

## Capítulo 10

# SECUESTRO DE CARBONO EN SUELOS DE LA RTP 105 CUETZALAN, MÉXICO, A TRAVÉS DE TELEDETECCIÓN Y MODELOS MIXTOS.

G. Linares Fleites<sup>1</sup>; E.Torres Trejo<sup>2</sup>; M. de L. Sandoval Solís<sup>3</sup>

<sup>1</sup>Posgrado Cs. Ambientales y Dpto. en Cs. Agrícolas

<sup>2</sup>Doctorado en Ciencias Ambientales

<sup>3</sup>Facultad de Ciencias de la Computación

Benemérita Universidad Autónoma de Puebla

Puebla, México

malosa1753@gmail.com

### ABSTRACT.

Carbon sequestration in soils is of great importance to the mitigation of climate change and its assessment, from the estimate of the concentration of carbon in soil, allows decision making in conservation and ecosystem management. This study is based on the digital processing of normalized vegetation index (NDVI) to estimate the degree of involvement of deforested areas through selected images coming from the satellite Landsat-5 TM and Landsat sensor 7 ETM +. The objective is to estimate, by linear mixed model, the relationship between NDVI and soil properties of the RTP region 105: Cuetzalan, defined as priority for conservation.

**Keywords:** climate change, vegetation indices, linear mixed regression, R Software

### RESUMEN.

El secuestro de carbono en suelos es de gran importancia para la mitigación del cambio climático; su evaluación, a partir de la estimación de la concentración del carbono en suelo, permite la toma de decisiones en la conservación y manejo de los ecosistemas. Este estudio está basado en el procesamiento digital del índice de vegetación normalizado (NDVI) que permite estimar los grados de afectación de áreas deforestadas a través de imágenes seleccionadas que proceden del satélite Lansat- 5 sensor TM y Lansat 7 ETM+. El objetivo es estimar, a través de modelo lineal mixto, las relaciones existentes entre NDVI y algunas propiedades del suelo de la región RTP 105: Cuetzalan, definida como prioritaria para la conservación.

## 1. INTRODUCCIÓN

Como resultado del aumento de concentraciones de gases de efecto de invernadero, existen evidencias científicas que sugieren que el clima global se verá alterado en este siglo. El incremento actual de la concentración de dióxido de carbono (CO<sub>2</sub>) atmosférico, asociado al cambio climático, puede ser parcialmente mitigado por el secuestro de carbono en suelos.

El secuestro de carbono es esencialmente el proceso de transformación del carbono del aire en carbono almacenado en el suelo. El dióxido de carbono es absorbido por las plantas a través del proceso de fotosíntesis, e incorporado en la planta como tejido vegetal. Cuando las plantas mueren, el carbono de las hojas, tallos y raíces se descompone en el suelo y se convierte en materia orgánica. Este proceso reduce los niveles de CO<sub>2</sub> en la atmósfera, disminuyendo las probabilidades del calentamiento global.

Puede afirmarse que el mayor responsable del cambio climático global es el CO<sub>2</sub>, que tiene entre sus fuentes emisoras la deforestación y la destrucción de los suelos [1]. El carbono orgánico del suelo es un gran y activo reservorio y los ecosistemas forestales pueden absorber cantidades significativas de CO<sub>2</sub>, por lo que hay un gran interés por incrementar el contenido de carbono en estos ecosistemas.

A pesar de la importancia del secuestro de carbono para la mitigación del cambio climático, su evaluación se encuentra muy limitada en muchas zonas del mundo, en particular, en los suelos del estado de Puebla, México.

Los métodos para estudiar los almacenes de carbono en el suelo pueden clasificarse, según el objetivo del estudio, en local, regional o global [2]. Nuestro objetivo es local y perseguimos encontrar ecuaciones que modelen el comportamiento de los contenidos de carbono orgánico en el suelo con respecto a la vegetación existente en la zona. En la actualidad se están ideando nuevas metodologías de investigación y se están utilizando nuevas herramientas de búsqueda de información. Los instrumentos de teledetección son útiles en estudios como el que presentamos, ya que permiten obtener información sobre la vegetación de la zona de interés. Esa información se puede relacionar posteriormente con las tradicionales mediciones *in situ*.

Una estrategia de mitigación llamada Reducción de Emisiones por Deforestación y Degradación de Bosques (REDD) se está desarrollando para hacer frente a las emisiones debidas a la pérdida de bosques en los países en desarrollo. REDD será el instrumento básico de cualquier acuerdo climático según el documento final de la 15<sup>a</sup> Conferencia de las Naciones Unidas de las Partes [3]. Sin embargo, la implementación de REDD presenta retos científicos, además de los políticos. Los métodos estadísticos y, en particular los modelos mixtos, pueden jugar un papel fundamental en la evaluación de los contenidos de carbono orgánico en suelo.

En el capítulo se hace un breve recuento de los aspectos de teledetección e índices de vegetación que pueden obtenerse a través de imágenes satelitales. Posteriormente, se relacionan los conceptos y procedimientos de los modelos lineales mixtos y se aplican para obtener las mejores ecuaciones que permitan estimar el porcentaje de carbono en suelo en la Región Terrestre Prioritaria 105, Cuetzalan, México. Finalmente, se brindan las conclusiones y las referencias.

## 2. TELEDETECCIÓN

El término teledetección es una traducción del inglés “*RemoteSensing*”, y se refiere, no sólo, a la captación de datos desde el aire o desde el espacio sino también a su posterior tratamiento [4]. En teledetección la adquisición de información puede realizarse a nivel orbital (satélites), suborbital (aviones) o terrestre (radiómetros portátiles). Las imágenes satelitales y las fotografías aéreas son los productos más difundidos y, en ocasiones, son de fácil acceso y gratis. A continuación se caracteriza el área de estudio y el material satelital utilizado.

