

CLASIFICACIÓN DIURNA DE LA ALTURA DE TOPE DE NUBES EN BASE A IMÁGENES GOES 16 EN TIEMPO REAL

Ernesto Marchesoni Seijo - Rodrigo Alonso Suarez - Matías Di Martino

Buenos Aires, 8 de noviembre 2022



Introducción

- El Laboratorio de Energía Solar de la UDELAR tiene como objetivo lograr un conocimiento detallado del recurso solar y del funcionamiento de aplicaciones de energía solar en el territorio de Uruguay y la región.
- Este trabajo se centra en el desarrollo de una serie de clasificadores automáticos de nubosidad en base a información satelital operativa, con el objetivo de discriminar la altura de tope de nube en tres macro categorías (altas, medias y bajas).

<i>Categoría</i>	<i>Descripción</i>
0 - Sin nubosidad	El píxel no supera en, por lo menos, un 40% al Brillo de Fondo característico o la altitud del tope es menor a 500 metros.
1 - Baja	Topes mayores a 500 y menores a 2000 metros
2 - Media	Topes mayores a 2000 y menores a 6000 metros
3 - Alta	Topes mayores a 6000 metros, sin límite superior.

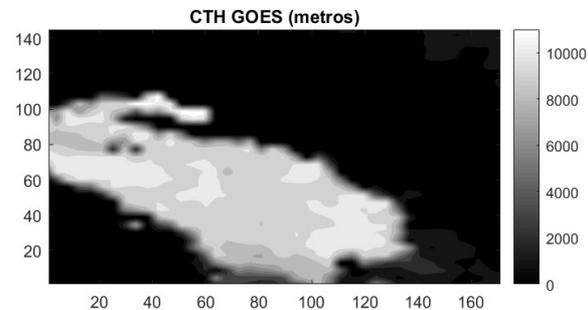
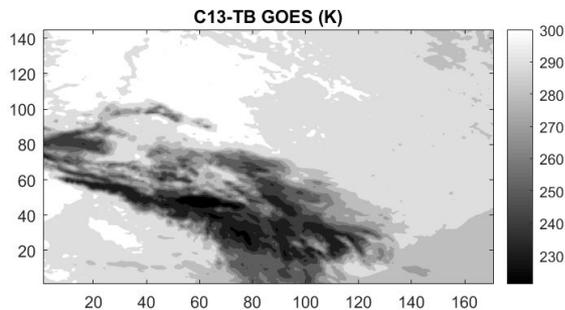
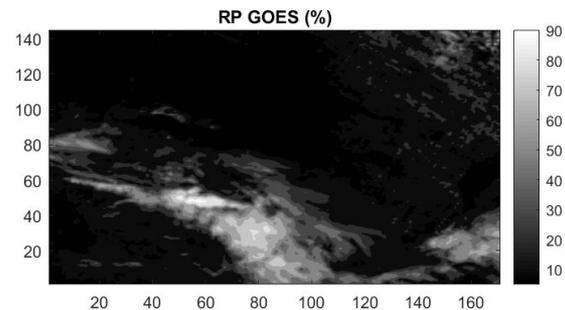
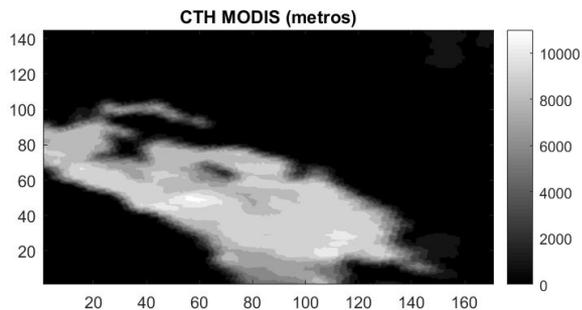
- El objetivo de la herramienta es asistir en el diagnóstico de predicciones operativas de radiación solar.

Desarrollo de los Clasificadores

- Se basan en técnicas empíricas simples y de aprendizaje automático, tales como Árboles de Decisión o Redes Neuronales.
- Se utilizan distintas propiedades analizadas y derivadas de la nubosidad vista por satélite, como su reflectividad, movimiento, temperatura de brillo, textura, gradientes, entre otras.
- Para entrenarlos, fue necesario confeccionar una base de datos etiquetada. La misma reúne la información a utilizar por los clasificadores, junto a una clasificación de nubosidad derivada de la altitud de nubosidad que provee el satélite TERRA con su espectrorradiómetro MODIS (CTH - MODIS), que sirve como verdad de referencia para el entrenamiento y validación de los clasificadores.
- A su vez, la base de datos incorpora una clasificación basada en el producto Cloud Top Height de GOES 16 (CTH - GOES16). Esta información es utilizada como base de desempeño para evaluar al clasificador generado.

