

# Análisis de regresión para el cálculo de irradiación solar en el departamento de Nariño (Colombia) utilizando imágenes satelitales Landsat y MODIS

XXX

Universidad de Nariño  
San Juan de Pasto, Colombia  
Email: xxx@mail.com

XXX

Universidad de Nariño  
San Juan de Pasto, Colombia  
Email: xxx@mail.com

XXX

Universidad de Nariño  
San Juan de Pasto, Colombia  
Email: xxx@mail.com

XXX

Universidad de Nariño  
San Juan de Pasto, Colombia  
Email: xxx@mail.com

**Abstract**—En este artículo se describen los componentes básicos de la investigación aplicada que tiene como objetivo construir un primer mapa del potencial de irradiancia solar en el departamento de Nariño (Colombia) a partir de imágenes satelitales Landsat y MODIS de libre acceso. Se analizan las diferentes bandas de las imágenes satelitales disponibles y su relación con bases de datos previas de irradiancia solar proporcionadas por 3TIER Inc. una compañía de Vaisala aplicando diferentes técnicas de regresión y obteniendo un modelo para la generación de mapas actualizados en el departamento de Nariño.

**Keywords**—*irradiance, solar, regression models analysis, Landsat, MODIS*

## I. INTRODUCCIÓN

Países como Alemania, España, Japón, Estados Unidos, Italia son las potencias mundiales en la generación de energía renovable mediante el uso de energía solar, este tipo de fuente de energía limpia ha adquirido gran importancia en el entorno investigativo en los últimos años y han surgido grandes expectativas para el futuro. Según estudios del Ministerio de Minas y Energía de Colombia, en el departamento de Nariño hay 15 municipios con cobertura eléctrica inferior al 80% [?]. Como nueva estrategia para enfrentar esta problemática se ha planteado la medición y estimación de potenciales energéticos en las zonas más viables de la región. Uno de los componentes a analizar es el potencial de biomasa para la generación eléctrica. Sin embargo, uno de los problemas que se plantea para la ubicación de lugares propicios es la ausencia de bases de datos actualizadas en el área de estudio que permitan su respectivo análisis.

Actualmente la revolución de fuentes energéticas y los cambios acelerados de las condiciones climáticas obliga a las naciones a plantear un cambio a los enfoques de generación de energía, para esta serie de cambios se requiere un análisis de todas y cada una de las fuentes de energía renovables; en algunas situaciones es pertinente resaltar las condiciones para optimizar aprovechamiento del potencial energético.

“Cada año el sol arroja 4000 veces más energía que la que se consume, lo que demuestra que esta fuente energética está aún infravalorada y desaprovechada en

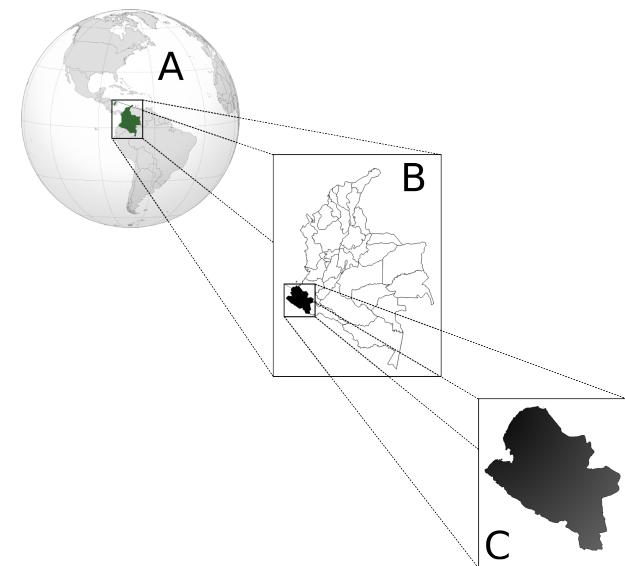


Figure 1. Localización área de estudio

relación a sus potenciales energéticos” [?]; la implementación para esta fuente energética es posible si un territorio cuenta una buena posición geográfica, economía sostenible y fenómenos climáticos regulados, estas condiciones favorables para la energía a base de paneles solares o energía térmica se encuentran presentes en muchas regiones de América Central y América del Sur, como ventaja adicional la energía solar puede aprovecharse directamente o almacenarse para un consumo posterior y no tiene riesgo de agotarse.

Este artículo está orientado a cumplir con los requerimientos necesarios para la generación de un modelo de predicción de energía solar y su extrapolación al resto del área de estudio.

