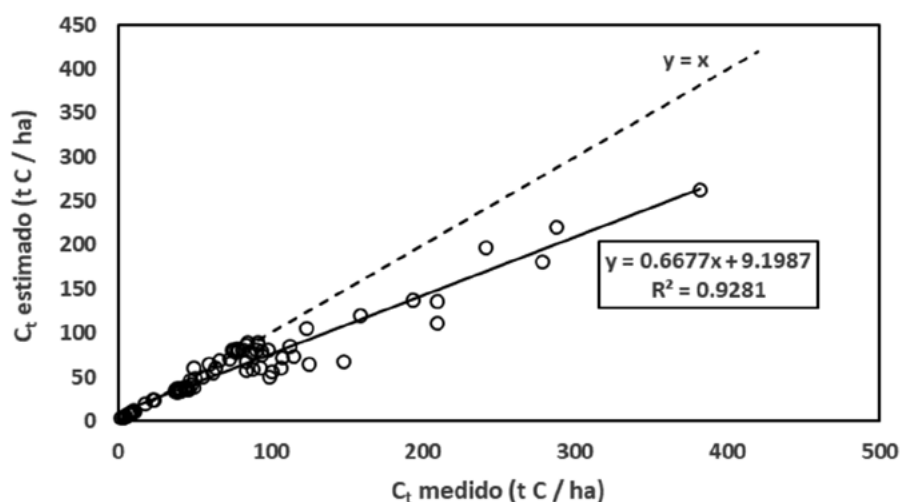


Figura 6. Resultados de la aplicación del modelo RothC para el caso de mediciones promedio.

Una alternativa interesante para establecer métricas relativas al estado estacionario, es usar el concepto de saturación del COS (Carter *et al.*, 2003; Gulde *et al.*, 2008; Matus *et al.*, 2016; Six *et al.*, 2002; Stewart *et al.*, 2007) de modo que el valor de COS no cambia al aumentar I. Diferentes autores proponen relaciones en COS a saturación, con la textura del suelo, por lo que es posible usar este enfoque, aunque la evidencia experimental no muestra relaciones únicas (Feng *et al.*, 2013; Hassink, 1997; Matus, 2021). Paz *et al.* (2016) proponen relaciones entre la textura y el COS máximo aparentemente saturado y el COS mínimo (aparentemente el COI) por tipos de USyV de México, por intervalo de variación y métricas asociadas, para el máximo y 0 para el mínimo, lo que puede usarse

para establecer qué tan cerca o lejos se está del estado estacionario, bajo el supuesto de que el COS máximo y el COS mínimo, están en este estado.

Hacia una estrategia armonizada para el segundo reporte del PMC

Mapas del COS e incertidumbres

Para armonizar los diferentes métodos de generación de mapas del carbono orgánico de los suelos, en lo siguiente se presentan los métodos de ML y geoestadística indicador Bayesiana, en donde este último se utiliza para fusionar ambos enfoques.

Métodos de aprendizaje automático (ML).

Los métodos de ML están orientados a la búsqueda de estructuras de relaciones lineales o no lineales, libres de hipótesis estadísticas con el uso de covariados, variables indirectas relacionadas con el COS para extraer patrones que reduzcan las métricas de incertidumbre definidas. Al respecto se ha utilizado una gran variedad de covariados (Grunwald, 2009; Lamichhane *et al.*, 2019; Rasel *et al.*, 2017), con diferentes marcos conceptuales de los cuales los principales para definir covariados son cuatro (Ma *et al.*, 2019): primero, el Modelo CLORPT de Jenny (1941) de los factores formadores del suelo (S), el cual se planteó producir conocimiento:

$$S = f(cl, o, r, p, t) \tag{1}$$

en donde cl = clima, o = organismos, r = relieve, p = material parental y t = tiempo.

El segundo modelo es de SCORPAN de McBratney *et al.* (2003):

$$S = f(s, c, o, r, p, a, n) \tag{2}$$

en donde S son atributos o propiedades del suelo, s es el suelo u otro atributo del suelo o clases, c = clima, o = organismos, incluida la vegetación, r = topografía (relieve), p = material parental, a = edad o tiempo, n = localización espacial o posición geográfica.

Cabe destacar que el modelo SCORPAN no se formuló para explicar la formación del suelo, sino que se trata de un modelo que es pragmático, que permite predecir propiedades del suelo o sus clases.

En tercer lugar, se encuentra el Modelo STEP-AEWBH (Grunwald *et al.*, 2011), en donde la propiedad o atributo del suelo (SA) es:

$$SA = f(s, t, e, p - a, e, w, h) \tag{3}$$

donde s = propiedad auxiliar del suelo, t = topografía, e = propiedades ecológicas, p = material parental, a = propiedades atmosféricas, w = propiedades hídricas, b = propiedades bióticas y h = forzamientos inducidos por humanos.

Finalmente, se encuentra el Modelo de Simonson (1959), para la evolución de los tipos de suelos:

$$S = f(\text{adición, remoción, traslocación, transformación}) \tag{4}$$

La necesidad de considerar conocimiento pedológico en los métodos de ML se ha enfatizado por diversos autores (Grunwald, 2009; Ma *et al.*, 2019; Wadoux *et al.*, 2020), lo cual se ha llevado a cabo agregando pseudo puntos de conocimiento experto (Hengl *et al.*, 2017) en la base de datos utilizada. El uso de mapas edafológicos permite extraer conocimiento de la cartografía de una propiedad o atributo del suelo (Bui, 2004; Ma *et al.*, 2019).

Generalmente, los métodos de ML no consideran la auto-correlación espacial, aunque al incorporar las coordenadas geográficas de los puntos de los covariados incorporan esta información en términos genéricos (Wadoux *et al.*, 2020). Por otro lado, los resultados definidos por métodos de ML enfatizan generalmente métricas de precisión y no las incertidumbres asociadas, por lo cual para establecer métricas de incertidumbre se han utilizado resultados de modelos múltiples disjuntos con diferentes conjuntos de covariados o modelos y técnicas de regresión de cuantiles, las cuales tienen algunas limitaciones.

Tal como se discutió anteriormente, en la Figura 1 los datos del COS de México y otras partes del mundo, tienen un amplio intervalo de medición por lo que es necesario implementar consideraciones espacio-tiempo en las estimaciones de los métodos de ML (Heuvelink *et al.*, 2020).

Los métodos de ML se han enfocado en el nuevo conocimiento (Bui, 2016; Molnar *et al.*, 2010; Wadoux *et al.*, 2020), para lo que resulta crítica la selección de covariados que sean plausibles en función del marco conceptual que se utilice y requiere del dominio del uso dado a ML (Roscher *et al.*, 2020). Wadoux *et al.* (2019) destacan que se pueden generar resultados precisos similares a los de ML con el uso de covariados que no tengan algo que ver con las propiedades o atributos del suelo definidos, por lo que es necesario tener precaución con los resultados.

Un aspecto crítico en el uso de métodos de ML es la interpretación y explicación de los resultados obtenidos (Molnar *et al.*, 2010; Roscher *et al.*, 2020; Wadoux *et al.*, 2020) que son función de los covariados, algoritmos usados y parámetros definidos internamente por los métodos ML, limitando su transparencia (modelos tipo caja negra). Esto también es aplicable a métodos de aprendizaje profundo (Deep learning; LeCun *et al.*, 2015). El uso de covariados plausibles y congruentes con el marco conceptual adoptado en un formato de planteamiento de hipótesis, define un camino a la utilización de ML.