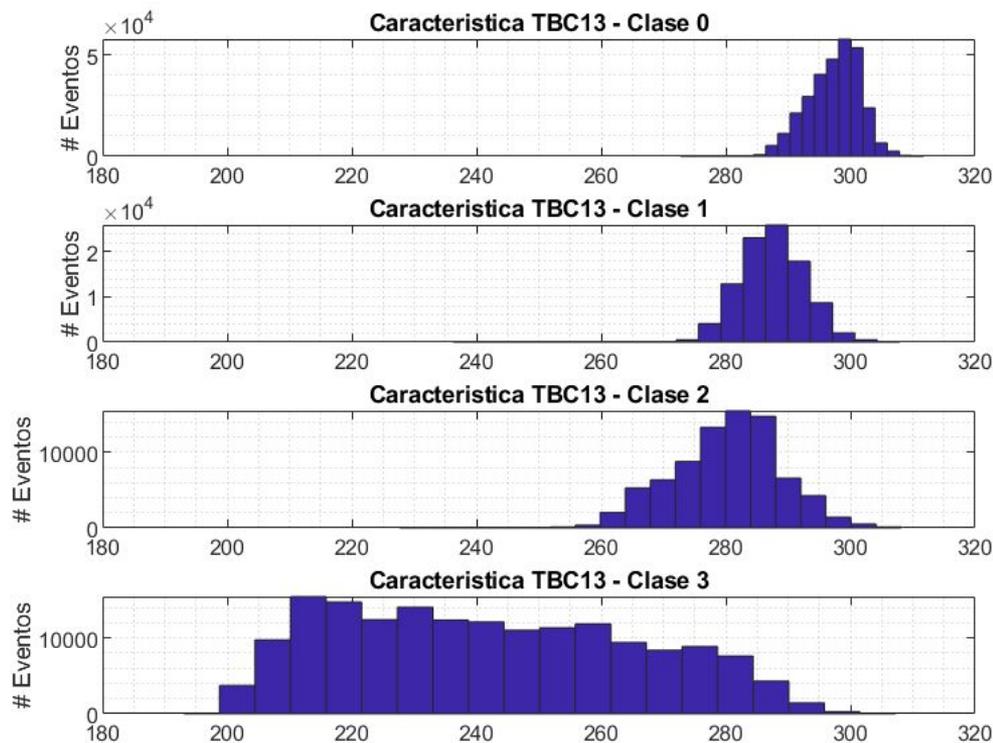
Ejemplos de exploración de los datos utilizados

1.



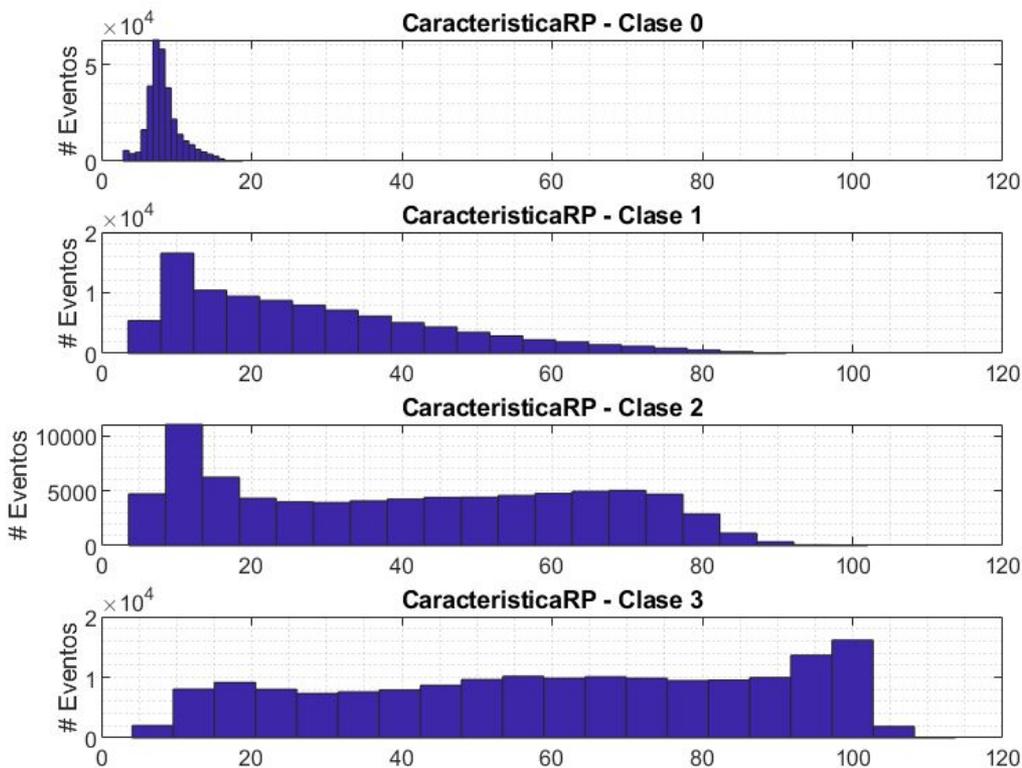
Ejemplos de exploración de los datos utilizados

