

Ministerio de Agricultura y Ganadería

Ministerio del Ambiente, Agua  
y Transición Ecológica



Organización de las Naciones Unidas  
para la Alimentación y la Agricultura



**2da  
edición**

Quito, 2021

Memoria Técnica

# Mapeo digital de Carbono orgánico en los suelos del Ecuador



República  
del Ecuador



Juntos  
lo logramos

**Ministerio de Agricultura y Ganadería**

Coordinación General de Información Nacional Agropecuaria  
Dirección de Generación de Geoinformación Agropecuaria

**Ministerio del Ambiente, Agua y Transición Ecológica**

Subsecretaría de Patrimonio Natural  
Subsecretaría de Cambio Climático  
Sistema Único de Información Ambiental

**Organización de las Naciones Unidas para la Alimentación y la Agricultura-FAO**

**Alianza Mundial por el Suelo**

**Proyecto Regional Andes Resilientes al Cambio Climático**

Ing. Pedro José Álava González  
**Ministro de Agricultura y Ganadería**

Ing. Gustavo Rafael Manrique Miranda  
**Ministro del Ambiente, Agua y Transición Ecológica**

Mat. Víctor Hugo Bucheli León  
**Coordinador General de Información Nacional Agropecuaria**

Mag. Karina Maribel Barrera Moncayo  
**Subsecretaria de Cambio Climático**

Ing. Glenda Givabel Ortega Sánchez  
**Subsecretaria de Patrimonio Natural**

Ing. Viviana Carolina Ruiz Villafuerte  
**Directora de Generación de Geoinformación Agropecuaria**

Econ. Agustín Zimmerman  
**Representante de la FAO en Ecuador**

### **Equipo técnico desarrollador del estudio**

Ing. Wilmer Jiménez

Punto Focal del Pilar 4 en Ecuador de la Alianza Mundial por el Suelo  
Unidad de Suelos MAG-CGINA-DGGA

Ing. Darwin Sánchez

Unidad de Suelos MAG-CGINA- DGGA

Ing. Pablo Caza

Especialista de Adaptación al Cambio Climático MAAE-DACC

Ing. Digner Jiménez

Especialista en Administración y Control Forestal MAAE-SUIA

Lic. Miguel Chinchero

Especialista en Calidad de los Recursos Naturales MAAE-SUIA

Ing. Ximena Herrera

Especialista Técnico en Monitoreo Forestal e Inventarios Forestales

Ing. Jorge Armijos

Especialista en Nivel de Referencia de Emisiones Forestales

Mag. Andrés Mogro

Coordinador Nacional en Ecuador del Proyecto Regional Andes Resilientes al Cambio  
Climático

Daphne Armas, PhD (c)

Consultora

## Mapeo digital de Carbono orgánico en los suelos del Ecuador. Segunda Edición.

### Memoria Técnica

#### Cita recomendada:

MAG, MAATE, FAO, GSP. 2021. *Mapeo digital de Carbono orgánico en los suelos del Ecuador. Segunda edición. Memoria Técnica*. Quito, Ecuador.

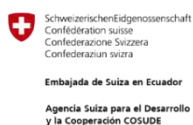
Fotografía de la portada: ©MAG/Wilmer Jiménez

Esta obra se encuentra disponible bajo la licencia de Creative Commons Attribution Non Commercial-ShareAlike 3.0 IGO (CC BY-NC-SA 3.0 IGO; <https://creativecommons.org/licenses/by-nc-sa/3.0/igo/legalcode>). Bajo los términos de esta licencia, esta obra podrá ser copiada, redistribuida y adaptada con fines no comerciales a condición de que la obra sea citada apropiadamente.

Con el apoyo financiero de:



Impulsado por:



Consortio facilitador:



## CONTENIDO

<b>1. ANTECEDENTES</b> .....	4
<b>2. OBJETIVOS</b> .....	6
<b>3. METODOLOGÍA</b> .....	6
<b>4. RESULTADOS Y DISCUSIÓN</b> .....	10
<b>5. CONCLUSIONES</b> .....	13
<b>6. RECOMENDACIONES</b> .....	14
<b>7. REFERENCIAS</b> .....	15
<b>8. ANEXOS</b> .....	16
<b>Anexo 1.</b> Preparación de la información .....	16
<b>Anexo 2.</b> Mapeo de carbono orgánico del suelo de Ecuador .....	36

## 1. ANTECEDENTES

La Convención de las Naciones Unidas de Lucha contra la Desertificación (CNULD) ha dado un paso importante hacia la neutralidad en la degradación de la tierra (NDT) en las decisiones adoptadas en su duodécima Conferencia de las Partes (COP), celebrada en Ankara - Turquía, en octubre de 2015 la cual prioriza alcanzar el objetivo 15.3 de los ODS, como un vehículo sólido para impulsar la implementación de la CNULD”.

Al mismo tiempo, la meta 15.3 es relevante para las otras Convenciones de Río: Convención sobre la Diversidad Biológica (CDB) y Convención Marco de las Naciones Unidas sobre el Cambio Climático (CMNUCC). Es así, que la implementación de NDT apuntala y cataliza el logro de otros ODS y sus objetivos relacionados. Es decir, el ODS 13 sobre cambio climático es particularmente relevante para la CMNUCC, ya que se evidencia múltiples relaciones y retroalimentaciones entre la tierra y el clima. Los objetivos relacionados con la biodiversidad en el ODS 15 muestran vínculos claros con el CDB y respalda muchos de los procesos que sustentan el funcionamiento del ecosistema de la tierra. Desarrollar acciones interconectadas que abarquen los intereses de las Convenciones de Río es vital para avanzar hacia NDT, especialmente a nivel nacional, donde será necesario sumar esfuerzos para el cumplimiento de las acciones.

El concepto de la neutralidad de la degradación de la tierra (NDT) se presentó oficialmente en 2011 en la CNULD en la Décima Conferencia de las Partes y luego ampliamente reconocida en los Estados Conferencia de Naciones sobre Desarrollo Sostenible (ONU, 2012). En 2015 la CNULD declaró el concepto de la NDT como parte de la agenda central del Convenio y se definió como: "un estado por el cual la cantidad y calidad de los recursos de la tierra, necesarios para apoyar las funciones del ecosistema y servicios y mejorar la seguridad alimentaria, se mantiene estable o aumenta dentro de escalas y ecosistemas temporales y espaciales especificados” (CNULD, 2015).

La información precisa sobre los indicadores de cambio y uso de la tierra, productividad neta primaria y reservas de carbono en el suelo es crítica para comprender las causas del cambio y para desarrollar políticas y estrategias efectivas para frenar y revertir la degradación de la tierra.

En el marco del último indicador, el Ecuador cuenta con el Mapa de Carbono Orgánico del Suelo (COS) hasta 30 cm de profundidad con su incertidumbre asociada, a una resolución espacial de 1 km, realizado en colaboración con la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO, por sus siglas en inglés) y el Pilar 4 de la Alianza Mundial por Suelo (AMS) del Ministerio de Agricultura y Ganadería (MAG) (Loayza, et al. 2020; MAG & FAO, 2018 ), sin embargo, en la

primera versión del mapa sólo se utilizaron datos de los perfiles de suelos del área agropecuaria del país, los cuales fueron levantados mediante el convenio MAGAP, IEE y SENPLADES, por ello, se ve la necesidad de actualizar la información del mapa (Versión 2), en función de que se conoce que existen datos que han sido generados por el MAAE a través de la implementación del Proyecto Evaluación Nacional Forestal, llevada a cabo en el periodo 2009 - 2013, en la que además de otras variables que hacen referencia al estado de los bosques, se recopiló información en territorio sobre variables asociados al cálculo de los contenidos de carbono orgánico del suelo. El uso articulado y conjunto de los datos generados y referidos en los procesos implementados por ambas carteras de estado permitirá obtener un modelo más preciso que nos ayude a la toma de decisiones y a la gestión de diferentes futuros productos. Hay que tener en cuenta sin embargo que esta información fue colectada mediante diferentes proyectos con objetivos y metodologías propias por lo cual se vio en la necesidad de realizar una homologación y análisis los datos previos para así tener una base de datos estandarizada con la cual se trabajó para la construcción del mapa de carbono orgánico de suelos en el Ecuador.

## 2. OBJETIVOS

### General:

Desarrollar la segunda versión del mapa de carbono orgánico en suelos utilizando la información levantada por el Ministerio del Ambiente y Agua y el Ministerio de Agricultura y Ganadería.

### Específicos:

- Desarrollar la versión 2 del mapa de carbono orgánico del suelo a través del uso de la base de datos homologada de la información levantada por el MAAE y MAG, a nivel nacional y provincial.
- Desarrollar un análisis de incertidumbre del mapa, a través de la metodología de la Alianza Mundial por el Suelo.

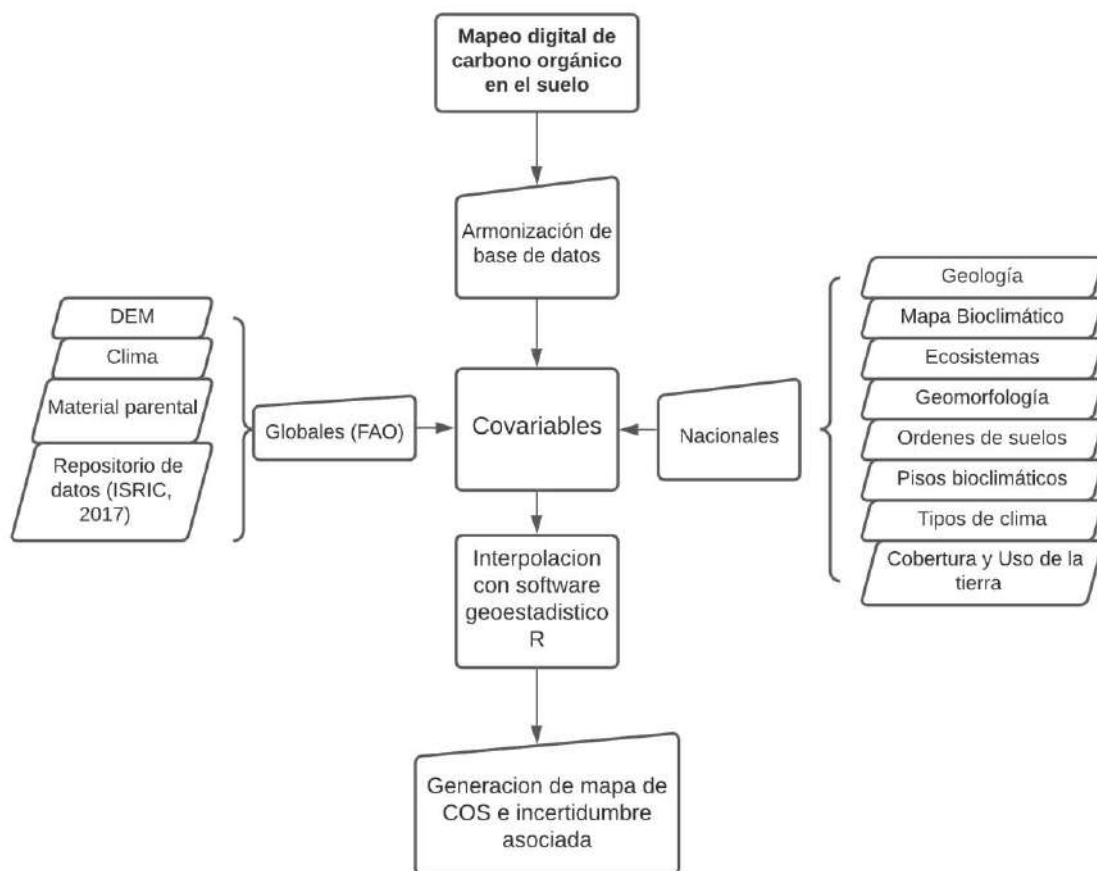
## 3. METODOLOGÍA

Para la elaboración de la versión 2 del mapa de carbono orgánico del suelo a través del uso de la base de datos homologada de la información levantada por el MAAE y MAG, a nivel nacional y provincial se aplicó la metodología Soil Organic Carbon Mapping CookBook, elaborada por la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO 2017, por sus siglas en inglés).





En la Figura 2, se muestra en general la metodología empleada para la generación del mapa de COS a nivel nacional. En esta fase se realizó el ajuste en la base de datos, preparación y generación de covariables: bioclimáticas, geológicas, tipos de clima, geomorfológicos, órdenes de suelos, uso y cobertura de la tierra y ecosistemas para Ecuador. Algunas de estas covariables son un conjunto de variables, un ejemplo de esto es el modelo digital de elevación (DEM) que luego de procesarlo nos da 17 nuevas variables. Con estos insumos se procesó la información mediante el software geoestadístico R con su interfaz RStudio y aplicando las herramientas de interpolación espacial de Rkriging, se obtuvo un modelo de regresión lineal que predice un valor medio del contenido de carbono en los suelos, obteniéndose el mapa. Igualmente, se realizó el mapa de incertidumbre, ya que es la cual mide la probabilidad de presencia del valor verdadero en un intervalo estadístico y así podemos identificar que zonas presentan valores más confiables de valores de COS y cuáles no.



**Figura 2.** Diagrama de proceso metodológico para el mapeo digital de carbono orgánico en el suelo a nivel nacional

Para este proceso se utilizaron dos scripts, el primero que se centra en la preparación de los datos para poder correr el modelo, y el segundo script está enfocado en correr el modelo en sí. En el anexo 1, se encuentra la descripción del primer script utilizado para la preparación de la información y en el Anexo 2 se encuentra la descripción del script utilizado para generar el modelo y la generación de la versión 2 del mapa de carbono orgánico en suelos del Ecuador y la incertidumbre asociada a él. Ambos documentos se los realizó mediante la herramienta R Markdown que es un tipo de documento de RStudio que permite integrar texto con código de R para así registrar de mejor manera que es lo que ocurre en cada parte.

Para la generación del modelo se hizo la selección de variables mediante pasos sucesivos (Stepwise Regression) donde en cada paso se introduce la variable independiente que no se encuentre ya en la ecuación y que tenga la mayor probabilidad, las variables ya introducidas en la ecuación de regresión pueden ser eliminadas del modelo (i.e. hacia atrás). El método termina cuando ya no hay más variables candidatas a ser incluidas o eliminadas. Para saber cuál es el mejor modelo se lo hizo con el criterio de información de Akaike AIC.

En la tabla 1 se pueden encontrar las variables que se tomaron en cuenta para el modelo

**Tabla 1.** Covariables ambientales para la estimación del COS.

COVARIABLE	Descripción	Detalle	Tipo de variables
x2mod3a		Feb/Mar	
Tx1mod3a	Valor medio de la temperatura en la superficie terrestre por periodos	Dic/Ene	Global
Tx5mod3a		Ago/Sep	
Tx6mod3a		Oct/Nov	
VerticalDistanceToChannelNetwork	Distancia vertical a la red de drenajes	En metros, derivada de la elevación del terreno	Nacional
Tdhmod3a	Temperatura en la superficie terrestre	Valor máximo	Global
Tdmmod3a		Valor medio	
G14esa3a	Cobertura y uso de la tierra	matorral (<5 m)	Nacional
Evmmod3a	Índices de vegetación basados en teledetección	Mean value of the monthly MODIS EVI	Global

#### 4. RESULTADOS Y DISCUSIÓN

Como productos finales creados en esta etapa tenemos la versión 2 del mapa nacional de carbono orgánico en los suelos el cual muestra que el contenido más bajo de carbono orgánico en los suelos de Ecuador fluctúa entre 6 t/ha que principalmente se localiza en las zonas costeras amazonia y en la provincia insular, posiblemente de puntos de muestreo en esas zonas. Mientras que en valores mayores de contenido de carbono se observó un aumento al número que alcanza en esta segunda versión con valores que llegan hasta 321 t/ha que es superior al alcanzado en la versión 1 que solo presentó un valor de hasta 255 t/ha, cabe mencionar que estos valores más altos se dan en zonas específicas, pero al igual que en la versión 1 están presentes mayormente en la región sierra (Figura 3).

