

Clasificación satelital de la altura de tope de nubes y su aplicación a la evaluación de predicciones numéricas de nubosidad

Informe de Trabajo Especial de Grado presentado por

Ernesto Marchesoni Seijo

Licenciatura en Ciencias de la Atmósfera
Facultad de Ingeniería - Facultad de Ciencias
Universidad de la República del Uruguay

Tutores

Dr. Ing. Rodrigo Alonso-Suárez
Dr. Ing. Gabriel Cazes Boezio

Montevideo, 15 de marzo de 2025



Clasificación satelital de la altura de tope de nubes y su aplicación a la evaluación de predicciones numéricas de nubosidad por Ernesto Marchesoni Seijo tiene licencia [CC Atribución 4.0](https://creativecommons.org/licenses/by/4.0/).

Página intencionalmente dejada en blanco.

Página intencionalmente dejada en blanco.

Resumen

Este trabajo trata sobre el desarrollo y validación de un clasificador de Altura de Tope de Nube (CTH) con el fin de emplearlo para seleccionar eventos de nubosidad de interés para su simulación utilizando el modelo de Predicción Numérica del Tiempo (del inglés: Numerical Weather Prediction) (NWP) denominado Weather Research and Forecasting (WRF). El clasificador de altura de toques nubosos se obtiene aplicando técnicas de aprendizaje automático sobre la base de imágenes satelitales GOES-16 y es capaz de clasificar los píxeles de cada imagen en cuatro categorías: "Sin Nubosidad", "Nubes Bajas"(de 500 a 2000 metros), "Nubes Medias"(de 2000 a 6000 metros) y "Nubes Altas"(más de 6000 metros).

Las técnicas de aprendizaje automático utilizadas son de distinta complejidad: un Árbol de Decisión simple, un Bosque Aleatorio y una Red Neuronal. Para el entrenamiento y validación de los algoritmos de aprendizaje automático se debió compilar una base de datos, compuesta por mediciones realizadas por el instrumento Advanced Baseline Imager (ABI) a bordo del satélite geostacionario Geostationary Operational Environmental Satellites 16 (GOES16) de la National Oceanic and Atmosphere Administration (NOAA). Estas mediciones son la radiancia observada por el satélite en distintas longitudes de onda para cada pixel observado por el satélite. Se utiliza esta información convertida a las cantidades satelitales de uso corriente (factor de reflectancia y temperatura de brillo), junto a procesamientos y cálculos realizados sobre estas cantidades.

Como verdad de referencia para el entrenamiento y validación del desempeño de los clasificadores obtenidos se utiliza el producto de CTH derivado del instrumento Moderate Resolution Imaging Spectroradiometer (MODIS), a bordo del satélite heliosincrónico de órbita polar TERRA del programa Earth Observing System (TERRA) de la National Aeronautics and Space Administration (NASA). Este producto ha sido validado en trabajos de la literatura contra medidas terrestres de ceilómetros, y ha sido utilizado como referencia en trabajos pre-existentes. Además, se incorpora al trabajo la propia estimación de CTH provista por el ABI como base de desempeño para la evaluación del clasificador. Dado el uso de información satelital de plataformas de observación distintas, fue necesario resolver la coincidencia espacial y temporal de las observaciones, con el fin de lograr una Base de Datos de Trabajo (BDT) consolidada.

El mejor clasificador obtenido a partir de las imágenes GOES16 mostró un desempeño similar a la clasificación derivada del producto CTH de ABI, mostrando ser una herramienta de similar desempeño. El producto CTH de ABI utiliza, además de las imágenes GOES16, información de NWP globales, estando disponible con frecuencia diezminutal (y utilizando más información que las imágenes de satélite). Este primer paso permitió mostrar un desempeño similar del clasificador implementado (utilizable además en tiempo real), conocer el desempeño típico de este tipo de herramientas en nuestra región, y explorar el uso de los algoritmos de aprendizaje automáticos para esta tarea y como herramienta de postproceso del producto CTH de ABI. Con este último enfoque se consiguió mejorar el desempeño de la clasificación, superando por la base de desempeño definida por la clasificación derivada del producto CTH de ABI.

Con el mejor clasificador desarrollado se seleccionaron 10 eventos de nubosidad categoría 1 dominante y 10 eventos de nubosidad categoría 3 dominante. Una porción de estos eventos, de características relevantes para la predicción numérica de nubosidad, fueron simulados y analizados tanto dinámica como termodinamicamente utilizando el modelo WRF.

Palabras clave: Clasificación de Nubosidad, Observación Satelital, Aprendizaje Automático, Predicción Numérica del Tiempo, GOES-16, WRF

Página intencionalmente dejada en blanco.

Índice general

1. Introducción	1
1.1. Antecedentes	1
1.2. Motivación	2
1.3. Objetivos del trabajo	3
1.4. Sobre este trabajo	3
2. Categorización, Teledetección y Pronóstico de Nubosidad	5
2.1. Tipos de nubes y sus propiedades físicas	5
2.2. Medida terrestre de la altura de nubes	7
2.3. Observación satelital de la nubosidad	8
2.3.1. Sistemas de observación pasivos	8
2.3.2. Sistemas de observación activos	15
2.3.3. Comparación entre sistemas activos y sistemas pasivos	17
2.4. Weather Research and Forecasting	18
2.5. Clasificación de nubosidad por aprendizaje automático	19
3. Construcción de la base de datos	23
3.1. Fuentes de información	23
3.2. Características primarias: imágenes del ABI GOES16	23
3.3. Características secundarias: derivadas del ABI GOES16	26
3.3.1. Movimiento horizontal	26
3.3.2. Promedios espaciales	27
3.3.3. Variabilidad espacial	27
3.3.4. Gradientes	27
3.4. Características astronómicas	28
3.5. Productos CTH	29
3.6. Concurrencia espacio temporal	30
3.7. Asignación de categorías de nubosidad	31
3.8. Resumen general de la BDT	32

