

Resultados propuestas de investigación.

Avances en el proyecto “Análisis de Oportunidades Energéticas con Fuentes Alternativas en el Departamento de Nariño - ALTERNAR”.

Componente: Biomasa

Grupo de Investigación en Ingeniería Eléctrica y Electrónica¹

¹Departamento de Ingeniería Electrónica
Universidad de Nariño

Agosto 2015

Agenda

- 1 Planteamiento del problema.
- 2 Trabajos relacionados.
- 3 La propuesta...
- 4 Ejecución de la propuesta...
 - Los datos...
 - Pre-procesamiento...
 - Modelado...
 - Implementación y visualización de resultados...

Agenda

- 1 Planteamiento del problema.
- 2 Trabajos relacionados.
- 3 La propuesta...
- 4 Ejecución de la propuesta...
 - Los datos...
 - Pre-procesamiento...
 - Modelado...
 - Implementación y visualización de resultados...

Planteamiento del problema

- Ubicación de potenciales lugares donde las condiciones de biomasa resulten apropiadas.
- No existen series de tiempo disponibles para la región.
- Estos dependen de un trabajo de campo en la zona (costoso).

Algunas alternativas

- Acceso a un amplio repositorio de imágenes satelitales (Landsat).
- Acceso a estudios previos de estimación de biomasa (2002-03).
- Disponibilidad de implementaciones para la generación de modelos de regresión.
- Diversos estudios han explorado el uso de imágenes satelitales para la estimación de biomasa.

Agenda

- 1 Planteamiento del problema.
- 2 Trabajos relacionados.
- 3 La propuesta...
- 4 Ejecución de la propuesta...
 - Los datos...
 - Pre-procesamiento...
 - Modelado...
 - Implementación y visualización de resultados...

Trabajos relacionados

- Klemas [2013] presenta una metodología para detectar cambios en los niveles de biomasa en diferentes zonas costeras de USA usando imágenes Lidar y regresión lineal.
- Baccini et al. [2008] usan imágenes MODIS y árboles de decisión para estimar el índice AGB (above-ground biomass) en una extensa área del África tropical.
- De manera similar, Mitchard et al. [2009] usan imágenes de radar para predecir AGB en cuatro reservas y parques naturales africanos clasificando diversos tipos de corteza terrestre.

Trabajos relacionados

- Muukkonen and Heiskanen [2007] usan imágenes MODIS y ASTER para estimar biomasa con el fin de levantar un inventario de captura de carbono.
- Powell et al. [2010] utilizan nuevas técnicas de regresión para la generación de modelos de biomasa utilizando esta vez imágenes Landsat.
- En Hall et al. [2006] se introduce el método bioSTRUCT el cual genera correlaciones entre las bandas de imágenes Landsat y valores nominales de AGB medidas en laboratorio.

Agenda

- 1 Planteamiento del problema.
- 2 Trabajos relacionados.
- 3 La propuesta...
 - 4 Ejecución de la propuesta...
 - Los datos...
 - Pre-procesamiento...
 - Modelado...
 - Implementación y visualización de resultados...

Objetivo

Calcular modelos de regresión de biomasa a partir de bandas de imágenes satelitales Landsat para la generación de mapas de biomasa en el departamento de Nariño.

Metodología

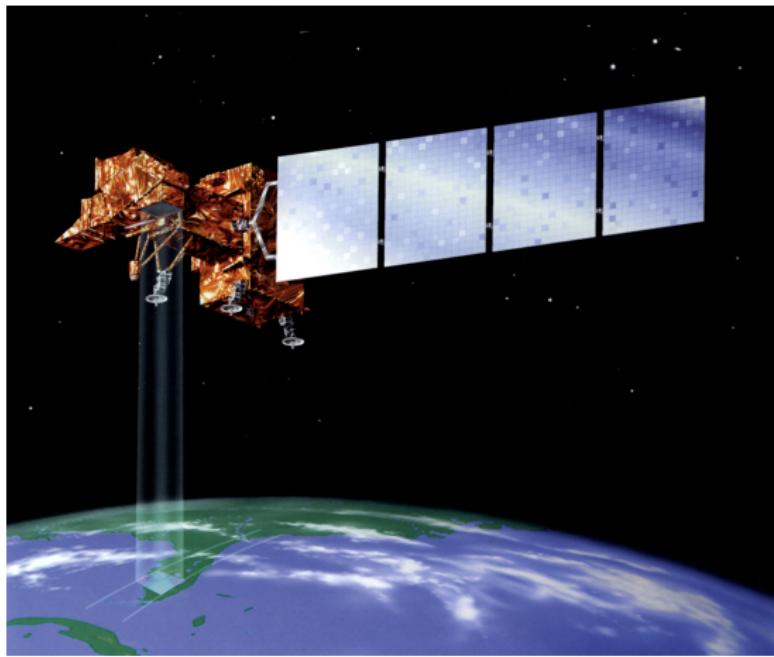
- Obtener un banco de imágenes satelitales del área de estudio desde los repositorios de libre acceso de Landsat.
- Seleccionar las técnicas de regresión y parámetros más apropiados.
- Aplicar las técnicas seleccionadas sobre los datos procesados para la obtención de los modelos de regresión.
- Extrapolar y visualizar los resultados obtenidos.