El área de estudio en la investigación fue el departamento de Nariño (Colombia) el cual está ubicado en el extremo sur occidental de Colombia, en la frontera con Ecuador con una extensión aproximada de 33.268 km, una población de 1,702 millones según el censo de 2013, su ubicación está en latitud 00° 31' 08" y 02° 41' 08" Norte, Longitud 76° 51' 19" y 79° 01' 34" Oeste.

## II. TRABAJOS RELACIONADOS

Dada la creciente demanda por la generación y apropiamiento de las energías limpias, el estudio de alternativas energéticas basadas en paneles solares también ha venido en aumento. Sin embargo, un componente clave para la implementación de una solución de este tipo es analizar de antemano las posibles ubicaciones con mayor potencial de generación eléctrica a base de radiación solar. Diversos estudios han consolidado el uso de imágenes satelitales de libre acceso como herramientas fundamentales para este propósito. [?] realizó la construcción de mapas solares de alta resolución comparando metodologías basadas en imágenes satelitales y modelamiento climático. Para estos modelos de radiación se contemplan condiciones propias del terreno como la latitud, el terreno, estación, hora del día y condición atmosférica(nubes, polvo, polucion, vapor de agua y efectos de la montaña).

[?] desarrolló el método HELIOSAT con el objetivo de estimar los niveles de radiación solar a partir de imágenes satelitales geoestacionarias, el metodo se encuentra implementado en un algoritmo que permite separar la radiación de los componentes atmosférico de la radiación en nubes para finalmente obtener la radiación superficial.[?] realiza una recopilación de métodos para la predicción de radiación solar también basados en imágenes por satélite o modelos climáticos, en este estudio se contempla la variación de la resolución espacial de la imagen satelital, tipos de sensores utilizados y la escala temporal.

[?] obtuvo gran éxito al proponer un nuevo modelo basado en redes neuronales apropiado para pronosticar el potencial solar a corto plazo bajo condiciones meteorológicas en constante cambio. [?] señala dos enfoques para obtener radiación neta diaria utilizando los productos de Kalpana VHRR y Oceansat OCM2 mediante el manejo de la radiación de onda larga resultante del flujo de ondas entrantes y salientes, este resultado es computarizado usando la ecuación de Stefan Boltzmann para corregir la humedad y la nubosidad, también se realiza un enfoque basado en la estimación mediante la computación de un cielo despejado con el flujo de ondas cortas de radiación que permiten obtener una radiación neta

[?] y [?] generan información de cobertura de suelo, y los métodos que permiten obtener más detalle conservando una fiabilidad aceptable sobre la región del Tancítaro, Michoacán y comprende bosques templados y tropicales secos, pastizales y áreas de cultivos. Enfocan el estudio en índices de vegetación mediante compuestos espectrales de 8 días e imágenes de reflectancia diarias que fueron evaluados por medio de dos metodologías; la máxima verosimilitud y redes neuronales, en cada una de estas se incorporaron dos tipos de datos auxiliares. Los resultados muestran que es posible obtener mapas confiables a partir de estos datos de baja resolución si se usan categorías generales.

[?] destaca la correlación existente entre la radiación solar incidente en la superficie terrestre y la radiación solar reflejada al espacio; Adicionalmente realiza un estudio comparativo en el análisis de imágenes satelitales MODIS respecto a imágenes satelitales provenientes de SRB, Earth

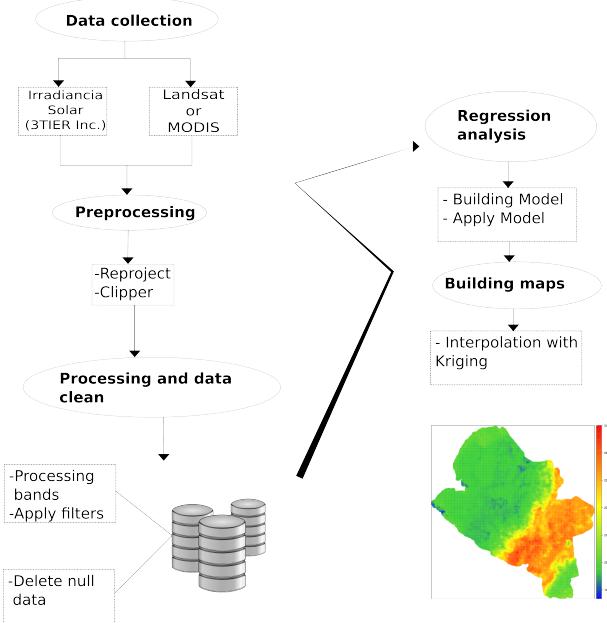


Figure 2. Metodología

Observing System (EOS) y Tropical Rainfall Measuring Mission (TRMM), en este estudio se destaca los métodos para estimar la radiación neta obtenida de ondas cortas medidas en la superficie.