3.9. Exploración de la BDT	32
4. Clasificación satelital de altura de topes nubosos	41
4.1. Herramientas de aprendizaje automático utilizadas	41
4.1.1. Árbol de decisión simple	41
4.1.2. Bosques aleatorios	42
4.1.3. Redes neuronales	43
4.2. Evaluación de desempeño de la clasificación	44
4.2.1. Validación cruzada	44
4.2.2. Matriz de confusión y métricas de desempeño	45
4.3. Desempeño de la clasificación	47
4.4. Desempeño de la clasificación incluyendo el CTH ABI	48
4.5. Análisis de sensibilidad	52
5. Selección y análisis de eventos	57
5.1. Selección de días	57
5.2. Setup del WRF	59
5.3. Análisis de casos	60
5.3.1. Evento del 21 de diciembre de 2019	60
5.3.2. Evento del 7 de enero de 2020	69
5.3.3. Evento del 4 de noviembre de 2019	78
5.3.4. Evento del 17 de febrero de 2020	87
6. Conclusiones y Perspectivas	97
A. Abreviaciones	105

Capítulo 1

Introducción

Este capítulo presenta los antecedentes más relevantes de este trabajo. Se presentan brevemente los antecedentes del desempeño del modelo WRF para el pronóstico de nubosidad, que motivó al desarrollo de la herramienta de clasificación generada en este trabajo. Se describen también algunos elementos sobre la clasificación de nubosidad derivada de observaciones por medios satelitales pasivos. Además, se presentan antecedentes de aplicación de técnicas de aprendizaje automático a la clasificación de nubosidad y el único antecedente regional de clasificación de nubosidad por satélite. Finalmente, se utiliza la descripción anterior para detallar la motivación y objetivos de este trabajo.

1.1. Antecedentes

La resolución de los procesos de Capa Límite Atmosférica (PBL) son un desafío actual en el campo del modelado numérico de la atmósfera. La escala de estos intercambios ha llevado históricamente a parametrizarlos utilizando distintos esquemas que buscan representar estos procesos. El desempeño del modelo WRF en el pronóstico de nubosidad está estrechamente ligado a su capacidad para manejar los procesos dentro de la PBL, por lo que la elección de esquemas de parametrización adecuados puede mejorar significativamente la precisión en la predicción de la altura de la PBL y, por consiguiente de la nubosidad (Mantovani Júnior et al., 2023).

En el trabajo de Porrini (2017) se evaluó el modelo WRF como herramienta de pronóstico de radiación solar en superficie en nuestro territorio. Uno de sus resultados fue mostrar la subestimación en la nubosidad pronosticada frente a observaciones derivadas de MODIS. Además, encontró que la Humedad Relativa (HR) pronosticada muestra una buena relación con los errores de sobrestimación de radiación, originados en la subestimación de la nubosidad. Estos casos, donde la HR es elevada en el tope de la PBL, se asociaron a un desvío sistemático en las predicciones de irradiancia. Particularmente, Porrini comprobó una subestimación de la nubosidad en todos los niveles de la tropósfera, pero especialmente en los niveles bajos. Como las nubes en niveles bajos muestran un albedo elevado, las predicciones de irradiancia en superficie son especialmente sensibles a esta subestimación. También, comprobó que la subestimación de la cobertura nubosa en niveles bajos, es del orden de 60 a 70 %, con respecto a la nubosidad detectada en imágenes MODIS. Un siguiente paso en este análisis es evaluar el desempeño de la predicción de nubosidad categorizando las distintas alturas de nubes, para obtener desempeños discriminados. Esto requiere la capacidad de clasificar por teledetección la altura de nubes sobre extensiones territoriales.

La nubosidad desempeña un papel crucial en el balance energético del sistema terrestre, principalmente reflejando la Radiación de Onda Corta (SWR) de origen solar, además de absorber y emitir Radiación de Onda Larga (LWR) de origen terrestre. Los diferentes tipos de nubes afectan de forma distinta a la radiación en la atmósfera, influenciando el clima y el tiempo de manera significativa (Li et al., 2007; Matuszko, 2012). Por esta razón, la caracterización de la nubosidad a partir de imágenes satelitales ha sido una de las principales áreas de investigación desde el lanzamiento de los primeros satélites meteorológicos (A. K. Heidinger et al., 2020; Huo et al., 2020). En particular, la altura de tope de nube (CTH) es un producto históricamente derivado de imágenes satelitales con fines de

pronóstico meteorológico y estudios climatológicos (Forsythe, 2007; Min et al., 2020).

La estimación de CTH derivada de instrumentos satelitales pasivos se basa en la opacidad de las nubes a la radiación infrarroja. Observando en la ventana atmosférica del infrarrojo (una porción del espectro electromagnético donde la atmósfera es básicamente transparente a la radiación infrarroja), las nubes aparecen con una Temperatura de Brillo Infrarroja (TB-IR) más fría que la superficie, destacándose en la escena. Una ventaja notable de considerar la ventana infrarroja es que permite la continuidad de la estimación a lo largo de las 24 horas del día, al no depender de la iluminación solar. Un procedimiento clásico es la asunción de que la temperatura de tope de nube determina su altura. Esto, si bien es cierto en términos generales, carece de precisión cuando se intenta clasificar todos los píxeles de una imagen o todos los tipos de nubosidad. Por tanto, la TB-IR no es la única variable necesaria para estimar la altura de tope de nubes y la metodología a desarrollar requiere mayor complejidad, trascendiendo largamente el establecimiento de una parametrización que relacione fenomenológicamente ambas variables.