Agenda

- 1 Planteamiento del problema.
- 2 Trabajos relacionados.
- 3 La propuesta...
- 4 Ejecución de la propuesta...
 - Los datos...
 - Pre-procesamiento...
 - Modelado...
 - Implementación y visualización de resultados...

Landsat 7

- Landsat 7 (L7) Enhanced Thematic Mapper Plus (ETM+)

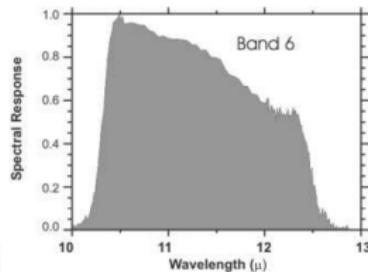
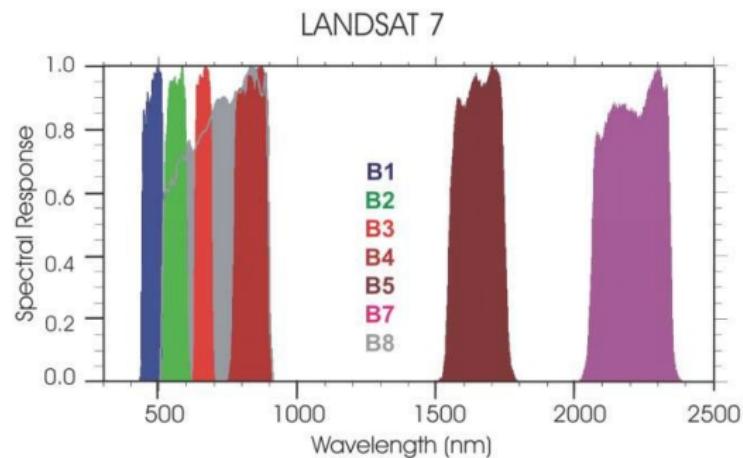


Landsat 7

- Landsat 7 (L7) Enhanced Thematic Mapper Plus (ETM+)
 - Disponible desde Abril de 1999.
 - Intervalo de repetición cada 16 días.
 - Una banda pancromática con resolución espacial de 15m.
 - Tres bandas visibles para el azul, verde, rojo a 30m.
 - Tres bandas infrarrojas a 30m.
 - Una banda térmica a 60m.

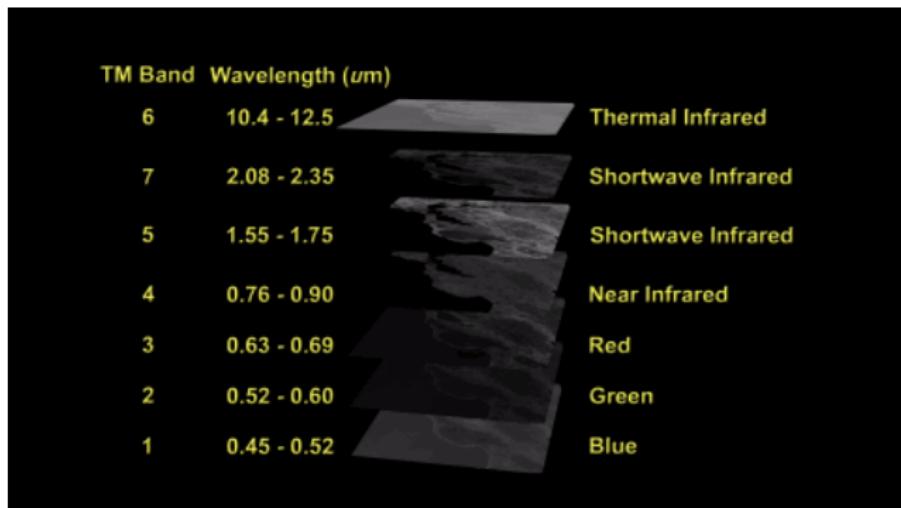
Landsat 7

- Landsat 7 (L7) Enhanced Thematic Mapper Plus (ETM+)