2.



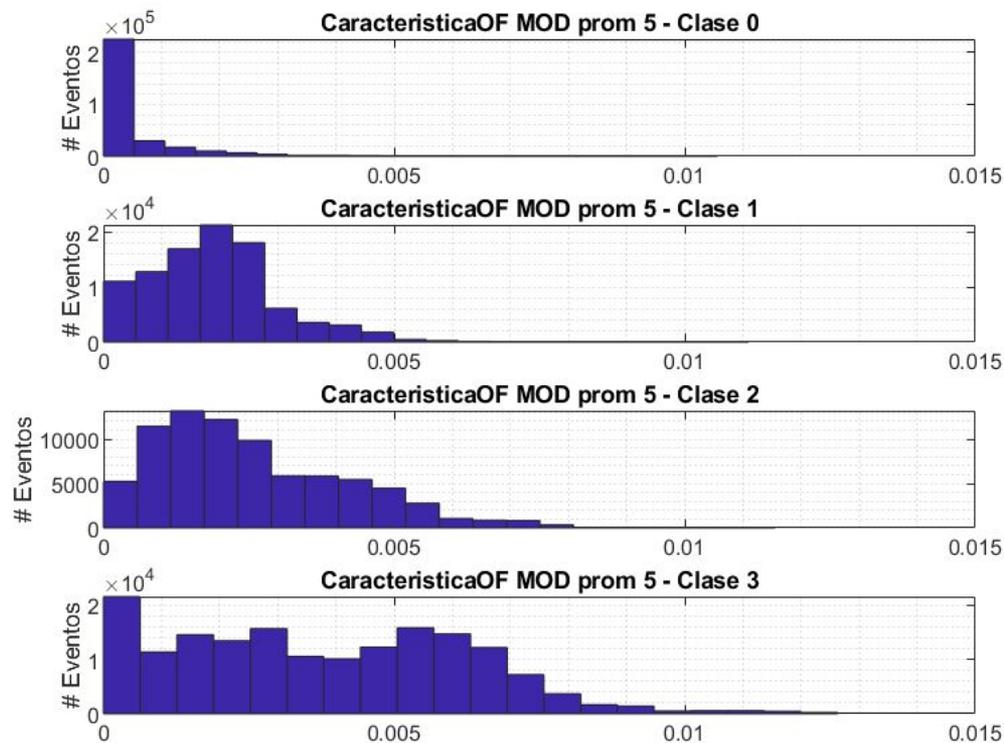
Ejemplos de exploración de los datos utilizados

3.