Tanto MODIS como ABI cuentan con productos CTH generados a partir de sus imágenes satelitales multiespectrales. Estos productos tienen base física, y parten de distinta información. En el caso de MODIS, se utiliza la gran capacidad de resolución espectral del satélite. En el caso de ABI, la capacidad espectral es más limitada. En ambos casos las metodologías incorporan además información de simulaciones NWP y Modelos de Transferencias Radiativas en la Atmósfera (RTM). La resolución espacial y temporal alcanzable es también distinta, así como la posibilidad de disponerlos o no en tiempo real. Lo primero depende de las capacidades de las imágenes colectadas por cada satélite y la metodología utilizada, siendo 1 km y 2-3 instancias por día para MODIS, y 10 km y cada 10 minutos para el caso de ABI. Lo segundo depende de la estructura de producto y distribución de cada agencia espacial y los insumos que requiera. En particular, los productos de esta categoría están disponibles con un retraso operativo de 1 día para MODIS y sin retraso para ABI.

En la región se tienen dos experiencias relevantes en el campo de la detección y clasificación de altura de nubosidad por satélite. En primer lugar, la línea de trabajo brasileña que se presenta en Ceballos y Negri (2022) que tiene como resultado el desarrollo y mejora a lo largo de los años de un sistema de clasificación de nubosidad operativo, a partir de la Reflectividad Planetaria (RP) del canal visible y la TB-IR del canal 4 (10.70 μm) obtenidas de las imágenes del satélite GOES12. Esta clasificación y sus actualizaciones utiliza el método de *clustering* con $K=30$ sobre 4 variables, las ya mencionadas RP y TB-IR y dos variables de textura derivadas de estas (Bottino y Ceballos, 2006). Si bien este clasificador fue desarrollado originalmente para imágenes de satélite GOES12 (actualmente fuera de servicio) ha sido rediseñado para su uso con las imágenes del instrumento ABI del GOES16. En Ceballos y Negri (2022) se presenta la metodología actualizada y se menciona que la base del clasificador continúa siendo similar al enfoque original. En particular destacan que el desempeño de esta técnica en la detección de nubosidad es similar al producto Cloud Mask (CM) de NOAA.

La siguiente referencia regional es el trabajo argentino de Yuchechen et al. (2020). En este estudio también se implementa un método de *clustering* basado en la combinación de los algoritmos *k-means* y *k-means++* sobre imágenes diarias de anomalías estandarizadas de la TB-IR del canal 4 (10.70 μm) obtenidas por el satélite GOES-13. El trabajo se centra en el período comprendido entre el 1 de diciembre de 2010 y el 30 de noviembre de 2016, tomando como área de interés el continente sudamericano y los océanos adyacentes.

Particularmente en Uruguay no existen a la fecha experiencias previas de clasificación de la nubosidad en base a imágenes satelitales, ya sea a través de modelado físico de la atmósfera o por metodologías de aprendizaje por computadora.

1.2. Motivación

Las herramientas de NWP, ya sean de alcance global o regional, predicen razonablemente bien la nubosidad asociada a sistemas convectivos de escala sinóptica, como los sistemas frontales o los sistemas convectivos de mesoescala. Sin embargo, las nubes bajas son sistemáticamente mal pronosticadas por los modelos de NWP (Markovic et al., 2008; Osborne et al., 2014; Yang y Kleissl, 2016). Como se mencionó en la sección 1.1, las predicciones numéricas operativas en Uruguay mues-

tran una subestimación sistemática de la nubosidad en todos los niveles, pero especialmente en los niveles bajos. Por esto, la plataforma de pronósticos numéricos empleada en Uruguay genera una sobrestimación ocasional, pero frecuente, de la irradiancia a nivel de superficie. Esto implica que las predicciones numéricas de radiación solar disponibles en nuestro país cuentan aún con gran espacio de mejora (Porrini et al., 2015; Teixeira-Branco et al., 2018), a diferencia de las predicciones de otras variables, como el viento (Arce et al., 2018) y las precipitaciones (Silveira et al., 2018).

El desarrollo de una herramienta local que permita conocer el tipo de nubosidad en tiempo casi real, así como clasificar largos periodos de tiempo pasado, es una contribución para el estudio y evaluación del pronóstico de radiación solar. También, es posible que las herramientas de estimación satelital (a tiempo real y retrospectivos) de la radiación solar descendente puedan verse beneficiadas al incorporar en sus modelos información del tipo de nubosidad presente o alguna de sus características, como la altura de tope o base nubosa. Este trabajo es, además, una primera aproximación al problema de la clasificación o estimación de características de nubes por satélite en Uruguay, generando experiencia local en la temática. Por último, resulta destacable la posibilidad de probar herramientas de aprendizaje automático para la obtención de productos satelitales operativos validados, teniendo como insumo principal las imágenes satelitales almacenadas y rutinariamente descargadas por el Laboratorio de Energía Solar de la Universidad de la República (UDELAR) (LES).

1.3. Objetivos del trabajo

Este trabajo busca generar y evaluar una herramienta que permita detectar y clasificar la presencia de nubosidad a distinta altura sobre el territorio de Uruguay en horas diurnas, que son las de interés para el pronóstico de la radiación solar. Una vez desarrollada y validada, la herramienta satelital es utilizada para detectar eventos con presencia de nubosidad con distinta altura de topes nubosos. Los eventos seleccionados serán analizados dinámica y termodinamicamente a partir de simulaciones obtenidas del modelo numérico de atmósfera WRF.