El estudio se sitúa en la Región Terrestre Prioritaria (RPT) 105. En [5] se define a una RTP como al “Área continental cuyas características físicas y biológicas la hacen particularmente importante desde el punto de vista de la biodiversidad”; son regiones que por sí mismas tienen calidad de prioritarias para preservar la biodiversidad bajo esquemas particulares de conservación, o de desarrollo sustentable. La RPT 105 se ubica en las coordenadas extremas: 19°46’23’’ y 20°11’55’’ de latitud norte y 97°09’17’’ a 97°38’36’’ de longitud oeste, con una superficie de 1,284 km<sup>2</sup> (Figura 1). Esta región se definió como prioritaria para la conservación debido a la existencia de bosques mesófilo de montaña, selva alta perennifolia en las partes bajas y bosques de encino, y bosque encino en asociación con pino. El aprovechamiento de los recursos naturales que se realiza en la RTP se relaciona con actividades forestales, agrícolas y/o pecuarias, generalmente asociadas a la deforestación, degradación o transformación del ecosistema.

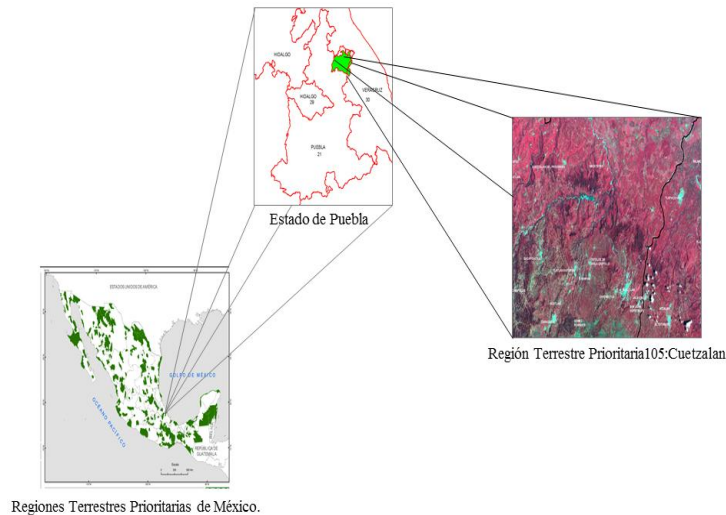


Figura 1.- Localización de la zona de la RTP 105: Cuetzalan.

El material satelital comprende: imágenes Landsat con los sensores MSS (MultiSpectral Scanner), TM (ThematicMapper), ETM+ (EnhancedThematicMapper Plus) que corresponden al satélite Landsat 5 y Landsat 7 respectivamente; la imagen TM corresponde al 3 de diciembre de 1994 y la ETM+ al 2 de noviembre del 2005, y fueron obtenidas en el sitio oficial del Servicio Geológico de los Estados Unidos (USGS, por sus siglas en inglés), <http://landsat.usgs.gov>.

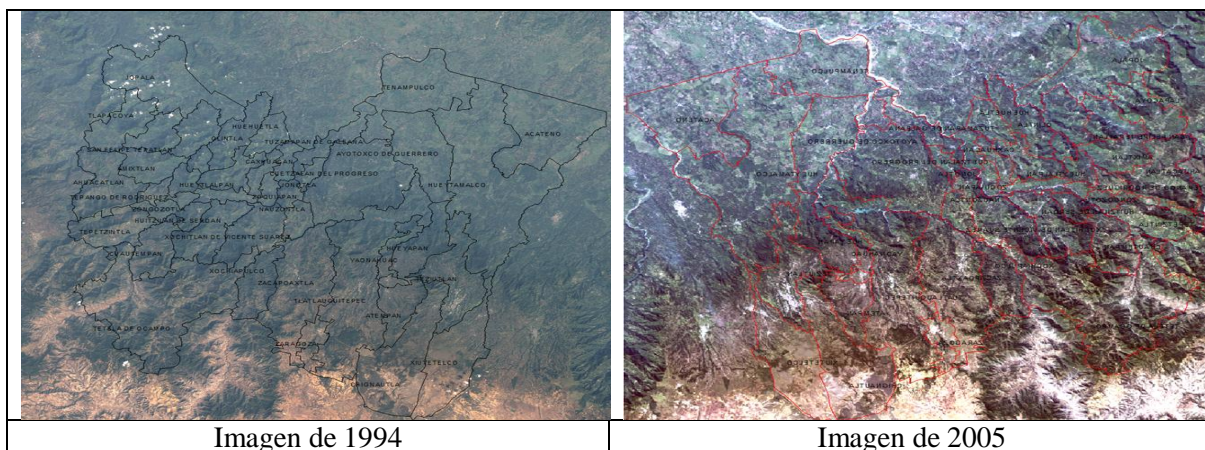


Figura 2. Imágenes de la zona de la RTP 105

A partir de estas imágenes (Figura 2) se llevó a cabo el procesamiento digital del Índice de Vegetación Normalizado (NDVI), que permite estimar el grado de afectación de áreas deforestadas, expresándolo en términos del número de píxeles.

### 3. ÍNDICES DE VEGETACIÓN

Los índices de vegetación son medidas cuantitativas que se basan en valores digitales útiles para estimar la biomasa o vigor vegetal. Usualmente el índice de vegetación es producto de varios valores espectrales que son sumados, divididos, o multiplicados en una forma diseñada para producir un simple valor que indique la cantidad o vigor de la vegetación dentro de un píxel. Valores altos de índices de vegetación, significan píxeles representativos para vegetación saludable; se aprovecha el comportamiento radiométrico de la vegetación sana en las diferentes bandas espectrales, especialmente entre el espectro visible y el infrarrojo cercano. En el espectro visible, en particular en la banda roja, los pigmentos de las hojas de las plantas absorben la mayor parte de la energía que reciben de sol, estas pigmentaciones apenas afectan al infrarrojo cercano [6].