[?] prueban diferentes algoritmos para la detección de aerosoles sobre la superficie terrestre contempladas dentro de imágenes satelitales MODIS, implementación del algoritmo Miller sobre las bandas de infrarrojo y adicionalmente se realiza una clasificación de los aerosoles detectado. [?] calcula la caída de potencial y fluctuaciones en la generación de energía a partir del seguimiento tridimensional de nubes y su impacto sobre paneles solares.

## III. METODOLOGÍA

La metodología usada para la construcción del mapa de radiación solar se la puede ver en la figura 2

### A. Obtención de datos

Se realiza el estudio de las imágenes satelitales más pertinentes para obtener la información, para la elección se tiene en cuenta 4 factores importantes *resolución espacial, resolución espectral, resolución radiométrica y resolución temporal* de la imagen satelital; Las imágenes de LandSat 7 son generadas cada 16 días y permiten adquirir datos de 8 bandas espectrales con una resolución de 15 x 15 metros por pixel para la banda 6, y 30 x 30 metros por pixel para las 7 bandas restantes, estas bandas tienen la capacidad de detectar diferentes indicadores para vegetación, reflectancia, temperatura, precipitación, nubosidad, etc. Por otra parte, las imágenes satelitales MODIS adquieren datos de 36 bandas espectrales con las cuales ofrece el producto MOD09GA con una resolución espacial aproximada de 500 x 500 metros en cada pixel, estos productos son generados diariamente y están diseñado para medir la reflectancia

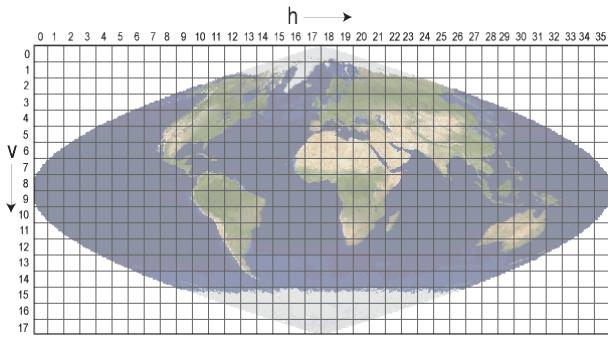


Figure 3. Grilla senoidal de MODIS

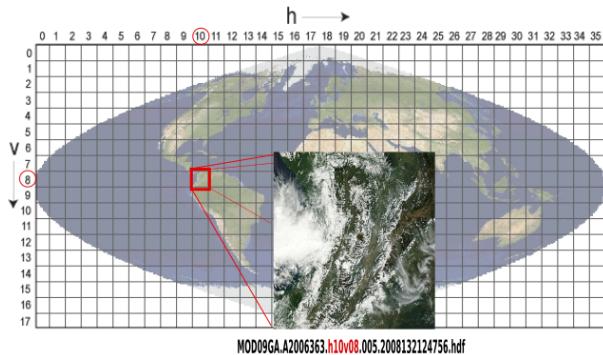
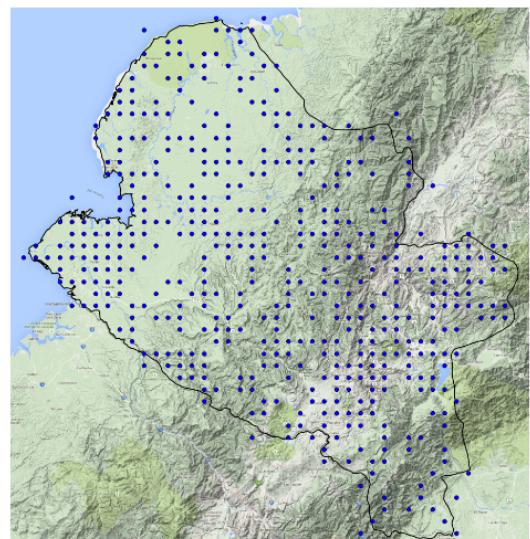