Para el desarrollo de la herramienta de clasificación de nubosidad se consideran técnicas de aprendizaje automático, por lo que es necesario la confección de una BDT que permita el entrenamiento y validación de estos clasificadores. El insumo de los clasificadores a desarrollar serán las imágenes obtenidas por el radiómetro ABI y se clasificará cada píxel de estas imágenes en cuatro categorías: sin nubosidad, nubosidad baja, nubosidad media y nubosidad alta, como ya se ha expresado.

En particular, se plantean los siguientes objetivos específicos para darle estructura al trabajo:

1. Adquirir experiencia en el uso y manejo de la información meteorológica satelital.
2. Utilizar en forma combinada información meteorológica proveniente de distintos satélites.
3. Adquirir experiencia en el uso y manejo del modelo WRF.
4. Desarrollar un clasificador de altura de tope nubes a partir de imágenes del satélite GOES16.
5. Utilizar el clasificador de nubosidad desarrollado para la selección y estudio de casos de nubosidad de interés para el diagnóstico de herramientas de predicción del recurso solar sobre el territorio de Uruguay.

1.4. Sobre este trabajo

Este trabajo describe el primer desarrollo en Uruguay destinado a estimar características de la nubosidad por satélite. Específicamente se atiende la clasificación de la altura de los topes nubosos en tres categorías de altura. Esto es, poder asignarle a cada píxel nuboso en una imagen satelital una categoría de altura según su tope se ubique a una altura baja, media o alta. No se busca en esta instancia generar una herramienta que permita estimar la altura de la nubosidad propiamente dicha o el tipo de nube presente, lo que son próximos pasos naturales para esta línea de desarrollo. El trabajo cuenta con una limitación de base, que es la no disponibilidad en Uruguay de medidas terrestres de ceilómetros de altas prestaciones. Esto ha sido suplido en este trabajo por el uso de otra herramienta

satelital validada en la literatura que es el producto MODIS. Por tanto, este trabajo en esencia genera una herramienta de clasificación satelital adaptada a GOES16. Los algoritmos desarrollados se comparan con el producto operativo de GOES-16 de altura de tope de nubes, obteniendo resultados similares y un poco mejores, incluso cuando el algoritmo operativo utiliza más información de entrada que las imágenes de satélite para su cometido. En una segunda etapa, se incorpora este producto operativo como entrada al problema de clasificación, obteniendo el clasificador de mejor desempeño de los desarrollados aquí. La herramienta final desarrollada permite clasificar la altura de topes nubosos sobre Uruguay cada 10 minutos y para cada píxel de la imagen satelital.

El resto de este manuscrito se organiza de la siguiente manera. El capítulo 2 presenta una síntesis de las temáticas relevantes para este trabajo, que son la clasificación de nubosidad, su teledetección, su predicción numérica y su clasificación mediante herramientas de aprendizaje automático. El capítulo 3 presenta el conjunto de datos a utilizar, describiendo el procesamiento necesario para confeccionar una base de datos utilizable para el resto del trabajo. Se presenta en este capítulo también un análisis exploratorio de la base de datos confeccionada. El capítulo 4 presenta muy brevemente los aspectos más importantes de los clasificadores de aprendizaje automático utilizados, los detalles para su implementación y evaluación, y los resultados de la clasificación. En este capítulo se decide el mejor clasificador implementado. El capítulo 5 desarrolla la aplicación de la herramienta al diagnóstico de aciertos y errores de la predicción de nubosidad por el WRF, junto con sus detalles de implementación. Finalmente, el capítulo 6 presenta brevemente las principales conclusiones y líneas abiertas que deja la realización de este trabajo.

Capítulo 2

Categorización, Teledetección y Pronóstico de Nubosidad

Este capítulo describe la clasificación histórica de la nubosidad para luego enfocarse en el estado del arte del reconocimiento y clasificación de nubosidad a través de medios satelitales. Se presentan productos derivados de sistemas de observación pasivos y activos. Además, se analiza la diferencia entre estos productos. También describe la habilidad de los NWP en el pronóstico de nubosidad y, en particular, se presentan características del WRF para este propósito. Por último, se introducen ejemplos de experiencias de aplicación de técnicas de aprendizaje automático a la clasificación de nubosidad.

2.1. Tipos de nubes y sus propiedades físicas

En el año 1803, Luke Howard publica su ensayo "The Modifications of Clouds" donde a partir de sus observaciones presenta un sistema de clasificación de nubosidad basado en su altura y organización (o forma) (Howard, 1803; WMO, 1956), incorporando esta clasificación en el estudio del estado y transiciones del tiempo. En una época donde la exploración de la atmósfera era incipiente, el trabajo de Howard propuso una solución elegante que fue ampliamente aceptada y permitió sistematizar la observación de la nubosidad, una tarea considerada infructuosa hasta ese momento. La Tabla 2.1 presenta los 10 géneros de nubosidad que estableció en esta clasificación, especificando los niveles troposféricos y el rango de altitudes asociado a cada género. De igual forma, la Figura 2.1 representa los 10 géneros morfológicamente y en sus niveles atmosféricos. Con el desarrollo de las comunicaciones, esta clasificación permitió informar la nubosidad presente y complementar así la descripción del estado del tiempo. Si bien fue originalmente desarrollada para climas templados, actualmente se emplea en las observaciones que realizan las estaciones meteorológicas de superficie en todo el mundo. Para esto se emplea la clasificación de Howard ampliada, que se mantiene actualizada en el Atlas Internacional de Nubes (AIN) de la Organización Meteorológica Mundial (OMM). Sin embargo, los 10 géneros históricos presentados por Howard siguen siendo la base descriptiva de la nubosidad.