Los contrastes de la imagen son utilizados para el cálculo de los índices de vegetación. La metodología se basa en el uso del Sistema de Información Geográfica (SIG) ArcMap 9.3, para el procesamiento digital e interpretación de las imágenes que proceden del satélite Landsat-5 sensor TM, que proporcionan imágenes de 8 bits en 7 y 8 bandas espectrales, tres de las cuales son visibles y se reconocen como Banda 1, Banda 2 y Banda 3 y cuatro infrarrojo, que son la Banda 4, Banda 5, Banda 6 y Banda 7. Las bandas 3 y 4, rojo visible e infrarrojo cercano, respectivamente, son adecuadas para la estimación de índices de vegetación, y facilitan establecer relaciones entre la densidad de vegetación alta y baja.

Los cálculos en este estudio corresponden al Índice de Vegetación Normalizado (NDVI). Este índice de vegetación se clasificó en dieciocho clases según su desviación estándar y, a partir de estas determinaciones, se realizó una reclasificación para obtener solo dos densidades de vegetación, la alta y la baja. La signatura espectral característica de la vegetación sana muestra un fuerte contraste entre la baja reflectividad en el rojo y la alta reflectividad en el infrarrojo. Esta diferencia es tanto mayor cuanto mayor es la densidad de la vegetación y mejor su estado fitosanitario. De esta forma para el cálculo del Índice de Vegetación Normalizado (NDVI) se utilizó la siguiente ecuación:

$$X = (B4 - B3) / (B4 + B3)$$

donde:

B4: Banda 4 de la imagen de Landsat TM (reflectividad en el infrarrojo cercano)

B3: Banda 3 de la imagen de Landsat TM (reflectividad en el rojo)

Este índice varía entre -1 y 1, lo que facilita su interpretación. Puede señalarse como umbral crítico para cubiertas vegetales un valor de NDVI de 0.1 y para la vegetación densa de 0.5. Cualquier valor negativo implica ausencia de vegetación por lo que deberían reclasificarse a cero [7]

El disponer de imágenes satelitales de la región permitió estimar la densidad en la cobertura vegetal en los años 1994 y 2005 [8]. En ambos años, el NDVI se encontró en un rango de 0.1 a 0.7, que nos indica que existe una densidad de vegetación que va desde baja hasta alta en la Región Terrestre Prioritaria 105: Cuetzalan. El índice permitió identificar la presencia de vegetación verde o sana, en la superficie. (Figura 3).

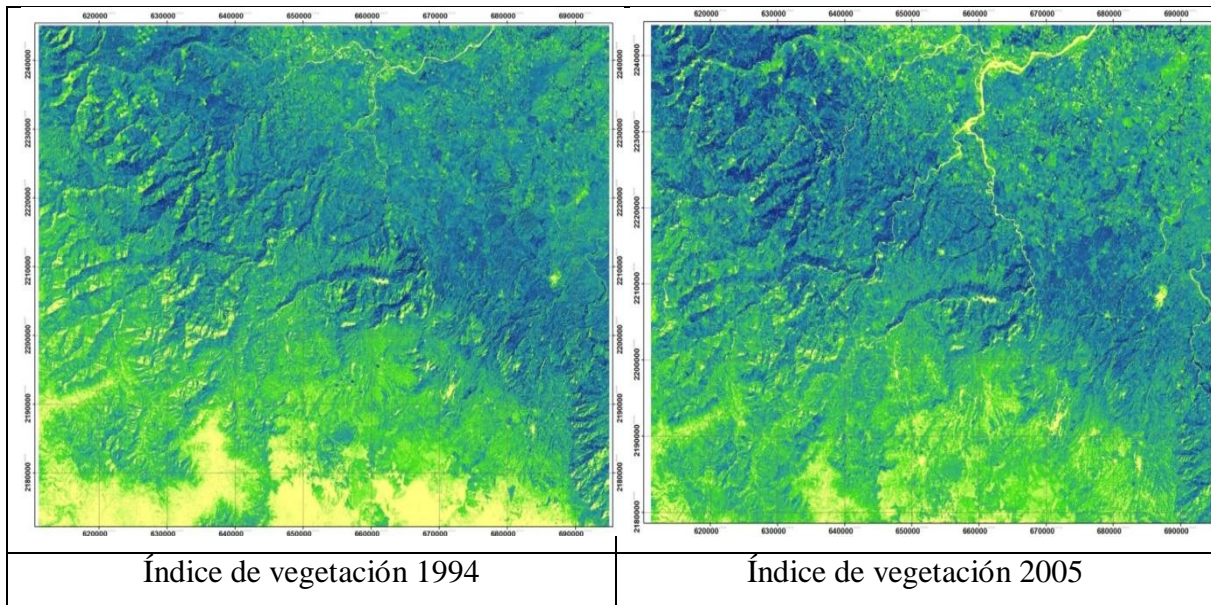


Figura 3. Imágenes de la zona de la RTP 105

#### 4. MODELO DE REGRESIÓN LINEAL MIXTO

Los modelos mixtos son usados cuando los datos tienen algún tipo de agrupación, como es el caso que nos ocupa, donde se tiene la realización de dos niveles de densidades de vegetación (baja y alta), determinadas por el índice de vegetación NDVI. Los modelos mixtos permiten tener coeficientes fijos y aleatorios (cuyos niveles son sólo una realización de todos los posibles niveles procedentes de una población) y varios términos de error. Estos modelos pueden ser una herramienta muy útil y se cuenta actualmente con buenas referencias tales como: [9],[10], [11] y [12]

La especificación de estos modelos puede plantearse como sigue:

Sea  $Y_{ki}$ : variable dependiente continua con realizaciones  $y_{ki}$ , con  $i$  observaciones en el grupo  $k$ , donde  $i = 1, \dots, n_k$  y  $k = 1, \dots, K$ .