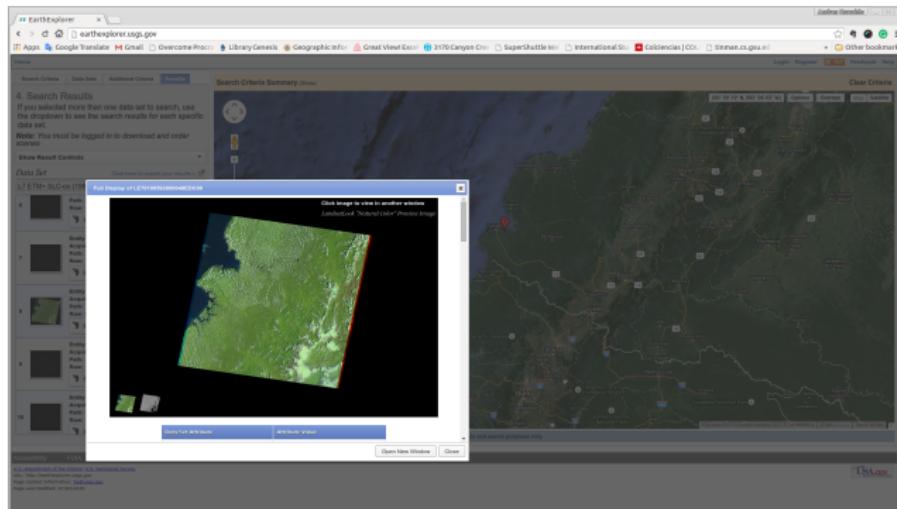
Landsat 7

- Landsat 7 (L7) Enhanced Thematic Mapper Plus (ETM+)



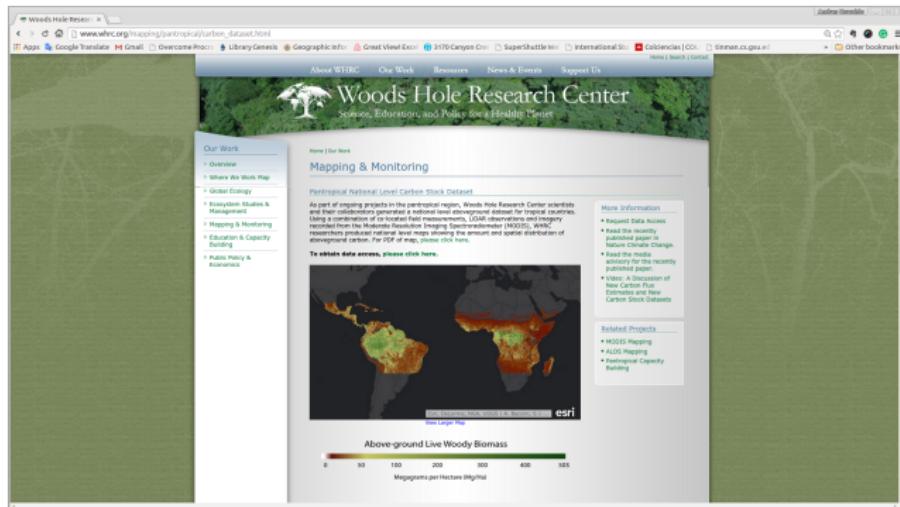
Landsat 7

- Earth Explorer



Pantropical National Level Carbon Stock Dataset

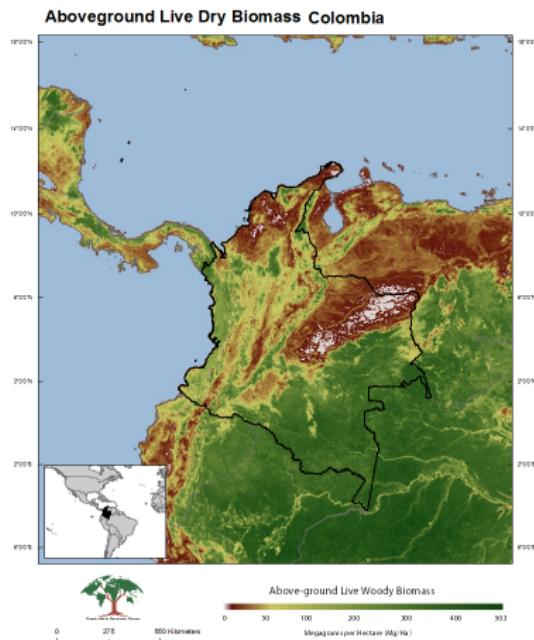
• The Woods Hole Research Center



[Baccini et al., 2012]