Figure 4. Ubicación en grilla senoidal de MODIS

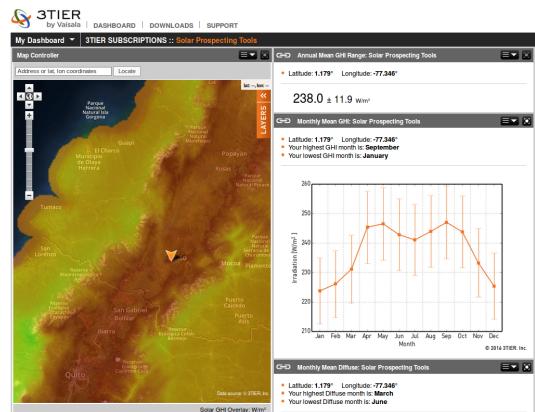
de la superficie terrestre[?][?], MOD09GA presenta 7 bandas que tienen relación directa con la delimitación de territorios, tipos de vegetación, incidencia de aerosoles, temperatura y reflectancia la cual está relacionada con la propiedad reflectiva de la vegetación y los aerosoles; mediante la propiedad reflectiva de la vegetación se pretende realizar una estimación para la radiación solar que se irá diría en una superficie.

Se descargaron 1362 imágenes satelitales de Landsat 7 desde el año 1999 hasta el año 2015, que cubren el departamento de Nariño, para cubrir todo el departamento fue necesario descargar las imágenes satelitales con los siguientes paths y rows: (009,059), (009,060), (010,058), (010,059), (011,059). Las imágenes satelitales de MODIS están ubicadas en una grilla senoidal de aproximadamente 10x10 grados cada sección como lo muestra la figura 3, esta grilla cuenta con un índice vertical y un índice horizontal que permiten ubicar la imagen satelital en un área determinada; el territorio colombiano se encuentra ubicado en la columna 10 (h10) en la fila 8 (v8) como lo muestra la figura 4, estos dos índices son necesarios para realizar la descarga de todas las imágenes satelitales comprendidas entre el año 2005 y 2015, en estos 11 años se descargaron 3912 imágenes satelitales.

En la obtención de datos también se utilizó el mapa de irradiancia solar proporcionado por VAISALA INC (3TIER), en el cual se tomaron 500 muestras bien distribuidas dentro del departamento de Nariño, como lo muestra la figura 5.



(a) Muestras Uniformemente distribuidas en el Departamento de Nariño



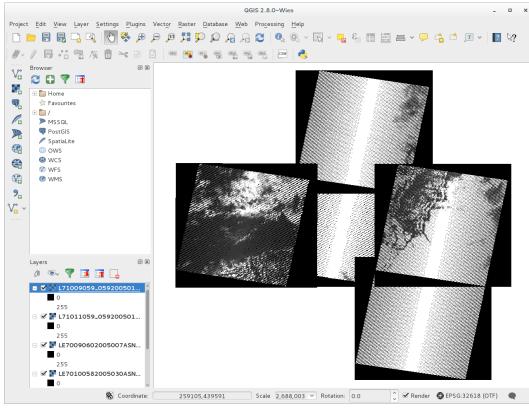
(b) Consulta de radiación en 3TIER

Figure 5. Muestreo de radiación solar

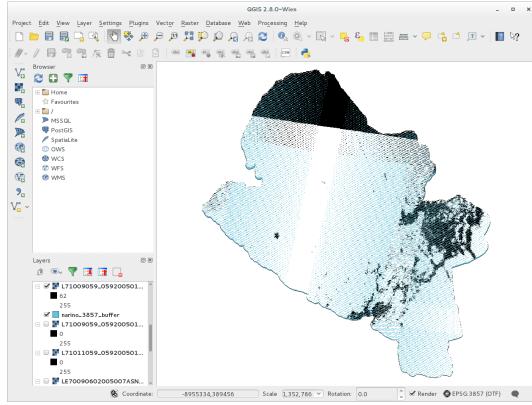
## B. Preprocesamiento

En la etapa de preprocesamiento para la imágenes satelitales Landsat 7, se reproyectó las imágenes obtenidas, debido a que las cinco imágenes que cubren el departamento de Nariño, están en distintos sistemas de coordenadas (EPSG:32618 y EPSG:32617) y se lo reproyectó al sistema EPSG:3857. Así como también se recortó las imágenes con el fin de únicamente tener el área que cubre el departamento de Nariño, como lo muestra la figura 6.

Para el caso de MODIS se convirtió el formato científico original de la imagen satelital(HDF) en un formato de imágenes de mapa de bits (TIFF) reprojectando el sistema de coordenadas espaciales original de EPSG:4326 a EPSG:3857. Posterior a la conversión de formato y reproyección, se recortaron las imágenes satelitales con el fin únicamente de tener el área que cubre el departamento de Nariño.