Los distintos tipos de nubes tienen diferentes propiedades físicas. En primer lugar, se destaca la fase termodinámica de las partículas de la nube, pudiendo ser partículas de agua líquida, hielo o una mezcla de ambas. También se puede estudiar el tamaño y la forma de las partículas que componen a la nubosidad, aunque estas varían, se pueden mencionar partículas con formas esféricas, lineales o irregulares, entre otras, (R. Zhang et al., 2023). Sus tamaños pueden variar desde los micrómetros en el caso de las pequeñas gotitas de agua líquida, hasta los centímetros en el caso de granizo de gran tamaño. También se considera el espesor geométrico de la nube, dado por la distancia vertical entre la base y el tope de la nube. Otra propiedad de gran interés es el espesor óptico de la nube, definido como el cociente entre la potencia radiante incidente y transmitida a través de la nube. De esta forma, cuanto mayor sea el espesor óptico, menor será la cantidad de potencia radiante transmitida a través de la nube. Para el caso del espesor óptico en el rango espectral visible, las nubes no absorben esta radiación pero sí la dispersan y reflejan.

Nivel	Género y abreviación	Extensión Vertical
Alto	Cirrus (Ci)	5-13 km
Alto	Cirrocumulus (Cc)	5-13 km
Alto	Cirrostratus (Cs)	5-13 km
Medio	Alto cumulus (Ac)	2-7 km
Medio	Altostratus (As)	2-7 km
Medio	Nimbostratus (Ns)	2-7 km
Bajo	Stratus (St)	Superficie a 2 km
Bajo	Stratocumulus (Sc)	Superficie a 2 km
Bajo	Cumulus (Cu)	Superficie a 2 km
Desarrollo Vertical	Cumulonimbus (Cb)	Superficie a 13 km

Tabla 2.1: Clasificación de Howards con los 10 géneros principales de nubes según nivel y organización. Esta clasificación ha sido utilizada desde el comienzo de las comunicaciones de las observaciones meteorológicas, permitiendo estandarizar la descripción de la nubosidad presente.

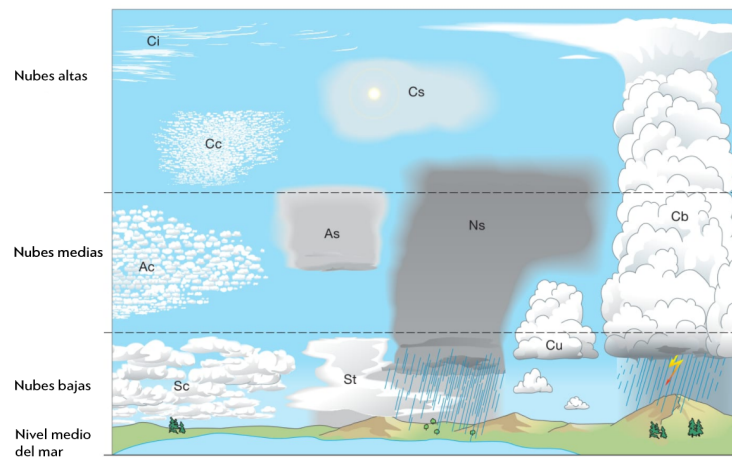


Figura 2.1: Los diez géneros en su nivel o niveles asignados. Tomada del Atlas Internacional de Nubes.

También se definen propiedades de los topos nubosos tales como su altura, CTH, que representa la altitud geopotencial del tope de la nubosidad observada. De igual forma se define la Presión de Tope de Nube (CTP) y la Temperatura de Tope de Nube (CTT), que representan respectivamente el nivel de presión en el que se encuentra el tope nuboso y su temperatura. Estas últimas tres propiedades son usualmente estimadas por medios satelitales. Por último, tanto en el contexto de la simulación numérica de la nubosidad, como en la observación satelital, se define la Fracción de Nubosidad (CLDFRA) como el porcentaje de cada píxel en imágenes satelitales o cada cuadro de cuadrícula en un modelo meteorológico o climático que está cubierto de nubes. De esta forma, una fracción de nubosidad de '1' significa que el píxel está completamente cubierto de nubes, mientras que una fracción de nubosidad de '0' representa un píxel totalmente libre de nubes.

A continuación se presentan las definiciones de los 10 géneros principales en base a lo descrito en el AIN y se desarrollan las propiedades físicas destacables de cada género:

Cirrus (Ci): Nubes altas y delgadas, compuestas por cristales de hielo, lo que les da un aspecto fibroso o sedoso. Generalmente aparecen en forma de filamentos o bandas blancas y tienen un espesor geométrico generalmente delgado. Su espesor óptico es bajo, lo que contribuye a su apariencia translúcida.

Cirrocumulus (Cc): Nubes altas que se presentan como pequeños copos o gránulos blancos sin sombras. Estas nubes están formadas por cristales de hielo y suelen disponerse en filas. Tienen un espesor geométrico relativamente delgado y un espesor óptico bajo, a menudo mostrando patrones de ondas.

Cirrostratus (Cs): Nubes altas que forman una capa delgada y blanquecina que cubre total

o parcialmente el cielo. Están compuestas principalmente por cristales de hielo y pueden producir halos alrededor del sol o la luna debido a la refracción de la luz. Tienen un espesor geométrico mayor que los cirrus y un espesor óptico moderado.