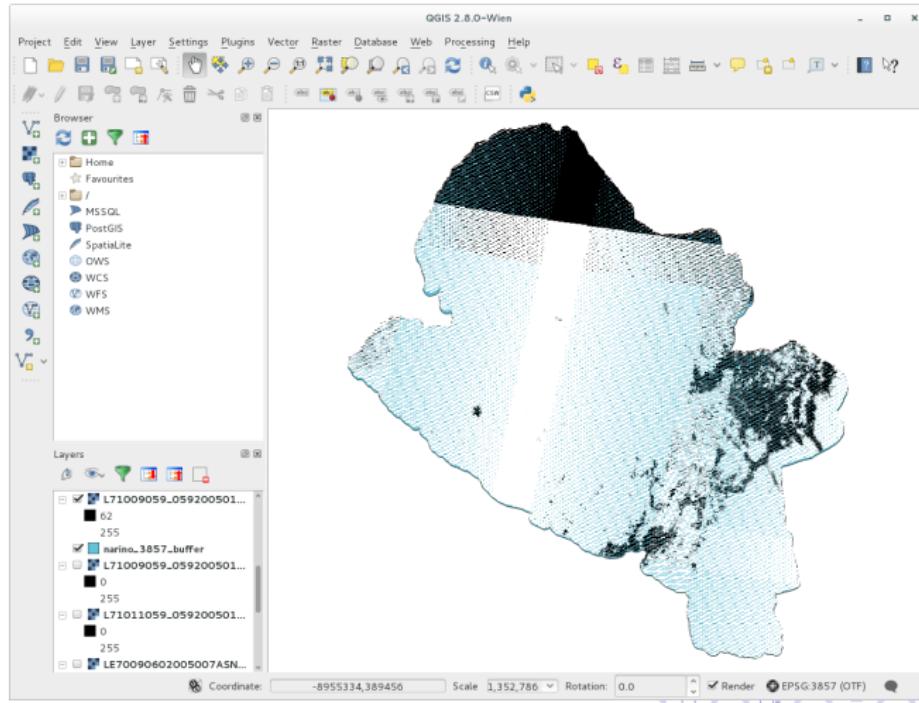
Pantropical National Level Carbon Stock Dataset

- The Woods Hole Research Center



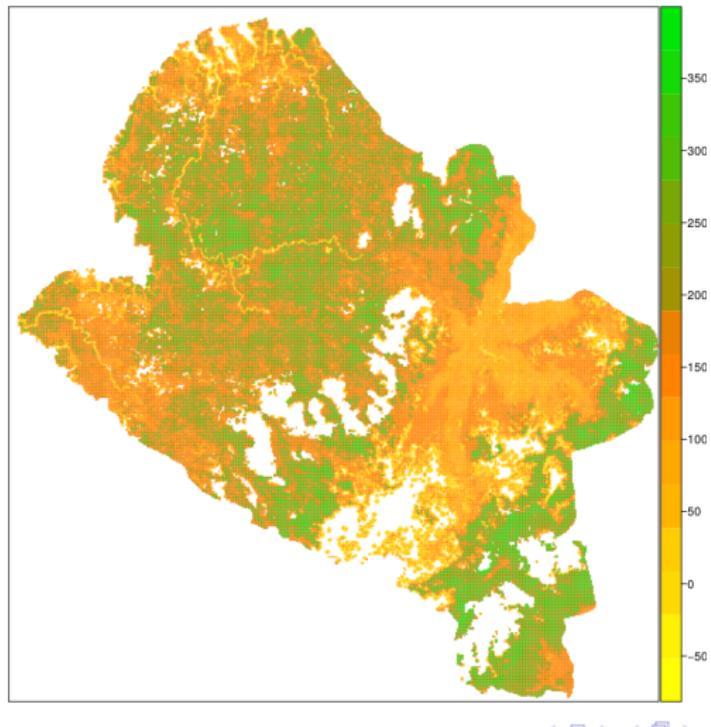
Landsat ETM+

- Combinación, recorte, cálculo de reflectancia y eliminación de píxeles nublados.