(a) Imágenes Satelitales de Nariño



(b) Imágenes recortadas de Nariño

Figure 6. Preprocesamiento

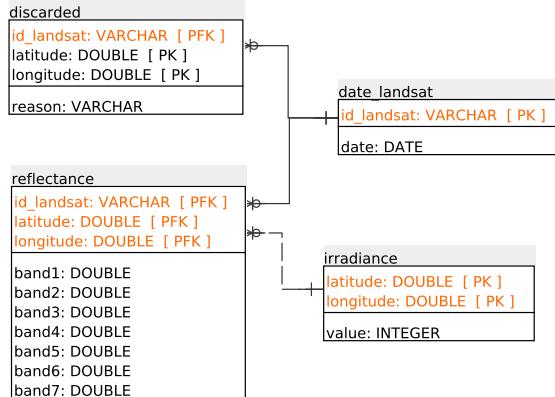


Figure 7. Modelo entidad-relacion Landsat

### C. Procesamiento y limpieza de datos

Se diseñaron dos base de datos, una para el sensor Landsat 7 y otra para el sensor MODIS.

El diseño de la base de datos para Landsat 7 lo muestra la figura 7, la cual tiene 4 tablas.

Tabla date\_landsat: en la cual se almacenan las fechas de las imágenes satelitales.

Table I. DATOS OBTENIDOS EN EL PROCESAMIENTO Y LIMPIEZA DE DATOS LANDSAT 7

Nombre	Valor	Detalle
Imágenes procesadas	1321	Imágenes de Nariño de 2000 a 2014
Datos Totales	51.076.512	Registros Totales desde año 2000 a 2014
Nube caliente	3.731.768	Registros de 2000 a 2014
Nube fria	27.827.009	Registros de 2000 a 2014
Ambiguo	11.987.340	Registros de 2000 a 2014
Datos Validos Reflectance	4.071.185	Registros de 2000 a 2014

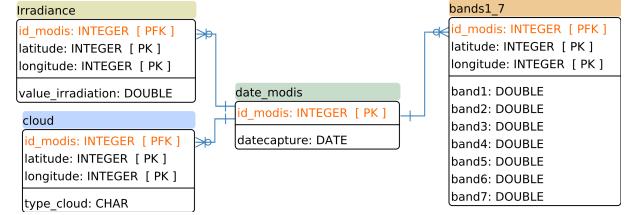


Figure 8. Modelo entidad-relacion MODIS

Tabla reflectance: en la cual se almacenan los datos capturados y convertidos en reflejante, de las bandas landsat (1 - 5,7) y la temperatura en grados kelvin de la banda 6.

Tabla discarded: en la cual se almacenan datos que fueron descartados, por varias razones, son nubes calientes, nubes frias o ambiguas.

Tabla irradiance: en la cual se almacenan los datos de irradiance muestreados de 3TIER.

Para procesar las imágenes y llenar la base de datos se realizó un Script, el cual captura el Digital Number de las imágenes satelitales y lo transforma en valor en reflectance. En este procesamiento de imágenes, se adicionó al Script unos filtros para para detección de nubes calientes, nubes frias, datos ambiguos como lo muestra el algoritmo propuesto por [?].

La tabla I muestra la relación de los datos obtenidos en este proceso.

El diseño de la base de datos para MODIS lo muestra la figura 8, la cual tiene 4 tablas.

Tabla date\_modis: en la cual se almacenan las fechas de las imágenes satelitales.

Tabla bands1\_7: se almacenan los valores de las bandas 1 a 7.

Tabla cloud: se almacenan las coordenadas donde hay nubes.

Tabla irradiance: en la cual se almacenan los datos de irradiance muestreados de 3TIER.

Para procesar las imágenes y llenar la base de datos se realizó un Script, el cual recorre las 7 bandas en formato TIFF obtenidas del producto MOD09GA, se almacenan datos cada 450 metros y se aplica el filtro de nubes según las especificaciones de las bandas [?].

La tabla II muestra la relación de los datos obtenidos en este proceso.

Table II. DATOS OBTENIDOS EN EL PROCESAMIENTO Y LIMPIEZA DE DATOS MODIS

Nombre	Valor	Detalle
Imágenes procesadas	3912	Imágenes de Nariño de 2005 a 2015
Datos Totales	565.722.468	Registros Totales desde año 2005 a 2015
Nubes	192.051.992	Registros de 2005 a 2015
Nube tipo 1	160.318.600	Registros de 2005 a 2015
Nube tipo 2	31.733.392	Registros de 2005 a 2015
Datos Válidos	373.670.476	Registros de 2005 a 2015