Métodos de geoestadística indicador Bayesiana.

La FAO (2018) ha propuesto el Kriging de regresión, además de métodos de ML, para uso en la generación de mapas de COS. Este tipo de Kriging tiene limitaciones importantes al no considerar covariados y otro tipo de conocimiento.

El problema de estimación espacial de la

geoestadística puede ponerse como la determinación de los pesos λ_i para estimar la variable Z en el punto X_0 (Figura 7):

$$Z^*(X_0) = \sum_{i=1}^n \lambda_i Z(X_i) \tag{5}$$

en donde el uso del signo * significa valor estimado (sin información).

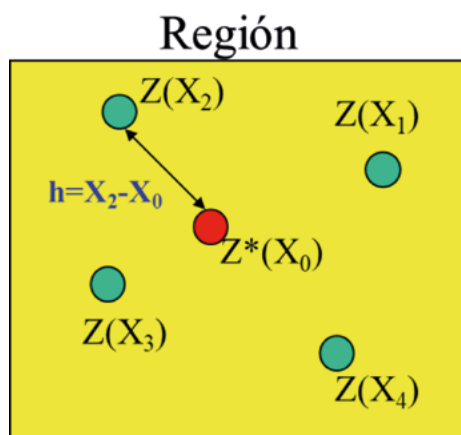


Figura 7. Problema de estimación geoestadístico.

Para determinar los pesos λ_i se utilizan funciones de estructura las cuales son momentos estadísticos de segundo orden de la distribución bivariada

$Z(x)-Z(x+h)$ (Figura 8), de variables separadas entre sí a una distancia h cualquiera:

$$\text{Variograma: } 2\gamma(h) = E[\{Z(X) - Z(X + h)\}^2] \tag{6}$$

$$\text{Covariograma: } C(h) = E[Z(X)Z(X + h)] - E[Z(X)]E[Z(X + h)] \tag{7}$$

donde $E[]$ es el operador de esperanza matemática promedio. Los diferentes modelos *teóricos* de covariogramas y variogramas se encuentran en Chiles y Delfiner (2012).

En esencia, las funciones variograma y covariograma hacen operativo el concepto de que muestras cercanas al punto de estimación tienen mayor correlación que las alejadas.

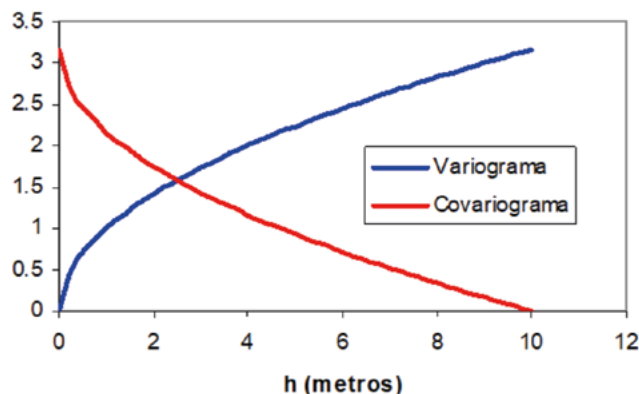


Figura 8. Función variograma y covariograma.

El problema de estimación denominado técnica Kriging en geoestadística se plantea como:

$$\text{Minimizar } E\{Z(X) - Z^*(X)\}^2 \quad (8)$$

$$\text{Sujeto a } \sum \lambda_i = 1$$

La solución del sistema 4 produce estimaciones precisas con varianza de estimación mínima e insesgadas, $E[Z^*(X)] = E[Z(X)]$, cumpliendo los requisitos del IPCC (2006). En caso de desconocimiento de la media del proceso, la solución del sistema 8 se denomina Kriging ordinario (KO).

Un problema similar de estimación, lo constituye el uso de una variable indirecta $Y(X)$ correlacionada con la de interés, en el sistema Co-Kriging ordinario, el cual estima los pesos de acuerdo con:

$$Z^*(X_0) = \sum_{i1=1}^{n1} \lambda_{i1} Z(X_{i1}) + \sum_{i2=1}^{n2} \lambda_{i2} Y(X_{i2}) \quad (9)$$

que al igual que en el Kriging ordinario, los pesos se obtienen usando un sistema similar al 8.

La desventaja del Co-Kriging es que requiere del conocimiento de los auto-variogramas de $Z(X)$ y $Y(X)$ y los variogramas cruzados entre $Z(X)$ y $Y(X)$, lo que limita sus aplicaciones dados los requerimientos de información.

Los interpoladores Kriging, y Co-Kriging, tienen la propiedad de que son exactos; esto es, la estimación en un punto donde existe una medición, genera el valor medido. Estos interpoladores son del tipo global ya que usan la información de toda la región de interés, sin discriminarla en valores particulares y generalmente producen estimaciones suavizadas y sobre estimaciones que no representan la variabilidad del fenómeno bajo estudio; además de que los errores de estimación son dependientes de la configuración geométrica del muestreo (Goovaerts, 1997). Según Journal (1983) un problema asociado a las técnicas Kriging es que su dominio de atracción es Gaussiano lo que introduce simetrías en forma intrínseca.

El uso del Kriging y sus variantes, aparte del dominio de atracción Gaussiano, produce estimaciones que no consideran lo local; es decir, usa el mismo proceso para valores altos de la variable de interés en zonas de influencia pequeñas; por ejemplo, mediciones del COS debajo de árboles y valores bajos en zonas de influencia altas; por ejemplo, mediciones del COS entre arbustos/ árboles en la sabana. Por esta razón, los valores altos “pepitas de oro” generan sobre estimaciones marcadas

al influenciar los valores alrededor de los *datos fuera del promedio*.

Para realizar estimaciones tipo local y no lineales, Journal (1983) propuso el uso de funciones indicadores:

$$I_{z_i(X)} = \begin{cases} 1, & \text{si } Z(X) \leq z_i \\ 0, & \text{si } Z(X) > z_i \end{cases} \quad (10)$$

donde z_i representa un valor umbral definido. El uso de funciones indicadores es utilizado por el Kriging indicador (Journal, 1983) para el sistema:

$$\text{Minimizar } E\{I(X) - I^*(X)\}^2 \quad (11)$$

$$\text{Sujeto a } \sum \lambda_i = 1$$

El uso de variogramas indicadores permite la estimación de la distribución de probabilidad completa en un punto dado (Figura 9) no sólo el promedio.

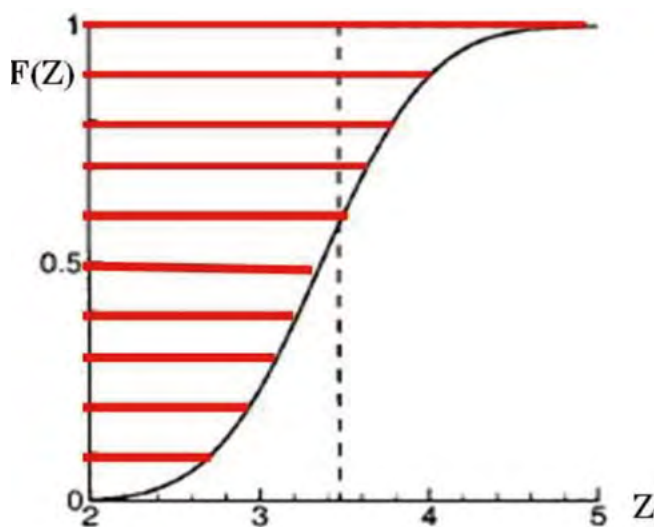


Figura 9. Estimación de la distribución de probabilidad de una variable en un punto dado, al discretizar los valores de Z para diferentes valores z_i y construir indicadores.