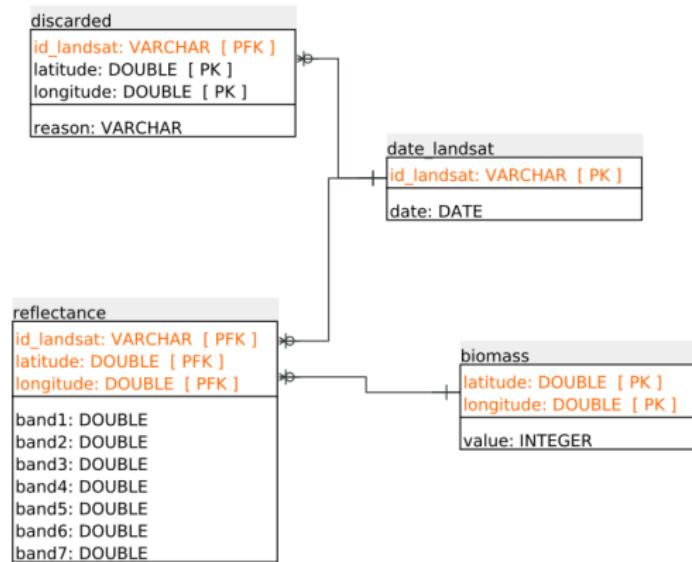
The Woods Hole Research Center

- Clip y eliminación de píxeles nulos



Gestión en bases de datos relacional

- Esquema de la base de datos procesada...



Selección de técnicas de regresión

- Conjunto de datos...

- 1321 imágenes procesadas.
- 15 años (2000 – 2014).
- 51'076.512 puntos disponibles.
- 4'071.185 puntos con mediciones válidas (no nublados).
- 110.776 puntos entre los años 2000-03 y con al menos 4 mediciones disponibles que se pueden asociar a un valor nominal de biomasa.

Selección de técnicas de regresión

- Conjunto de datos...

latitude double precision	longitude double precision	band1 double precision	band2 double precision	band3 double precision	band4 double precision	bands double precision	band6 double precision	band7 double precision	bio double precision
-8722800	247050	0.110937746345502	0.0875227658094195	0.0556337494328216	0.33029247542471	0.134253622827108	296.598717092975	0.0439186339886883	254
-8704350	105300	0.0967749840594034	0.085948010988303	0.0572605753238298	0.321224804610565	0.150461495271173	295.946567320797	0.0557410515933595	168
-8704350	106650	0.0993748880848757	0.0895403252217237	0.0630206501421573	0.307954216462809	0.16214589756336	297.327673875994	0.0631214121826714	173
-8704350	107100	0.0952935184949335	0.076991162562075	0.0492759134986458	0.300531123663615	0.127141196939211	295.189411906665	0.0441264420814832	179
-8704350	108000	0.0962442654219356	0.0858564249088067	0.05609319040898	0.334557951148325	0.142164438055948	296.511854384231	0.0497366530298198	198
-8704350	108450	0.094820730393922	0.0747968021115459	0.0493442879668408	0.270454932106202	0.115572099941545	295.179374825633	0.040582189824492	198
-8704350	108900	0.09771196802748732	0.082511508747281	0.0542984726848289	0.336980102105092	0.153876466665768	295.211691801885	0.05498436065704398	218
-8703900	105300	0.0906218618344076	0.0718891187714392	0.0463693674466194	0.246640497058423	0.0993145652817392	294.140032241724	0.0357541763067319	166
-8703900	105750	0.0983272152287389	0.0887660351898503	0.0577130293405062	0.340376502078177	0.150859547388612	294.94209758228	0.056397558576962	178
-8703900	106200	0.0948581564172237	0.0797555169512569	0.0512201246391938	0.325409534896812	0.130672634906235	294.691774134883	0.0456230664918918	185
-8703900	106650	0.097774305066167	0.0871658399740855	0.0593463947026036	0.311078329715316	0.132469497835564	295.359112230295	0.0486500946835153	204
-8703900	107100	0.0913114916172413	0.0739439490724913	0.0474426048817328	0.26136192310183	0.11324189856262	294.363006306459	0.0412751790631618	210
-8703900	107550	0.0925818640461121	0.0747063690400352	0.048628222426379	0.268398763394159	0.113046733754567	294.467300459837	0.0400754177935147	181
-8703900	108000	0.0958883300380165	0.0815729121283403	0.0534706974402068	0.333018425714928	0.14474620212802	295.843465247976	0.0532878034064126	192
-8703450	105300	0.0958393948746795	0.0832557535478574	0.0552854491016998	0.345099607594272	0.158444532443521	295.98844045644	0.0591904302573667	162
-8703450	105750	0.0929254311847555	0.07690217572268802	0.048820867746753	0.309881197690928	0.120018646602964	293.565716345121	0.0433772534346381	192
-8703450	106200	0.0956063738930535	0.080618383147213	0.05335901959775701	0.315742949163939	0.129438632697063	295.031431871104	0.0486050503317672	220
-8703450	107550	0.0894969048551368	0.0830394916840209	0.0529779066664658	0.377919703081002	0.15862281873806	295.44951849403	0.056215539391293	214