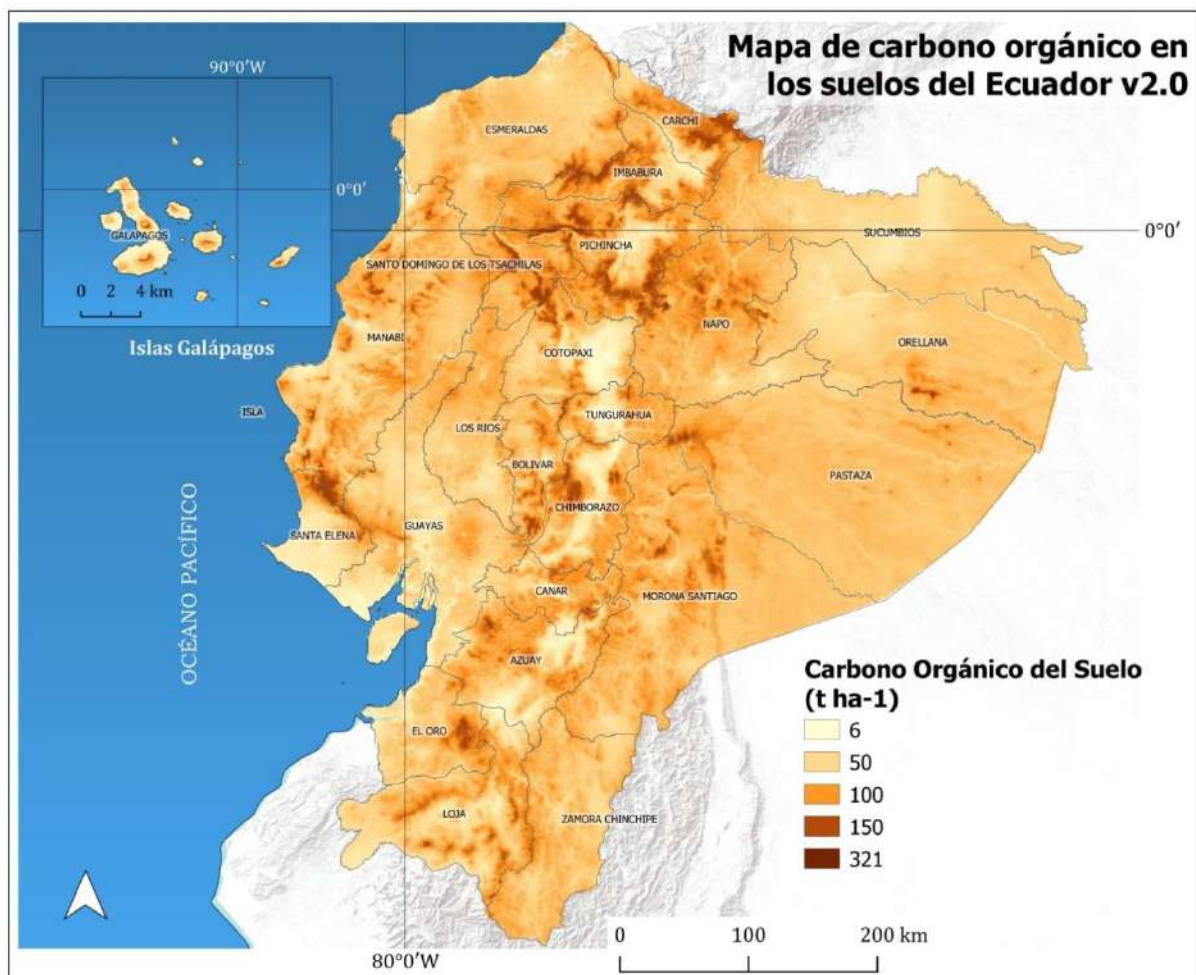
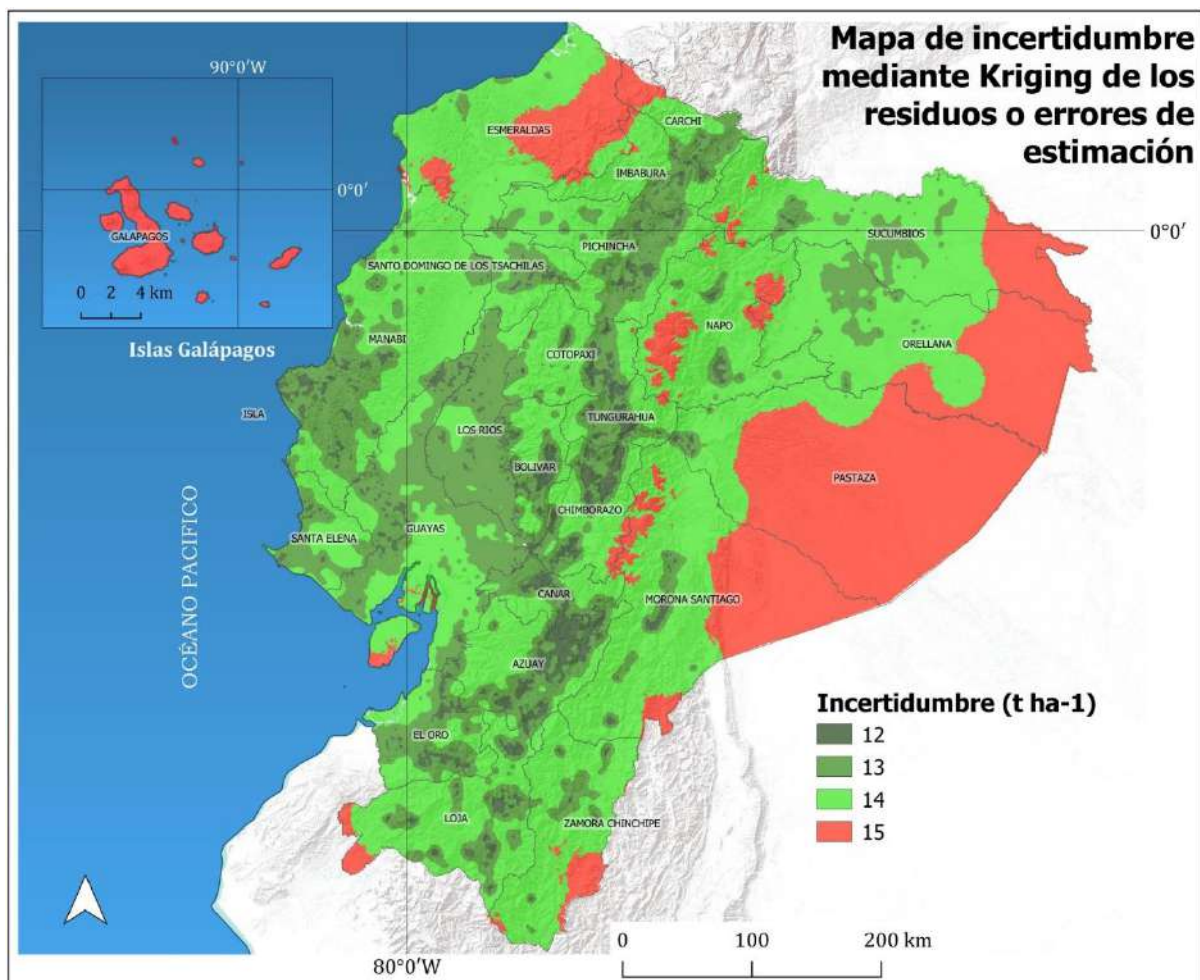


Figura 3. Mapa de carbono orgánico en los suelos del Ecuador V02

El mapa de incertidumbre (Figura 4) ilustra claramente el efecto de la calidad de muestreo en mapeo digital de suelos. La mayoría de las muestras (Figura. 1) están distribuidas en suelos agrícolas, por lo que los niveles de incertidumbre en esas áreas son aceptables. En áreas pobremente muestreadas como en la región de la Amazonía, Islas Galápagos y la provincia de Esmeraldas presentan un error de entre  $\pm 12.8$  y  $\pm 13.56$  la incertidumbre es alta, indicando que no es posible confiar en las predicciones, esto también se encontró en la versión 1 del mapa. En general el rango máximo de incertidumbres es entre  $\pm 12.22$  y  $\pm 15.18$  t ha<sup>-1</sup>, que son valores un poco más altos comparado con la versión anterior, pero esto puede deberse a los nuevos datos utilizados los cuales se encontraban esparcidos por zonas que antes no estaban representadas.



**Figura 4.** Mapa de Incertidumbre mediante Kriging de los residuos o errores de estimación

Se realizó el análisis de COS en cada una de las 24 provincias (Tabla 2) donde podemos ver que las provincias que presentan los valores de las medias de COS más bajos son Galápagos (48,60 t/ha), Guayas (57,48 t/ha) y Sucumbíos con 57,74 t/ha, las dos primeras era de esperarse por la falta de puntos de muestreo pero en la zona de la provincia de Guayas vemos que la cantidad de puntos es representativa y que la incertidumbre en esta zona está entre los valores más bajos (Figura 4) lo cual hace que este dato de COS muy fiable. Otro dato que hay que mencionar es el de la Provincia de Esmeralda la cual no tiene muchos datos, pero el valor de COS que presenta no es tan bajo (67,21 t/ha) como de las otras tres provincias antes mencionadas. Al hablar de valores de COS altos la provincia con la media más alta es Santo Domingo de los Tsáchilas con 100,40 t/ha seguido por Carchi (96,40 t/ha) y Pichincha (96,11 t/ha), las tres provincias como era de esperarse están muy bien representadas con respecto a datos y abarcan zonas con gran representatividad de diversos cultivos.

**Tabla 2.** Concentración de carbono orgánico en suelos t/ha en cada una de las 24 provincias del Ecuador.

PROVINCIA	MÍNIMO	MÁXIMO	MEDIA	MEDIANA	DES. ESTAN.	VARIANZA	OBLICUIDAD	CURTOSIS
AZUAY	8	184	72,62	75,0	27,67	765,93	-0,40	2,65
BOLIVAR	25	176	82,78	77,0	25,07	628,51	0,79	3,32
CAÑAR	17	170	73,68	75,0	24,91	620,79	0,08	2,80
CARCHI	20	215	96,40	98,0	35,51	1261,58	0,23	2,62
CHIMBORAZO	14	197	75,64	77,5	31,14	969,97	0,20	2,81
COTOPAXI	6	205	63,64	62,0	30,56	934,18	0,63	3,70
EL ORO	20	191	71,36	66,0	26,44	699,54	1,17	4,64
ESMERALDAS	13	227	67,21	64,0	18,48	341,72	1,87	9,83
GALÁPAGOS	16	182	48,60	40,0	27,08	733,38	1,15	5,43
GUAYAS	9	321	57,48	55,0	21,79	475,15	1,84	16,37
IMBABURA	16	197	88,45	89,0	31,22	974,93	0,15	2,85
LOJA	10	165	69,20	66,0	24,61	605,72	0,46	2,78
LOS RÍOS	25	216	66,14	65,0	15,17	230,18	0,80	8,87
MANABÍ	17	261	75,64	72,0	23,55	554,94	1,26	7,00
MORONA SANTIAGO	26	232	75,58	72,0	16,84	283,60	1,78	11,74
NAPO	23	190	81,96	80,0	21,15	447,61	0,49	3,83
ORELLANA	14	222	61,91	61,0	16,00	256,13	2,33	17,90
PASTAZA	27	210	69,06	67,0	11,53	133,07	3,08	20,97
PICHINCHA	16	248	96,11	94,0	32,82	1077,74	0,46	3,92
SANTA ELENA	11	212	59,57	47,0	35,38	1252,22	1,11	3,70
SANTO DOMINGO DE LOS TSÁCHILAS	41	265	100,40	93,0	33,01	1090,13	1,80	6,87
SUCUMBÍOS	15	197	57,74	57,0	14,36	206,39	0,97	5,84
TUNGURAHUA	12	181	72,74	79,0	27,39	750,57	-0,27	2,65
ZAMORA	33	127	68,14	67,0	13,13	172,56	0,49	3,45

## 5. CONCLUSIONES

La predicción espacial del contenido de COS hasta 30 cm de profundidad en el Ecuador fue posible usando las técnicas de mapeo digital de suelos con la cual se pudo generar la versión 2 del mapa de predicción de COS, junto a mapas de incertidumbre, ambos a una resolución de 1 km. La estimación de la incertidumbre es una evaluación esencial de la calidad de las predicciones.

Se pudo obtener esta segunda versión gracias a la disponibilidad de nuevos datos de suelos tomados en zonas que antes no estaban presentes lo cual dio información nueva para poder generar un modelo más acorde a la realidad, pero se sigue haciendo evidente la falta de datos en algunas zonas. En esta nueva versión se obtuvo un valor de  $1.72 \text{ Pg} \pm 0.12$  total de COS a nivel nacional, lo cual al comparar con lo obtenido en la versión 1 que presentó  $1.63 \text{ Pg} \pm 0.38$  denota un aumento. Estos mapas son referenciales ya que no toman en cuenta la variación temporal de carbono orgánico del suelo ni las variables que intervienen específicamente en cada bioma.

La selección de variables al igual que los datos se lo hizo para un contexto nacional y para una resolución de 1km, es por esto que los datos podrían variar si se los compara con datos a una resolución más fina o para un ecosistema específico.

La información de este estudio será útil para la identificación de zonas potenciales para el secuestro de carbono del suelo, así como para la demarcación de las zonas más vulnerables. También apoyará a los tomadores de decisiones en la formulación de políticas para el manejo de recursos naturales en los diferentes ecosistemas.

## 6. RECOMENDACIONES

- Se recomienda la actualización constante y la mejora de la calidad de los modelos de COS especialmente en países con alta heterogeneidad en sus ecosistemas, como el Ecuador. Adicionalmente es necesario un monitoreo pues el contenido de carbono en los suelos está sujeto al manejo de este.
- Se recomienda en futuras investigaciones plantear un diseño de muestreo probabilísticos que permitan aumentar la capacidad predictiva de los modelos e involucrar la variación muestral dentro de la estimación de la incertidumbre.
- Se recomienda realizar análisis más puntuales para diferentes ecosistemas (biomas), haciendo un modelo más fino, con variables específicas para cada ecosistema y así sacar valores reales que ayuden a la gestión puntual de estas zonas.
- Se recomienda para futuras actualizaciones del mapa, si es posible tener una fase de campo donde se pueda validar el modelo, enfocado en zonas con datos atípicos.

## 7. REFERENCIAS

- CNULD (Convención de las Naciones Unidas de Lucha contra la Desertificación), (2015). *Climate change and desertification: Anticipating, assessing & adapting to future change in drylands*. Impulse report for 3rd SC-UNCCD. Ed. UNCCD co-edition of Agropolis International. ISBN: 978-2-35682-379-3 Montpellier, France.
- FAO, (2017). *Soil Organic Carbon Mapping Cookbook*, <http://www.fao.org/3/ab901e.pdf> (accessed 11 January 2018).
- IEE (Instituto Espacial Ecuatoriano), (2015). *Proyecto de Generación de Geoinformación para la Gestión del Territorio Nacional a escala "...", socioeconómico, vulnerabilidad y riesgos, elaborada a nivel cantonal*.
- Loayza, (2020). V., Sevilla, V., Olivera, C., Guevara, M., Olmedo, G., Vargas, R., Oyonarte, C., y Jiménez, W. *Mapeo digital de carbono orgánico en suelos de Ecuador. Ecosistemas 29(2):1852*.
- MAATE, (2014). *Evaluación Nacional Forestal del Ecuador (ENF)*.
- MAG & FAO, (2018). *Mapeo Digital de Carbono Orgánico en los Suelos del Ecuador*. Memoria técnica. Ministerio de Agricultura y Ganadería del Ecuador - Coordinación General del Sistema de Información Nacional, Organización de las Naciones Unidas para la Alimentación y la Agricultura. Quito, Ecuador.
- MAGAP, (2015). *Levantamiento de Cartografía Temática escala 1:25.000, lote 1 y 2*. Sistema Nacional de Información y Gestión de Tierras Rurales e Infraestructura Tecnológica (SIGTIERRAS).
- ONU, (2012). *La Convención de las Naciones Unidas de Lucha contra la Desertificación en la Décima Conferencia de las Partes*.
- UNFCCC, (2011). *Decision 2/CMP.7 of the Conference of the Parties serving as the meeting of the Parties to the Kyoto Protocol*, adopted by the 17th Conference of the Parties of the UNFCCC meeting in Durban.



## 8. ANEXOS

### Anexo 1. Preparación de la información

```
## Establecemos el directorio de trabajo  
setwd("C:/Users/usuario/Desktop/SCRIPT2/Ecuador")
```

#### Instalación de paquetes

```
library(sp)  
library(rgdal)  
library(raster)  
library(GSIF)  
library(aqp)  
library(plyr)  
library(gstat)  
library(tiff)  
  
##adicionales##  
  
library(lattice)  
library(ggplot2)  
  
library(car)  
library(caret)  
library(knitr)  
library(shapefiles)  
library(splines)  
library(plotKML)  
library(ggmap)
```

#### Cargar funciones adicionales

```
## Support functions  
load("DSM_supportfunctions.RData")  
Cargar datos de suelos (base de datos)  
# Cargamos datos de suelos  
dat=read.csv("BD_COMPLETA.csv", sep=";")  
dat=data.frame(dat)
```

```
str(dat)
```

```
## 'data.frame': 13251 obs. of 7 variables:  
## $ ID : chr "CG1-P003" "CG1-P003" "CG1-P004" "CG1-P004" ...
```

```
## $ latitude : chr "-2,23" "-2,23" "-2,22" "-2,22" ...
## $ longitude: chr "-79,5" "-79,5" "-79,58" "-79,58" ...
## $ top      : int 0 0 0 0 0 0 0 0 0 ...
## $ bottom   : int 30 30 30 30 30 30 30 30 30 ...
## $ SOC      : chr "24,36" "26,68" "26,22" "22,56" ...
## $ BLD      : chr "1,4" "1,49" "1,39" "1,53" ...

dat_subset=data.frame(id=dat$ID,
                      dat[,2:3],
                      dat[,4:5],
                      SOC=dat$SOC,
                      BLD=dat[,7])

str(dat_subset)

## 'data.frame': 13251 obs. of 7 variables:
## $ id : chr "CG1-P003" "CG1-P003" "CG1-P004" "CG1-P004" ...
## $ latitude : chr "-2,23" "-2,23" "-2,22" "-2,22" ...
## $ longitude: chr "-79,5" "-79,5" "-79,58" "-79,58" ...
## $ top : int 0 0 0 0 0 0 0 0 0 ...
## $ bottom : int 30 30 30 30 30 30 30 30 30 ...
## $ SOC : chr "24,36" "26,68" "26,22" "22,56" ...
## $ BLD : chr "1,4" "1,49" "1,39" "1,53" ...
```

```
summary(dat_subset)
```

```
##      id          latitude      longitude      top
## Length:13251 Length:13251 Length:13251 Min. :0
## Class :character Class :character Class :character 1st Qu.:0
## Mode :character Mode :character Mode :character Median :0
##                                         Mean :0
##                                         3rd Qu.:0
##                                         Max. :0
##      bottom      SOC          BLD
## Min. : 2.0 Length:13251 Length:13251
## 1st Qu.:20.0 Class :character Class :character
## Median :30.0 Mode :character Mode :character
## Mean :24.8
## 3rd Qu.:30.0
## Max. :30.0
```

Guarda el nuevo archivo en formato .csv

```
### Vamos a grabar estos datos en un archivo csv
write.csv2(dat_subset, 'alldata1.csv', row.names = F)

## Se puede abrir el archivo alldata.csv con Excel y revisar los datos o aumentar
nuevos datos si es necesario
## Luego guardamos el archivo como alldata.csv y continuamos

## Cargamos los datos
dat_subset <- read.csv2('alldata1.csv')
dat_subset$id <- paste(dat_subset$id, dat_subset$latitude, dat_subset$longitude, sep
= "_")
summary(dat_subset)
```

```
##      id          Latitude      Longitude      top
## Length:13251      Min.    :-4.980      Min.    :-81.00      Min.    :0
## Class :character  1st Qu.:-2.920      1st Qu.:-79.71      1st Qu.:0
## Mode  :character  Median :-1.700      Median :-79.08      Median :0
##                               Mean  :-1.786      Mean   :-79.08      Mean   :0
##                               3rd Qu.:-0.770      3rd Qu.:-78.57      3rd Qu.:0
##                               Max.   : 1.385      Max.   :-75.46      Max.   :0
##      bottom      SOC          BLD
## Min.    : 2.0      Min.    : 0.00      Min.    :0.24
## 1st Qu.:20.0      1st Qu.: 13.69      1st Qu.:0.97
## Median :30.0      Median : 23.78      Median :1.08
## Mean   :24.8      Mean   : 31.32      Mean   :1.13
## 3rd Qu.:30.0      3rd Qu.: 40.43      3rd Qu.:1.38
## Max.   :30.0      Max.   :277.03      Max.   :1.69
```

### Revisión de datos

```
## Revisamos los datos por problemas como datos faltantes
# Profundidades incompletas, etc
# If we find problem we should open the data again in Excel,
# correct the wrong data and go back to line 37
dat_subset[is.na(dat_subset$bottom),5] <- dat_subset[is.na(dat_subset$bottom),4] + 10
checkdata(dataframe = dat_subset)
```

```
##                               id Latitude Longitude top bottom SOC BLD rejected
## 12964 44052_0.2326_-77.5126  0.2326 -77.5126  0    30  0 0.64    TRUE
```

```
## Eliminamos datos incompletos
dat_subset <- dat_subset[complete.cases(dat_subset[,1:6]),]
```

```
dat_subset$SOC <- as.numeric(dat_subset$SOC)
```

Estimación de densidad aparentes, esta parte no se la realizo ya que en la base de datos ya teníamos ese dato, pero se lo deja en el script ya que puede ser necesario para otros proyectos no se conste con este dato. Esto se lo hace siguiendo las recomendaciones del Cookbook de la FAO.

```
# Available methods:
# Saini_1996, Drew_1973, Jeffrey_1979, Grigal_1989, Adams_1973,
# Honeyset_Ratkowsky_1989
dat_subset$BLD <- estimateBD(dat_subset$SOC, method="Grigal_1989")
summary(dat_subset)
```

Calculo de stock de carbono orgánico COS debe estar en g/kg y la densidad aparente en kg/m3

```
# CRF in percentage
OCSKGM <- OCSKGM(ORCDRC = dat_subset$SOC, BLD = dat_subset$BLD*1000, CRFVOL = 0,
HSIZE = dat_subset$bottom - dat_subset$top)
```

Se crea dos campos nuevos en la base de datos el Carbono y el error

```
dat_subset$OCSKGM <- OCSKGM
dat_subset$meaERROR <- attr(OCSKGM, "measurementError")
```

Se convierte la base de datos en una colección de perfiles de suelos que tienen una profundidad determinada y se multiplica el Carbono por 100 para luego poder calcular el COS.