Con el uso del enfoque de funciones indicadoras, Journal (1986) propuso el Kriging blando para integrar información *dura* que son mediciones de la variable de interés con *blanda* que son mediciones indirectas con otras variables correlacionadas o covariados a través del Co-Kriging indicador, con mayores requerimientos de información que el sistema normal, tales como auto-variogramas y variogramas cruzados indicadores.

Alabert (1987), bajo la perspectiva del Teorema de Bayes, planteó el desarrollo del (Co-) Kriging

indicador Bayesiano, como un esquema para integrar información blanda, a través de un proceso de actualizar probabilidades *a priori*, con nueva información para estimar probabilidades *a posteriori*.

Para caracterizar la incertidumbre con el uso de covariados, se definen las probabilidades p_1 y p_2 , las cuales pueden estimarse como se muestra en la Figura 10. La información de la decisión se captura por la función $B(z_i) = p_1 - p_2$, $p_1 > p_2$ si $p_1 = 1$ ($p_2 = 0$), entonces $B(z_i) = 1$ y la relación entre $Z(X)$ y $Y(X)$ es una relación lineal perfecta. Si $p_1 = p_2$, entonces $B(z_i) = 0$ y no existe una relación (coeficiente de determinación = 0) entre $Z(X)$ y $Y(X)$.

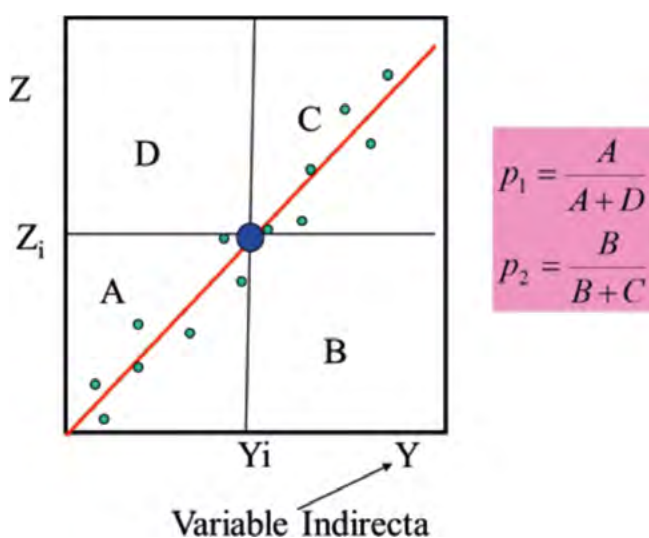


Figura 10. Estimación de las probabilidades p_1 y p_2 , donde A, B, C y D se refieren al conteo de los datos en cada área rectangular definida.

Las estimaciones del Kriging indicador Bayesiano (KiB) requieren de un sistema Co-Kriging, el cual puede simplificarse con el uso de diferentes hipótesis Markov-Bayes (Journel, 1999).

El uso del Kriging indicador Bayesiano necesita que las variables medidas, directas e indirectas que sean formateadas en forma adecuada (Zhu y Journel, 1993). La información blanda, marco de la teoría de probabilidad clásica puede representar: (a) un intervalo de valores, (b) condición de valores menores o iguales a valor, (c) condición de valores mayores a un valor, (d) una distribución de probabilidad asociada a un tipo o clase. Desde la perspectiva de la teoría generalizada de la incertidumbre (Zadeh, 2006) los diferentes formatos de los datos para el KiB, definen restricciones basadas en conocimiento que se deben cumplir, por lo que el

conocimiento se introduce en formato probabilístico, aun cuando es posible utilizar otras formulaciones (Klir, 2006).

En geoestadística se contempla el uso de simulaciones estocásticas condicionadas de funciones indicadoras (Alabert, 1987) como una forma equivalente a realizar estimaciones de distribuciones de probabilidad usando Kriging indicador. Al igual que el Kriging, las simulaciones son condicionadas a los datos duros y blandos para reproducir la incertidumbre asociada.

El uso de información blanda multi-fuente puede analizarse mediante el KiB o simulaciones en forma desintegrada, o bien agregando las fuentes de información blanda covariada en una sola fuente (e.g. resultados de métodos ML). Desde esta perspectiva, los métodos de ML pueden ser en un análisis exploratorio para definir covariados (Bui, 2016). En lo general, el uso de métodos de ML puede integrarse en el KiB como otra fuente más de datos y con incertidumbre asociada, función $B(z)$, lo que amplía su uso.

En la generación de mapas del COS es práctica común mezclar datos con diferentes fechas de muestreo (Guevara *et al.*, 2020b), por lo que no se considera al tiempo en forma explícita. Las estimaciones espacio-tiempo con técnicas geoestadísticas, es una práctica relativamente común (Kyriakidis y Journel, 1999; Montero *et al.*, 2015), en donde el tiempo se considera en las funciones de estructura espacio-tiempo (Kyriakidis y Journel, 1999). Los desarrollos que se muestran para las técnicas geoestadísticas espaciales pueden expandirse sin mayor problema al caso de las estimaciones en espacio-tiempo con el KiB.

La teoría geoestadística discutida previamente, se basa en estadísticas de dos puntos separados por una distancia h cualquiera, por lo que está limitada a momentos de segundo orden tal como el variograma y covariograma y, por lo tanto, no puede reproducir estadísticas multi-punto (MP) (Journel, 2005). La extensión de las estadísticas de dos puntos a las multi-punto permite definir medidas de conectividad entre objetos (Krishman y Journel, 2003) para reproducir en mejor forma un determinado paisaje o agrupación de objetos. El problema con la geoestadística MP es que requiere información que difícilmente está disponible en las aplicaciones. Ante esta limitación se ha propuesto el uso de imágenes de entrenamiento (e.g. mapas de USyV o tipos de suelo o propiedades del suelo; Bui, 2004) para extraer la información para la caracterización MP y realizar simulaciones estocásticas

condicionadas (Journal, 2005; Ortiz y Deutsch, 2004). Con la integración de información blanda multi-fuente (Journal, 2002) usando la hipótesis de permanencia de las razones de actualización asociadas al Teorema de Bayes es posible incorporar esta información en las simulaciones estocásticas MP (Liu *et al.*, 2005; Ortiz y Deutsch, 2004).

Una limitación, para el caso de las estimaciones geoestadísticas MP del COS, es que los mapas de USyV son dinámicos; aunque es posibles utilizar los tipos de suelos y sus propiedades asociadas para reproducir los patrones espaciales y temporales del COS en los objetos caracterizados.

Dinámica del COS y estado estacionario

Para entender la propuesta de cuantificar el estado estacionario del COS, es necesario introducir el modelo COLPOS (Paz *et al.*, 2016) de la distribución del carbono orgánico por tamaño de las fracciones físicas del suelo, bajo la condición de dispersión total del suelo. La Figura 11 muestra la distribución del carbono orgánico total (C_t), el particulado (C_p) y el órgano mineral (C_m), donde $C_p = C_t - C_m$. El COI es el carbono orgánico inerte y C_{50} es el carbono orgánico asociado a la fracción arcilla más limo.