Selección de técnicas de regresión

- Paquete rminer [Cortez, 2010], Cortez [2015]

Algoritmos

- ctree – conditional inference tree
- rpart – decision tree
- kknn – k-nearest neighbor
- mlp – multilayer perceptron with one hidden layer
- mlpe – multilayer perceptron ensemble
- ksvm – support vector machine
- randomForest – random forest algorithm
- mr – multiple regression
- mars – multivariate adaptive regression splines
- cubist – M5 rule-based model
- pcr – principal component regression
- plsr – partial least squares regression
- cppls – canonical powered partial least squares

Selección de técnicas de regresión

- Paquete rminer [Cortez, 2010], Cortez [2015]

Métricas de validación

- SAE – sum absolute error/deviation (regression, "<", [0,Inf[)).
- MAE – mean absolute error (regression, "<", [0,Inf[)).
- RAE – relative absolute error (regression, "<", [0 %,Inf[)).
- RMSE – root mean squared error (regression, "<", [0,Inf[)).
- COR – correlation (regression, ">", [-1,1]).
- R2 – coefficient of determination R^2 (regression, ">", squared pearson correlation coefficient: [0,1]).

Selección de técnicas de regresión

- Resultados...

Disponible en <http://grias.udesar.edu.co:9191/resources/Energy/RAG.pdf>

1.5. Conjunto de 1009 datos y 35 muestras

	SAE	MAE	RAE	RMSE	COR	R2
crtree	10406.58225	30.88007	65.04650	40.02893	0.69401	0.48165
rpart	10197.95826	30.26100	63.74249	39.37592	0.70520	0.49730
kknn	9147.51425	27.14396	57.17667	36.86581	0.74955	0.56182
mlp	9000.95905	26.70908	56.26063	35.73037	0.76809	0.58996
mlpe	8791.27839	26.08688	54.95001	33.91572	0.79244	0.62796
ksvm	8463.32951	25.11374	52.90017	34.67996	0.79832	0.63731
randomForest	8931.29430	26.50236	55.82519	34.93439	0.77899	0.60683
mr	10591.61645	31.42913	66.20305	39.10858	0.71055	0.50488
mars	8842.91866	26.24011	55.27279	33.96852	0.79161	0.62665
cubist	9012.54150	26.74345	56.33302	35.70576	0.77611	0.60235
pcr	10337.63121	30.67546	64.61552	38.59290	0.72023	0.51873
plsr	10337.63121	30.67546	64.61552	38.59290	0.72023	0.51873
cppls	10337.63121	30.67546	64.61552	38.59290	0.72023	0.51873

Cuadro 5: Métricas de modelos analizados con 35 muestras

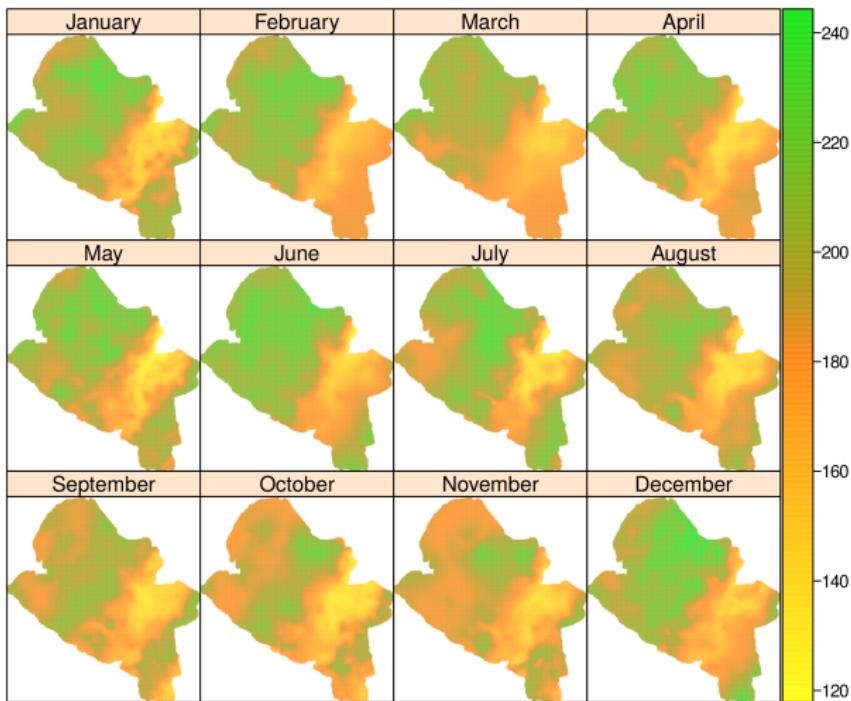
Implementación del modelo

- ① Aplicar el modelo obtenido a los puntos disponibles.
- ② Seleccionar una muestra representativa (~5000 muestras).
- ③ Aplicar técnicas de interpolación espacial (kriging [Bivand et al., 2013]) a una resolución de 450m.