**#Convertimos a Soil profile collection**

```
dat_subset[dat_subset$OCSKGM>100,]
```

```
## [1] id Latitude Longitude top bottom SOC BLD
## [8] OCSKGM meaERROR
## <0 rows> (or 0-length row.names)
```

```
dat_aqp <- dat_subset
write.csv(dat_aqp, 'aqp_VF_8.csv')
summary(dat_aqp)
```

```
## id Latitude Longitude top
## Length:13251 Min. :-4.980 Min. :-81.00 Min. :0
## Class :character 1st Qu.: -2.920 1st Qu.: -79.71 1st Qu.:0
## Mode :character Median : -1.700 Median : -79.08 Median :0
## Mean : -1.786 Mean : -79.08 Mean :0
## 3rd Qu.: -0.770 3rd Qu.: -78.57 3rd Qu.:0
## Max. : 1.385 Max. : -75.46 Max. :0
## bottom SOC BLD OCSKGM
## Min. : 2.0 Min. : 0.00 Min. :0.24 Min. : 0.000
## 1st Qu.:20.0 1st Qu.: 13.69 1st Qu.:0.97 1st Qu.: 3.854
## Median :30.0 Median : 23.78 Median :1.08 Median : 6.417
## Mean :24.8 Mean : 31.32 Mean :1.13 Mean : 7.585
## 3rd Qu.:30.0 3rd Qu.: 40.43 3rd Qu.:1.38 3rd Qu.: 9.988
## Max. :30.0 Max. :277.03 Max. :1.69 Max. :84.948
## meaERROR
## Min. :0.214
## 1st Qu.:2.160
## Median :3.110
## Mean :3.084
## 3rd Qu.:4.210
## Max. :8.610
```

```
depths(dat_aqp) <- id ~ top + bottom
site(dat_aqp) <- ~ Longitude + Latitude
coordinates(dat_aqp) <- ~ Longitude + Latitude
```

Para estandarizar la profundidad a los 30 cm, estimar y rellenar la información faltante de los perfiles del suelo dentro de esta profundidad, se usa la aplicación spline.

**## mpsspline**

```
library(GSIF)
try(OCSKGM30 <- mpsspline(dat_aqp, 'OCSKGM', d = t(c(0,30))))
```

## Fitting mass preserving splines per profile...

Se genera un nuevo campo `OCSKGM30` donde con el cual se calculará el COS

```

dat_subset <- data.frame(id = dat_aqp@site$id,
  latitude = dat_aqp@sp@coords[,2],
  longitude = dat_aqp@sp@coords[,1],
  OCSKGM30 = OCSKGM30$var.std[,1])

dat_subset <- dat_subset[complete.cases(dat_subset),]

dat_subset <- dat_subset[dat_subset$OCSKGM < 100,]
#dat_subset
summary(dat_subset)

##      id          latitude      longitude      OCSKGM30
## Length:12165      Min.   :-4.970      Min.   :-81.00      Min.   : 0.000
## Class :character  1st Qu.:-2.930      1st Qu.:-79.71      1st Qu.: 4.022
## Mode  :character  Median :-1.720      Median :-79.07      Median : 6.594
##                               Mean   :-1.828      Mean   :-79.08      Mean   : 7.735
##                               3rd Qu.:-0.840      3rd Qu.:-78.57      3rd Qu.:10.063
##                               Max.    : 1.385      Max.    :-75.46      Max.    :84.948

campo OCSKGM30 se lo multiplica por 10 para obtener toneladas por hectarea

### se multiplica por 10 para obtener toneladas por hectarea#####
dat_subset_2 <- dat_subset
summary(dat_subset_2)

##      id          latitude      longitude      OCSKGM30
## Length:12165      Min.   :-4.970      Min.   :-81.00      Min.   : 0.000
## Class :character  1st Qu.:-2.930      1st Qu.:-79.71      1st Qu.: 4.022
## Mode  :character  Median :-1.720      Median :-79.07      Median : 6.594
##                               Mean   :-1.828      Mean   :-79.08      Mean   : 7.735
##                               3rd Qu.:-0.840      3rd Qu.:-78.57      3rd Qu.:10.063
##                               Max.    : 1.385      Max.    :-75.46      Max.    :84.948

dat_subset_2$OCSKGM <- (dat_subset_2$OCSKGM = dat_subset_2$OCSKGM*10)
#dat_subset_2
summary(dat_subset_2)

##      id          latitude      longitude      OCSKGM30
## Length:12165      Min.   :-4.970      Min.   :-81.00      Min.   : 0.000
## Class :character  1st Qu.:-2.930      1st Qu.:-79.71      1st Qu.: 4.022
## Mode  :character  Median :-1.720      Median :-79.07      Median : 6.594
##                               Mean   :-1.828      Mean   :-79.08      Mean   : 7.735
##                               3rd Qu.:-0.840      3rd Qu.:-78.57      3rd Qu.:10.063
##                               Max.    : 1.385      Max.    :-75.46      Max.    :84.948
##      OCSKGM
## Min.   : 0.00
## 1st Qu.: 40.22
## Median : 65.94
## Mean   : 77.35
## 3rd Qu.:100.63
## Max.   :849.48

colnames(dat_subset_2)[(5)]<- "OCSTONHA"
#dat_subset_2
summary(dat_subset)

```

```
##      id      Latitude      Longitude      OCSKGM30
## Length:12165      Min.   :-4.970      Min.   :-81.00      Min.   : 0.000
## Class :character  1st Qu.:-2.930      1st Qu.:-79.71      1st Qu.: 4.022
## Mode  :character  Median :-1.720      Median :-79.07      Median : 6.594
##      Mean   :-1.828      Mean   :-79.08      Mean   : 7.735
##      3rd Qu.:-0.840      3rd Qu.:-78.57      3rd Qu.:10.063
##      Max.   : 1.385      Max.   :-75.46      Max.   :84.948
```

```
summary(dat_subset_2)
```

```
##      id      Latitude      Longitude      OCSKGM30
## Length:12165      Min.   :-4.970      Min.   :-81.00      Min.   : 0.000
## Class :character  1st Qu.:-2.930      1st Qu.:-79.71      1st Qu.: 4.022
## Mode  :character  Median :-1.720      Median :-79.07      Median : 6.594
##      Mean   :-1.828      Mean   :-79.08      Mean   : 7.735
##      3rd Qu.:-0.840      3rd Qu.:-78.57      3rd Qu.:10.063
##      Max.   : 1.385      Max.   :-75.46      Max.   :84.948
```

```
##      OCSTONHA
## Min.   : 0.00
## 1st Qu.: 40.22
## Median : 65.94
## Mean   : 77.35
## 3rd Qu.:100.63
## Max.   :849.48
```

```
dat_subset <- dat_subset_2
```

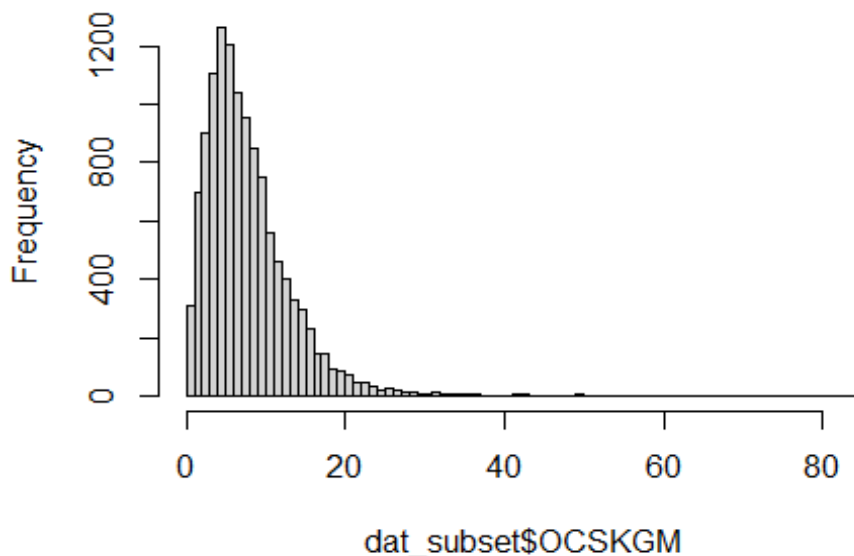
```
summary(dat_subset)
```

Se hace un análisis descriptivo de COS en La base de datos

```
### Analisis descriptivo
```

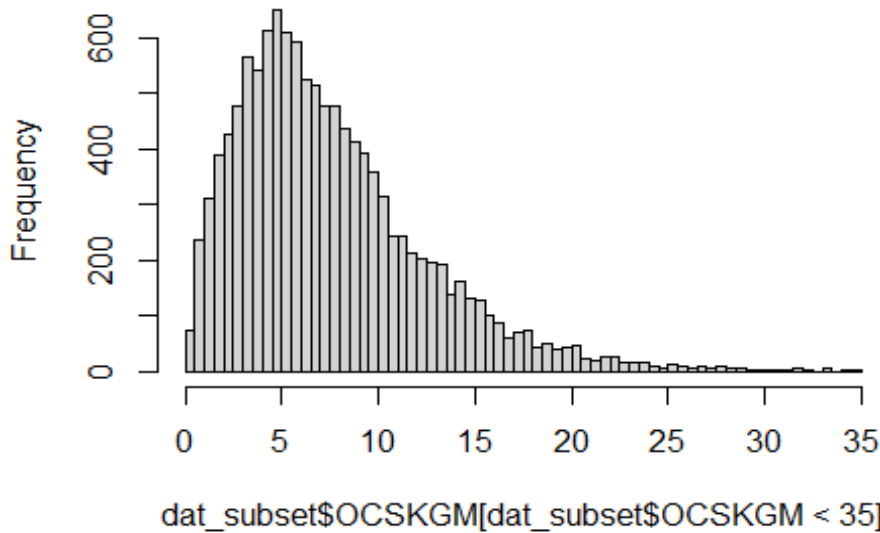
```
hist(dat_subset$OCSKGM, breaks = 100)
```

**Histogram of dat\_subset\$OCSKGM**



```
hist(dat_subset$OCSKGM[dat_subset$OCSKGM < 35], breaks = 100)
```

### Histogram of dat\_subset\$OCSKGM[dat\_subset\$OCSKGM < 35]



```
summary(dat_subset$OCSKGM)
```

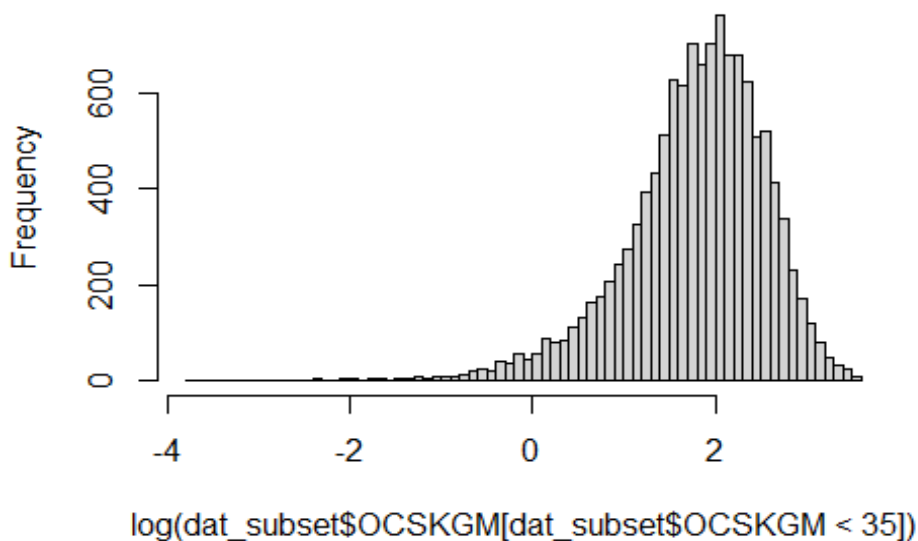
*Estadística descriptiva básica del COS*

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	0.000	4.022	6.594	7.735	10.063	84.948

*Histograma del Logaritmo de COS*

```
hist(log(dat_subset$OCSKGM[dat_subset$OCSKGM < 35]), breaks = 100)
```

### Histogram of log(dat\_subset\$OCSKGM[dat\_subset\$OCSKGM < 35])



Preparación de las covariables nacionales. Se tiene una carpeta con las capas en formato .tif se lo llama y se lo guarda en formato .rds para que sea más fácil trabajar con ellas. Adicionalmente de las grafica para estar seguros que es la capa deseada.

```
### Preparacion de las covariables ####

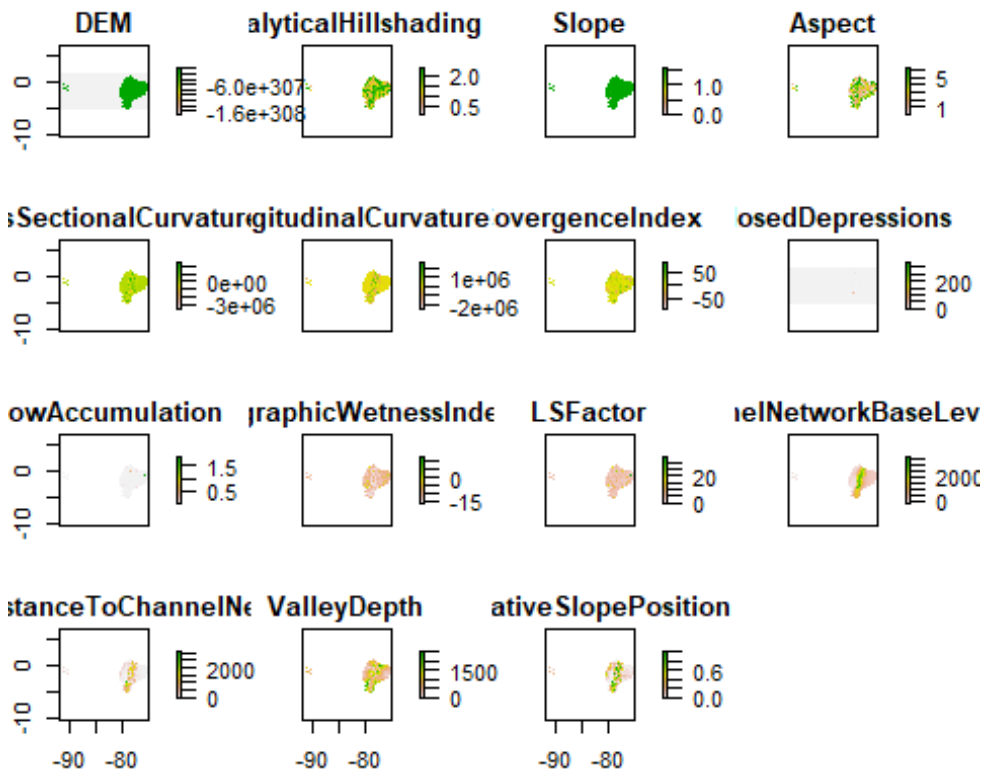
## Create a folder 'covariables' and put the covariates there...

topo <- stack('ECUtopo/ECUtopo.tif')
namesTopo <- readRDS('ECUtopo/namesTOPO.rds')
names(topo)

## [1] "ECUtopo.1" "ECUtopo.2" "ECUtopo.3" "ECUtopo.4" "ECUtopo.5"
## [6] "ECUtopo.6" "ECUtopo.7" "ECUtopo.8" "ECUtopo.9" "ECUtopo.10"
## [11] "ECUtopo.11" "ECUtopo.12" "ECUtopo.13" "ECUtopo.14" "ECUtopo.15"

names(topo) <- namesTopo
plot(topo)

## Warning in plot.window(...): Internal(pretty()): intervalo muy grande..
## corregido
```

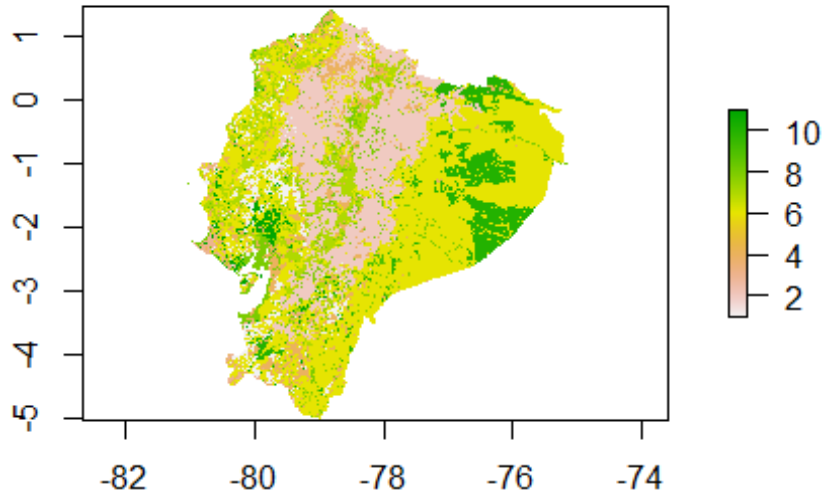


```
#Suelos <- readGDAL(file.choose())
Suelos <- stack('Covariables/Suelos_ordenes.tif')
#plot(Suelos, main="ORDENES DE SUELOS")
#projection(Suelos)
#names(Suelos)
saveRDS(Suelos, file='Suelos.rds')
```



```
Suelos=readRDS("SueLos.rds")  
plot(Suelos, main="ORDENES DE SUELOS W")
```

### ORDENES DE SUELOS W

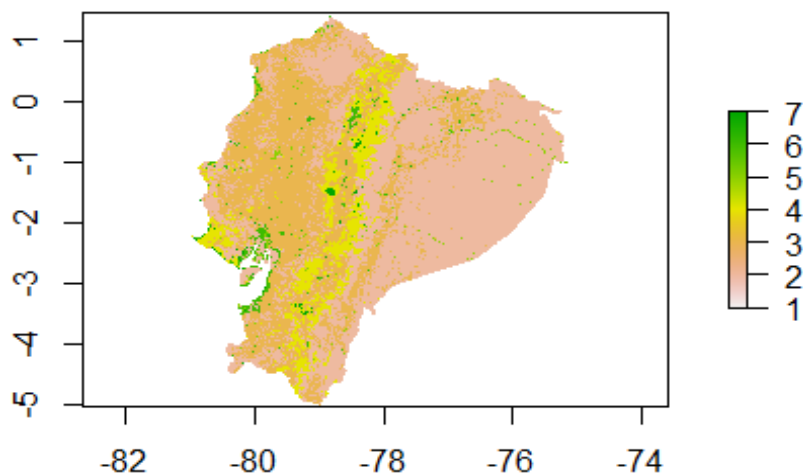


```
names(Suelos)
```

```
## [1] "SueLos_ordenes"
```

```
#cobertura <- readGDAL(file.choose())  
cobertura <- stack('Covariables/Cobertura_Uso_100.tif')  
#plot(cobertura, main="COBERTURA Y USO DE LA TIERRA")  
#names(cobertura)  
#projection(cobertura)  
saveRDS(cobertura, file='cobertura.rds')  
cobertura=readRDS("cobertura.rds")  
plot(cobertura, main="COBERTURA Y USO DE LA TIERRA W")
```