Altostratus (As): Nubes medias que aparecen como masas o capas blancas o grises, generalmente con sombras. Estas nubes contienen gotas de agua líquida y, ocasionalmente, cristales de hielo. Su estructura es a menudo ondulada o de rodillos, con un espesor geométrico y óptico moderados, permitiendo que el sol se vea parcialmente.

Altostratus (As): Nubes medias que forman una capa grisácea o azulada, cubriendo el cielo total o parcialmente. Estas nubes pueden contener una mezcla de gotas de agua líquida y cristales de hielo. Tienen un espesor geométrico medio a grueso y un espesor óptico moderado a alto, a veces ocultando el sol.

Nimbostratus (Ns): Nubes bajas y densas, de color gris oscuro. Están asociadas con precipitaciones continuas y persistentes, como lluvia o nieve. Estas nubes contienen principalmente gotas de agua líquida y posibles cristales de hielo, y tienen un gran espesor geométrico y un espesor óptico muy alto.

Stratus (St): Nubes bajas que forman una capa uniforme de color gris. Están compuestas principalmente por pequeñas gotas de agua líquida y pueden producir llovizna. Su espesor geométrico es relativamente delgado, pero tienen un espesor óptico alto, lo que puede bloquear completamente el sol.

Stratocumulus (Sc): Nubes bajas que forman una capa de grandes masas o rodillos de color blanco o gris. Estas nubes contienen gotas de agua líquida y posibles cristales de hielo. Su espesor geométrico y óptico es variable, lo que a menudo resulta en una apariencia de cielo cubierto con espacios claros entre las nubes, rara vez produciendo precipitación.

Cumulus (Cu): Nubes de desarrollo vertical con apariencia de montículos o torres. Están formadas por gotas de agua líquida en las partes bajas y cristales de hielo en las partes altas. Tienen un desarrollo vertical significativo con un espesor geométrico y óptico variable, típicas en días soleados y pueden evolucionar a nubes de tormenta.

Cumulonimbus (Cb): Nubes de gran desarrollo vertical que pueden extenderse desde niveles bajos hasta niveles altos de la tropósfera. Son nubes de tormenta que pueden producir lluvias fuertes, actividad eléctrica, granizo y vientos intensos. Contienen desde pequeñas gotas de agua hasta grandes cristales de hielo y granizo, con un gran espesor geométrico y un espesor óptico muy alto.

Para finalizar, es necesario resaltar que la clasificación de nubosidad que recoge el AIN (apenas presentada) tiene su origen en el trabajo de descripción y clasificación de Howard realizado hace más de 200 años. Aún así, esta clasificación (mejorada) continúa siendo utilizada en la meteorología, climatología y ciencias de la atmósfera contemporáneas.

2.2. Medida terrestre de la altura de nubes

Existen tres instrumentos principales que permiten medir la altura de la nubosidad desde la superficie terrestre: el Radio Detection And Ranging (RADAR) meteorológico, el radiosondeo y el Laser Detection And Ranging (LIDAR). A partir de estos se puede obtener la altura de base, altura de tope y extensión vertical de la nubosidad. En el caso del radar meteorológico, este es un instrumento fundamental para la observación de la atmósfera. Emite ondas de radio o microondas que se reflejan en las partículas de las nubes, permitiendo determinar su ubicación y la reflectividad que devuelven. Existen varios tipos de radares meteorológicos. Los trabajos de [Doviak et al. \(1979\)](#) y [Council et al. \(2002\)](#) sirvieron de referencia en esta descripción. En particular, el radar de simple polarización emite ondas de radio y mide la reflectividad devuelta por las partículas junto con el tiempo de retorno del eco. Los radares de doble polarización, emiten y reciben ondas polarizadas (señal horizontal y señal vertical), lo que permite distinguir mejor entre diferentes tipos de hidrometeoros y estimar la forma y orientación de estas partículas. Los radares Doppler, por su parte, además de medir la reflectividad, aprovechan el efecto Doppler para detectar el cambio en la frecuencia de las ondas debido al movimiento de las partículas, permitiendo así estimar la velocidad radial

del viento y detectar fenómenos como mesociclones y hasta tornados. De esta forma, los radares son capaces de determinar la ubicación de la nubosidad, extensión vertical y ciertas características de las partículas que las componen. Por otra parte, el radiosondeo se obtiene mediante el lanzamiento de un globo equipado con instrumentos que miden diversas variables atmosféricas, como la temperatura, la humedad y la presión, a diferentes alturas, proporcionando un perfil vertical detallado de la atmósfera. En particular, en un sondeo atmosférico es posible estimar el nivel de condensación por ascenso, es decir, la base de la nubosidad. En cuanto a los sistemas LIDAR, emiten pulsos láser y miden el tiempo de retorno de los pulsos reflejados por las partículas atmosféricas. De esta forma pueden determinar la posición vertical (altura) de la base de las nubes con gran precisión. Los sistemas LIDAR también pueden ser polarizados y aprovechar el efecto Doppler. El trabajo de [McGill et al. \(2002\)](#) presenta una descripción de estos sistemas enfocados en la teledetección de nubosidad. La base física del funcionamiento de los sistemas RADAR y LIDAR es la misma. Sin embargo, en los sistemas RADAR es de interés conocer la reflectividad devuelta por el blanco circunstancial y así inferir cualitativamente la densidad de partículas encontradas en la exploración. En el caso de los sistemas LIDAR, estos se enfocan en conocer la posición del objetivo, por lo que principalmente interesa determinar el tiempo de retorno de la señal reflejada. En la meteorología operativa de apoyo a la actividad aeronáutica es usual la instalación de LIDAR optimizados, que son denominados ceilómetros.