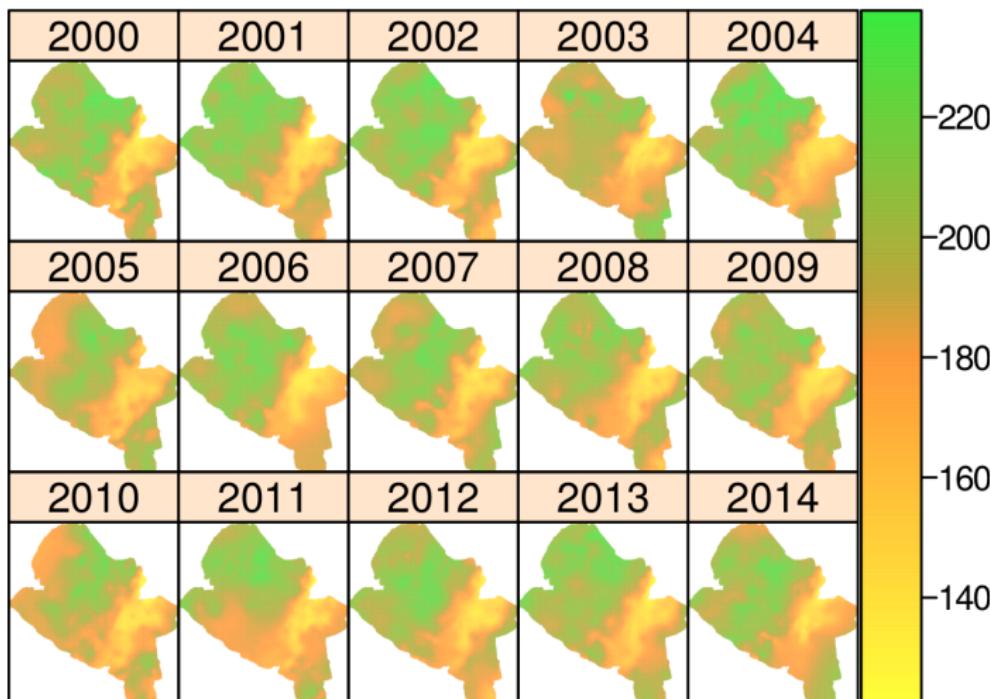
```
library(rgdal)
library(mapproj)
library(gstat)
library(sp)
library(geoR)

#Crear un modelo del variograma
mod<-vgm(psill=var(data$Value),model="Sph",range=sqrt(areaSpatialGrid(study.grid))/4,nugget=0)
#Ajustar el variograma
fit_ols<-fit.variogram(variogram(Value~1,data),model=mod,fit.method=6)
#Aplicar interpolación
map<-krige(Value~1,data,model=fit_ols,newdata=study.grid)
#Escribir el mapa a disco
writeGDAL(map["var1.pred"], fname=nameGeotiff, drivername = "GTiff", type = "Float32", mvFlag = 255)
```