### COBERTURA Y USO DE LA TIERRA W



```
names(cobertura)

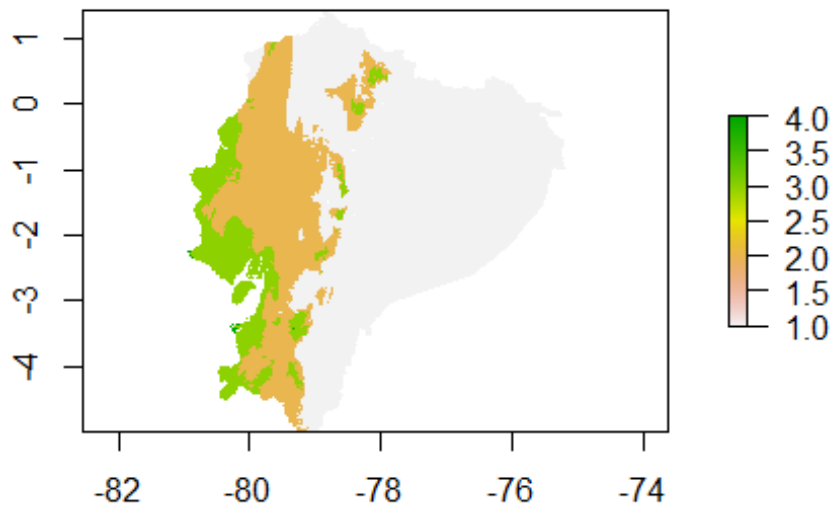
## [1] "Cobertura_Uso_100"

#bioclimatico <- readGDAL(file.choose())
bioclimatico <- stack('Covariables/Bioclimatico.tif')
#plot(bioclimatico, main="BIOCLIMATICO")
#names(bioclimatico)
projection(bioclimatico)

## [1] "+proj=longlat +datum=WGS84 +no_defs"

saveRDS(bioclimatico, file='bioclimatico.rds')
bioclimatico=readRDS("bioclimatico.rds")
plot(bioclimatico, main="BIOCLIMATICO W")
```

## BIOCLIMATICO W

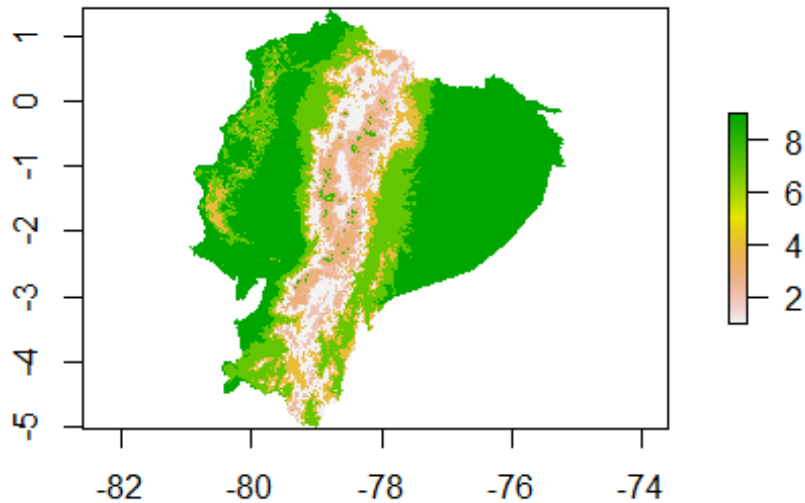


```
names(bioclimatico)

## [1] "Bioclimatico"

#pisosclimaticos <- readGDAL(file.choose())
pisosclimaticos <- stack('Covariables/Pisos_bioclimatico.tif')
#plot(pisosclimaticos, main="PISOS BIOCLIMATICOS")
#names(pisosclimaticos)
#projection(pisosclimaticos)
saveRDS(pisosclimaticos, file='pisosclimaticos.rds')
pisosclimaticos=readRDS("pisosclimaticos.rds")
plot(pisosclimaticos, main="PISOS CLIMATICOS W")
```

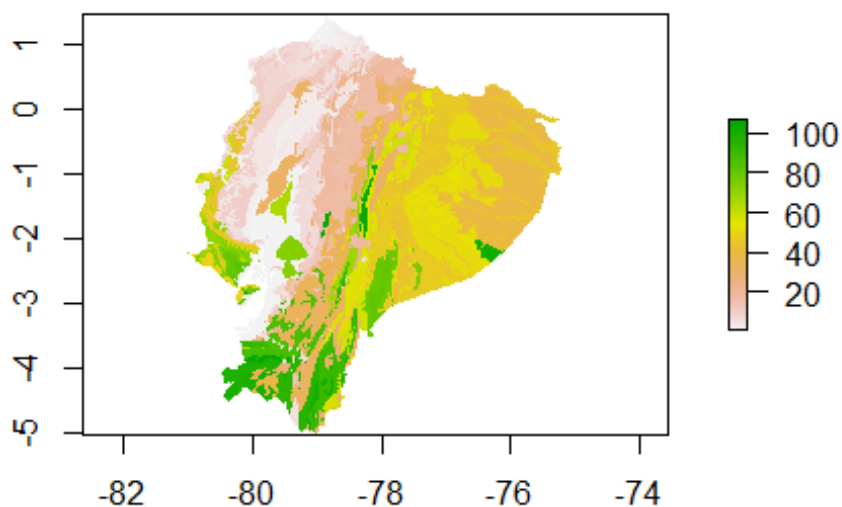
### PISOS CLIMATICOS W



```
names(pisosclimaticos)
## [1] "Pisos_bioclimatico"

#geologia <- readGDAL(file.choose())
geologia <- stack('Covariables/Geologia.tif')
#plot(geologia, main="GEOLOGIA")
#names(geologia)
#projection(geologia)
saveRDS(geologia, file='geologia.rds')
geologia=readRDS("geologia.rds")
plot(geologia, main="GEOLOGIA W")
```

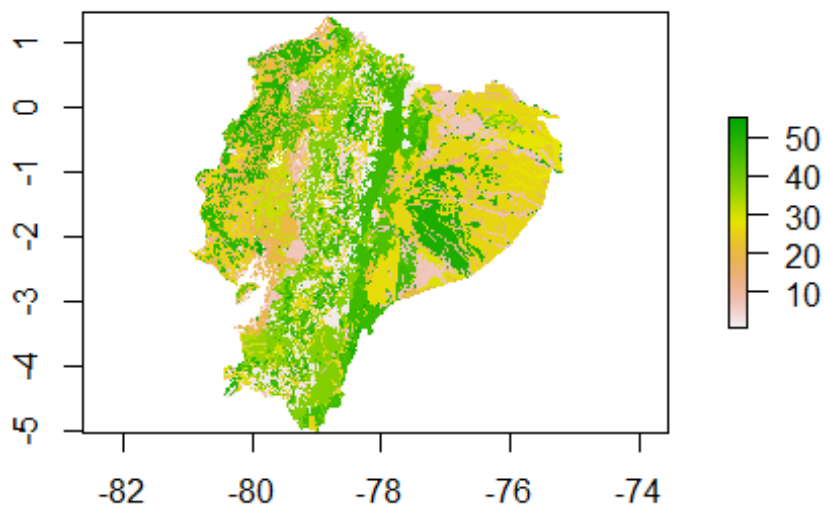
### GEOLOGIA W



```
names(geologia)
```

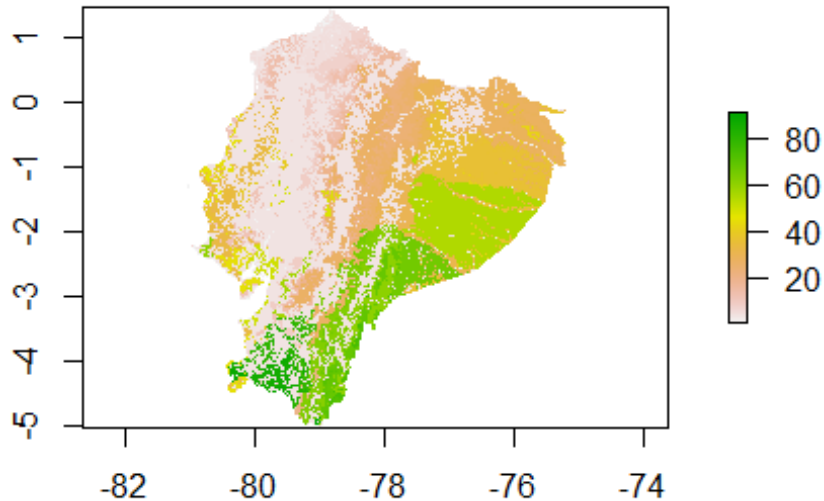
```
## [1] "Geologia"  
  
#geomorfologia <- readGDAL(file.choose())  
geomorfologia <- stack('Covariables/Geomorfologia.tif')  
#plot(geomorfologia, main="GEOMORFOLOGIA")  
#names(geomorfologia)  
#projection(geomorfologia)  
saveRDS(geomorfologia, file = 'geomorfologia.rds')  
geomorfologia=readRDS("geomorfologia.rds")  
plot(geomorfologia, main="GEOMORFOLOGIA W")
```

## GEOMORFOLOGIA W



```
names(geomorfologia)  
  
## [1] "Geomorfologia"  
  
#ecosistemas <- readGDAL(file.choose())  
ecosistemas <- stack('Covariables/Ecosistemas.tif')  
#plot(ecosistemas, main="ECOSISTEMAS")  
#names(ecosistemas)  
#projection(ecosistemas)  
saveRDS(ecosistemas, file = 'ecosistemas.rds')  
ecosistemas=readRDS("ecosistemas.rds")  
plot(ecosistemas, main="ECOSISTEMAS W")
```

## ECOSISTEMAS W

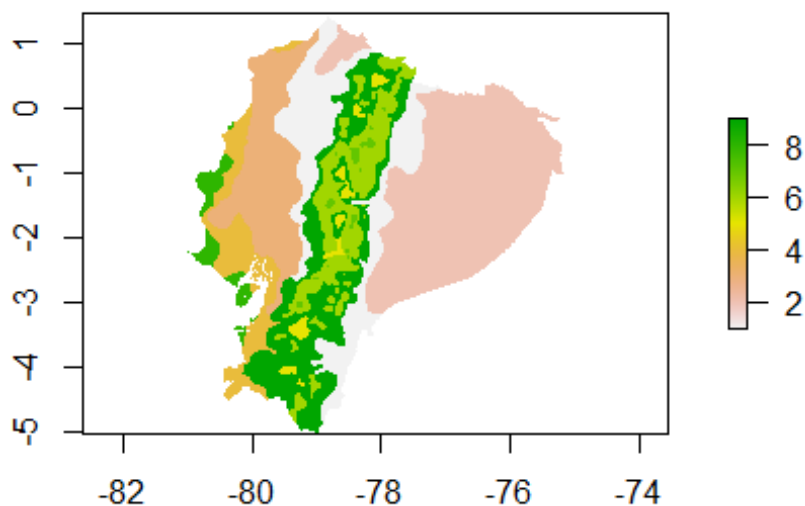


```
names(ecosistemas)

## [1] "Ecosistemas"

#tiposclima <- readGDAL(file.choose())
tiposclima<- stack('Covariables/Tipos_clima.tif')
#plot(tiposclima, main="TIPOS DE CLIMA")
#names(tiposclima)
#projection(tiposclima)
saveRDS(tiposclima, file = 'tiposclima.rds')
tiposclima=readRDS("tiposclima.rds")
plot(tiposclima, main="TIPOS DE CLIMA W")
```

## TIPOS DE CLIMA W



```
names(tiposclima)

## [1] "Tipos_clima"

cov <- stack('ECU_worldgridsCOVS/ECU_worldgridsCOVS.tif')
namesCov <- readRDS('ECU_worldgridsCOVS/worldgridsCOVS_names.rds')
names(cov)

## [1] "ECU_worldgridsCOVS.1" "ECU_worldgridsCOVS.2" "ECU_worldgridsCOVS.3"
## [4] "ECU_worldgridsCOVS.4" "ECU_worldgridsCOVS.5" "ECU_worldgridsCOVS.6"
## [7] "ECU_worldgridsCOVS.7" "ECU_worldgridsCOVS.8" "ECU_worldgridsCOVS.9"
## [10] "ECU_worldgridsCOVS.10" "ECU_worldgridsCOVS.11" "ECU_worldgridsCOVS.12"
## [13] "ECU_worldgridsCOVS.13" "ECU_worldgridsCOVS.14" "ECU_worldgridsCOVS.15"
## [16] "ECU_worldgridsCOVS.16" "ECU_worldgridsCOVS.17" "ECU_worldgridsCOVS.18"
## [19] "ECU_worldgridsCOVS.19" "ECU_worldgridsCOVS.20" "ECU_worldgridsCOVS.21"
## [22] "ECU_worldgridsCOVS.22" "ECU_worldgridsCOVS.23" "ECU_worldgridsCOVS.24"
## [25] "ECU_worldgridsCOVS.25" "ECU_worldgridsCOVS.26" "ECU_worldgridsCOVS.27"
## [28] "ECU_worldgridsCOVS.28" "ECU_worldgridsCOVS.29" "ECU_worldgridsCOVS.30"
## [31] "ECU_worldgridsCOVS.31" "ECU_worldgridsCOVS.32" "ECU_worldgridsCOVS.33"
## [34] "ECU_worldgridsCOVS.34" "ECU_worldgridsCOVS.35" "ECU_worldgridsCOVS.36"
## [37] "ECU_worldgridsCOVS.37" "ECU_worldgridsCOVS.38" "ECU_worldgridsCOVS.39"
## [40] "ECU_worldgridsCOVS.40" "ECU_worldgridsCOVS.41" "ECU_worldgridsCOVS.42"
## [43] "ECU_worldgridsCOVS.43" "ECU_worldgridsCOVS.44" "ECU_worldgridsCOVS.45"
## [46] "ECU_worldgridsCOVS.46" "ECU_worldgridsCOVS.47" "ECU_worldgridsCOVS.48"
## [49] "ECU_worldgridsCOVS.49" "ECU_worldgridsCOVS.50" "ECU_worldgridsCOVS.51"
## [52] "ECU_worldgridsCOVS.52" "ECU_worldgridsCOVS.53" "ECU_worldgridsCOVS.54"
## [55] "ECU_worldgridsCOVS.55" "ECU_worldgridsCOVS.56" "ECU_worldgridsCOVS.57"
## [58] "ECU_worldgridsCOVS.58" "ECU_worldgridsCOVS.59" "ECU_worldgridsCOVS.60"
## [61] "ECU_worldgridsCOVS.61" "ECU_worldgridsCOVS.62" "ECU_worldgridsCOVS.63"
## [64] "ECU_worldgridsCOVS.64" "ECU_worldgridsCOVS.65" "ECU_worldgridsCOVS.66"
## [67] "ECU_worldgridsCOVS.67" "ECU_worldgridsCOVS.68" "ECU_worldgridsCOVS.69"
## [70] "ECU_worldgridsCOVS.70" "ECU_worldgridsCOVS.71" "ECU_worldgridsCOVS.72"
## [73] "ECU_worldgridsCOVS.73" "ECU_worldgridsCOVS.74" "ECU_worldgridsCOVS.75"
## [76] "ECU_worldgridsCOVS.76" "ECU_worldgridsCOVS.77" "ECU_worldgridsCOVS.78"
## [79] "ECU_worldgridsCOVS.79" "ECU_worldgridsCOVS.80" "ECU_worldgridsCOVS.81"
## [82] "ECU_worldgridsCOVS.82" "ECU_worldgridsCOVS.83" "ECU_worldgridsCOVS.84"
## [85] "ECU_worldgridsCOVS.85" "ECU_worldgridsCOVS.86" "ECU_worldgridsCOVS.87"
## [88] "ECU_worldgridsCOVS.88" "ECU_worldgridsCOVS.89" "ECU_worldgridsCOVS.90"
## [91] "ECU_worldgridsCOVS.91" "ECU_worldgridsCOVS.92" "ECU_worldgridsCOVS.93"
## [94] "ECU_worldgridsCOVS.94" "ECU_worldgridsCOVS.95" "ECU_worldgridsCOVS.96"
## [97] "ECU_worldgridsCOVS.97" "ECU_worldgridsCOVS.98" "ECU_worldgridsCOVS.99"
## [100] "ECU_worldgridsCOVS.100" "ECU_worldgridsCOVS.101" "ECU_worldgridsCOVS.102"
## [103] "ECU_worldgridsCOVS.103" "ECU_worldgridsCOVS.104" "ECU_worldgridsCOVS.105"
## [106] "ECU_worldgridsCOVS.106" "ECU_worldgridsCOVS.107" "ECU_worldgridsCOVS.108"
## [109] "ECU_worldgridsCOVS.109" "ECU_worldgridsCOVS.110" "ECU_worldgridsCOVS.111"
## [112] "ECU_worldgridsCOVS.112" "ECU_worldgridsCOVS.113" "ECU_worldgridsCOVS.114"
## [115] "ECU_worldgridsCOVS.115" "ECU_worldgridsCOVS.116" "ECU_worldgridsCOVS.117"
## [118] "ECU_worldgridsCOVS.118"
```

```
#names(cov) <- namesCov
```

Se hace un remuestreo a las covariables para que todas tengan la misma proyección y sistema de coordenadas

```
topo<-resample(topo, Suelos)
cov<-resample(cov, Suelos)
cobertura<-resample(cobertura, Suelos)
bioclimatico<-resample(bioclimatico, Suelos)
```

```

pisosclimaticos<-resample(pisosclimaticos, Suelos)
geologia<-resample(geologia, Suelos)
geomorfologia<-resample(geomorfologia, Suelos)
ecosistemas<-resample(ecosistemas, Suelos)
tiposclima<-resample(tiposclima, Suelos)

```

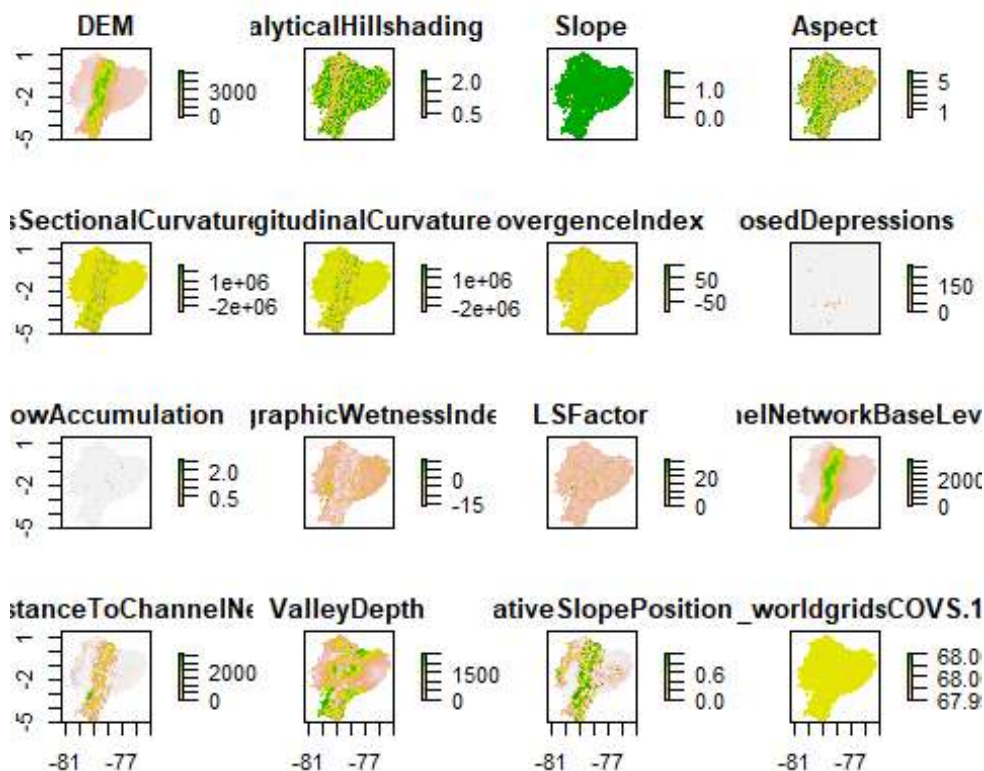
Se une a todas las covariables en un solo archivo

```

COV <- stack(topo, cov, Suelos, cobertura, bioclimatico, pisosclimaticos, geologia,
geomorfologia, ecosistemas, tiposclima)
#COV <- stack(topo, cov, Suelosw, coberturaw, bioclimaticow, pisosclimaticosw,
geologiaw, geomorfologiaw, ecosistemasw, tiposclimaw)