- *Componente estocástico*:  $Y_{ki}$  se distribuye normal con media  $\mu_{ki}$  y varianza  $\sigma^2$ .
- El vector  $q$ -dimensional de efectos aleatorios,  $b_k$ , se restringe a tener media cero y por tanto se caracteriza completamente por la matriz de varianza y covarianza  $\Psi$ , simétrica y semidefinida positiva.
- *Componente sistemático*:  $\mu_{ki} = X_{ki}\beta + Z_{ki}b_k$

donde

$X_{ki}$  es el arreglo ( $n_k \times p \times K$ ) de variables explicatorias conocidas (no aleatorias) de efectos fijos,

$\beta$  es el vector  $p$ -dimensional de coeficientes de efectos fijos,

$Z_{ki}$  es el arreglo ( $n_k \times q \times K$ ) de variables explicatorias de efectos aleatorios conocidos, y  $b_k$  es el vector  $q$ -dimensional de efectos aleatorios.

Para la evaluación del secuestro de carbono en la zona bajo estudio, especificaremos varios modelos mixtos en los que el efecto aleatorio se representa por los niveles bajo y alto del índice de vegetación NVDI (Indveg). Para los efectos fijos utilizaremos, según el modelo especificado, las variables densidad aparente (DA) y porcentaje de nitrógeno total (Ntot), elegidas según criterios de especialistas en edafología y estudio exploratorios previos. La variable respuesta de todos los modelos es el porcentaje de carbono en suelo (Corg). El procesamiento estadístico se realiza en lenguaje R (Bates *et al*, 2011).

En forma general, los modelos estudiados son:

Modelo 1:

$$y_{ki} = \mu + b_k + \varepsilon_{ki}$$

donde  $\mu$  es la media poblacional fija y la parte  $b_k + \varepsilon_{ki}$  incluye los parámetros aleatorios. Obsérvese que el error residual del modelo lineal se divide en dos partes mutuamente independientes: un efecto del grupo aleatorio para el grupo  $k$  y un residuo aleatorio para el individuo  $i$  del grupo  $k$ . Se supone que los efectos de grupos y los residuos son variables independientes y normalmente distribuidas. Este es el modelo más simple de los modelos de efectos mixtos, se le conoce como modelo de componentes de la varianza y se usa cuando no se tienen otros predictores. Los parámetros de este modelo son  $\mu$ ,  $\sigma_b^2$  y  $\sigma^2$ .

El modelo de componentes de la varianza se generaliza a situaciones donde  $\mu$  se reemplaza por una función lineal de predictores. En nuestro caso tendremos el

Modelo 2:

$$y_{ki} = \beta_0 + \beta_1 X_{ki}^{(1)} + b_k + \varepsilon_{ki}$$

y el

Modelo 3:

$$y_{ki} = \beta_0 + \beta_1 x_{ki}^{(1)} + \beta_2 x_{ki}^{(2)} + b_k + \varepsilon_{ki}$$

Después de especificar un modelo con cierta parte sistemática y estructura para la parte aleatoria, el próximo paso es estimar los parámetros del modelo. Los dos métodos de estimación, más frecuentemente utilizados en la práctica, son el de mínimos cuadrados y el de máxima verosimilitud. En la literatura citada anteriormente, se recomiendan dos variantes de la máxima verosimilitud: el criterio máximo verosímil y el criterio máximo verosímil restringido, a los que nos referiremos por sus siglas en inglés como ML y REML respectivamente. El método ML puede usarse para estimar tanto los coeficientes  $\beta$  como los parámetros que especifican la estructura de varianza y covarianza residual, sin embargo, los estimadores de la estructura de varianza y covarianza residual son sesgados. El método REML brinda estimadores sólo para los parámetros que especifican la estructura de varianza y covarianza residual y estos estimadores son insesgados al menos en ciertos casos especiales y generalmente son menos sesgados que los estimadores [12].

El método de estimación que se utiliza en el paquete estadístico *lme4* en el lenguaje R, y que hemos usado aquí, es: primero estimar los parámetros de varianza y covarianza usando REML y posteriormente utilizar estimadores mínimos cuadrados generalizados (GLS, por sus siglas en inglés) para los parámetros de la parte sistemática. Sin embargo, debe destacarse que el método ML puro es favorable en situaciones especiales donde se requieran pruebas de hipótesis para comparar modelos.

También se utilizan los criterios de información para la comparación de modelos. Los más comúnmente utilizados son el criterio de Información de Akaike (AIC) y el criterio de información bayesiano (BIC). En la literatura consultada para estos y otros criterios de selección de modelos que aparecen en las salidas del programa *lme4*, se elige como mejor modelo aquel con los valores más pequeños del criterio.

Como es bien conocido, el ajuste de un modelo requiere de la comprobación de las suposiciones. Las suposiciones básicas de un modelo lineal son que el modelo es correcto, las observaciones son independientes y las varianzas son iguales. Además, para hacer pruebas de hipótesis sobre un modelo donde se ha utilizado estimadores mínimos cuadrados se necesita normalidad en los residuos y, en la estimación ML, la normalidad debe asumirse en la etapa de ajuste del modelo. El lenguaje R brinda herramientas para producir gráfico que permiten explorar las suposiciones del modelo.

#### 4.1 Modelos Lineales Mixtos En 1994

La tabla de datos para ajustar modelos en este período contaba con 30 observaciones agrupadas según dos niveles del factor índice de vegetación (Indveg). Las tablas 1 y 2 muestran los estimadores REML y ML de los tres modelos analizados para la RTP 105 según la imagen satelital del año 1994.

En el modelo 1, según los estimadores REML, la media de las medias de los niveles del índice de vegetación es 5.6376 y los estimadores de los componentes de la varianza son 0.04009 para  $\sigma_b^2$  y 11.38805 para la varianza de los residuos. Los estimadores ML muestran resultados similares con media de 5.652 y desviación estándar residual de 3.32.