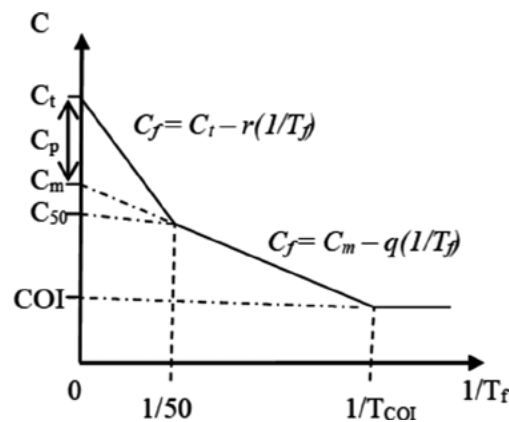


Figura 11. Distribución del carbono orgánico por fracciones (C_f) con relación al tamaño de las fracciones (T_f) del modelo COLPOS.

El almacén órgano-mineral tiene una capacidad C_{SM} (saturación potencial del carbono órgano-mineral) y el particulado C_{SP} (saturación potencial del carbono orgánico particulado).

En términos de un suelo particular, fraccionamiento físico, la relación entre el carbono orgánico particulado y el total está dado por $C_p = C_t - C_m$ (Figura 11).

La Figura 12, caso de estudio que se discute en Paz *et al.* (2016) muestra que a medida que la vegetación está menos perturbada, es decir en estado natural o cercano a él, la pendiente de las líneas rectas tiende a 1.0 y el intercepto con el eje C_t es mayor. Al incrementarse C_t , C_p se incrementa (el total). Las relaciones lineales pueden plantearse como:

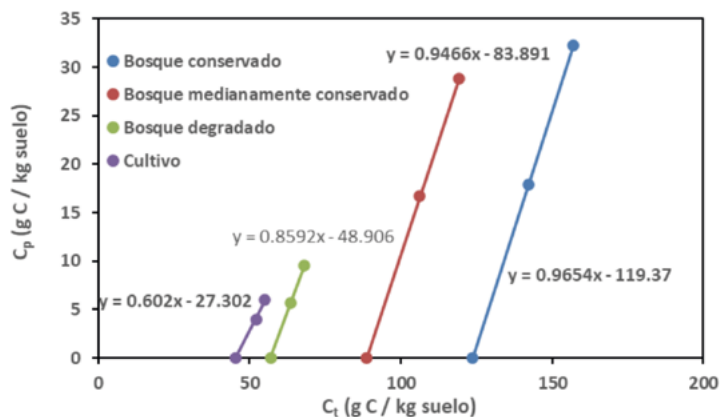


Figura 12. Relación experimental entre C_p y C_t de las fracciones de C_p para andosoles con diferentes grados de perturbación y uso del suelo.

$$C_p = -A + BC_t \quad (13)$$

donde los parámetros A y B están definidos como;

$$\frac{A}{B} = C_{50} \quad (14)$$

$$B = \frac{C_p}{C_t - C_{50}} = \frac{C_t - C_m}{C_t - C_{50}} \quad (15)$$

La relación A/B define el punto de intersección de las líneas rectas con el eje C_t . La pendiente B mide el grado de saturación de los suelos, en donde un valor cercano a 1.0 implica que los almacenes órgano-mineral y particulado no tienen capacidad para absorber más C. En esta situación $C_{SM} = C_{50}$.

En la Figura 13 se muestra un análisis preliminar de la base de datos de suelos mexicanos, con dispersión total, discutidos en Paz *et al.* (2016).

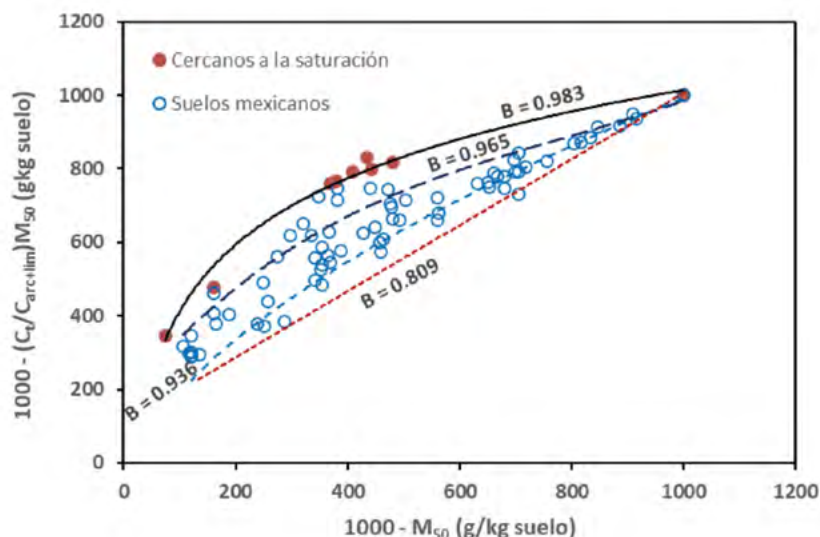


Figura 13. Relación experimental textura y la relación $C_t/C_{arc+lim}$, en donde se muestran suelos mexicanos cercanos a la saturación, además de otros valores.

De acuerdo con la Figura 13, si se conoce la textura del suelo y la razón C_t/C_{50} , es posible conocer el valor de B y evaluar la cercanía o alejamiento al estado estacionario ($B = 1.0$). Desde el punto de vista de conocimiento de dominio y las discusiones previas, un modelo conceptual general para estimar el COS consiste de dos pasos. El primero es estimar el potencial de saturación del carbono órgano-mineral para una clase de USyV, para una textura del suelo y, el segundo, consiste estimar la razón C_t/C_{50} . Para las aplicaciones de generar escenarios es necesario el desarrollo de un modelo de la dinámica del COS que se parametrizó con los datos definidos.

COMENTARIOS FINALES

Los elementos propuestos para el desarrollo de una estrategia y su marco teórico y metodológico, plantea una hoja de ruta para nuevos desarrollos y

requerimientos de información para la implementación de modelos de generación de escenarios de acciones de intervención o de otro tipo.

La versión metodológica de uso del Kriging indicador Bayesiano, permite la integración de información multi-fuente, incluida la generada por métodos de aprendizaje de máquinas, lo que permite un marco teórico general para armonizar diferentes enfoques en uso. Asimismo, el planteamiento de la saturación de almacenes del carbono orgánico para definir el estado estacionario del COS permite la parametrización directa de modelos de la dinámica del COS y el uso de modelos como RothC sin hipótesis implícitas.

Un punto importante de considerar es la superficie de los suelos en México, ya que las estimaciones consideran el área proyectada sobre una superficie horizontal y no la asociada a la orografía del país.

La consideración de la escala en las estimaciones del COS no se abordó en este trabajo, pero forma parte de las líneas de desarrollo a explorar en el futuro.