```

```
plot(COV)
```



```

#summary(COV)
#writeRaster(COV, file = "COVARIABLEStodas.tif")

#rasterize <- stack()

#list.files (pattern='tif')
#COV <- COV[[idx]]
# Lo convertimos a un SpatialGridDataFrame
COV <- stack('COVARIABLEStodas.tif')
#saveRDS(COV, file='namesCOVARIABLEStodas.rds')
namesCov <- readRDS('namesCOVARIABLEStodas.rds')
names(COV)

```

```
## [1] "COVARIABLEStodas.1" "COVARIABLEStodas.2" "COVARIABLEStodas.3"
## [4] "COVARIABLEStodas.4" "COVARIABLEStodas.5" "COVARIABLEStodas.6"
## [7] "COVARIABLEStodas.7" "COVARIABLEStodas.8" "COVARIABLEStodas.9"
## [10] "COVARIABLEStodas.10" "COVARIABLEStodas.11" "COVARIABLEStodas.12"
## [13] "COVARIABLEStodas.13" "COVARIABLEStodas.14" "COVARIABLEStodas.15"
## [16] "COVARIABLEStodas.16" "COVARIABLEStodas.17" "COVARIABLEStodas.18"
## [19] "COVARIABLEStodas.19" "COVARIABLEStodas.20" "COVARIABLEStodas.21"
## [22] "COVARIABLEStodas.22" "COVARIABLEStodas.23" "COVARIABLEStodas.24"
## [25] "COVARIABLEStodas.25" "COVARIABLEStodas.26" "COVARIABLEStodas.27"
## [28] "COVARIABLEStodas.28" "COVARIABLEStodas.29" "COVARIABLEStodas.30"
## [31] "COVARIABLEStodas.31" "COVARIABLEStodas.32" "COVARIABLEStodas.33"
## [34] "COVARIABLEStodas.34" "COVARIABLEStodas.35" "COVARIABLEStodas.36"
## [37] "COVARIABLEStodas.37" "COVARIABLEStodas.38" "COVARIABLEStodas.39"
## [40] "COVARIABLEStodas.40" "COVARIABLEStodas.41" "COVARIABLEStodas.42"
## [43] "COVARIABLEStodas.43" "COVARIABLEStodas.44" "COVARIABLEStodas.45"
## [46] "COVARIABLEStodas.46" "COVARIABLEStodas.47" "COVARIABLEStodas.48"
## [49] "COVARIABLEStodas.49" "COVARIABLEStodas.50" "COVARIABLEStodas.51"
## [52] "COVARIABLEStodas.52" "COVARIABLEStodas.53" "COVARIABLEStodas.54"
## [55] "COVARIABLEStodas.55" "COVARIABLEStodas.56" "COVARIABLEStodas.57"
## [58] "COVARIABLEStodas.58" "COVARIABLEStodas.59" "COVARIABLEStodas.60"
## [61] "COVARIABLEStodas.61" "COVARIABLEStodas.62" "COVARIABLEStodas.63"
## [64] "COVARIABLEStodas.64" "COVARIABLEStodas.65" "COVARIABLEStodas.66"
## [67] "COVARIABLEStodas.67" "COVARIABLEStodas.68" "COVARIABLEStodas.69"
## [70] "COVARIABLEStodas.70" "COVARIABLEStodas.71" "COVARIABLEStodas.72"
## [73] "COVARIABLEStodas.73" "COVARIABLEStodas.74" "COVARIABLEStodas.75"
## [76] "COVARIABLEStodas.76" "COVARIABLEStodas.77" "COVARIABLEStodas.78"
## [79] "COVARIABLEStodas.79" "COVARIABLEStodas.80" "COVARIABLEStodas.81"
## [82] "COVARIABLEStodas.82" "COVARIABLEStodas.83" "COVARIABLEStodas.84"
## [85] "COVARIABLEStodas.85" "COVARIABLEStodas.86" "COVARIABLEStodas.87"
## [88] "COVARIABLEStodas.88" "COVARIABLEStodas.89" "COVARIABLEStodas.90"
## [91] "COVARIABLEStodas.91" "COVARIABLEStodas.92" "COVARIABLEStodas.93"
## [94] "COVARIABLEStodas.94" "COVARIABLEStodas.95" "COVARIABLEStodas.96"
## [97] "COVARIABLEStodas.97" "COVARIABLEStodas.98" "COVARIABLEStodas.99"
## [100] "COVARIABLEStodas.100" "COVARIABLEStodas.101" "COVARIABLEStodas.102"
## [103] "COVARIABLEStodas.103" "COVARIABLEStodas.104" "COVARIABLEStodas.105"
## [106] "COVARIABLEStodas.106" "COVARIABLEStodas.107" "COVARIABLEStodas.108"
## [109] "COVARIABLEStodas.109" "COVARIABLEStodas.110" "COVARIABLEStodas.111"
## [112] "COVARIABLEStodas.112" "COVARIABLEStodas.113" "COVARIABLEStodas.114"
## [115] "COVARIABLEStodas.115" "COVARIABLEStodas.116" "COVARIABLEStodas.117"
## [118] "COVARIABLEStodas.118" "COVARIABLEStodas.119" "COVARIABLEStodas.120"
## [121] "COVARIABLEStodas.121" "COVARIABLEStodas.122" "COVARIABLEStodas.123"
## [124] "COVARIABLEStodas.124" "COVARIABLEStodas.125" "COVARIABLEStodas.126"
## [127] "COVARIABLEStodas.127" "COVARIABLEStodas.128" "COVARIABLEStodas.129"
## [130] "COVARIABLEStodas.130" "COVARIABLEStodas.131" "COVARIABLEStodas.132"
## [133] "COVARIABLEStodas.133" "COVARIABLEStodas.134" "COVARIABLEStodas.135"
## [136] "COVARIABLEStodas.136" "COVARIABLEStodas.137" "COVARIABLEStodas.138"
## [139] "COVARIABLEStodas.139" "COVARIABLEStodas.140" "COVARIABLEStodas.141"
```

```
#names(COV) <- namesCov
```

*Unimos Los datos con La información de todas Las covariables para hacer una matriz de regresión, se Le agrega el límite del ecuador para delimitar el área.*

```
##### Juntamos covariables con los datos #####
# Convert to spatial points df and project
dat_subset_sp <- dat_subset
coordinates(dat_subset_sp) <- ~ Longitude + Latitude
## ADM limits
```



```
Lim=readRDS("EcuLim2014.rds")
plot(Lim, main="ECUADOR")
```



```
# Show the first 10 rows
```

```
head(Lim@data, 10)
```

#DPA_PROVIN	DPA_DESPRO	DESCRIP_1
#0	<NA>	<NA> ISLAS
#1	<NA>	<NA> ZONA EN ESTUDIO: AZUAY - GUAYAS
#2	<NA>	<NA> ZONA EN ESTUDIO: JUVAL (CANAR - CHIMBORAZO)
#3	<NA>	<NA> ZONA EN ESTUDIO: MATILDE ESTHER (GUAYAS - LOS RIOS)
#4	<NA>	<NA> ZONA EN ESTUDIO: SANTA ROSA DE AGUA CLARA (BOLIVAR - GUAYAS)
#5	01	AZUAY <NA>
#6	02	BOLIVAR <NA>
#7	03	CANAR <NA>
#8	04	CARCHI <NA>
#9	05	COTOPAXI <NA>
	AREA_M2	AREA_HA
#0	11016303	1101.630
#1	607856307	60785.631
#2	606841217	60684.122
#3	23693800	2369.380

```
#4 37145347 3714.535
```

```
#5 7872637757 787263.776
```

```
#6 3897940328 389794.033
```

```
#7 3135005882 313500.588
```

```
#8 3782566754 378256.675
```

```
#9 6166650079 616665.008
```

```
plot(lim,
```

```
col = "grey93",
```

```
border="grey")
```



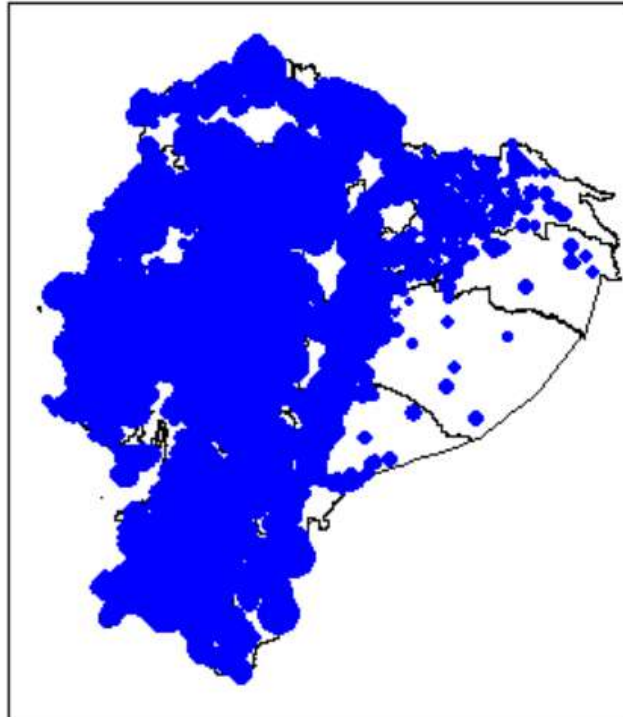
```
bubble(dat_subset_sp, "OCSKGM30", sp.layout = lim,
```

```
col = "blue",
```

```
border="blue",
```

```
main="OCSKGM30 ECUADOR")
```

## OCSKGM30 ECUADOR



```
dat_subset <- cbind(dat_subset, extract(COV, dat_subset_sp))
summary(dat_subset)
```

```
#id          Latitude      Longitude      OCSKGM30
#Length:12165      Min.    :-4.970  Min.    :-81.00  Min.    : 0.000
#Class :character  1st Qu.: -2.930  1st Qu.: -79.71  1st Qu.: 4.022
#Mode  :character  Median  :-1.720  Median  :-79.07  Median  : 6.594
# OCSTONHA      COVARIABLEStodas.1 COVARIABLEStodas.2 COVARIABLEStodas.3
#Min.    : 0.00      Min.    : 0.186  Min.    :0.2799  Min.    :0.06066
#1st Qu.: 40.22     1st Qu.: 244.257  1st Qu.:1.0322  1st Qu.:1.57026
#Median  : 65.94     Median  :1021.997  Median  :1.4606  Median  :1.57064
#COVARIABLEStodas.4 COVARIABLEStodas.5 COVARIABLEStodas.6 COVARIABLEStodas.7
#Min.    :0.06972   Min.    :-2655931  Min.    :-2947282  Min.    :-72.411
#1st Qu.:2.11278   1st Qu.: -240549  1st Qu.: -394035  1st Qu.: -7.537
#Median  :3.45359   Median  : -22228  Median  : -89379  Median  : -1.443
#COVARIABLEStodas.8 COVARIABLEStodas.9 COVARIABLEStodas.10 COVARIABLEStodas.11
```

```
#Min. : -0.00012 Min. :0.000070 Min. :-16.028 Min. : 0.3152
#1st Qu.: -0.00001 1st Qu.:0.000189 1st Qu.: -13.402 1st Qu.: 3.8127
#Median : 0.00000 Median :0.000506 Median :-11.906 Median : 4.6815
#COVARIABLEStodas.12 COVARIABLEStodas.13 COVARIABLEStodas.14 COVARIABLEStodas.15
#Min. : -52.73 Min. : 0.00 Min. : 0.7418 Min. :0.00000
#1st Qu.: 122.53 1st Qu.: 21.87 1st Qu.: 608.5974 1st Qu.:0.01933
#Median : 810.53 Median : 129.11 Median :1013.8849 Median :0.11682
#COVARIABLEStodas.16 COVARIABLEStodas.17 COVARIABLEStodas.18 COVARIABLEStodas.19
#Min. :68 Min. : -0.574 Min. : 2951 Min. : 462.3
#1st Qu.:68 1st Qu.: 234.979 1st Qu.: 9323 1st Qu.:3578.8
#Median :68 Median :1006.426 Median :10669 Median :4204.9
```

#identificar mediante el summary anterior las variables con muchos NA's altos, luego quitar#

```
dat_subset$g06esa3a <- NULL
```

```
dat_subset$g04esa3 <- NULL
```

```
dat_subset$g06esa3a <- NULL
```

Guardar la información de puntos y covariables como matriz de regresión

```
dat <- dat_subset[complete.cases(dat_subset),]
```

```
write.csv(dat, 'EcuadorRegMatrix.csv')
```

## Anexo 2. Mapeo de carbono orgánico del suelo de Ecuador

```
# SOIL ORGANIC CARBON MAPPING ECUADOR
```

```
#REGRESSION-KRIGING#
```

Establecemos el directorio donde se encuentra la información origen y se guardará todos los archivos creados

```
## Establecemos el directorio de trabajo  
setwd("C:/Users/usuario/Desktop/SCRIPT2/Ecuador")
```

### Instalación de paquetes

```
## Load required packages  
library(raster)
```

```
library(car)
```

```
library(rgdal)
```

```
library(gstat)
```

```
library(caret)
```

```
library(reshape)
```

```
library(sp)
```

```
Load("DSM_supportfunctions.RData")
```

```
### Preparamos los datos ###
```

Cargamos la matriz de regresión con los puntos y covariables

```
# Cargamos los datos de los splines  
dat <- read.csv("EcuadorRegMatrix.csv")  
names(dat)
```

```
## [1] "X" "id"  
## [3] "Latitude" "Longitude"  
## [5] "OCSKGM30" "OCSKGM30"  
## [7] "OCSKGM" "DEM"  
## [9] "AnalyticalHillshading" "Slope"  
## [11] "Aspect" "CrossSectionalCurvature"  
## [13] "LongitudinalCurvature" "CovergenceIndex"  
## [15] "ClosedDepressions" "FlowAccumulation"  
## [17] "TopographicWetnessIndex" "LSFactor"  
## [19] "ChannelNetworkBaseLevel" "VerticalDistanceToChannelNetwork"  
## [21] "ValleyDepth" "RelativeSlopePosition"  
## [23] "cntgad3a" "DEMSRE3a"  
## [25] "etmnts3a" "evmmod3a"  
## [27] "evsmod3a" "g01esa3a"  
## [29] "g01igb3a" "g02esa3a"  
## [31] "g02igb3a" "g03esa3a"  
## [33] "g04esa3a" "g04igb3a"  
## [35] "g05esa3a" "g06esa3a"
```

## [37] "g07esa3a"	"g08esa3a"
## [39] "g09esa3a"	"g10esa3a"
## [41] "g10igb3a"	"g11esa3a"
## [43] "g11igb3a"	"g12esa3a"
## [45] "g12igb3a"	"g13esa3a"
## [47] "g14esa3a"	"g15esa3a"
## [49] "g16esa3a"	"g17esa3a"
## [51] "g18esa3a"	"g19esa3a"
## [53] "g20esa3a"	"g21esa3a"
## [55] "g22esa3a"	"gabhws3a"
## [57] "gacgem3a"	"gachws3a"
## [59] "galhws3a"	"ganhws3a"
## [61] "garhws3a"	"gathws3a"
## [63] "gchhws3a"	"gclhws3a"
## [65] "gcmhws3a"	"gcrhws3a"
## [67] "geaisg3a"	"gflhws3a"
## [69] "gfrhws3a"	"gglhws3a"
## [71] "ggyhws3a"	"ghshws3a"
## [73] "gkshws3a"	"glcesa3a"
## [75] "glcjrc3a"	"glphws3a"
## [77] "glvhws3a"	"glwwwf3a"
## [79] "glxhws3a"	"gnthws3a"
## [81] "gphhws3a"	"gplhws3a"
## [83] "gpthws3a"	"gpzhws3a"
## [85] "grghws3a"	"gschws3a"
## [87] "gsnhws3a"	"gsthws3a"
## [89] "gumhws3a"	"gvrhws3a"
## [91] "iflgre3a"	"inmsre3a"
## [93] "inssre3a"	"l01igb3a"
## [95] "l02igb3a"	"l03igb3a"
## [97] "l04igb3a"	"l05igb3a"
## [99] "l06igb3a"	"l07igb3a"
## [101] "l08igb3a"	"l09igb3a"
## [103] "l10igb3a"	"l11igb3a"
## [105] "l12igb3a"	"l13igb3a"
## [107] "l14igb3a"	"l15igb3a"
## [109] "l16igb3a"	"l3pobi3b"
## [111] "lamm3a"	"lasm3a"
## [113] "lmbgsh3a"	"lmtgsh3a"
## [115] "ln1dms3a"	"ln2dms3a"
## [117] "lnmdms3a"	"opisre3a"
## [119] "px1wcl3a"	"px2wcl3a"
## [121] "px3wcl3a"	"px4wcl3a"
## [123] "SLPSRT3a"	"smkisir3a"
## [125] "tdhmod3a"	"tdlmod3a"
## [127] "tdmmod3a"	"tdsmod3a"
## [129] "tnhmod3a"	"tnlmod3a"
## [131] "tnmmod3a"	"tnsm3a"
## [133] "twisre3a"	"tx1mod3a"
## [135] "tx2mod3a"	"tx3mod3a"
## [137] "tx4mod3a"	"tx5mod3a"
## [139] "tx6mod3a"	"wmkmod3a"
## [141] "SueLos_ordenes"	"Cobertura_Us0_100"
## [143] "Bioclimatico"	"Pisos_bioclimatico"
## [145] "Geologia"	"Geomorfologia"
## [147] "Ecosistemas"	"Tipos_clima"
## [149] "DEM.1"	"AnalyticalHillshading.1"
## [151] "Slope.1"	"Aspect.1"
## [153] "CrossSectionalCurvature.1"	"LongitudinalCurvature.1"