Tabla1: Modelos Lineales Mixtos ajustados por el criterio REML ['lmerMod']

<b>Modelo 1:</b> Corg ~ 1 + (1   Indveg)	<b>Modelo 2 :</b> Corg ~ 1 + DA + (1   Indveg)	<b>Modelo 3.</b> Corg ~ 1 + DA + Ntot + (1 Indveg)
Convergencia REML :156.3	Convergencia REML : 135.7	Convergencia REML : 134.3
<b>Efectos aleatorios</b> Grupo VarianzaDesEst Interc. 0.04009 0.2002 Residuo 11.38805 3.3746	<b>Efectos aleatorios</b> Grupo VarianzaDesEst Interc. 0.7467 0.8641 Residuo 6.5074 2.5510	<b>Efectos aleatorios</b> Grupo Varianza DesEst Interc. 1.100 1.049 Residuo 6.603 2.570
Numero de observaciones: 30 Grupo: Indveg con 2 niveles	Numero de observaciones: 30 Grupo: Indveg con 2 niveles	Numero de observaciones: 30 Grupo: Indveg con 2 niveles
<b>Efectos fijos</b> EstimErrEst t Interc 5.6376 0.6365 8.86	<b>Efectos fijos</b> EstimErrEst t Interc 16.330 2.466 6.62 DA -14.451 3.094 -4.67	<b>Efectos fijos</b> EstimErrEst t Interc 15.9727 2.5720 6.21 DA -14.3590 3.1212 -4.6 Ntot 0.4175 0.6615 0.63
	<b>Correlación de efectos fijos:</b> (Interc) DA -0.946	<b>Correlación de efectos fijos:</b> (Interc) DA DA -0.923 Ntot -0.202 0.053

En el modelo 2, la media por niveles del índice de vegetación es (16.330 -14.451 “DA”) según los estimadores REML, y (16.51 -14.37 “DA”) según los estimadores ML. Se destaca la fuerte influencia de la densidad aparente sobre el secuestro de carbono. Obsérvese que la constante del modelo 2 varía alrededor de 16.330 con una desviación estándar de 2.5510 lo que implica que el nivel de la recta varía según la densidad baja o alta del índice de vegetación. Los resultados mostrados con la estimación REML se reiteran con los estimadores ML.

En el modelo 3, los estimadores REML y ML también muestran resultados similares. La media por niveles del índice de vegetación es (15.9727 -14.3590 “DA” + 0.4175“Ntot”), según REML, y (16.3356 -14.3045 “DA” +0.2418 “Ntot”) según ML. Como en el modelo 2, se destaca la fuerte influencia de la densidad aparente sobre el secuestro de carbono y se aprecia que el efecto del nitrógeno total sobre los contenidos de carbono orgánico en el suelo no es significativo.

Teniendo en cuenta los criterios de selección de modelos, el modelo 2 puede considerarse el mejor de los tres modelos dado que los valores del AIC y del BIC son más pequeños. Sin embargo, para corroborar esta decisión, se llevó a cabo la prueba de razón de



verosimilitud; debido a que el p-valor es grande (Tabla 3) no debiéramos rechazar al modelo 2 a favor del modelo 3; preferimos el modelo 2 más parsimonioso y lo consideraremos para el análisis final.

Tabla2: Modelos Lineales Mixtos ajustados por Máxima Verosimilitud (ML) ['lmerMod']

<b>Modelo 1:</b> Corg ~ 1 + (1   Indveg)	<b>Modelo 2 :</b> Corg ~ 1 + DA + (1   Indveg)	<b>Modelo 3.</b> Corg ~ 1 + DA + Ntot + (1 Indveg)
AIC      BIC      logLik 163.135 167.338 -78.5676	AIC      BIC      logLik 148.5451 154.1499 -70.2726	AIC      BIC      logLik 150.399 157.4054 -70.1997
<b>Efectos aleatorios</b> Grupo    DesvEst Intercepto    0.00 Residuo            3.32	<b>Efectos aleatorios</b> Grupo    DesvEst Intercepto    0.000 Residuo            2.518	<b>Efectos aleatorios</b> Grupo    DesvEst Intercepto    0.000 Residuo    2.512
Numero de observaciones: 30 Grupo: Indveg con 2 niveles	Numero de observaciones: 30 Grupo: Indveg con 2 niveles	Numero de observaciones: 30 Grupo: Indveg con 2 niveles
<b>Efectos fijos</b> Estimador Intercepto    5.652	<b>Efectos fijos</b> Estimador Intercepto    16.51 DA                -14.37	<b>Efectos fijos</b> Estimador Intercepto 16.3356 DA                -14.3045 Ntot 0.2418

Tabla 3: Comparación de modelos

Modelos: Modelo 2: Corg ~ 1 + DA + (1   Indveg) Modelo 3: Corg ~ 1 + DA + Ntot + (1   Indveg) Df AIC BIC Chisq Chi DfPr(>Chisq) Modelo2 : 4 148.54 154.15 Modelo3:5 150.40 157.41 0.1457 1 0.7027
--

Después de varias exploraciones para detectar datos aberrantes, decidimos tomar como ajuste del modelo el que aparece en la tabla 4 con 28 observaciones. Los gráficos que se muestran en la figura 4 permiten comprobar el cumplimiento de las suposiciones y considerar que este modelo es adecuado para predecir el contenido de carbono orgánico en suelos según la imagen satelital del año 1994 y los estudios de propiedades del suelo realizado en ese período de tiempo.