LITERATURA CITADA

- Alabert, F. 1987. Stochastic Imaging of Spatial Distributions using Hard and Soft Information. MSc Thesis. Stanford University. Stanford, CA.
- Banco Mundial, CIAT y CATIE. 2014. Agricultura climáticamente inteligente en México. Serie de perfiles de agricultura climáticamente inteligente para América Latina. Washington, D.C. 12 p.
- Batjes, N. H., E. Ribeiro, A. van Oostrum, J. Leenaars, T. Hengl and J. M. de Jesús. 2017. WoSIS: providing standardized soil profile data for the world. *Earth System Science Data* 9:1-14.
- Beven, K. and J. Freer. 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *J. Hydrol* 249:11-2.
- Bui, E. N. 2004. Soil survey as a knowledge system. *Geoderma* 120:17-26.
- Bui, E. N. 2016. Data-driven critical zone science: a new paradigm. *Science of Total Environment* 568:587-593.
- Carter, M. R., D. A. Angers, E. G. Gregorich and M. A. Bolinder. 2003. Characterizing organic matter retention for surface soils in eastern Canada using density and particle size fractions. *Can. J. Soil Sci* 83:11-23.
- Caruso, T., F. T. De Vries, R. D. Bardgett and J. Lehmann. 2018. Soil organic carbon dynamics matching ecological equilibrium theory. *Ecology and Evolution* 8:11169-11178.
- Chiles, J. P. and P. Delfiner. 2012. *Geostatistics, Modeling Spatial Uncertainty*. Second Edition. John Wiley and Sons. Inc New Jersey. 726 p.
- Coleman, K. and S. D. Jenkinson. 1996. RothC-26.3: a model of the turnover of carbon in soil. pp.237-246. *In*: Powlson D. S., P. Smith and J. U. Smith (eds.). *Evaluation of Soil Organic Matter Models using Existing Long-term Datasets*. Springer-Verlag. Berlin.
- CONAFOR. 2009. Manual y Procedimientos para el Muestreo de Campo (Re-Muestreo 2009). Inventario Nacional Forestal y de Suelos. Comisión Nacional Forestal. Guadalajara, Jalisco, México. 140 p.
- CONAFOR-SEMARNAT. 2014. National Forest reference emission level proposal Mexico. Report prepared for United Nations Framework Convention on Climate Change. Comisión Nacional Forestal – Secretaría del Medio Ambiente y Recursos Naturales. México. 48 p.
- CONAFOR-SEMARNAT. 2015. National Forest reference emission level proposal Mexico (modified version). Report prepared for United Nations Framework Convention on Climate Change. Comisión Nacional Forestal – Secretaría del Medio Ambiente y Recursos Naturales. México. 68 p.
- CONAFOR-SEMARNAT. 2020a. Nivel de Referencia de Emisiones Forestales de México (2007-2016). Comisión Nacional Forestal – Secretaría del Medio Ambiente y Recursos Naturales. México. 101 p.
- CONAFOR-SEMARNAT. 2020b. Nivel de Referencia de Emisiones Forestales de México (2007-2016), Anexo Metodológico. Comisión Nacional Forestal – Secretaría del Medio Ambiente y Recursos Naturales. México. 94 p.
- Covalada, S., F. Paz y A. Ranero. 2016. Carbono edáfico en Chiapas: planteamiento de políticas públicas de mitigación de emisiones. *Terra Latinoamericana* 34:97-112.
- Covalada, S., F. Paz y A. Ranero. 2018. Escenarios de mitigación de emisiones para el sector rural del Estado de Chiapas utilizando modelos de estados y transiciones. *Madera y Bosques* 24: e2401897. DOI:10.21829/myb.2018.2401897
- Covalada-Ocón, S., F. Paz-Pellat, A. Ranero-Puig, M. I. Marín-Sosa, M. Casiano-Domínguez, B. de Jong, J. D. Etchevers-Barra y A. S. Velázquez-Rodríguez. 2019. Capítulo 24: Escenarios asociados al ciclo del carbono y sus interacciones: Estado de Chiapas. pp. 573-602. *En*: Paz-Pellat, F., J. M. Hernández-Ayón., R. Sosa-Ávalos y A. S. Velázquez-Rodríguez. *Estado del Ciclo del Carbono en México: Agenda Azul y Verde*. Programa Mexicano del Carbono. Texcoco, Estado de México. ISBN 978-607-96490-7-4.
- Cruz-Cárdenas, G., L. López-Mata, C. A. Ortiz-Solorio, J. L. Villaseñor, E. Ortiz, J. T. Silva and F. Estrada-Godoy. 2014. Interpolation of Mexican soil properties at a scale of 1:1,000,000. *Geoderma* 213:29-35.
- de Jong, B., C. Anaya, O. Masera, M. Olguin, F. Paz, J. Etchevers, R. Martínez, G. Guerrero and C. Balbontin. 2010. Greenhouse gas emissions between 1993 and 2002 from land-use change and forestry in Mexico. *Forest Ecology and Management* 260:1689-1701.
- de Jong, B., M. Olguin, F. Rojas, V. Maldonado, F. Paz, J. Etchevers, C. O. Cruz y J. A. Argumedo. 2009. Inventario nacional de emisiones de gases invernadero 1990-2006. Uso del suelo, cambio de uso del suelo y silvicultura. Reporte preparado para el Instituto Nacional de Ecología. Distrito Federal, México. 119 p.
- de Jong, B., O. Masera, R. D. Martínez, F. Paz, M. Olguin, C. Anaya, C. Balbontin, M. Motolinía y G. Guerrero. 2006. Inventario nacional de emisiones de gases invernadero 1993-2002. Uso del suelo, cambio de uso del suelo y bosques. Reporte preparado para el Instituto Nacional de Ecología. D.F., México. 78 p.
- Dimassi, B., B. Guenet, N. P. A. Saby, F. Munoz, M. Bardy, F. Millet and M. P. Martin. 2018. The impacts of century model initialization scenarios on soil organic carbon dynamics simu-