## [155]	"CovergenceIndex.1"	"ClosedDepressions.1"
## [157]	"FlowAccumulation.1"	"TopographicWetnessIndex.1"
## [159]	"LSFactor.1"	"ChannelNetworkBaseLevel.1"
## [161]	"VerticalDistanceToChannelNetwork.1"	"ValleyDepth.1"
## [163]	"RelativeSlopePosition.1"	"cntgad3a.1"
## [165]	"DEMSRE3a.1"	"etmmts3a.1"
## [167]	"evmod3a.1"	"evsmmod3a.1"
## [169]	"g01esa3a.1"	"g01igb3a.1"
## [171]	"g02esa3a.1"	"g02igb3a.1"
## [173]	"g03esa3a.1"	"g04esa3a.1"
## [175]	"g04igb3a.1"	"g05esa3a.1"
## [177]	"g06esa3a.1"	"g07esa3a.1"
## [179]	"g08esa3a.1"	"g09esa3a.1"
## [181]	"g10esa3a.1"	"g10igb3a.1"
## [183]	"g11esa3a.1"	"g11igb3a.1"
## [185]	"g12esa3a.1"	"g12igb3a.1"
## [187]	"g13esa3a.1"	"g14esa3a.1"
## [189]	"g15esa3a.1"	"g16esa3a.1"
## [191]	"g17esa3a.1"	"g18esa3a.1"
## [193]	"g19esa3a.1"	"g20esa3a.1"
## [195]	"g21esa3a.1"	"g22esa3a.1"
## [197]	"gabhws3a.1"	"gacgem3a.1"
## [199]	"gachws3a.1"	"galhws3a.1"
## [201]	"ganhws3a.1"	"garhws3a.1"
## [203]	"gathws3a.1"	"gchhws3a.1"
## [205]	"gclhws3a.1"	"gcmhws3a.1"
## [207]	"gcrhws3a.1"	"geaisg3a.1"
## [209]	"gflhws3a.1"	"gfrhws3a.1"
## [211]	"gglhws3a.1"	"ggyhws3a.1"
## [213]	"ghshws3a.1"	"gkshws3a.1"
## [215]	"glcesa3a.1"	"glcjrc3a.1"
## [217]	"glphws3a.1"	"glvhws3a.1"
## [219]	"glwwwf3a.1"	"glxhws3a.1"
## [221]	"gnthws3a.1"	"gphhws3a.1"
## [223]	"gplhws3a.1"	"gpthws3a.1"
## [225]	"gpszws3a.1"	"grghws3a.1"
## [227]	"gschws3a.1"	"gsnhws3a.1"
## [229]	"gsthws3a.1"	"gumhws3a.1"
## [231]	"gvrhws3a.1"	"iflgre3a.1"
## [233]	"inmsre3a.1"	"inssre3a.1"
## [235]	"l01igb3a.1"	"l02igb3a.1"
## [237]	"l03igb3a.1"	"l04igb3a.1"
## [239]	"l05igb3a.1"	"l06igb3a.1"
## [241]	"l07igb3a.1"	"l08igb3a.1"
## [243]	"l09igb3a.1"	"l10igb3a.1"
## [245]	"l11igb3a.1"	"l12igb3a.1"
## [247]	"l13igb3a.1"	"l14igb3a.1"
## [249]	"l15igb3a.1"	"l16igb3a.1"
## [251]	"l3pobi3b.1"	"lammod3a.1"
## [253]	"lasm3a.1"	"lmbgsh3a.1"
## [255]	"lmtgsh3a.1"	"ln1dms3a.1"
## [257]	"ln2dms3a.1"	"lnmdms3a.1"
## [259]	"opisre3a.1"	"px1wcl3a.1"
## [261]	"px2wcl3a.1"	"px3wcl3a.1"
## [263]	"px4wcl3a.1"	"SLPSRT3a.1"
## [265]	"smkisir3a.1"	"tdhmod3a.1"
## [267]	"tdlmod3a.1"	"tdmmod3a.1"
## [269]	"tdsmod3a.1"	"tnhmod3a.1"
## [271]	"tnlmod3a.1"	"tnmmod3a.1"

```
## [273] "tnsmod3a.1"          "twisre3a.1"
## [275] "tx1mod3a.1"         "tx2mod3a.1"
## [277] "tx3mod3a.1"         "tx4mod3a.1"
## [279] "tx5mod3a.1"         "tx6mod3a.1"
## [281] "wmkmod3a.1"         "Suelos_ordenes.1"
## [283] "Cobertura_Uso_100.1" "Bioclimatico.1"
## [285] "Pisos_bioclimatico.1" "Geologia.1"
## [287] "Geomorfologia.1"    "Ecosistemas.1"
## [289] "Tipos_clima.1"
```

### Marcamos

```
#dat$suelos <- as.factor(dat$suelos)
```

Revisamos que datos están en cada variables

```
str(dat)
```

```
## 'data.frame': 6486 obs. of 289 variables:
## $ X : int 3 7 8 16 18 19 28 37 38 39 ...
## $ id : chr "11024_-4.3103_-79.7849_-4.3103_-79.7849_-4.3103_-79.7849" "11029_-4.2795_-79.5272_-4.2795_-79.5272_-4.2795_-79.5272" "11030_-4.0448_-79.6682_-4.0448_-79.6682_-4.0448_-79.6682" "22006_-1.9994_-80.7392_-1.9994_-80.7392_-1.9994_-80.7392" ...
## $ Latitude : num -4.31 -4.28 -4.04 -2 -2.32 ...
## $ Longitude : num -79.8 -79.5 -79.7 -80.7 -80.8 ...
## $ OCSKGM30 : num 7.49 8.97 4.71 4.13 2.51 ...
## $ OCSTONHA : num 74.9 89.7 47.1 41.3 25.1 ...
## $ OCSKGM : num 74.9 89.7 47.1 41.3 25.1 ...
## $ DEM : num 1383.4 1775.4 1215.3 33 45.7 ...
## $ AnalyticalHillshading : num 1.136 0.813 0.875 1.746 0.891 ...
## $ Slope : num 1.57 1.57 1.57 1.57 1.57 ...
## $ Aspect : num 4.78 5.55 5.08 3.63 5.06 ...
## $ CrossSectionalCurvature : num -594164 -213178 -308407 -27874 51475
...
## $ LongitudinalCurvature : num 268066 283156 -477482 7494 25001 ...
## $ CoverageIndex : num -18.18 -1.61 -1.36 -11.03 28.16 ...
## $ ClosedDepressions : num -2.65e-05 3.48e-05 -4.48e-05 -7.38e-07 -4.73e-07 ...
## $ FlowAccumulation : num 1.44e-03 1.24e-04 5.01e-04 8.42e-05 9.28e-05 ...
## $ TopographicWetnessIndex : num -9.94 -14.46 -13.1 -12.35 -11.35 ...
## $ LSFactor : num 7.42 3.46 4.74 3.1 3.23 ...
## $ ChannelNetworkBaseLevel : num 769.6 1131.03 916.14 8.82 12.66 ...
## $ VerticalDistanceToChannelNetwork : num 613.8 644.4 299.1 24.2 33 ...
## $ ValleyDepth : num 1236 857 1565 549 673 ...
## $ RelativeSlopePosition : num 0.3318 0.4291 0.1605 0.0422 0.0467 ...
## $ cntgad3a : int 68 68 68 68 68 68 68 68 68 ...
## $ DEMSRE3a : num 1336 1785.5 1203.8 34.6 46.8 ...
## $ etmnts3a : num 9731 12671 9613 5267 4404 ...
## $ evmmod3a : num 3647 4107 3325 2476 2220 ...
## $ evsmmod3a : num 1095 913 1160 1260 1077 ...
## $ g01esa3a : int 0 0 0 0 0 0 0 0 0 ...
## $ g01igb3a : num 12.69 6.47 7.34 13.38 9.98 ...
## $ g02esa3a : num 0 0 0 0 0 0 0 0 0 ...
## $ g02igb3a : num 12.59 6.47 9 14 14 ...
## $ g03esa3a : num 48.99 5.63 1.72 0 0 ...
## $ g04esa3a : num 0.484 0 33.628 0 0 ...
## $ g04igb3a : num 8.98 8.41 5.92 13.69 13.07 ...
## $ g05esa3a : num 37 0 15.4 93.8 0 ...
```



```

## $ g06esa3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ g07esa3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ g08esa3a : int 0 0 0 0 0 0 0 0 0 0 ...
## $ g09esa3a : int 0 0 0 0 0 0 0 0 0 0 ...
## $ g10esa3a : int 0 0 0 0 0 0 0 0 0 0 ...
## $ g10igb3a : num 8.62 8.86 8.38 13.02 13.21 ...
## $ g11esa3a : num 0 0 0 5.76 0 ...
## $ g11igb3a : num 8.62 9 8.38 13.02 10.93 ...
## $ g12esa3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ g12igb3a : num 8.87 6.97 8.38 12.68 12 ...
## $ g13esa3a : num 13.5 94.4 49.3 0 0 ...
## $ g14esa3a : num 0 0 0 0.483 0 ...
## $ g15esa3a : num 0 0 0 0 100 0 0 0 0 ...
## $ g16esa3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ g17esa3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ g18esa3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ g19esa3a : int 0 0 0 0 0 0 0 0 0 0 ...
## $ g20esa3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ g21esa3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ g22esa3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ gabhws3a : int 0 0 0 0 0 0 0 0 0 0 ...
## $ gacgem3a : num 83.7 55.3 69.1 92.5 15.5 ...
## $ gachws3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ galhws3a : num 30 30 0 0 0 0 0 0 0 0 ...
## $ ganhws3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ garhws3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ gathws3a : int 0 0 0 0 0 0 0 0 0 0 ...
## $ gchws3a : int 0 0 0 0 0 0 0 0 0 0 ...
## $ gclhws3a : int 0 0 0 0 0 0 0 0 0 0 ...
## $ gcmhws3a : num 70 70 0 0 0 50 50 30 30 30 ...
## $ gcrhws3a : int 0 0 0 0 0 0 0 0 0 0 ...
## $ geaisg3a : num 38 18 58.5 72 91 ...
## $ gflhws3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ gfrhws3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ gglhws3a : num 0 0 0 0 0 30 30 0 0 0 ...
## $ ggyhws3a : int 0 0 0 0 0 0 0 0 0 0 ...
## $ ghshws3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ gkshws3a : int 0 0 0 0 0 0 0 0 0 0 ...
## $ glcesa3a : num 35.9 130 93 41.9 150 ...
## $ glcjrc3a : num 17.1 14 12.3 13 14 ...
## $ glphws3a : num 0 0 30 0 0 0 0 0 0 0 ...
## $ glvhws3a : num 0 0 0 0 0 20 20 0 0 0 ...
## $ glwwwf3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ glxhws3a : int 0 0 0 0 0 0 0 0 0 0 ...
## $ gnthws3a : int 0 0 0 0 0 0 0 0 0 0 ...
## $ gphws3a : num 0 0 0 0 0 0 0 70 70 70 ...
## $ gplhws3a : num 0 0 0 70 70 0 0 0 0 0 ...
## $ gpthws3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ gpzhws3a : int 0 0 0 0 0 0 0 0 0 0 ...
## $ grghws3a : num 0 0 70 30 30 0 0 0 0 0 ...
## $ gschws3a : int 0 0 0 0 0 0 0 0 0 0 ...
## $ gsnhws3a : int 0 0 0 0 0 0 0 0 0 0 ...
## $ gsthws3a : int 0 0 0 0 0 0 0 0 0 0 ...
## $ gumhws3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ gvrhws3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ iflgre3a : num 0 0 0 0 0 0 0 0 0 0 ...
## $ inmsre3a : num 35.4 35.6 35 34 33 ...
## $ inssre3a : num 22.3 23.2 21.9 20.7 20.7 ...
## $ l01igb3a : num 0 0 0 0 0 0 0 0 0 0 ...

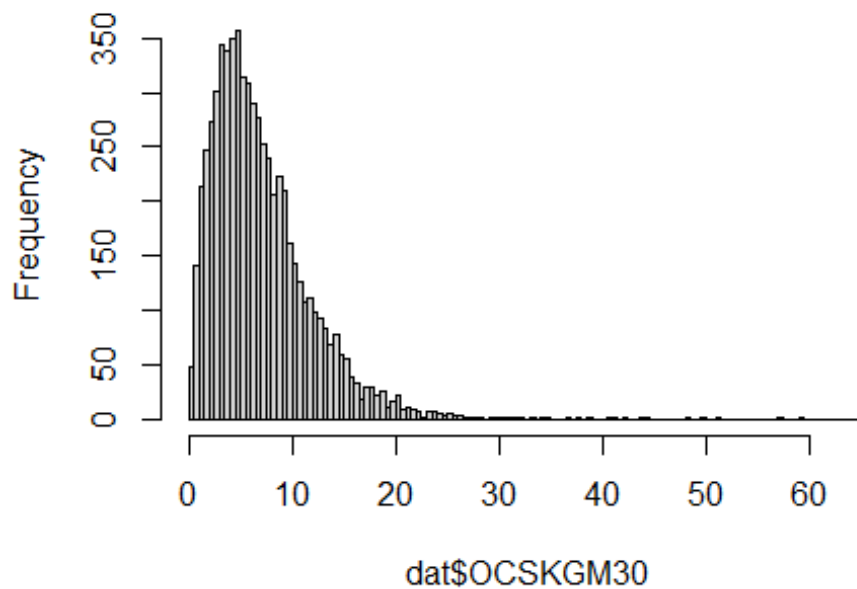
```

```
## $ l02igb3a : num 0 0 0 0 3.25 ...  
## $ l03igb3a : num 0 0 0 0 0 0 0 0 0 ...  
## $ l04igb3a : num 0 0 18.9 0 0 ...  
## $ l05igb3a : num 0 23.5 0 0 0 ...  
## $ l06igb3a : num 5 0 0 0 0 ...  
## [list output truncated]
```

Se hace un análisis descriptivo de COS en la base de datos

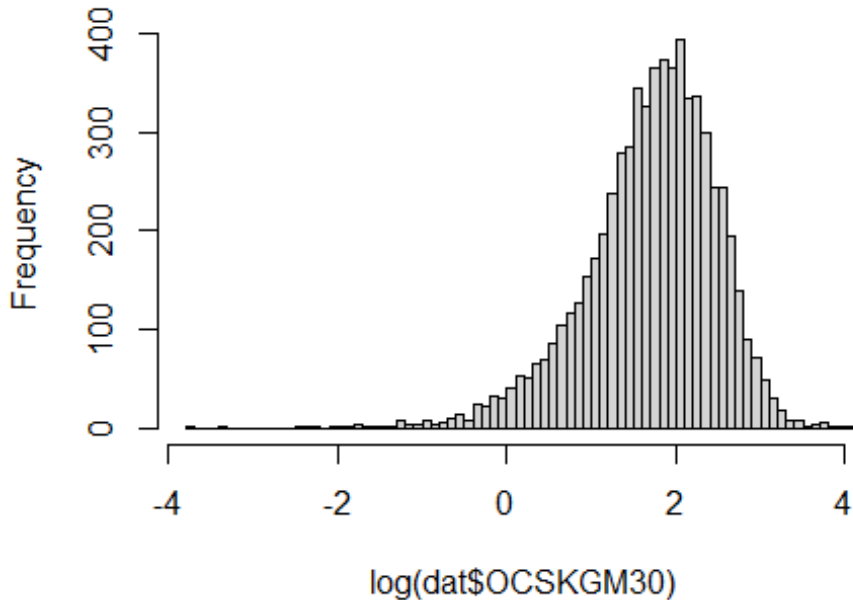
```
hist(dat$OCSKGM30, breaks=100)
```

**Histogram of dat\$OCSKGM30**



```
## Transform data  
hist(Log(dat$OCSKGM30), breaks=100)
```

## Histogram of log(dat\$OCSKGM30)



```
## Recreamos el objeto con la ubicacion de los puntos
```

```
dat_sp <- dat
```

```
coordinates(dat_sp) <- ~ Longitude + Latitude
```