Tabla 4: Modelo mixto para estimar CO<sub>2</sub> en suelo. (Año1994)

<b>Modelo 2 :</b> Corg ~ 1 + DA + (1   Indveg)
---

Criterio de convergencia de REML: 103.8

Residuos escalados:

Min	1Q	Mediana	3Q	Max
-1.9120	-0.6466	-0.1830	0.5156	2.0618

Efectos aleatorios:

Grupo	Varianza	Desv.Est.
Indveg (Intercepto)	0.2274	0.4769
Residuo	2.7751	1.6659

Número de observaciones: 28,  
Grupo: Indveg, 2

Efectos fijos:

Estimador	Error Est.	Valor de t
(Intercepto)	14.367	1.652
DA	-12.345	2.089

Correlación of Efectos Fijos:  
(Intercepto)  
DA -0.957

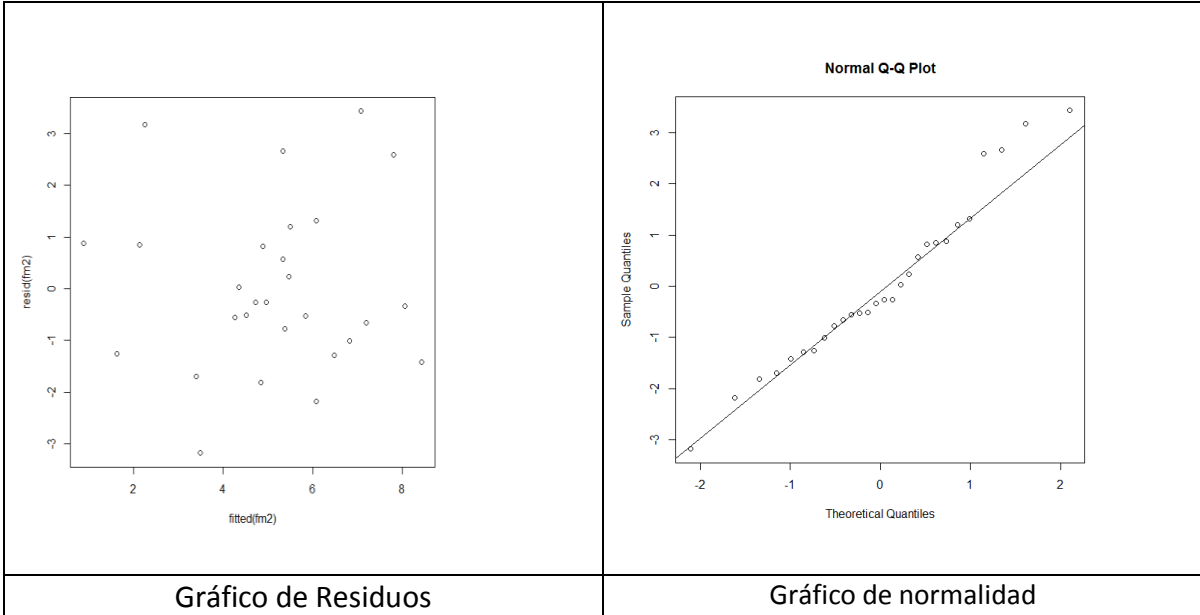


Figura 4: Gráficos para checar suposiciones (Año 1994)

## 4.2 Modelos Lineales Mixtos En 2005

La tabla de datos para ajustar modelos en este período contaba con 51 observaciones agrupadas según dos niveles del factor índice de vegetación (Indveg). Las tablas 5 y 6 muestran los estimadores REML y ML de los tres modelos analizados para la RTP 105 según la imagen satelital del año 2005.

En el modelo 1, según los estimadores REML, la media de las medias de los niveles del índice de vegetación es 4.1639 y los estimadores de los componentes de la varianza son cero para  $\sigma_b^2$  y 10.95 para la varianza de los residuos. Los estimadores ML muestran unamedia de 4.164y desviación estándar residual de 3.27, esto es, prácticamente resultados iguales.

En el modelo 2, la media por niveles del índice de vegetación es (13.404 -12.272 “DA”) según los estimadores REML, y (13.40 -12.27 “DA”) según los estimadores ML. Se destaca la fuerte influencia de la densidad aparente sobre el secuestro de carbono. Obsérvese que la constante del modelo 2 varía alrededor de 13.404 con una desviación estándar casi nula lo que implica que la recta no varía según la densidad baja o alta del índice de vegetación. Los resultados mostrados con la estimación REML se reiteran con los estimadores ML.

En el modelo 3, los estimadores REML y ML también muestran resultados similares. La media por niveles del índice de vegetación es (12.6722 -11.9713 “DA” + 1.3496“Ntot”), según REML, y lo mismo según ML. Como en el modelo 2, se destaca la fuerte influencia de la densidad aparente sobre el secuestro d carbono. Se aprecia que el efecto del nitrógeno total sobre los contenidos de carbono orgánico en el suelo es significativo.

Tabla 5: Modelos Lineales Mixtos ajustados por el criterio REML ['lmerMod']

<b>Modelo 1:</b> Corg ~ 1 + (1   Indveg)	<b>Modelo 2 :</b> Corg ~ 1 + DA + (1   Indveg)	<b>Modelo 3.</b> Corg ~ 1 + DA + Ntot + (1 Indveg)
Convergencia REML : 265.5	Convergencia REML : 247.4	Convergencia REML : 242.4
<b>Efectos aleatorios</b> Grupo VarianzaDesEst Interc. 0.00 0.000 Residuo 10.95 3.309	<b>Efectos aleatorios</b> Grupo VarianzaDesEst Interc. 0.00 0.000 Residuo 8.439 2.905	<b>Efectos aleatorios</b> Grupo Varianza DesEst Interc. 4.742e-14 2.178e-07 Resid7.942e+00 2.818e+00
Numero de observaciones: 51 Grupo: Indveg con 2 niveles	Numero de observaciones: 51 Grupo: Indveg con 2 niveles	Numero de observaciones: 51 Grupo: Indveg con 2 niveles
<b>Efectos fijos</b> EstimErrEst t Interc4.1639 0.4634 8.986	<b>Efectos fijos</b> EstimErrEst t Interc 13.404 2.354 5.693 DA -12.272 3.080 -3.984	<b>Efectos fijos</b> EstimErrEst t Interc12.6722 2.3126 5.480 DA -11.9713 2.9915 -4.002 Ntot1.3496 0.6691 2.017

	<b>Correlación de efectos fijos:</b> (Intercepto) DA -0.985	<b>Correlación de efectos fijos:</b> (Intercepto) DA DA -0.979 Ntot -0.157 0.050
--	---	---

Tabla 6: Modelos Lineales Mixtos ajustados por Máxima Verosimilitud (ML) ['lmerMod']

<b>Modelo 1:</b> Corg ~ 1 + (1   Indveg)	<b>Modelo 2 :</b> Corg ~ 1 + DA + (1   Indveg)	<b>Modelo 3.</b> Corg ~ 1 + DA + Ntot + (1   Indveg)
AIC BIC logLik271.7833 277.5787 - 132.8916	AIC BIC logLik259.4699 267.1972 - 125.7350	AIC BIC logLik 257.3208 266.9799 - 123.6604
<b>Efectos aleatorios</b> Grupo DesvEst Intercepto 0.000 Residuo 3.277	<b>Efectos aleatorios</b> Grupo DesvEst Intercepto 0.000 Residuo 2.848	<b>Efectos aleatorios</b> Grupo DesvEst Intercepto 0.000 Residuo 2.734
Numero de observaciones: 51 Grupo: Indveg con 2 niveles	Numero de observaciones: 51 Grupo: Indveg con 2 niveles	Numero de observaciones: 51 Grupo: Indveg con 2 niveles
<b>Efectos fijos</b> Estimador Intercepto 4.164	<b>Efectos fijos</b> Estimador Intercepto 13.40 DA -12.27	<b>Efectos fijos</b> Estimador Intercepto12.67 DA -11.97 Ntot1.35

Teniendo en cuenta los criterios de selección de modelos, el modelo 3 puede considerarse el mejor de los tres modelos dado que los valores del AIC y del BIC son más pequeños. Para corroborar esta decisión, se llevó a cabo la prueba de razón de verosimilitud (Tabla 7); debido a que el p-valor es pequeño (menor a un nivel de significación  $\alpha = 0.05$ ) se debe rechazar al modelo 2 a favor del modelo 3. El modelo 3 más será considerado para el ajuste final.

Tabla 7: Comparación de modelos

Modelos: Modelo 2: Corg ~ 1 + DA + (1   Indveg) Modelo 3: Corg ~ 1 + DA + Ntot + (1   Indveg) DfAIC BIClogLik Chi DfPr(>Chisq)
---

Modelo 2: 4 259.47 267.20 -125.73 Modelo 3: 5 257.32 266.98 -123.66 1 0.04166 *
--

Después de varias exploraciones para detectar datos aberrantes, decidimos tomar como ajuste del modelo el que aparece en la tabla 8 con 40 observaciones. Los gráficos que se muestran en la figura 5 permiten comprobar el cumplimiento de las suposiciones y considerar que este modelo es adecuado para predecir el contenido de carbono orgánico en suelos según la imagen satelital del año 2005 y los análisis de propiedades del suelo realizado en ese período de tiempo.

Tabla 8. Modelo mixto para estimar CO<sub>2</sub> en suelo. (Año2005)

<b>Modelo 3 :</b>			
<b>Corg ~ 1 + DA + Ntot + (1   Indveg)</b>			
Criterio de convergencia de REML: 155.9			
Efectos aleatorios:			
Grupo		Variance	Std.Dev.
Indveg (Intercepto)	0.2423		0.4922
Residuo	3.3017		1.8170
Número de observaciones: 40			
Grupo: Indveg, 2			
Efectos fijos:			
	Estimador	ErrorEst.	Valor de t
(Intercepto)	7.2509	1.9512	3.716
DA	-5.6929	2.3768	-2.395
Ntot 0.7968	0.4443	1.793	
Correlación of Efectos Fijos:			
(Intr) DA			
DA	-0.968		
Ntot	-0.096	0.021	

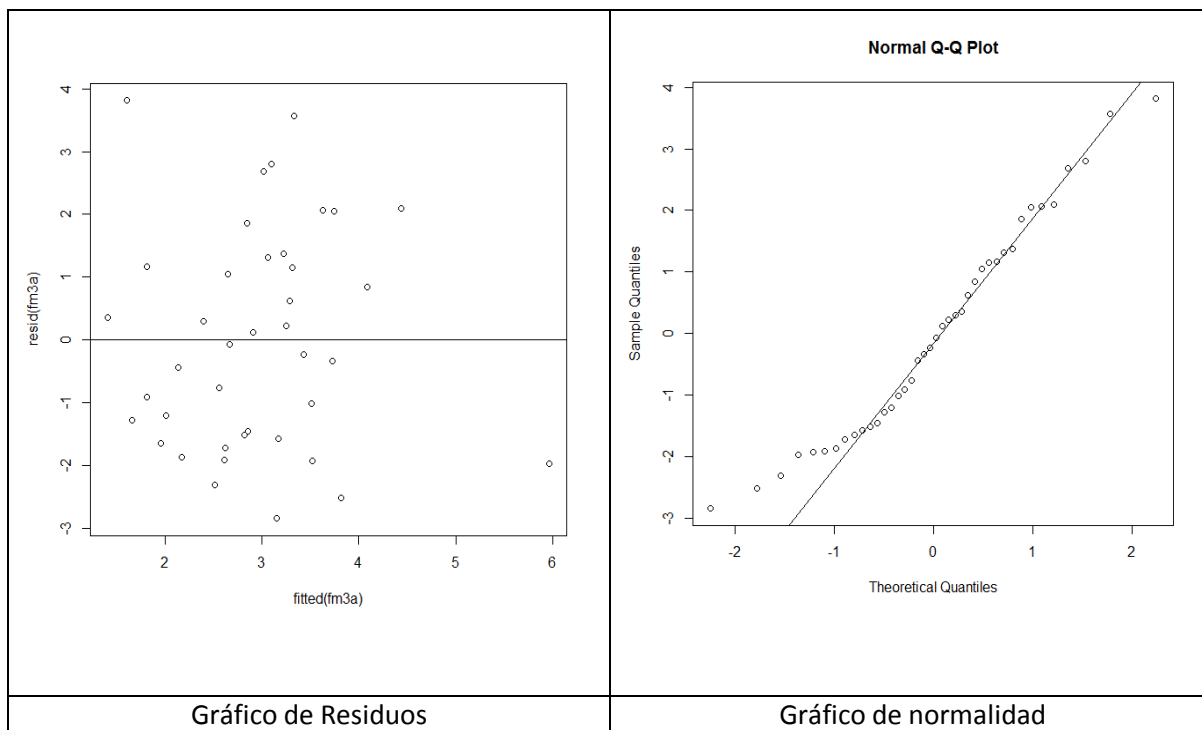


Figura 5: Gráficos para checar suposiciones (Año 2005).

## 5. CONCLUSIONES

El estudio de índices de vegetación en la Región Terrestre Prioritaria 105: Cuetzalan, Puebla, México, mediante el uso de sensores remotos, permitió contar con valores del índice de vegetación NDVI, y con ellos, visualizar el contraste entre dos clases de densidad de vegetación, la baja y la alta. Bajo la consideración que estos dos niveles son una muestra de la población de niveles del índice NDVI, se procedió a ajustar diferentes modelos de regresión lineal mixta.

Se identificaron dos modelos de regresión mixtos, uno por cada imagen satelital, que pueden ser útiles en la evaluación del carbono orgánico en suelo en la región considerada. Para la imagen del año 1994, el modelo estuvo en función de un único predictor (la densidad aparente de esos suelos), mientras que para la imagen del año 2005, el modelo con dos predictores (densidad aparente y porcentaje de nitrógeno) mostró las mejores características.

Profundizar en la modelación estadística para obtener buenas estimaciones y predicciones del carbono orgánico en suelos, es de suma importancia para aumentar el conocimiento sobre esta forma de mitigación del cambio climático: el secuestro de carbono.

## REFERENCIAS

[1] MASERA, O. and SHEINBAUM, C. (2004). **Mitigación de emisiones de carbono y prioridades de desarrollo nacional.** En: *Cambio climático: una visión desde México.*

Compiladores: Martínez, J. y Fernández, A. Instituto Nacional de Ecología. Secretaría del Medio Ambiente y Recursos Naturales. México, pp. 255-368.

[2] RODEGHIERO, M, HEINEMEYER, A., BELLAMY, P. (2010). **Determination of soil carbon stocks and changes.** En: *Soil Carbon Dynamics. An Integrated Methodology*. Kutsch, W. L. Bahn, M. and Heinemeyer, A. Cambridge University Press, UK, 49-75.

[3] ARCIDIACONO-BÁRSONYA, C., CIAISA, P., VIOVYA, N. VUICHARD, N. (2011). *REDD Mitigation*. Elsevier BV. **Procedia Environmental Sciences** 6. 50 – 59.

[4] ZHENG, D.; RADEMACHER, J.; CHEN, J.; CROW, T.; BRESEE, M.; LE MOINE, J. y RYU, S. (2004). *Estimating aboveground biomass using Landsat 7 ETM+ data across a managed landscape in northern Wisconsin, USA*. **Remote Sensing of Environment**. v.93, p.402– 411.

[5] ARRIAGA, L.; J.M. ESPINOZA; C. AGUILAR; E. MARTÍNEZ; L. GÓMEZ y E. LOA (Eds). (2000). **Regiones Terrestres Prioritarias de México. Comisión Nacional para el conocimiento y uso de la biodiversidad**, México.

[6] CHUVIECO, E. (2007). *Earth Observation and Global Change*. Nueva York: Springer. Geosciences.

[7] ROUSE, J.W., HAAS, R.H., SCHELL, J.A., DEERING, D.W. and HARLAN, J.C., (1974). Monitoring the vernal advancement of retrogradation of natural vegetation. NASA/GSFC, Type III, Final Report, Greenbelt, MD, USA, 371 p.

[8] TORRES, E.; LINARES, G.; TENORIO, M.G.; PEÑA, R.; CASTELÁN, R.; RODRÍGUEZ, A. (2014). *Obtención de índices de vegetación para estimar densidades de vegetación y cambios de Uso de Suelo en la Región Terrestre Prioritaria 105: Cuetzalan, Puebla, México*. ACEPTADO en **Revista Iberoamericana de Ciencias**. ISSN 2334-2501.

[9] PINHEIRO, J.C. and BATES, D.M. (2000). *Mixed-effects Models in S and S-PLUS*. New York: Springer-Verlag. 528 p

[10] DEMIDENKO, E. (2013). *Mixed models. Theory and applications with R*. 2<sup>nd</sup> edition. Wiley, New York, USA.

[11] GALECKI, A., and BURZYKOWSKI, T. (2013). *Linear mixed-effects models using R. A step by step approach*. 1<sup>st</sup> edition. Springer, New York, 542 p.

[12] MEHTÄTALO, L. (2013). *Linear mixed-effects models with examples in R*. University of Eastern Finland. 158 p.

[13] BATES, D., MAECHLER, M and BOLKER, B. (2011). *lme4: Linear mixed-effects models using Eigen and Eigen*. R package version 0.999375-42.