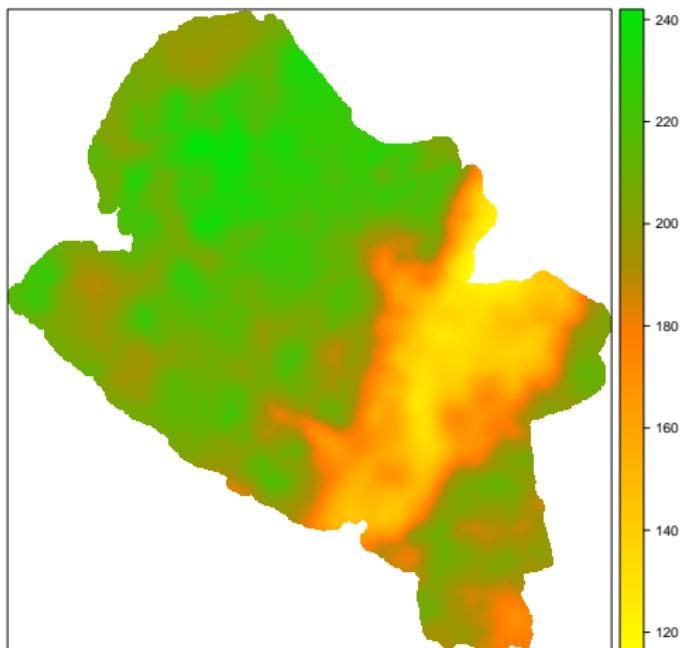
Visualización por meses



Visualización por años



Visualización General



También disponibles en

<http://geoalternar.udnar.edu.co/>

- A. Baccini, N. Laporte, S. J. Goetz, M. Sun, and H. Dong. A first map of tropical africa's above-ground biomass derived from satellite imagery. *Environmental Research Letters*, 3(4):045011, October 2008. ISSN 1748-9326. doi: 10.1088/1748-9326/3/4/045011. URL <http://iopscience.iop.org/1748-9326/3/4/045011>.
- A. Baccini, S. J. Goetz, W. S. Walker, N. T. Laporte, M. Sun, D. Sulla-Menashe, J. Hackler, P. S. A. Beck, R. Dubayah, M. A. Friedl, S. Samanta, and R. A. Houghton. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change*, 2(3):182–185, January 2012. ISSN 1758-678X, 1758-6798. doi: 10.1038/nclimate1354. URL <http://tiles.arcgis.com/tiles/tWF28TdwCm3ZaV7m/arcgis/rest/services/PantropicBiomassFull/MapServer>.
- Roger S. Bivand, Edzer Pebesma, and Virgilio Gómez-Rubio. *Applied Spatial Data Analysis with R*. Springer New York, New York, NY, 2013. ISBN 978-1-4614-7617-7 978-1-4614-7618-4. URL <http://link.springer.com/10.1007/978-1-4614-7618-4>.
- Paulo Cortez. Data mining with neural networks and support vector machines using the R/rminer tool. In *Advances in Data Mining. Applications and Theoretical Aspects*, pages 572–583. Springer, 2010. URL http://link.springer.com/chapter/10.1007/978-3-642-14400-4_44.
- Paulo Cortez. A tutorial on using the rminer R package for data mining tasks. report, Universidade do Minho. Escola de Engenharia, July 2015. URL <http://repositorium.sdum.uminho.pt/handle/1822/36210>.
- R.J. Hall, R.S. Skakun, E.J. Arsenault, and B.S. Case. Modeling forest stand structure attributes using landsat ETM+ data: Application to mapping of aboveground biomass and stand volume. *Forest Ecology and Management*, 225(1-3):378–390, April 2006. ISSN 03781127. doi: 10.1016/j.foreco.2006.01.014. URL <http://linkinghub.elsevier.com/retrieve/pii/S0378112706000235>

- Victor Klemas. Remote sensing of coastal wetland biomass: An overview. *Journal of Coastal Research*, 290:1016–1028, September 2013. ISSN 0749-0208, 1551-5036. doi: 10.2112/JCOASTRES-D-12-00237.1. URL
<http://jcronline.org/doi/abs/10.2112/JCOASTRES-D-12-00237.1>.
- E. T. A. Mitchard, S. S. Saatchi, I. H. Woodhouse, G. Nangendo, N. S. Ribeiro, M. Williams, C. M. Ryan, S. L. Lewis, T. R. Feldpausch, and P. Meir. Using satellite radar backscatter to predict above-ground woody biomass: A consistent relationship across four different african landscapes. *Geophysical Research Letters*, 36(23):n/a-n/a, 2009. ISSN 1944-8007. doi: 10.1029/2009GL040692. URL
<http://onlinelibrary.wiley.com/doi/10.1029/2009GL040692/abstract>.
- P. Muukkonen and J. Heiskanen. Biomass estimation over a large area based on standwise forest inventory data and ASTER and MODIS satellite data: A possibility to verify carbon inventories. *Remote Sensing of Environment*, 107(4):617–624, April 2007. ISSN 0034-4257. doi: 10.1016/j.rse.2006.10.011. URL
<http://www.sciencedirect.com/science/article/pii/S003442570600407X>.
- Scott L. Powell, Warren B. Cohen, Sean P. Healey, Robert E. Kennedy, Gretchen G. Moisen, Kenneth B. Pierce, and Janet L. Ohmann. Quantification of live aboveground forest biomass dynamics with landsat time-series and field inventory data: A comparison of empirical modeling approaches. *Remote Sensing of Environment*, 114(5):1053–1068, May 2010. ISSN 0034-4257. doi: 10.1016/j.rse.2009.12.018. URL
<http://www.sciencedirect.com/science/article/pii/S0034425709003745>.