*Se hace un análisis de correlación entre las covariables*

```
## Ajustamos un modelo de regresion lineal multiple
```

```
## Pruebas de modelos #####
```

```
#### Analisis de correlacion
```

```
names(dat_sp@data)
```

```
COR <- cor(as.matrix(dat_sp@data[,3]), as.matrix(dat_sp@data[, -c(1:3, 133)]))
```

```
## Warning in cor(as.matrix(dat_sp@data[, 3]), as.matrix(dat_sp@data[, -c(1:3, :  
## the standard deviation is zero
```

```
COR
```

```
##      OCSKGM DEM AnalyticalHillshading Slope Aspect  
## [1,]      1      1 NaN                -0.01880716 0.06140301 0.04726006  
##      CrossSectionalCurvature LongitudinalCurvature CoverageIndex  
## [1,]                0.1214281                0.1537979                0.1076731  
##      ClosedDepressions FlowAccumulation TopographicWetnessIndex LSFactor  
## [1,]      -0.07619768      -0.03890132                -0.1259765 -0.07548198  
##      ChannelNetworkBaseLevel VerticalDistanceToChannelNetwork ValleyDepth  
## [1,]                -0.05301802                0.2560132      -0.1895565  
##      RelativeSlopePosition cntgad3a DEMSRE3a etmnts3a evmmod3a evsmod3a  
## [1,]                0.2775459      NA 0.03587897 0.001053781 0.2767067 0.0952547  
##      g01esa3a g01igb3a g02esa3a g02igb3a g03esa3a g04esa3a  
## [1,]      NA -0.07564532 0.09027824 -0.08116583 0.1106826 -0.03904765
```

```

##      g04igb3a  g05esa3a  g06esa3a  g07esa3a  g08esa3a  g09esa3a  g10esa3a
## [1,] -0.09654848 0.09482903 0.01562746 0.004700417      NA      NA      NA
##      g10igb3a  g11esa3a  g11igb3a  g12esa3a  g12igb3a  g13esa3a
## [1,] -0.1021599 -0.1969757 -0.1173455 -0.01263234 -0.1144924 0.07960095
##      g14esa3a  g15esa3a  g16esa3a  g17esa3a  g18esa3a  g19esa3a
## [1,] -0.2238118 -0.04820654 0.002224984 0.01650669 0.01505738      NA
##      g20esa3a  g21esa3a  g22esa3a  gabhws3a  gacgem3a  gachws3a
## [1,] 0.01788399 -0.01595079 -0.004884024      NA 0.01155267 -0.07988801
##      galhws3a  ganhws3a  garhws3a  gathws3a  gchhws3a  gclhws3a  gcmhws3a
## [1,] 0.03551554 0.176973 -0.09041792      NA      NA      NA -0.00781756
##      gcrhws3a  geaisg3a  gflhws3a  gfrhws3a  gglhws3a  ggyhws3a
## [1,]      NA -0.01538151 0.0458282 -0.01483839 -0.04142244      NA
##      ghshws3a  gkshws3a  glcesa3a  glcjrc3a  glphws3a  glvhws3a
## [1,] -0.01271958      NA -0.1909645 0.04634427 -0.06318894 0.00471045
##      glwwwf3a  glxhws3a  gnthws3a  gphhws3a  gplhws3a  gpthws3a  gpzhws3a
## [1,] -0.01907376      NA      NA 0.01026758 -0.07211271 -0.01483839      NA
##      grghws3a  gschws3a  gsnhws3a  gsthws3a  gumhws3a  gvrhws3a  iflgre3a
## [1,] -0.1805161      NA      NA      NA 0.1111965 -0.1274298 -0.01623027
##      inmsre3a  inssre3a  l01igb3a  l02igb3a  l03igb3a  l04igb3a
## [1,] 0.03728952 0.03534979 0.001317087 0.0916482 0.003434706 0.0151673
##      l05igb3a  l06igb3a  l07igb3a  l08igb3a  l09igb3a  l10igb3a
## [1,] 0.07764048 -0.06171173 -0.124861 0.07829523 0.02994008 -0.02359794
##      l11igb3a  l12igb3a  l13igb3a  l14igb3a  l15igb3a  l16igb3a
## [1,] -0.003475584 -0.01153793 -0.1447873 -0.02562095 -0.01242638 -0.03478315
##      l3pobi3b  lammod3a  lasmod3a  lmbqsh3a  lmtgsh3a  ln1dms3a
## [1,] 0.009592703 0.1051425 0.1600325 0.002538524 -0.002538524 0.1571156
##      ln2dms3a  lnmdms3a  opisre3a  px1wcl3a  px2wcl3a  px3wcl3a  px4wcl3a
## [1,] 0.1068743 -0.1550986 0.05806077 0.1531903 0.1723895 0.06300121 0.02936554
##      SLPsRT3a  smkisir3a  tdhmod3a  tdlmod3a  tdmmod3a  tdsmod3a
## [1,] 0.07048761 -0.1420914 -0.2810612 -0.07867261 -0.2131153 -0.1090987
##      tnhmod3a  tnlmod3a  tnmmmod3a  tnsmod3a  twisre3a  tx1mod3a
## [1,] -0.05331798 -0.01532287 -0.03057952 -0.02484648 -0.1439232 -0.2403158
##      tx3mod3a  tx4mod3a  tx5mod3a  tx6mod3a  wmkmod3a  Suelos_ordenes
## [1,] -0.187954 -0.1779113 -0.2330123 -0.2297682 -0.0239136 -0.2025279
## Cobertura_Usos_100 Bioclimatico Pisos_bioclimatico Geologia
## [1,] -0.08820134 -0.1493179      0.01522015 -0.04112401
## Geomorfologia Ecosistemas Tipos_clima DEM.1 AnalyticalHillshading.1
## [1,] -0.04034017 0.02846129 -0.05304596      NaN      -0.01880716
## Slope.1 Aspect.1 CrossSectionalCurvature.1 LongitudinalCurvature.1
## [1,] 0.06140301 0.04726006      0.1214281      0.1537979
## CovergenceIndex.1 ClosedDepressions.1 FlowAccumulation.1
## [1,] 0.1076731 -0.07619768 -0.03890132
## TopographicWetnessIndex.1 LSFactor.1 ChannelNetworkBaseLevel.1
## [1,] -0.1259765 -0.07548198 -0.05301802
## VerticalDistanceToChannelNetwork.1 ValleyDepth.1 RelativeSlopePosition.1
## [1,] 0.2560132 -0.1895565 0.2775459
## cntgad3a.1 DEMSRE3a.1 etmnts3a.1 evmmod3a.1 evsmod3a.1 g01esa3a.1
## [1,]      NA 0.03587897 0.001053781 0.2767067 0.0952547      NA
## g01igb3a.1 g02esa3a.1 g02igb3a.1 g03esa3a.1 g04esa3a.1 g04igb3a.1
## [1,] -0.07564532 0.09027824 -0.08116583 0.1106826 -0.03904765 -0.09654848
## g05esa3a.1 g06esa3a.1 g07esa3a.1 g08esa3a.1 g09esa3a.1 g10esa3a.1
## [1,] 0.09482903 0.01562746 0.004700417      NA      NA      NA
## g10igb3a.1 g11esa3a.1 g11igb3a.1 g12esa3a.1 g12igb3a.1 g13esa3a.1
## [1,] -0.1021599 -0.1969757 -0.1173455 -0.01263234 -0.1144924 0.07960095
## g14esa3a.1 g15esa3a.1 g16esa3a.1 g17esa3a.1 g18esa3a.1 g19esa3a.1
## [1,] -0.2238118 -0.04820654 0.002224984 0.01650669 0.01505738      NA
## g20esa3a.1 g21esa3a.1 g22esa3a.1 gabhws3a.1 gacgem3a.1 gachws3a.1
## [1,] 0.01788399 -0.01595079 -0.004884024      NA 0.01155267 -0.07988801
## galhws3a.1 ganhws3a.1 garhws3a.1 gathws3a.1 gchhws3a.1 gclhws3a.1

```

```
## [1,] 0.03551554 0.176973 -0.09041792 NA NA NA
## gcmhws3a.1 gcrhws3a.1 geaisg3a.1 gflhws3a.1 gfrhws3a.1 gglhws3a.1
## [1,] -0.00781756 NA -0.01538151 0.0458282 -0.01483839 -0.04142244
## ggyhws3a.1 ghshws3a.1 gkshws3a.1 glcesa3a.1 glcjrca.1 glphws3a.1
## [1,] NA -0.01271958 NA -0.1909645 0.04634427 -0.06318894
## glvhws3a.1 glwwf3a.1 glxhws3a.1 gnthws3a.1 gphhws3a.1 gplhws3a.1
## [1,] 0.00471045 -0.01907376 NA NA 0.01026758 -0.07211271
## gpthws3a.1 gpzhws3a.1 grghws3a.1 gschws3a.1 gsnhws3a.1 gsthws3a.1
## [1,] -0.01483839 NA -0.1805161 NA NA NA
## gumhws3a.1 gvrhws3a.1 iflgre3a.1 inmsre3a.1 inssre3a.1 l01igb3a.1
## [1,] 0.1111965 -0.1274298 -0.01623027 0.03728952 0.03534979 0.001317087
## l02igb3a.1 l03igb3a.1 l04igb3a.1 l05igb3a.1 l06igb3a.1 l07igb3a.1
## [1,] 0.0916482 0.003434706 0.0151673 0.07764048 -0.06171173 -0.124861
## l08igb3a.1 l09igb3a.1 l10igb3a.1 l11igb3a.1 l12igb3a.1 l13igb3a.1
## [1,] 0.07829523 0.02994008 -0.02359794 -0.003475584 -0.01153793 -0.1447873
## l14igb3a.1 l15igb3a.1 l16igb3a.1 l3pobi3b.1 lammod3a.1 lasmod3a.1
## [1,] -0.02562095 -0.01242638 -0.03478315 0.009592703 0.1051425 0.1600325
## lmbgsh3a.1 lmtgsh3a.1 ln1dms3a.1 ln2dms3a.1 lnmdms3a.1 opisre3a.1
## [1,] 0.002538524 -0.002538524 0.1571156 0.1068743 -0.1550986 0.05806077
## px1wcl3a.1 px2wcl3a.1 px3wcl3a.1 px4wcl3a.1 SLPsRT3a.1 smkisir3a.1
## [1,] 0.1531903 0.1723895 0.06300121 0.02936554 0.07048761 -0.1420914
## tdhmod3a.1 tdlmod3a.1 tdmmod3a.1 tdsmod3a.1 tnhmod3a.1 tnmod3a.1
## [1,] -0.2810612 -0.07867261 -0.2131153 -0.1090987 -0.05331798 -0.01532287
## tnmmod3a.1 tnsmod3a.1 twisre3a.1 tx1mod3a.1 tx2mod3a.1 tx3mod3a.1
## [1,] -0.03057952 -0.02484648 -0.1439232 -0.2403158 -0.2901936 -0.187954
## tx4mod3a.1 tx5mod3a.1 tx6mod3a.1 wmkmod3a.1 Suelos_ordenes.1
## [1,] -0.1779113 -0.2330123 -0.2297682 -0.0239136 -0.2025279
## Cobertura_Uso_100.1 Bioclimatico.1 Pisos_bioclimatico.1 Geologia.1
## [1,] -0.08820134 -0.1493179 0.01522015 -0.04112401
## Geomorfologia.1 Ecosistemas.1 Tipos_clima.1
## [1,] -0.04034017 0.02846129 -0.05304596

x <- subset(melt(COR), value != 1 | value != NA)
x <- x[with(x, order(-abs(x$value))),]
#as.character(x$X2[1:10])
x[1:10,]

## X1 X2 value
## 270 1 tx2mod3a.1 -0.2901936
## 120 1 tdhmod3a -0.2810612
## 260 1 tdhmod3a.1 -0.2810612
## 17 1 RelativeSlopePosition 0.2775459
## 157 1 RelativeSlopePosition.1 0.2775459
## 21 1 evmmod3a 0.2767067
## 161 1 evmmod3a.1 0.2767067
## 15 1 VerticalDistanceToChannelNetwork 0.2560132
## 155 1 VerticalDistanceToChannelNetwork.1 0.2560132
## 129 1 tx1mod3a -0.2403158

idx <- as.character(x$X2[1:20])

dat2 <- dat[c('OCSKGM30', idx, 'Latitude', 'Longitude')]
names(dat2)

## [1] "OCSKGM30" "tx2mod3a.1"
## [3] "tdhmod3a" "tdhmod3a.1"
## [5] "RelativeSlopePosition" "RelativeSlopePosition.1"
## [7] "evmmod3a" "evmmod3a.1"
## [9] "VerticalDistanceToChannelNetwork" "VerticalDistanceToChannelNetwork.1"
```

```
## [11] "tx1mod3a"          "tx1mod3a.1"
## [13] "tx5mod3a"          "tx5mod3a.1"
## [15] "tx6mod3a"          "tx6mod3a.1"
## [17] "g14esa3a"          "g14esa3a.1"
## [19] "tdmmod3a"          "tdmmod3a.1"
## [21] "Suelos_ordenes"   "Latitude"
## [23] "Longitude"
```

```
dat2$Suelos_ordenes<-NULL
dat2$RelativeSlopePosition<-NULL
dat2$RelativeSlopePosition.1<-NULL
```

```
dat2[dat$OCSKGM30 == 0, 1] <- NA
```

### Ajuste de modelos con una porción de los datos

```
## Ahora, a diferencia del ejercicio 4, el modelo lo ajustamos solo con datos.model
## y no con todos los datos...!
```

```
modelo.MLR <- lm(Log(OCSKGM30) ~ . - Latitude - Longitude, data = dat2)
```

```
summary(modelo.MLR)
```

```
##
## Call:
## lm(formula = log(OCSKGM30) ~ . - Latitude - Longitude, data = dat2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4.5436 -0.3881  0.0640  0.4671  2.4651
##
## Coefficients: (8 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.327e+00  1.399e-01  9.486 < 2e-16 ***
## tx2mod3a.1   -6.510e-02  6.047e-03 -10.766 < 2e-16 ***
## tdhmod3a     -3.857e-02  4.168e-03  -9.253 < 2e-16 ***
## tdhmod3a.1   NA              NA         NA         NA
## evmmod3a     2.613e-04  1.318e-05  19.826 < 2e-16 ***
## evmmod3a.1   NA              NA         NA         NA
## VerticalDistanceToChannelNetwork  4.199e-04  3.444e-05  12.192 < 2e-16 ***
## VerticalDistanceToChannelNetwork.1 NA              NA         NA         NA
## tx1mod3a     1.469e-02  5.569e-03  2.638  0.00835 **
## tx1mod3a.1   NA              NA         NA         NA
## tx5mod3a     -2.130e-02  4.490e-03  -4.744  2.14e-06 ***
## tx5mod3a.1   NA              NA         NA         NA
## tx6mod3a     4.675e-02  4.585e-03  10.196 < 2e-16 ***
## tx6mod3a.1   NA              NA         NA         NA
## g14esa3a     -3.151e-03  4.501e-04  -7.001  2.80e-12 ***
## g14esa3a.1   NA              NA         NA         NA
## tdmmod3a     3.893e-02  6.366e-03  6.116  1.02e-09 ***
## tdmmod3a.1   NA              NA         NA         NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.684 on 6476 degrees of freedom
## Multiple R-squared:  0.2344, Adjusted R-squared:  0.2333
## F-statistic: 220.3 on 9 and 6476 DF, p-value: < 2.2e-16
```

Análisis anova del modelo hecho con los datos, sales las covariables con mejores valores

`anova(modelo.MLR)`

```
## Analysis of Variance Table
##
## Response: Log(OCSKGM30)
##
## Df Sum Sq Mean Sq F value Pr(>F)
## tx2mod3a.1 1 283.96 283.96 606.972 < 2.2e-16 ***
## tdhmod3a 1 67.66 67.66 144.623 < 2.2e-16 ***
## evmmod3a 1 343.25 343.25 733.711 < 2.2e-16 ***
## VerticalDistanceToChannelNetwork 1 79.68 79.68 170.311 < 2.2e-16 ***
## tx1mod3a 1 33.01 33.01 70.558 < 2.2e-16 ***
## tx5mod3a 1 3.84 3.84 8.198 0.004207 **
## tx6mod3a 1 61.94 61.94 132.393 < 2.2e-16 ***
## g14esa3a 1 36.77 36.77 78.590 < 2.2e-16 ***
## tdmmod3a 1 17.50 17.50 37.406 1.015e-09 ***
## Residuals 6476 3029.66 0.47
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Se hace una selección de variables por sucesión de paso, en cada paso se introduce la variable independiente que no se encuentre ya en la ecuación y que tenga la probabilidad para F más pequeña (i.e. hacia adelante). El método termina cuando ya no hay más variables candidatas a ser incluidas o eliminadas

## Hacemos selección de variables por stepwise

`modelo.MLR.step <- step(modelo.MLR, direction="both")`

```
## Start: AIC=-4917.11
## Log(OCSKGM30) ~ (tx2mod3a.1 + tdhmod3a + tdhmod3a.1 + evmmod3a +
## evmmod3a.1 + VerticalDistanceToChannelNetwork +
VerticalDistanceToChannelNetwork.1 +
## tx1mod3a + tx1mod3a.1 + tx5mod3a + tx5mod3a.1 + tx6mod3a +
## tx6mod3a.1 + g14esa3a + g14esa3a.1 + tdmmod3a + tdmmod3a.1 +
## Latitude + Longitude) - Latitude - Longitude
##
##
## Step: AIC=-4917.11
## Log(OCSKGM30) ~ tx2mod3a.1 + tdhmod3a + tdhmod3a.1 + evmmod3a +
## evmmod3a.1 + VerticalDistanceToChannelNetwork +
VerticalDistanceToChannelNetwork.1 +
## tx1mod3a + tx1mod3a.1 + tx5mod3a + tx5mod3a.1 + tx6mod3a +
## tx6mod3a.1 + g14esa3a + g14esa3a.1 + tdmmod3a
##
##
## Step: AIC=-4917.11
## Log(OCSKGM30) ~ tx2mod3a.1 + tdhmod3a + tdhmod3a.1 + evmmod3a +
## evmmod3a.1 + VerticalDistanceToChannelNetwork +
VerticalDistanceToChannelNetwork.1 +
## tx1mod3a + tx1mod3a.1 + tx5mod3a + tx5mod3a.1 + tx6mod3a +
## tx6mod3a.1 + g14esa3a + tdmmod3a
##
##
## Step: AIC=-4917.11
```

```
## Log(OCSKGM30) ~ tx2mod3a.1 + tdhmod3a + tdhmod3a.1 + evmmod3a +
##   evmmod3a.1 + VerticalDistanceToChannelNetwork +
VerticalDistanceToChannelNetwork.1 +
##   tx1mod3a + tx1mod3a.1 + tx5mod3a + tx5mod3a.1 + tx6mod3a +
##   g14esa3a + tdmmod3a
##
##
## Step: AIC=-4917.11
## Log(OCSKGM30) ~ tx2mod3a.1 + tdhmod3a + tdhmod3a.1 + evmmod3a +
##   evmmod3a.1 + VerticalDistanceToChannelNetwork +
VerticalDistanceToChannelNetwork.1 +
##   tx1mod3a + tx1mod3a.1 + tx5mod3a + tx6mod3a + g14esa3a +
##   tdmmod3a
##
##
## Step: AIC=-4917.11
## Log(OCSKGM30) ~ tx2mod3a.1 + tdhmod3a + tdhmod3a.1 + evmmod3a +
##   evmmod3a.1 + VerticalDistanceToChannelNetwork +
VerticalDistanceToChannelNetwork.1 +
##   tx1mod3a + tx5mod3a + tx6mod3a + g14esa3a + tdmmod3a
##
##
## Step: AIC=-4917.11
## Log(OCSKGM30) ~ tx2mod3a.1 + tdhmod3a + tdhmod3a.1 + evmmod3a +
##   evmmod3a.1 + VerticalDistanceToChannelNetwork + tx1mod3a +
##   tx5mod3a + tx6mod3a + g14esa3a + tdmmod3a
##
##
## Step: AIC=-4917.11
## Log(OCSKGM30) ~ tx2mod3a.1 + tdhmod3a + tdhmod3a.1 + evmmod3a +
##   VerticalDistanceToChannelNetwork + tx1mod3a + tx5mod3a +
##   tx6mod3a + g14esa3a + tdmmod3a
##
##
## Step: AIC=-4917.11
## Log(OCSKGM30) ~ tx2mod3a.1 + tdhmod3a + evmmod3a +
VerticalDistanceToChannelNetwork +
##   tx1mod3a + tx5mod3a + tx6mod3a + g14esa3a + tdmmod3a
##
##
##
##          Df Sum of Sq   RSS   AIC
## <none>                    3029.7 -4917.1
## - tx1mod3a                1    3.257 3032.9 -4912.1
## - tx5mod3a                1   10.531 3040.2 -4896.6
## - tdmmod3a                1   17.500 3047.2 -4881.7
## - g14esa3a                1   22.930 3052.6 -4870.2
## - tdhmod3a                1   40.055 3069.7 -4833.9
## - tx6mod3a                1   48.636 3078.3 -4815.8
## - tx2mod3a.1              1   54.221 3083.9 -4804.1
## - VerticalDistanceToChannelNetwork 1   69.545 3099.2 -4771.9
## - evmmod3a                1  183.890 3213.6 -4536.9

summary(modelo.MLR.step)

##
## Call:
## lm(formula = Log(OCSKGM30) ~ tx2mod3a.1 + tdhmod3a + evmmod3a +
##   VerticalDistanceToChannelNetwork + tx1mod3a + tx5mod3a +
##   tx6mod3a + g14esa3a + tdmmod3a, data = dat2)
##
```

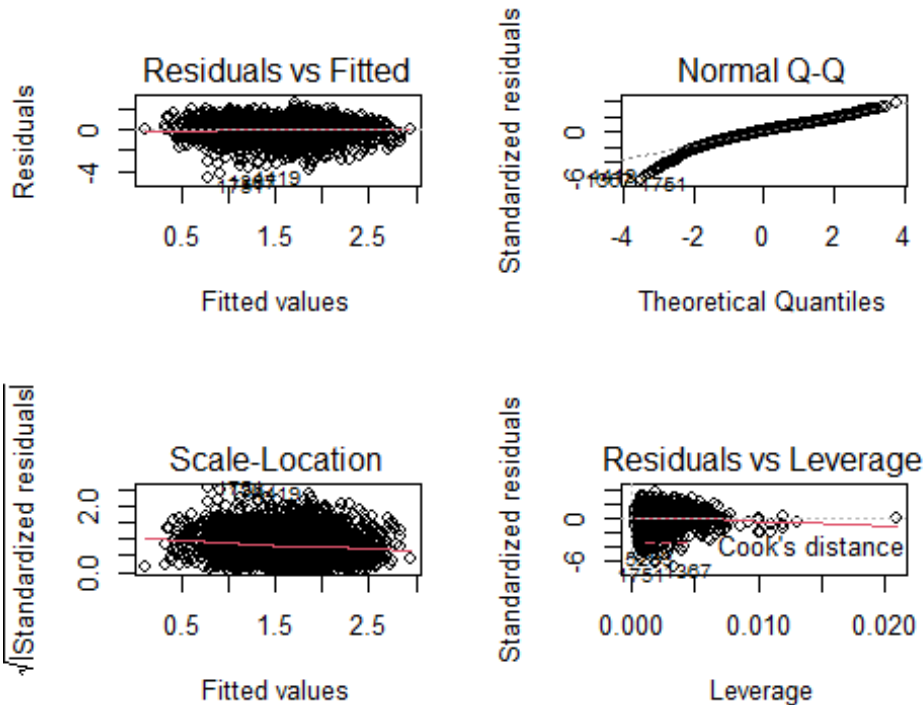