De los tres instrumentos descriptos, el LIDAR es el más preciso y el más adecuado para la determinación de la altura nubosa en un sitio en particular, permitiendo en muchos casos determinar la ubicación vertical de hasta 3 capas nubosas. En cambio, el RADAR explora una gran región y esta exploración se ve afectada por la curvatura de la Tierra. A medida que el haz del RADAR se aleja de su ubicación de origen la exploración se limita a niveles cada vez más altos de la atmósfera, ignorando que sucede en los niveles inferiores de esta para esas ubicaciones. Además, existe una porción de atmósfera sobre el sitio de emplazamiento del RADAR, conocida como cono de silencio, donde el instrumento no puede realizar mediciones. Por otra parte, un radiosondeo permite conocer la estructura vertical de la atmósfera representativa de una región, dependiendo de los vientos presentes al momento del lanzamiento el instrumento puede alejarse hasta 200 km de la ubicación de origen.

Actualmente Uruguay no cuenta con lanzamiento de radiosondas ni RADAR meteorológico. A la fecha, solamente se cuenta con un ceilómetro de apoyo a las estimaciones de altura de nubosidad realizadas en la estación meteorológica de Instituto Uruguayo de Meteorología (INUMET) ubicada en el Aeropuerto Internacional de Carrasco. Estas estimaciones de altura de nubosidad no fueron consideradas en este trabajo debido a problemas con su disponibilidad, pero no se descarta incorporarlas en trabajos futuros.

2.3. Observación satelital de la nubosidad

Los sistemas de teledetección satelital de la nubosidad y estimación de sus propiedades pueden clasificarse en dos tipos: los pasivos y los activos. Esta clasificación también se sostiene para elementos de medición terrestre, dado que en esencia diferencia el tipo de instrumento con que se realizan las mediciones. Los sistemas activos generan una señal que interactúa con el objeto, y obtienen información a partir de esa interacción, usualmente a través de un procesamiento de la señal reflejada. En cambio, los sistemas pasivos no emiten una señal para interactuar con el objeto, sino que sólo registran su emisión o la reflexión en dicho objeto de una fuente externa, en general, conocida con cierto grado de detalle (como el caso de la radiación solar incidente) o modelable físicamente (como el caso de la emisión térmica del sistema Tierra-Atmósfera). Las siguientes subsecciones describen ejemplos de instrumentos de ambos tipos a bordo de misiones satelitales.

2.3.1. Sistemas de observación pasivos

Los instrumentos pasivos miden la radiación reflejada o emitida por un objeto. Los instrumentos capaces de medir radiación electromagnética se denominan radiómetros, y pueden medir la radiación recibida en su sensor dentro de un ancho de banda de longitudes de onda o con resolución espectral (espectroradiómetro). Los instrumentos utilizados para el sensoramiento remoto aprovechan las porciones del espectro electromagnético en los rangos visible (entre 0.4 y 0.7 μm), infrarrojo cercano

y de onda corta (entre 0.7 y 3.0 μm), infrarrojo térmico (entre 3.0 y 15.0 μm , lo que incluye el infrarrojo medio y de onda larga), y de microondas (entre 1 mm y 1 m), siendo éste último el menos frecuente. Utilizando estos instrumentos es posible medir ciertas propiedades físicas de la nubosidad a partir de su interacción con la SWR de origen solar y la LWR de origen terrestre.

A continuación se presentan técnicas satelitales que se utilizan para reconocer nubosidad y determinar su altura. Esta descripción no pretende ser exhaustiva sino introducir conceptos que permitan comprender la base física de estas medidas. El trabajo de [Ackerman et al. \(1998\)](#) profundiza en el desarrollo de estas técnicas y su uso en distintas misiones de observación del sistema terrestre a lo largo del tiempo. Dicho esto y como se mencionó en la sección de antecedentes (1.1), las nubes son en general opacas a la LWR (radiación infrarroja). Por esto, si se observa en la ventana infrarroja (regiones del espectro cercanas a los 11 μm), las nubes se destacan con una TB-IR más fría que la superficie. También, se puede considerar la RP del canal visible, y definir umbrales característicos de RP en ausencia de nubosidad. Al superarse estos umbrales, el píxel en consideración puede designarse como nuboso. Luego de identificado un píxel nuboso, es posible determinar aproximadamente su nivel vertical utilizando la técnica conocida como *IR-Window* ([Baum et al., 2012](#)). Esta técnica es una de las primeras que fueron aplicadas a la estimación de altura de nubosidad. Utiliza un perfil vertical de temperatura como apoyo y asume que la TB-IR en la longitud de onda de 11 μm representa la CTT. De esta forma encuentra el nivel vertical del tope de la nubosidad detectada. Sin embargo, esta técnica tiene errores conocidos, por ejemplo, las nubes del tipo Cirrus, que se caracterizan por ser altamente transmisivas a la SWR y al mismo tiempo presentar opacidad a la LWR, son sistemáticamente asignadas a niveles bajos de la atmósfera con esta técnica ([Stephens y Webster, 1981](#); [Fu y Liou, 1993](#); [Hansen et al., 1997](#)). También con esta técnica es posible confundir nubes bajas con cuerpos de agua congelados o regiones nevadas. En busca de mejorar la técnica *IR-Window* derivada de umbrales fijos, se han implementado umbrales específicos a cada punto de grilla, más ajustados a la gran cobertura espacial que ofrece la observación satelital ([Isaacs et al., 1986](#)). Es oportuno señalar que en la sección 3 del presente trabajo se desarrolla la técnica denominada "filtro de Brillo de Fondo" que se utilizó para identificar píxeles nubosos. Esta técnica es un ejemplo de umbrales adaptativos en la identificación de nubosidad.