- lation in French long-term experiments. *Geoderma* 311:25-36.
- Etchevers, J. D. 2013. Estado actual de la investigación de los almacenes de carbono en México. pp. 27-31. En: Paz, F., M. Bazán y V. Saynes (eds.). *Dinámica del Carbono en el Suelo 2012. Serie Avances Temáticos del Ciclo del Carbono y sus Interacciones. Programa Mexicano del Carbono en colaboración con la Sociedad Mexicana de Ciencias del Suelo*. Texcoco, Estado de México, México.
- Etchevers, J. D., O. Maser, C. Balbontín, D. Gómez, A. Monterroso, R. Martínez, M. Acosta, M. Martínez and C. Ortiz. 2006. Soil carbon sequestration in Mexico and Central America (Biome A). p.p. 119-146. *In*: R. Lal., C., C. Cerri., M. Bernoux., J. Etchevers and E. Cerri (eds.). *Carbon Sequestration in Soils of Latin America*. The Haworth Press, Inc. New York, USA.
- Falloon, P., P. Smith, K. Coleman and S. Marshall. 1998. How important is inert organic matter for predictive soil carbon modelling using the Rothamsted carbon model? *Soil Biology and Biochemistry* 32:433-436.
- FAO. 2017. Soil organic carbon, the hidden potential. Food and Agriculture Organization of the United Nations. Rome, Italy. 90 p.
- FAO. 2018. Soil Organic Carbon Mapping Cookbook. Second Edition. Food and Agriculture Organization of the United Nations. Rome, Italy. 180 p.
- FAO. 2019. Recarbonization of global soils. A tool to support the implementation of the Koronivia joint work on agriculture. Food and Agriculture Organization of the United Nations. Rome, Italy. 12 p.
- FAO. 2020a. Technical Specifications and Country Guidelines for Global Organic Carbon Sequestration Potential Map (GSOCseq). Food and Agriculture Organization of the United States. Rome, Italy. 33 p.
- FAO. 2020b. A protocol for measurement, monitoring, reporting and verification of soil organic carbon in agricultural landscapes. – GSOC-MRV Protocol. Food and Agriculture Organization of the United Nations. Rome, Italy. 140 p.
- FAO-SAGARPA. 2012. Línea Base del Programa de Sustentabilidad de los Recursos Naturales, Subíndice de Uso Sustentable del Suelo – Metodología de Cálculo. México. 53 p.
- Feng, W., A. F. Plante and J. Six. 2013. Improving estimates of maximal organic carbon stabilization by fine soil particles. *Biogeochemistry* 112:81-93.
- Foerid, B., P. H. Bellamy, A. Holden and G. J. D. Kirk. 2012. On the initialization of soil carbon models and its effects on model predictions for England and Wales. *European Journal of Soil Science* 63:32-41.
- González-Molina, L., E. Espitia-Rangel, J. Pineda-Pineda, E. Muñoz-Reyes, M. G. Irizar-Garza y A. Ayala-Garay. 2020. Potencial de secuestro de carbono orgánico en quinua simulado con el modelo RothC-26.3. *Revista Mexicana de Ciencias Agrícolas* 11:789-799.
- González-Molina, L., J. D. Etchevers-Barra, F. Paz-Pellat y J. M. González-Camacho. 2010. El uso potencial del modelo RothC-26.3 en inventarios nacionales de carbono orgánico del suelo en México. pp: 229-248. En: García-Oliva F. y M. B. Turrión Nieves (eds.). *Materia Orgánica Edáfica y Captura de Carbono en Sistemas Iberoamericanos*. SIFyQA Ministerio de Ciencia e Innovación, Salamanca, España.
- González-Molina, L., M. Acosta-Mireles, F. Carillo-Anzures y A. Rueda-Sánchez. 2018. Simulación de los cambios de carbono orgánico del suelo en especies tropicales arbóreas con el modelo RothC 26.3. *Interciencia* 43:269-274.
- Goovaerts, P. 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press. New York, USA. 488 p.
- Gottschalk, P., J. U. Smith, M. Wattenbach, J. Bellarby, E. Stehfest, N. Arnell, T. J. Osborn, C. Jone and P. Smith. 2012. How will organic carbon stocks in mineral soils evolve under future climate? *Global projections using RothC for a range of climate change scenarios*. *Biogeosciences* 9:3151–3171.
- Grunwald, S. 2009. Multi-criteria characterization of recent digital soil mapping and modeling approaches. *Geoderma* 152:195-207.
- Grunwald, S., J. A. Thompson and J. L. Boettinger. 2011. Digital soil mapping and modeling at continental scales: finding solutions for global issues. *Soil Science Society of America Journal* 75:1201-1213.
- Guevara, M. and R. Vargas. 2020a. Soil organic carbon estimates and uncertainty at 1-m Depth across Mexico. 1999-2009. ORNL DAAC, Oak Ridge, Tennessee, USA. DOI: <https://doi.org/10.3334/ORNLDAAAC/1754>
- Guevara, M. y R. Vargas. 2021. Predicción de carbono orgánico en los suelos de México a un metro de profundidad y 90m de resolución espacial. *Terra Latinoamericana* 39:1-19.
- Guevara, M., C. Arroyo, N. Brunzell, C. O. Cruz, G. Domke, J. Equihua, J. Etchevers, D. Hayes, T. Hengl, A. Ibelle, K. Johnson, B. de Jong, Z. Libohova, R. Llamas, L. Nave, J. L. Ornelas, F. Paz, R. Ressler, A. Schwartz, A. Victoria, S. Wills and R. Vargas. 2020c. Soil organic carbon across Mexico and the conterminous United States (1991-2010). *Global Biogeochemical Cycles* 34: e2019GB006219. DOI:10.1029/2019GB006219
- Guevara, M., C. E. Arroyo-Cruz, N. Brunzell, C. O. Cruz-Gaisardo, G. M. Domke, J. Equihua, J. Etchevers, D. J. Hayes, T. Hengl, A. Ibelle, K. Johnson, B. de Jong, Z. Libohova, R. Llamas, L. Nave, J. L. Ornelas, F. Paz, R. Ressler, A. Schwartz, S. Wills and R. Vargas. 2020b. Soil Organic Carbon Estimates for 30-cm Depth, Mexico and Conterminous USA, 1991-2011. ORNL DAAC, Oak Ridge, Tennessee, USA. DOI: <https://doi.org/10.3334/ORNLDAAAC/1737>
- Guevara, M., G. Federico-Olmedo, E. Stell, Y. Yigini, Y. Aguilar-Duarte, C. Arellano-Hernández, G. E. Arévalo, C. E. Arro-

- yo-Cruz, A. Bolivar, S. Bunning, N. Bustamante-Cañas, C. O. Cruz-Gaistardo, F. Dávila, M. Dell-Acqua, A. Encina, H. Figueredo-Tacona, F. Fontes, J. A. Hernández-Herrera, A. R. Ibelle-Navarro, V. Loayza, A. M. Manueles, F. Mendoza-Jara, C. Olivera, R. Osorio-Hermosilla, J. Rodríguez-Rodríguez, R. Roopnarine, A. Rosales-Ibarra, K. A. Rosales-Riveiro, G. A. Schulz, A. Spence, G. M. Vásquez, R. R. Vargas and R. Vargas. 2018. No silver bullet for digital soil mapping: country-specific soil organic carbon estimates across Latin America. *Soil* 4:173-193.
- Gulde, S., H. Chung, W. Amelung, C. Chang and J. Six. 2008. Soil carbon saturation controls labile and stable carbon pools dynamics. *Soil Sci. Soc. Am. J.* 72:605-612.
- Hassink, J. 1997. The capacity of soils to preserve organic C and N by their association with clay and silt particles. *Plant and Soil* 191:77-87.
- Hengl, T., J. Mendes-de Jesus, G. B. M. Heuvelink, M. Ruiperez-Gonzalez, M. Kilibarda, A. Blagotic, W. Shangguan, M. N. Wright, X. Geng, B. Bauer-Marschallinger, M. A. Guevara, R. Vargas, R. A. MacMillan, N. H. Batjes, J. G. B. Leenaars, E. Ribeiro, I. Wheeler, S. Mantel and B. Kempen. 2017. Soil-Grids250m: global gridded soil information based in machine learning. *Plos One* 12: e0169748. DOI: 10.1371/journal.pone.0169748
- Heuvelink, G. B. M., M. E. Angelini, L. Poggio, Z. Bai, N. H. Batjes, R. van den Bosch, D. Bossio, S. Estella, J. Lehmann, G. F. Olmedo and J. Sanderman. 2020. Machine learning in space and time for modelling soil organic carbon change. *Soil Science*. DOI: 10.1111/ejss.12998
- INEGI. 2004. Información nacional sobre perfiles de suelo Versión 1.2. Instituto Nacional de Estadística, Geografía e Informática, Aguascalientes, México. <https://www.inegi.org.mx/temas/edafologia/> (Consulta: marzo 12, 2020).
- IPCC. 2006. Agriculture, forestry and other land use IPCC Guidelines for National Greenhouse Gas Inventories. *In: Eggleston H. S., L. Buendia, K. Miwa, T. Ngara and K. Tanabe (eds.)*. Prepared by the National Greenhouse Gas Inventories Programme. Institute for Global Environmental Strategies. Hayama, Japan.
- Jenny, H. 1941. Factors of soil formation. A System of Quantitative Pedology. McGraw Hill. USA. 119 p.
- Journel, A. G. 1983. Nonparametric estimation of spatial distributions. *Mathematical Geology* 15:445-468.
- Journel, A. G. 1986. Constrained interpolation and qualitative information – the soft kriging approach. *Mathematical Geology* 18:269-286.
- Journel, A. G. 1999. Markov models for cross-covariances. *Mathematical Geology* 31:955-964.
- Journel, A. G. 2002. Combining knowledge from diverse sources: an alternative to traditional independence hypotheses. *Mathematical Geology* 34:573-596.
- Journel, A. G. 2005. Beyond covariance: the advent of multiple-point geostatistics. pp. 225-233. *In: Lewangthong O. and C.V. Deutsch (eds.)*. Geostatistics Banff 2004. Springer. Netherlands.
- Klir, G. J. 2006. Uncertainty and Information: Foundations of Generalized Information Theory. John Wiley and Sons. Inc. 1551 p.
- Körschens, M., A. Weigel and E. Schulz. 1998. Turnover of soil organic matter (SOM) and long-term balances – tools for evaluating sustainable productivity of soils. *Z. Pflanzenernähr. Bodenk* 161:409-424.
- Krishnan, S. and A.G. Journel. 2003. Spatial connectivity: from variogram to multiple-point measures. *Mathematical Geology* 35:915-925.
- Kyriakidis, P. and A. G. Journel. 1999. Geostatistical space-time models: review. *Mathematical Geology* 31:651-684.
- Lal, R. 2009. Challenges and opportunities in soil organic matter research. *European Journal of Soil Science* 60:158-169.
- Lal, R. 2016. Beyond COP 21: potential and challenges of the “4 per Thousand” initiative. *Journal of Soil and Water Conservation* 71:20A-25A.
- Lamichhane, S., L. Kumar and B. Wilson. 2019. Digital soil mapping algorithms and covariates for soil organic carbon mapping and their implications: a review. *Geoderma* 352:395-413.
- LeCun, Y., Y. Bengio and G. Hinton. 2015. Deep learning. *Nature* 521:436-444.
- Liu, Y., A. Harding, R. Gilbert and A. Journel. 2005. A workflow for multiple-point geostatistical simulation. pp. 245-254. *In: Lewangthong, O. and C.V. Deutsch (eds.)*. Geostatistics Banff 2004. Springer. Netherlands.
- Ludwig, B., H. Kellin, N. Lingan and L. Xuejun. 2010. Modelling the dynamics of organic carbon in fertilization and tillage experiments in the North China Plain using the Rothamsted Carbon Model-initialization and calculation of C inputs. *Plant soil* 332: 193-206.
- Ma, Y., B. Minasny, B. P. Malone and A. B. McBratney. 2019. Pedology and digital soil mapping (DSM). *European Journal of Soil Science* 70:216-235.
- MAAF. 2015. The 4 0/00 initiative. Ministry of Agri-food and Forestry of France, France. <https://www.4p1000.org/>. (Consulta: marzo 28, 2020).
- Marschmann, G. L., H. Pagel, P. Kögler and T. Streck. 2019. Equifinality, sloppiness, and emergent structures of mechanistic soil biogeochemical models. *Environmental Modelling and software*. DOI: 10.116/j.envsoft.2019.104518
- Matus, F. 2021. Fine silt and clay content are the main factor defining maximal C and N accumulation in soils: a meta-analysis. *Scientific Reports* 11: 6438. DOI: <https://doi.org/10.1038/s41598-021-84821-6>
- Matus, F., E. Garrido, C. Hidalgo, F. Paz, J. Etchevers, C. Merino

- and A. Báez. 2016. Carbon saturation in the silt and clay particles in soils with contrasting mineralogy. *Terra Latinoamericana* 34:311-319.
- McBratney, A. B., M. L. Mendonca-Santos and B. Minasny. 2003. On digital soil mapping 17:3-52.
- Meersmans, J., M. P. Martin, E. Lacerre, T. G. Orton, S. De Baets, M. Gourrat, N. P. A. Saby, J. Wetterlind, A. Bispo, T. A. Quine and D. Arrouays. 2013. Estimating soil carbon input in France: an inverse modelling approach. *Pedosphere* 23:422-436.
- Minasny, B., B. P. Malone, A. B. McBratney, D. A. Angers, D. Arrouays, A. Chambers, V. Chaplot, Z. S. Chen, K. Cheng, B. S. Das, D. J. Field, A. Gimona, C. B. Hedley, S. Y. Hong, B. Mandal, B. P. Marchant, M. Martin, B. G. McConkey, V. L. Mulder, S. O'Rourke, A. C. Richer-de-Forges, I. Odeh, J. Pardanian, K. Paustian, G. Pan, L. Poggio, I. Savin, V. Stolbovov, U. Stockmann, Y. Sulaeman, C. C. Tsui, T. G. Vagen, B. van Wesemael and L. Winowiecki. 2017. Soil carbon 4 per mille. *Geoderma* 292:59-86.
- Molnar, C., G. Casalicchio and B. Bisch. 2010. Interpretable machine learning – a brief history, state-of-the-art and challenges. Department of Statistics, LMU Munich, Germany.
- Montero, J. M., G. Fernández-Avilés and J. Mateu. 2015. Spatial and spatio-temporal geostatistical modeling and Kriging. John Wiley and Sons Ltd. United Kingdom. 357 p.
- Ogle, S. M., F. J. Breidt and K. Paustian. 2005. Agricultural management impacts on soil organic carbon storage under moist and dry climatic conditions of temperate and tropical regions. *Biogeochemistry* 72:87-121.
- Ortiz, J. M. and C. V. Deutsch. 2004. Indicator simulation accounting for multiple-point statistics. *Mathematical Geology* 36:545-5665.
- Parton, W. J., J. W. B. Stewart and C. V. Cole. 1988. Dynamic of C, N, P and S in grassland soil: a model. *Biogeochemistry* 5: 109-131.
- Paz, F. y J. Etchevers. 2016. Distribución a profundidad del carbono orgánico en los suelos de México. *Terra Latinoamericana* 34:339-355.
- Paz, F., M. I. Marín, M. Bolaños-González, J. D. Etchevers, B. de Jong, J. Herrera y A. S. Velázquez-Rodríguez. 2020. La urgente necesidad de México de contar con datos de actividad a escala apropiada para el sector agricultura, forestería y otros usos del suelo. *Elementos para Políticas Públicas* 4:61-73.
- Paz, F., S. Covaleda, C. Hidalgo, J. Etchevers y F. Matus. 2016. Modelación simple y operativa de la distribución del carbono orgánico por fracciones físicas en los suelos. *Terra Latinoamericana* 34:321-337.
- Paz-Pellat, F., J. M. Hernández-Ayón, R. Sosa-Ávalos y A. S. Velázquez-Rodríguez. 2019a. Estado del Ciclo del Carbono en México: Agenda Azul y Verde. Programa Mexicano del Carbono. Texcoco, Estado de México. ISBN 978-607-96490-7-4.
- Paz-Pellat, F. y A. Velázquez-Rodríguez. 2020. Una nota de precaución para los meta análisis de la dinámica del carbono orgánico de los suelos. pp. 476-481. En: Hernández, J. M., M. Rojo, M. Fuentes, A. Velázquez y M. Bolaños (eds.). Estado Actual del Conocimiento del Ciclo del Carbono y sus Interacciones en México: Síntesis a 2020. Texcoco, Estado de México, México.
- Paz-Pellat, F. y A. S. Velázquez-Rodríguez. 2018. Base de datos de perfiles de suelos en México. *Elementos para Políticas Públicas* 2:210-235.
- Paz-Pellat, F. y C. A. Ortiz-Solorio. 2019a. Base de datos de la evaluación de la degradación de los suelos en México. *Elementos para Políticas Públicas* 3:51-56.
- Paz-Pellat, F., A. S. Velázquez-Rodríguez, J. D. Etchevers-Barra, C. I. Hidalgo-Moreno, M. Bolaños-González, B. de Jong, S. Covaleda-Ocón, M. Fuentes-Ponce, G. Vela-Correa, F. García-Oliva, Mario Guevara y R. Vargas. 2019b. Capítulo 20: Suelos. pp. 436-468. En: Paz-Pellat, F., J. M. Hernández-Ayón, R. Sosa-Ávalos y A. S. Velázquez-Rodríguez. Estado del Ciclo del Carbono en México: Agenda Azul y Verde. Programa Mexicano del Carbono. Texcoco, Estado de México. ISBN 978-607-96490-7-4.
- Paz-Pellat, F., C. Balbontín-Nesvara, J. Etchevers-Barra, M. Martínez-Menes y C. Ortiz-Solorio. 2010. Análisis multifractal del carbono en los suelos 3. Estimaciones escalantes. *Terra Latinoamericana* 28:89-95.
- Paz-Pellat, F., J. Argumedo-Espinoza, C. O. Cruz-Gaistardo, J. D. Etchevers y B. de Jong. 2016. Distribución espacial y temporal del carbono orgánico del suelo en los ecosistemas terrestres de México. *Terra Latinoamericana* 34:289-310.
- Poepflau, C., C. Vos and A. Don. 2017. Soil organic carbon stocks are systematically overestimated by misuse of the parameters bulk density and rock fragment content. *Soil* 3:61-66.
- Post, W. M., R. Izaurralde, L. K. Mann and N. B. Bliss. 2001. Monitoring and verifying changes of organic carbon in soil. *Climate Change* 51:73-99.
- Ramcharan, A., T. Hengl, D. Beaudette and S. Wills. 2018. A soil bulk density pedotransfer function based on machine learning: a case study with the NCSS soil characterization database. *Soil Sc. Soc. Am. J.* 81:1279-1287.
- Rasel, S.M.M., T.A. Groen., Y.A. Hussin and I.J. Diti. 2017. Proxies for soil organic carbon derive from remote sensing. *International Journal of Applied Earth Observation and Geoinformation* 59:157-166.
- Roscher, R., B. Bohn, M. F. Duarte and J. Garcke. 2020. Explainable machine learning for scientific insights and discoveries. *IEEE Access* 8: 42200-42216. DOI:10.1109/ACCESS.2020.2976199.
- Rühlmann, J. 1999. A new approach to estimating the pool of stable organic matter in soil using data from long-term field experiments. *Plant and Soil* 213:149-160.