```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4.5436 -0.3881  0.0640  0.4671  2.4651
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.327e+00  1.399e-01   9.486 < 2e-16 ***
## tx2mod3a.1  -6.510e-02  6.047e-03 -10.766 < 2e-16 ***
## tdhmod3a    -3.857e-02  4.168e-03  -9.253 < 2e-16 ***
## evmmod3a    2.613e-04  1.318e-05  19.826 < 2e-16 ***
## VerticalDistanceToChannelNetwork  4.199e-04  3.444e-05  12.192 < 2e-16 ***
## tx1mod3a    1.469e-02  5.569e-03   2.638  0.00835 **
## tx5mod3a   -2.130e-02  4.490e-03  -4.744  2.14e-06 ***
## tx6mod3a    4.675e-02  4.585e-03  10.196 < 2e-16 ***
## g14esa3a   -3.151e-03  4.501e-04  -7.001  2.80e-12 ***
## tdmmod3a    3.893e-02  6.366e-03   6.116  1.02e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.684 on 6476 degrees of freedom
## Multiple R-squared:  0.2344, Adjusted R-squared:  0.2333
## F-statistic: 220.3 on 9 and 6476 DF,  p-value: < 2.2e-16

anova(modelo.MLR.step)

## Analysis of Variance Table
##
## Response: Log(OCSKGM30)
##              Df Sum Sq Mean Sq F value    Pr(>F)
## tx2mod3a.1    1  283.96   283.96  606.972 < 2.2e-16 ***
## tdhmod3a      1   67.66    67.66  144.623 < 2.2e-16 ***
## evmmod3a      1  343.25   343.25  733.711 < 2.2e-16 ***
## VerticalDistanceToChannelNetwork  1   79.68    79.68  170.311 < 2.2e-16 ***
## tx1mod3a      1   33.01    33.01   70.558 < 2.2e-16 ***
## tx5mod3a      1    3.84     3.84    8.198  0.004207 **
## tx6mod3a      1   61.94    61.94  132.393 < 2.2e-16 ***
## g14esa3a      1   36.77    36.77   78.590 < 2.2e-16 ***
## tdmmod3a      1   17.50    17.50   37.406  1.015e-09 ***
## Residuals    6476 3029.66     0.47
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

par(mfrow=c(2,2))
plot(modelo.MLR.step)
```



```
par(mfrow=c(1,1))
```

```
#Falta de multicolinealidad en las variables x: podemos comprobar esto mediante
#el calculo de los Factores de Inflacion de la Varianza (FIVs)
```

```
library(car)
```

```
vif(modelo.MLR.step)
```

```
##          tx2mod3a.1          tdhmod3a
##          6.384951          4.537743
##          evmmod3a VerticalDistanceToChannelNetwork
##          1.974899          2.352436
##          tx1mod3a          tx5mod3a
##          6.831080          4.544193
##          tx6mod3a          g14esa3a
##          6.077359          1.559314
##          tdmmod3a
##          9.468623
```

```
#Variables problematicas tienen sqrt(FIV) > 2
```

```
sqrt(vif(modelo.MLR.step))
```

```
##          tx2mod3a.1          tdhmod3a
##          2.526846          2.130198
##          evmmod3a VerticalDistanceToChannelNetwork
##          1.405311          1.533765
##          tx1mod3a          tx5mod3a
##          2.613634          2.131711
##          tx6mod3a          g14esa3a
##          2.465230          1.248725
##          tdmmod3a
##          3.077113
```

Se elimina las variables con valores mayores a 2 en el sqrt(FIV)

```
modelo.MLR.step <- update(modelo.MLR.step, . ~ . - tx2mod3a.1
- tx1mod3a - tx5mod3a - tx6mod3a - tdmmod3a-tdmmod3a
- tdhmod3a)
```

### Se busca valores atípicos

#Vamos usar la prueba de Bonferroni para valores atipicos:  
`outlierTest(modelo.MLR.step)`

```
##      rstudent unadjusted p-value Bonferroni p
## 1751 -6.849772      8.0779e-12  5.2393e-08
## 1367 -6.234128      4.8287e-10  3.1319e-06
## 4419 -5.785131      7.5836e-09  4.9188e-05
## 1290 -5.429497      5.8565e-08  3.7985e-04
## 990  -5.059451      4.3204e-07  2.8022e-03
## 1289 -5.056136      4.3958e-07  2.8511e-03
## 4068 -4.809591      1.5466e-06  1.0031e-02
## 1284 -4.768391      1.8976e-06  1.2308e-02
## 5293 -4.766367      1.9167e-06  1.2432e-02
## 6192 -4.593232      4.4468e-06  2.8842e-02
```

# Lo convertimos a un SpatialGridDataFrame  
`topo <- stack('ECUtopo/ECUtopo.tif')`

`namesTopo <- readRDS('ECUtopo/namesTOPO.rds')`

`names(topo)`

```
## [1] "ECUtopo.1" "ECUtopo.2" "ECUtopo.3" "ECUtopo.4" "ECUtopo.5"
## [6] "ECUtopo.6" "ECUtopo.7" "ECUtopo.8" "ECUtopo.9" "ECUtopo.10"
## [11] "ECUtopo.11" "ECUtopo.12" "ECUtopo.13" "ECUtopo.14" "ECUtopo.15"
```

`names(topo) <- namesTopo`

`cov <- stack('ECU_worldgridsCOVS/ECU_worldgridsCOVS.tif')`

`namesCov <- readRDS('ECU_worldgridsCOVS/worldgridsCOVS_names.rds')`

`names(cov)`

```
## [1] "ECU_worldgridsCOVS.1" "ECU_worldgridsCOVS.2" "ECU_worldgridsCOVS.3"
## [4] "ECU_worldgridsCOVS.4" "ECU_worldgridsCOVS.5" "ECU_worldgridsCOVS.6"
## [7] "ECU_worldgridsCOVS.7" "ECU_worldgridsCOVS.8" "ECU_worldgridsCOVS.9"
## [10] "ECU_worldgridsCOVS.10" "ECU_worldgridsCOVS.11" "ECU_worldgridsCOVS.12"
## [13] "ECU_worldgridsCOVS.13" "ECU_worldgridsCOVS.14" "ECU_worldgridsCOVS.15"
## [16] "ECU_worldgridsCOVS.16" "ECU_worldgridsCOVS.17" "ECU_worldgridsCOVS.18"
## [19] "ECU_worldgridsCOVS.19" "ECU_worldgridsCOVS.20" "ECU_worldgridsCOVS.21"
## [22] "ECU_worldgridsCOVS.22" "ECU_worldgridsCOVS.23" "ECU_worldgridsCOVS.24"
## [25] "ECU_worldgridsCOVS.25" "ECU_worldgridsCOVS.26" "ECU_worldgridsCOVS.27"
## [28] "ECU_worldgridsCOVS.28" "ECU_worldgridsCOVS.29" "ECU_worldgridsCOVS.30"
## [31] "ECU_worldgridsCOVS.31" "ECU_worldgridsCOVS.32" "ECU_worldgridsCOVS.33"
## [34] "ECU_worldgridsCOVS.34" "ECU_worldgridsCOVS.35" "ECU_worldgridsCOVS.36"
## [37] "ECU_worldgridsCOVS.37" "ECU_worldgridsCOVS.38" "ECU_worldgridsCOVS.39"
## [40] "ECU_worldgridsCOVS.40" "ECU_worldgridsCOVS.41" "ECU_worldgridsCOVS.42"
## [43] "ECU_worldgridsCOVS.43" "ECU_worldgridsCOVS.44" "ECU_worldgridsCOVS.45"
## [46] "ECU_worldgridsCOVS.46" "ECU_worldgridsCOVS.47" "ECU_worldgridsCOVS.48"
## [49] "ECU_worldgridsCOVS.49" "ECU_worldgridsCOVS.50" "ECU_worldgridsCOVS.51"
```

```
## [52] "ECU_worldgridsCOVS.52" "ECU_worldgridsCOVS.53" "ECU_worldgridsCOVS.54"
## [55] "ECU_worldgridsCOVS.55" "ECU_worldgridsCOVS.56" "ECU_worldgridsCOVS.57"
## [58] "ECU_worldgridsCOVS.58" "ECU_worldgridsCOVS.59" "ECU_worldgridsCOVS.60"
## [61] "ECU_worldgridsCOVS.61" "ECU_worldgridsCOVS.62" "ECU_worldgridsCOVS.63"
## [64] "ECU_worldgridsCOVS.64" "ECU_worldgridsCOVS.65" "ECU_worldgridsCOVS.66"
## [67] "ECU_worldgridsCOVS.67" "ECU_worldgridsCOVS.68" "ECU_worldgridsCOVS.69"
## [70] "ECU_worldgridsCOVS.70" "ECU_worldgridsCOVS.71" "ECU_worldgridsCOVS.72"
## [73] "ECU_worldgridsCOVS.73" "ECU_worldgridsCOVS.74" "ECU_worldgridsCOVS.75"
## [76] "ECU_worldgridsCOVS.76" "ECU_worldgridsCOVS.77" "ECU_worldgridsCOVS.78"
## [79] "ECU_worldgridsCOVS.79" "ECU_worldgridsCOVS.80" "ECU_worldgridsCOVS.81"
## [82] "ECU_worldgridsCOVS.82" "ECU_worldgridsCOVS.83" "ECU_worldgridsCOVS.84"
## [85] "ECU_worldgridsCOVS.85" "ECU_worldgridsCOVS.86" "ECU_worldgridsCOVS.87"
## [88] "ECU_worldgridsCOVS.88" "ECU_worldgridsCOVS.89" "ECU_worldgridsCOVS.90"
## [91] "ECU_worldgridsCOVS.91" "ECU_worldgridsCOVS.92" "ECU_worldgridsCOVS.93"
## [94] "ECU_worldgridsCOVS.94" "ECU_worldgridsCOVS.95" "ECU_worldgridsCOVS.96"
## [97] "ECU_worldgridsCOVS.97" "ECU_worldgridsCOVS.98" "ECU_worldgridsCOVS.99"
## [100] "ECU_worldgridsCOVS.100" "ECU_worldgridsCOVS.101" "ECU_worldgridsCOVS.102"
## [103] "ECU_worldgridsCOVS.103" "ECU_worldgridsCOVS.104" "ECU_worldgridsCOVS.105"
## [106] "ECU_worldgridsCOVS.106" "ECU_worldgridsCOVS.107" "ECU_worldgridsCOVS.108"
## [109] "ECU_worldgridsCOVS.109" "ECU_worldgridsCOVS.110" "ECU_worldgridsCOVS.111"
## [112] "ECU_worldgridsCOVS.112" "ECU_worldgridsCOVS.113" "ECU_worldgridsCOVS.114"
## [115] "ECU_worldgridsCOVS.115" "ECU_worldgridsCOVS.116" "ECU_worldgridsCOVS.117"
## [118] "ECU_worldgridsCOVS.118"
```

```
names(cov) <- namesCov
```

```
COV <- stack(topo, cov)
```

```
COV <- COV[[idx]]
```

```
## Warning in .local(x, ...): invalid layer names omitted
```

```
names(COV)
```

```
## [1] "tdhmod3a" "RelativeSlopePosition"
## [3] "evmmod3a" "VerticalDistanceToChannelNetwork"
## [5] "tx1mod3a" "tx5mod3a"
## [7] "tx6mod3a" "g14esa3a"
## [9] "tdmmod3a"
```

```
# Project point data
```

```
dat_sp@proj4string <- COV@crs
```

```
dat_sp <- spTransform(dat_sp, CRS("+init=epsg:32717"))
```

```
# project covariates to WGS84 UTM 21S
```

```
COV <- projectRaster(COV, crs = CRS("+init=epsg:32717"), method='ngb')
```

```
COV.sp <- as(COV, "SpatialGridDataFrame")
```

MODELAMIENTO POR RKRIGING

```
### RK model
#install.packages(automap)
library(automap)

dat_sp <- dat_sp[dat_sp$OCSKGM30 != 0,]

dat_sp$Suelos_ordenes <-NULL
names(dat_sp)

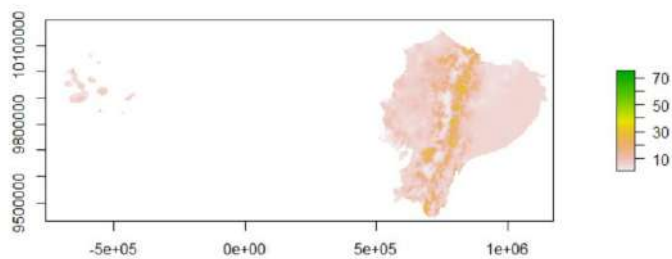
start <- Sys.time()
OCS.krige <- autoKrige(formula = as.formula(modelo.MLR.step$call$formula),
  input_data = dat_sp,
  new_data = COV.sp,
  verbose = TRUE,
  block = c(1000, 1000),
  model = c("Sph", "Exp"))

print(Sys.time() - start)

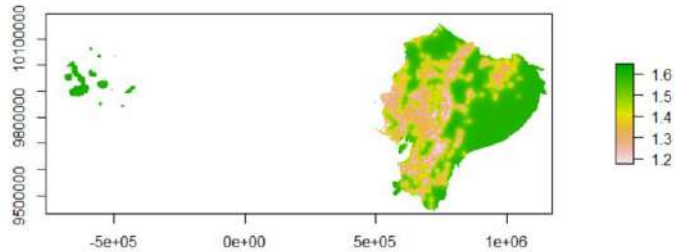
## Time difference of 15.821567 hours

# Convert prediction and standard deviation to rasters
# and back-transform the vlaues

RKprediction <- exp(raster(OCS.krige$krige_output[1]))
RKpredsd <- exp(raster(OCS.krige$krige_output[3]))
plot(RKprediction)
```



```
plot(RKpredsd)
```



```
summary(RKprediction)
summary(RKpredsd)

# Remove bad values
values(RKprediction)[values(RKprediction) < 0] <- NA
values(RKprediction)[values(RKprediction) > 100] <- NA
values(RKpredsd)[values(RKpredsd) > 10] <- NA

summary(RKprediction)
summary(RKpredsd)

#
plot(RKprediction)
plot(RKpredsd)

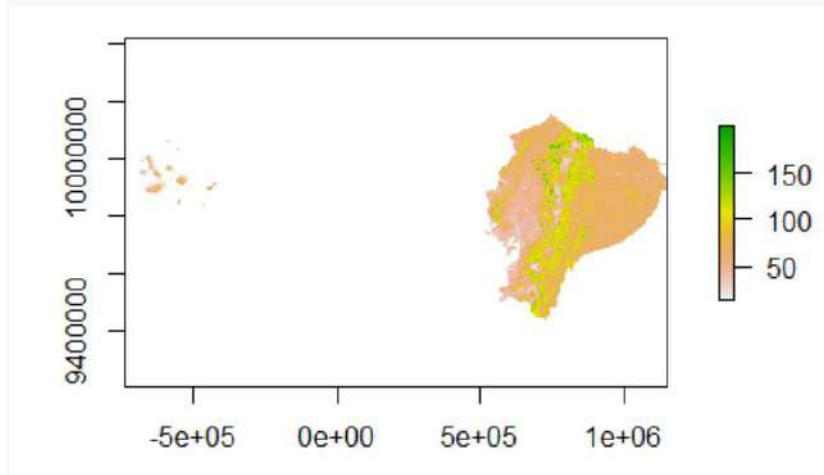
## Save results
writeRaster(RKprediction, filename = "ECU_OCS_RK31.tif")
writeRaster(RKpredsd, filename = "ECU_OCS_RKpredsd31.tif")

#Libreria necesaria
library(raster)
#importas tu raster
r1 <- raster('ECU_OCS_RK31.tif')
```

```
#haces la operacion
```

```
r2 <- r1 *10
```

```
plot(r2)
```



```
#y guardas el raster generado
```

```
writeRaster(r2, 'RKmapa_carbonoPor10ECU31.tif')
```

```
#### Uncertainty estimation
```

```
# using LOO cross-validation
```

```
dat_sp = dat_sp[which(!duplicated(dat_sp@coords)), ]
```

```
OCS.krige.cv <- autoKrige.cv(formula = as.formula(modelo.MLR.step$call$formula),  
                             input_data = dat_sp, nfold = 5)
```

```
summary(OCS.krige.cv)
```

```
#libreria necesaria
```

```
library(raster)
```

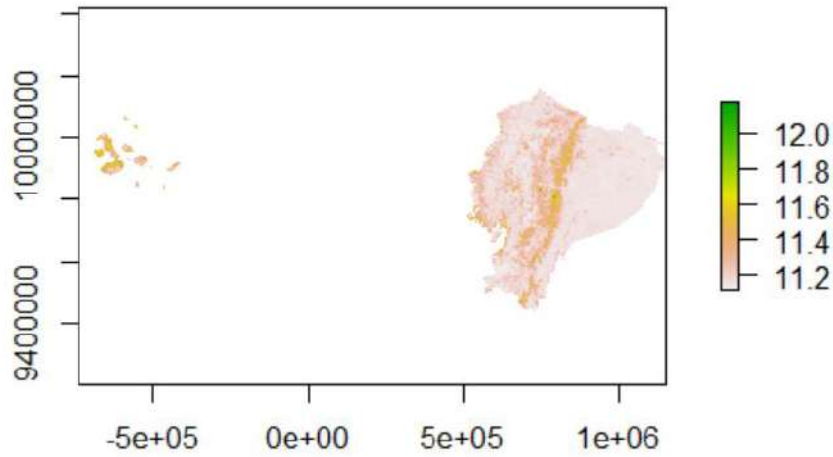
```
#importas tu raster
```

```
r3 <- raster ('ECU_OCS_RKpredsd31.tif')
```

```
#haces la operacion
```

```
r4 <- r3 *10
```

```
plot(r4)
```



```
#y guardas el raster generado
writeRaster(r4, 'RKpredsmapa_carbonoPor10ECU31.tif', overwrite=TRUE)
# Save results as *.tif files
writeRaster(RKprediction, filename = "results/MKD_OCSKGM_RK.tif",
overwrite = TRUE)
writeRaster(RKpredsd, filename = "results/MKD_OCSKGM_RKpredsd.tif",
overwrite = TRUE)
# Save the model
saveRDS(model.MLR.step, file="results/RKmodel.Rds")
```





@AgriculturaEcuador



@agricultura.ec



@AgriculturaEc

## Ministerio de Agricultura y Ganadería



República  
del Ecuador



**Gobierno**  
del Encuentro

Juntos  
lo logramos