Table III. MÉTRICAS DE MODELOS ANALIZADOS CON DATOS DEL SENSOR LANDSAT 7

	SAE	MAE	RAE	RMSE	COR	R2
crtree	771.66185	5.32181	33.37397	8.89544	0.85991	0.73945
rpart	819.97501	5.65500	35.4634	9.23174	0.84774	0.71865
kknn	583.3615	4.02318	25.23008	6.19161	0.93584	0.87580
mlp	558.43603	3.85128	24.15206	5.49114	0.94968	0.90189
mlpe	<b>461.93253</b>	<b>3.18574</b>	<b>19.97834</b>	<b>4.73616</b>	<b>0.96292</b>	<b>0.92721</b>
ksvm	574.76656	3.96391	24.85835	5.71528	0.94664	0.89613
randomForest	663.70528	4.57728	28.70490	6.89480	0.92117	0.84856
mr	752.19550	5.18756	32.53206	6.75745	0.92222	0.85049
mars	680.67053	4.69428	29.43864	6.34212	0.93186	0.86837
cubist	538.20590	3.71176	23.27712	6.34056	0.93141	0.86752
pls	748.89239	5.16478	32.38920	6.76538	0.92208	0.85023
cppls	748.89239	5.16478	32.38920	6.76538	0.92208	0.85023

#### D. Análisis de regresión

El análisis de regresión se realizó tomando los valores agrupados de las bandas del sensor landsat 7 entre el año 2000 a 2014 y el valor de irradiancia solar que fue muestreada en 3TIER, para obtener el mejor modelo se evaluó 13 modelos y 6 métricas, de la misma manera este proceso se lo realizó para los datos del sensor MODIS entre el año 2005 a 2015.

Para encontrar el mejor modelo se utilizó la biblioteca de código abierto rminer presentada por [?] para la herramienta R.

En la tabla III se muestra las métricas de los modelos analizados y el mejor modelo escogido para los datos obtenidos del sensor landsat 7 y en la tabla IV se muestran para el sensor MODIS.

Según las métricas SAE, MAE, RAE, RMSE, COR y  $R^2$  evaluadas en los 13 modelos, se encuentra que el modelo más óptimo para extraer los datos de irradiación es “mlpe” (multilayer perceptron ensemble) este modelo se aplica a todos los datos obtenidos y se encuentra el valor de irradiación.

#### E. Construcción de mapas

Para la construcción de mapas del potencial solar se utilizó el método Kriging que provee una solución al problema de la estimación basada en un modelo continuo de variación espacial estocástica, el objetivo de Kriging es el de estimar el valor de una variable aleatoria, Z, en uno o más puntos no muestreados o sobre grandes bloques.

Table IV. MÉTRICAS DE MODELOS ANALIZADOS CON DATOS DE MODIS

	SAE	MAE	RAE	RMSE	COR	R2
crtree	1360.88975	9.13349	60.07076	13.06545	0.64264	0.41299
rpart	1373.21922	9.21624	60.61499	14.14182	0.60074	0.36089
kknn	920.48535	6.17775	40.63096	10.31934	0.79280	0.62854
mlp	485.60361	3.25908	21.43493	4.51435	0.96284	0.92705
mlpe	<b>443.11836</b>	<b>2.97395</b>	<b>19.55960</b>	<b>3.97730</b>	<b>0.97157</b>	<b>0.94394</b>
ksvm	823.63424	5.52775	36.35587	8.12712	0.87861	0.77195
randomForest	1159.07988	7.77906	51.16271	11.04659	0.75129	0.56444
mr	597.58282	4.01062	26.37778	5.36391	0.94694	0.89670
mars	618.82261	4.15317	27.31532	5.47634	0.94475	0.89255
cubist	597.29188	4.00867	26.36494	5.36508	0.94688	0.89658
pls	597.29188	4.00867	26.36494	5.36508	0.94688	0.89658
cppls	597.29188	4.00867	26.36494	5.36508	0.94688	0.89658