- Salas-Aguilar, V. M. y F. Paz-Pellat. 2017. Desarrollo de una base de datos geomorfológica nacional. *Elementos para Políticas Públicas* 1:173-182.
- Salas-Aguilar, V. M. y F. Paz-Pellat. 2018. Desarrollo de una base de datos climática nacional anual, 1989-2012: resolución 250m. *Elementos para Políticas Públicas* 2:19-32.
- Sanderman, J. and J. A. Baldock. 2010. Accounting for soil carbon sequestration in national inventories: a soil scientist's perspective. *Environmental Research Letters* 5:1-6. DOI:10.1088/1748-9326/5/3/034003
- Segura, M. A., P. Sánchez, C. A. Ortiz y M. C. Gutiérrez. 2004. Carbono orgánico de los suelos de México. *Terra Latinoamericana* 23:21-28.
- SEMARNAT-CP. 2002. Evaluación de la degradación de los suelos causada por el hombre en la República Mexicana a escala 1:250,000. Secretaría de Medio Ambiente y Recursos Naturales. Memoria Nacional, D. F., México. [https://apps1.semarnat.gob.mx:8443/dgeia/informe_resumen14/03_suelos/3_2.html#:~:text=Semarnat%20y%20CP,%2C%20escala%201%3A%20250%20000.&text=En%20la%20degradaci%C3%B3n%20qu%C3%ADmica%20y,\(Figura%203.2.4\).](https://apps1.semarnat.gob.mx:8443/dgeia/informe_resumen14/03_suelos/3_2.html#:~:text=Semarnat%20y%20CP,%2C%20escala%201%3A%20250%20000.&text=En%20la%20degradaci%C3%B3n%20qu%C3%ADmica%20y,(Figura%203.2.4).) (Consulta: mayo 17, 2020).
- Simonson, R. W. 1959. Outline of a generalized theory of soil genesis. *Soil Science Society Proceedings* 23:152-156.
- Six, J., R. T. Conant, E. A. Paul and K. Paustian. 2002. Stabilization mechanisms of soil organic matter: implications for C-saturation of soils. *Plant and Soil* 241:155-176.
- Stewart, C. E., K. Paustian, R. T. Conant, A. F. Plante and J. Six. 2007. Soil carbon saturation: concept, evidence and evaluation. *Biogeochemistry* 86:19-31.
- Taghizadeh-Toosi, A., B. T. Christensen, N. J. Hutchings, J. Vejlin, T. Kätterer, M. Glendinning and J. E. Olesen. 2014. C-TOOL: a simple model for simulating whole-profile carbon storage in temperate agricultural soils. *Ecological Modelling* 292:11-25.
- The World Bank. 2012a. Carbon sequestration in agricultural soils. Economic and Sector Work. Report Number: 67395-GLB. Washington, D.C. 85 p.
- The World Bank. 2012b. Enhancing carbon stocks and reducing CO₂ emissions in agriculture and natural resources management projects. Toolkit. Washington, D.C. 115 p.
- Vargas, R., D. Alcaraz-Segura, R. Birdsey, N. A. Brunzell, C. O. Cruz-Gaistardo, B. de Jong, J. Etchevers, M. Guevara, D. J. Hayes, K. Johnson, H. W. Loescher, F. Paz, Y. Ryu, Z. Sanchez-Mejia and K. P. Toledo-Gutierrez. 2017. Enhancing interoperability to facilitate implementation of REDD+: case study of Mexico. *Carbon Management* 4:1-9.
- Vargas, R., F. Paz and B. de Jong. 2013. Quantification of forest degradation and belowground carbon dynamics: ongoing challenges for monitoring, reporting and verification activities for REDD+. *Carbon Management* 4:579-582.
- Wadoux, A. M. J. C., A. Samuel-Rosa, L. Poggio and V. L. Mulder. 2019. A note on knowledge discovery and machine learning in digital soil mapping. *Soil Science* 71:133-136.
- Wadoux, A. M. J. C., B. Minasny and A. B. McBratney. 2020. Machine learning for digital soil mapping: Applications, challenges and suggested solutions. *EARTH-Science Reviews* 210. DOI: 10.1016/j.earscirev.2020.103359
- Wang, Y. P., C. M. Trudinger and I. G. Enting. 2009. A review of applications of model-data fusion to studies of terrestrial carbon fluxes at different scales. *Agricultural and Forest Meteorology* 14:1829-1842.
- Zadeh, L. A. 2006. Generalized Theory of Uncertainty: Principal Concepts and Ideas. *Computational Statistics and Data Analysis* 51:15-46.
- Zhu, H. and Journal A. G. 1993. Formatting and integrating soft data: stochastic imaging via the Markov-Bayes algorithm. *In*: Soares A. (ed.). *Geostatistics Tróia '92. Quantitative Geology and Geostatistics* 5:1-12. DOI: https://doi.org/10.1007/978-94-011-1739-5_1