El trabajo de [Chahine \(1974\)](#) presenta un método basado en el estudio de la absorción del canal del Dióxido de Carbono (CO_2). Esta técnica actualmente es conocida como *CO2-Slicing* y se basa en la absorción diferencial con la altura que se observa en la porción de la ventana infrarroja, que se ve afectada por la presencia del CO_2 en la atmósfera. El desarrollo teórico de [Chahine](#) puede encontrarse recopilado en [Chahine \(1982\)](#). En particular, la técnica *CO2-Slicing* aplicada a una capa de nubosidad (single layer) puede desarrollarse siguiendo el trabajo de [Menzel et al. \(2008\)](#), donde se detalla que para un elemento nuboso en el Campo de Visión (FOV) del satélite, la radiancia observada $R(\nu)$ en la frecuencia ν puede expresarse como:

$$R(\nu) = (1 - NE) R_{\text{clr}}(\nu) + NE R_{\text{bcd}}(\nu, P_c), \quad (2.1)$$

donde $R_{\text{clr}}(\nu)$ es la radiancia de cielo claro, $R_{\text{bcd}}(\nu, P_c)$ es la radiancia de la nube en el nivel P_c , N es la fracción del FOV ocupada por la nube y E es la emisividad de la nube. En este caso, la radiancia de la nube puede expresarse como:

$$R_{\text{bcd}}(\nu, P_c) = R_{\text{clr}}(\nu) - \int_{P_c}^{P_s} \tau(\nu, p), \frac{dB[\nu, T(p)]}{dp} dp \quad (2.2)$$

donde P_s es la presión de superficie, P_c es la presión en el tope de la nube, $\tau(\nu, p)$ es la transmitancia de radiación de frecuencia ν emitida entre el nivel de presión atmosférica (p) hasta el tope de la atmósfera ($p = 0$) y $B[\nu, T(p)]$ es la radiancia de cuerpo negro de la Ley de Planck para una frecuencia ν y temperatura $T(p)$. De esta forma, el sumando negativo en la Ec. (2.2) representa la afectación en la radiancia de cielo claro ocasionada por la nube. Entonces, para dos frecuencias espectrales ν_1 y ν_2 observando el mismo FOV se tendrá:

$$\frac{R(\nu_1) - R_{\text{clr}}(\nu_1)}{R(\nu_2) - R_{\text{clr}}(\nu_2)} = \frac{NE_1 \int_{P_s}^{P_c} \tau(\nu_1, p) \frac{dB[\nu_1, T(p)]}{dp} dp}{NE_2 \int_{P_s}^{P_c} \tau(\nu_2, p) \frac{dB[\nu_2, T(p)]}{dp} dp}. \quad (2.3)$$

Eligiendo ν_1 y ν_2 cercanas se puede asumir que E_1 y E_2 son iguales. De esta forma es posible encontrar P_c , pero para esto es necesario contar con perfiles verticales de temperatura y transmitancia para las dos frecuencias consideradas. Estos perfiles son obtenidos de salidas de NWP y RTM. Luego, conociendo el nivel P_c , es decir, el valor CTP, es posible calcular CTH utilizando nuevamente la ayuda de perfiles verticales generados por NWP. El trabajo de [Smith y Platt \(1978\)](#) comprobó esta técnica contra medidas de LIDAR y radiosondeos, volviéndose la técnica base de la estimación de CTP y CTH con medios satelitales.

Otra técnica usual en este campo es la conocida como *Split-Window*. Este análisis utiliza dos o más bandas infrarrojas adyacentes. Así, no solo las TB-IR en cada banda son indicadoras de nubosidad y permiten estimar un nivel vertical para esta, sino que las diferencias de TB-IR entre dos bandas contiguas permiten mejorar la estimación del nivel vertical de la nubosidad detectada ([Purbantoro et al., 2018](#)). Esta técnica fue sugerida en el trabajo de [Liou \(1974\)](#) y comprobada en [Inoue \(1985\)](#) y [Saunders y Kriebel \(1988\)](#), donde se lograron detectar nubes del tipo Cirrus aplicando esta técnica, al evaluar la diferencia entre las TB-IR en 11 μm y 12 μm .

Las tres técnicas presentadas: *IR-Window*, *CO2-Slicing* y *Split-Window* se han refinado a lo largo de los años pero siguen siendo la base de la identificación de nubosidad y determinación de su altura a partir de la observación satelital pasiva. A continuación se presentan dos ejemplos destacables de sistemas satelitales de observación pasiva, de los cuales se obtienen estimaciones CTH. La base física de las estimaciones derivadas de estos instrumentos es una combinación de las técnicas presentadas en los párrafos anteriores, con el apoyo de NWP y RTM.

Moderate Resolution Imaging Spectroradiometer

MODIS es un instrumento clave en el sistema de observación terrestre ([Menzel y Strabala, 1997](#)). Existen dos de estos instrumentos, embarcados en los satélites de órbita baja y heliosíncronicos TERRA y AQUA del programa Earth Observing System (AQUA). La Tabla 2.2 presenta las 36 bandas espectrales de observación de MODIS. La información recolectada en estas bandas espectrales sirve de insumo para la generación de productos del estado de la atmósfera y en particular de la nubosidad. Al día de hoy, MODIS ofrece una resolución horizontal de 250 metros (en las bandas 1 y 2), 500 metros (en las bandas 3 a 7) y 1000 metros (en las bandas 8 a 36). Los satélites TERRA y AQUA tienen un tiempo de revisita medio de 16 días, es decir, cada 16 días se tienen imágenes MODIS sobre el mismo punto en el nadir de estos satélites. