

# CLASIFICACIÓN DIURNA DE LA ALTURA DE TOPE DE NUBES EN BASE A IMÁGENES GOES16 EN TIEMPO REAL

Ernesto Marchesoni Seijo<sup>1</sup>, Rodrigo Alonso-Suárez<sup>1</sup>, Matías Di Martino<sup>2</sup>  
[emarchesoni@fing.edu.uy](mailto:emarchesoni@fing.edu.uy), [r.alonso.suarez@gmail.com](mailto:r.alonso.suarez@gmail.com)

<sup>1</sup> Laboratorio de Energía Solar (LES), Universidad de la República, Uruguay

<sup>2</sup> Duke University Durham, NC 27705, USA

**Palabras clave:** Clasificación automática, nubosidad, imágenes satelitales GOES16.

## 1) INTRODUCCIÓN Y OBJETIVOS

En este trabajo se desarrolla un clasificador automático de nubosidad en base a información satelital operativa para discriminar la altura de tope de nube en tres macro categorías (altas, medias y bajas). El objetivo de la herramienta es asistir el diagnóstico de predicciones operativas de radiación solar. Los clasificadores implementados se basan en algoritmos de aprendizaje automático y toman como características, es decir información de entrada, a las imágenes espectrales del satélite geoestacionario GOES16 y cálculos derivados de estas.

La referencia para el entrenamiento y validación de los algoritmos desarrollados en este trabajo viene dada por la clasificación derivada del producto Cloud Top Height del espectrofotómetro MODIS (CTH-MODIS, Platnick et al., 2015), a bordo del satélite polar y heliosincrónico TERRA, dado que ha sido validado con observaciones terrestres. A su vez, se tuvo en cuenta la clasificación del producto operativo Cloud Top Height basado en GOES16 (CTH-GOES16, GOES-R Algorithm Working Group), como línea de base de desempeño para la validación de los clasificadores. La clasificación buscada considera para cada píxel de la imagen las siguientes categorías: sin nubes, nubosidad baja, nubosidad media y nubosidad alta. La categoría “sin nubes” no entra en consideración de los clasificadores dado que es posible diferenciarla con bajo error a partir de la información del canal visible del GOES16, a través de la utilización de un modelo de fondo (Alonso-Suárez et al., 2014). El alcance de las categorías se resume en la Tabla I.

| <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.                                                                          |

Tabla I. Descripción del criterio de las categorías de nubosidad utilizado.

## 2) BASE DE DATOS DE TRABAJO

La diferencia entre las plataformas de observación satelital (un satélite geoestacionario y otro de órbita polar) implicó resolver la sincronía de las observaciones en el espacio y en el tiempo. Se homogeneizaron las grillas espaciales (con distinta base en ambas plataformas y distinta resolución), eligiendo la grilla de GOES16 como soporte. Al final de este proceso se

logró reunir información correlativa de 55 días en grillas de 145 x 171 píxeles, lo que totaliza 1363725 muestras en la Base de Datos de Trabajo (BDT). La información reunida compilada se compone 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, como la espacialidad de los patrones de nubosidad o su movimiento. 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 cada píxel, además de los correspondientes valores de CTH-GOES16 y CTH-MODIS.

### 3) DESCRIPCIÓN DE LOS CLASIFICADORES IMPLEMENTADOS Y SU EVALUACIÓN DE DESEMPEÑO

Los clasificadores implementados se basan en tres tecnologías de aprendizaje automático que operan sobre esta 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>                   |

*Tabla II. Descripción de los clasificadores utilizados.*

Para la evaluación del desempeño de los clasificadores se utilizó la técnica de Validación Cruzada (VC), donde se reserva un 90% de las imágenes (49 imágenes) para entrenamiento y el 10% restante para evaluación (6 imágenes), seleccionando estos conjuntos de forma aleatoria unas 100 veces sobre la totalidad de la BDT. De esta forma, 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, mayores al real. La métrica elegida fue 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.

### 4) DESEMPEÑOS OBTENIDOS

Los resultados para los tres clasificadores se muestran en la Tabla III. Se presenta el desempeño promedio y el desvío estándar de la técnica VC. El CTH-GOES16, al ser un producto que no se entrena con la BDT, se evalúa sobre todo el conjunto de datos. Se puede observar que el desempeño de los clasificadores es competitivo con el desempeño del CTH-GOES16, estando el clasificador RNDF marginalmente (1%) por debajo de la Base de Desempeño. Esto nos da una referencia de desempeño alcanzable con información operativa. Se debe destacar en este punto, que el CTH-GOES16 utiliza también insumos del modelado numérico de la atmósfera para realizar su clasificación, utilizando por tanto más información que las imágenes operativas del satélite GOES16.

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

*Tabla III. Desempeños alcanzados con la mejor configuración de clasificadores implementada y desempeño del CTH-GOES16 (Base de Desempeño).*

El buen desempeño logrado por los clasificadores con información GOES16 habilita a explorar su utilidad como información de entrada a los algoritmos, siendo por tanto la herramienta de clasificación un postproceso del CTH que busca mejorar la línea de desempeño del 83%. Al replantear el problema, se incorpora el CTH-GOES16 como una característica extra a los clasificadores analizados, además de las 29 características anteriores. Esta familia de clasificadores se denominó CTHs modificados y su desempeño se presenta en la Tabla IV. Como se observa, con esta estrategia el clasificador basado en RNDF alcanzó el 87% de exactitud media, el mejor desempeño que hemos obtenido hasta el momento.

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

*Tabla IV. Indicadores de desempeño alcanzados con la incorporación de CTH-GOES16 como característica.*

## REFERENCIAS

**Aicardi, D., Musé, P., y Alonso-Suárez, R., 2022.** A comparison of satellite cloud motion vectors techniques to forecast intra-day hourly solar global horizontal irradiation. *Solar Energy*, 233, 46-60.

**Alonso-Suárez, R., Abal, G., Siri, R., y Muse, P., 2014.** Satellite-derived solar irradiation map for Uruguay. *Energy Procedia*, 57, 1237-1246.

**Farnebäck, G., 2003.** Two-frame motion estimation based on polynomial expansion. In Bigun, J. and Gustavsson, T., editors, *Image Analysis*, pages 363-370, Berlin, Heidelberg. Springer Berlin Heidelberg.

**GOES-R Algorithm Working Group, and GOES-R Series Program Office (2018).** NOAA GOES-R Series Advanced Baseline Imager (ABI) Level 2 Cloud Top Height (ACHA). NOAA National Centers for Environmental Information.

**Platnick, S., Ackerman, S., King, M., Wind, G., Meyer, K., Menzel, P., Frey, R., Holz, R., Baumy, B., y Yang, P., 2015.** MODIS atmosphere L2 cloud product (06\_L2), NASA MODIS Adaptive Processing System, Goddard Space Flight Center.