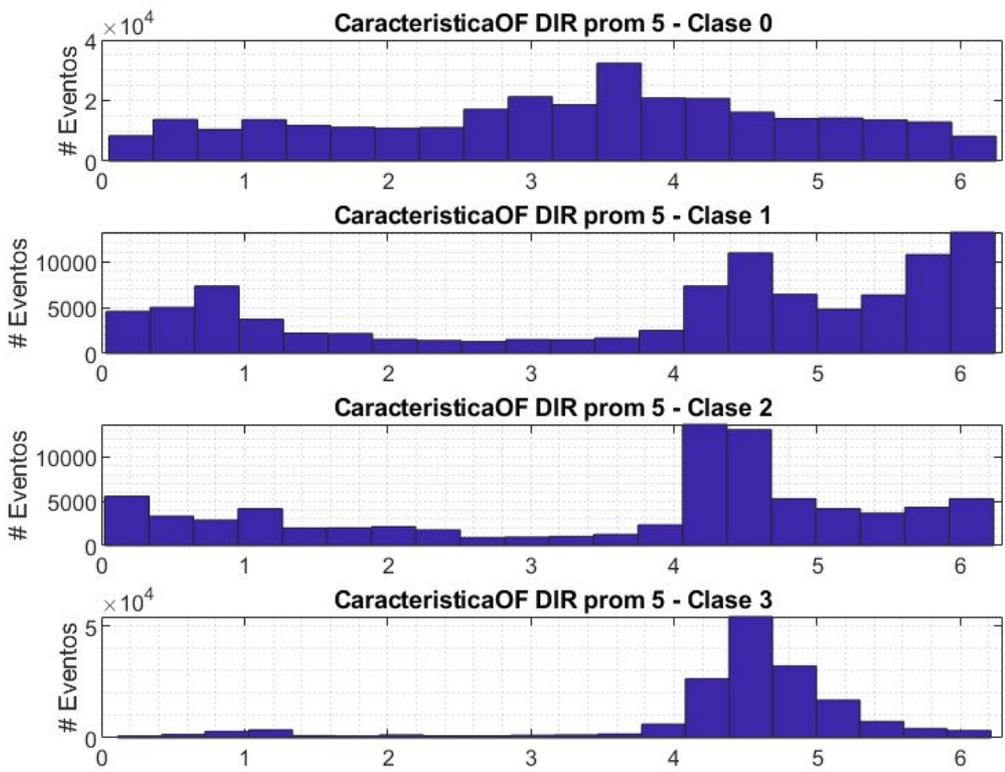
Ejemplos de exploración de los datos utilizados

4.



Ejemplos de exploración de los datos utilizados

5.



Base de Datos de Trabajo

- Se logró reunir información correlativa de 55 instantes en grillas de 145 x 171 píxeles, lo que totaliza 1363725 entradas.
- La información reunida se conforma por los datos primarios de los canales 2 (Visible), 13, 14, 15 y 16 (Infrarrojos) y de cálculos derivados de estos canales, buscando representar distintas propiedades de la imagen.
- Para la estimación de movimiento, se aplicó la técnica de flujo óptico de Farneback et al. (2003) específicamente ajustada para la región y este tipo de imágenes en Aicardi et al. (2022)
- Se consideraron un total de 29 variables diferentes que caracterizan a cada píxel, además de los correspondientes valores de CTH-GOES16 y CTH-MODIS.

Descripción de los Clasificadores Implementados

- Se basan en tres tecnologías de aprendizaje automático que operan sobre la base de datos: un árbol de decisión simple (DTC), árboles de decisión aleatorios (RNDF) y una red neuronal simple (MLP).

<i>Clasificador</i>	<i>Abreviación</i>	<i>Tecnología</i>
<i>Decision Tree Classifier</i>	<i>DTC</i>	<i>Árbol de Decisión</i>
<i>Random Forest</i>	<i>RNDF</i>	<i>Árboles de Decisión Aleatorios</i>
<i>Multilayer Perceptron</i>	<i>MLP</i>	<i>Red Neuronal</i>

Evaluación de Desempeño de los Clasificadores

- Se utilizó la técnica de Validación Cruzada (VC), reservando un 90% de las imágenes (49 imágenes) para entrenamiento y el 10% restante para evaluación (6 imágenes), se realizó 100 veces la selección de estos conjuntos, de forma aleatoria. Así, se aseguró que una misma imagen no entregue píxeles al conjunto de entrenamiento y validación en simultáneo, lo que llevaría a reportes de desempeño incorrectos.
- Se eligió como indicador de desempeño a la exactitud de clasificación promedio, es decir el promedio de las exactitudes obtenidas en las tres categorías. Dicha exactitud se obtiene como el cociente entre las muestras bien clasificadas por la discriminación binaria de cada categoría (Verdaderos Positivos y Verdaderos Negativos) sobre el total de instancias a clasificar.

$$E = \frac{VP + VN}{VP + VN + FP + FN}$$

Resultados Obtenidos

<i>Clasificador</i>	<i>Exactitud Media (%)</i>	<i>Desviación Estándar de Exactitud (%)</i>
RNDF	82	7
MLP	81	8
DTC	78	9
CTH-GOES16	83	8

Resultados Obtenidos (enfoque CTH Mejorado)

<i>Clasificador</i>	<i>Exactitud Media (%)</i>	<i>Desviación Estándar de Exactitud (%)</i>
CTH - RNDF	87	6
CTH - MLP	86	6
CTH - DTC	83	8
CTH-GOES16	83	8

CLASIFICACIÓN DIURNA DE LA ALTURA DE TOPE DE NUBES EN BASE A IMÁGENES GOES 16 EN TIEMPO REAL

Ernesto Marchesoni Seijo - Rodrigo Alonso Suarez - Matías Di Martino

Buenos Aires, 8 de noviembre 2022