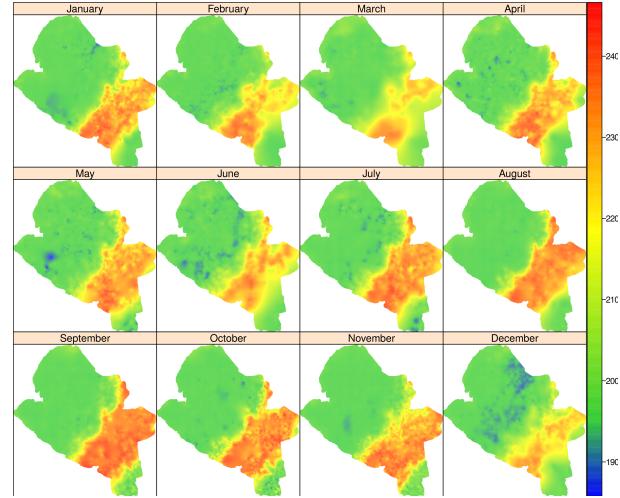


Figure 9. Mapas solar por meses datos landsat

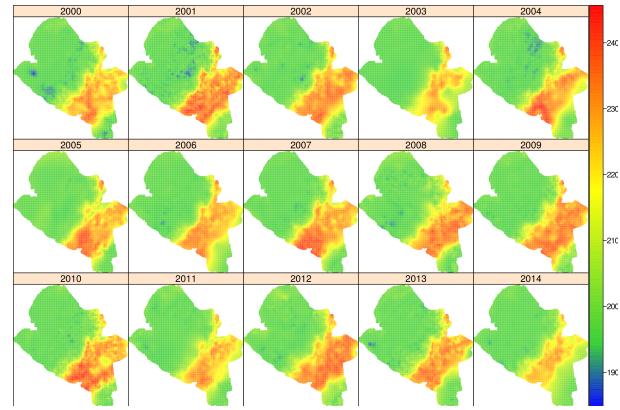


Figure 10. Mapas solar por años datos landsat

El método Kriging recibe como entrada datos de la muestra, y una malla dependiendo de la resolución que se quiera obtener, por ello los datos de muestra se obtuvieron aplicando el modelo obtenido en el análisis de regresión a datos agrupados en cada punto por mes, año y uno general entre el año 2000 a 2014 para el sensor landsat 7 y 2005 a 2015 para el sensor MODIS; y se construyó una malla con puntos regulares espaciados cada 450 metros.

En las figuras 9, 10 y 11 se muestra los mapas obtenidos por meses, años y general entre el año 2000 a 2014 respectivamente con datos del sensor landsat 7.

En las figuras 12, 13 y 14 se muestra los mapas obtenidos por meses, años y general entre el año 2005 a 2015 respectivamente con datos del sensor MODIS.

#### F. Análisis entre mapas del sensor landsat 7 y MODIS

Se realizó un análisis de correlación y coeficiente de determinación entre los mapas resultantes con los datos de Landsat7 y MODIS. Se tomo los mapas por años del 2005 al 2015, por meses y el mapa general como lo muestra la tabla V

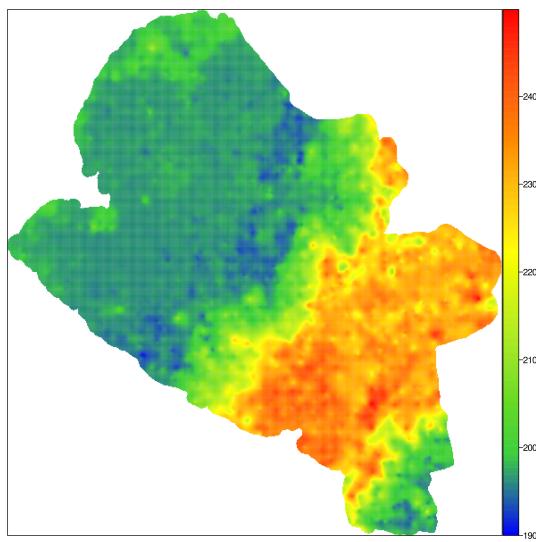


Figure 11. Mapas solar general años 2000-2014 datos landsat

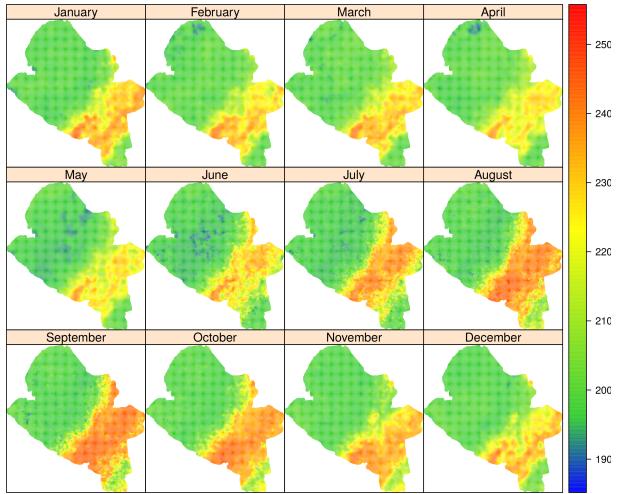


Figure 12. Mapas solar por meses datos MODIS

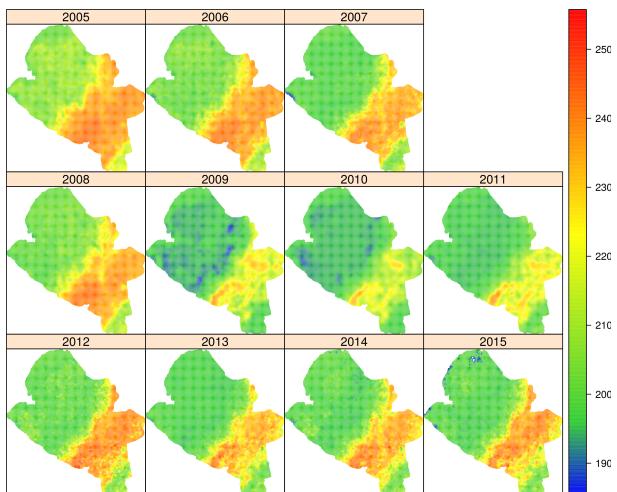


Figure 13. Mapas solar por años datos MODIS

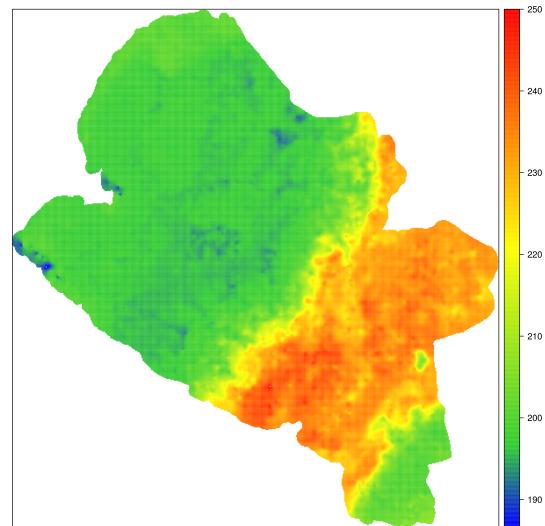


Figure 14. Mapas solar general años 2005-2015 datos MODIS

Table V. MÉTRICAS DE MODELOS ANALIZADOS CON DATOS DE MODIS

Map	COR	R2
2005	0.85594	0.73263
2006	0.90754	0.82363
2007	0.92884	0.86275
2008	0.90594	0.82073
2009	0.89443	0.80000
2010	0.88594	0.78488
2011	0.88233	0.77851
2012	0.92056	0.84743
2013	0.92982	0.86456
2014	0.93167	0.86800
January	0.89182	0.79535
February	0.89131	0.79444
March	0.90943	0.82706
April	0.90930	0.82683
May	0.91770	0.84218
June	0.90244	0.81439
July	0.90434	0.81783
August	0.91527	0.83772
September	0.91810	0.84290
October	0.92340	0.85266
November	0.87242	0.76111
December	0.86934	0.75576
General	0.94179	0.88696

#### IV. CONCLUSIONES

Se construyeron mapas energéticos con el componente solar en el departamento de Nariño, usando los sensores Landsat 7 y MODIS.

Las imágenes satelitales son una gran fuente de información debido a la capacidad de almacenar gran cantidad de registros históricos para diferentes tipos de datos, estos datos poco a poco están siendo utilizados por organizaciones para determinar características terrestres, fenómenos naturales, condiciones de los mares, características de la vegetación, etc. Por esta razón el uso de imágenes satelitales en la investigación da resultados aproximados y a bajo costo, teniendo en cuenta el costo que puede implicar hacer muestreo en campo.

Se construye una metodología para la construcción de mapas energéticos con el potencial de radiación solar, el cual se lo puede aplicar en zonas donde no tengan

estaciones climáticas con los sensores Landsat o MODIS.

En la construcción del mejor modelo, con ambos sensores Landsat 7 y MODIS se tuvo un  $R^2$  por encima del 90%, teniendo en cuenta que las imágenes landsat 7 se obtienen cada 16 días y las de MODIS son diarias, sería de mayor provecho usar el sensor MODIS para hacer estudios posteriores con series de tiempo.

Tanto la correlación como el  $R^2$  entre los mapas construidos a partir del sensor Landsat 7 y MODIS es buena, teniendo en cuenta que en el mapa general se tiene una correlación del 94% y un  $R^2$  del 88%.

#### AGRADECIMIENTOS

Al Sistemas General de Regalías por haber financiado esta investigación en el marco de proyecto “Análisis de Oportunidades Energéticas con Fuentes Alternativas en el Departamento de Nariño” ejecutado por el programa de Ingeniería Electrónica de la Universidad de Nariño.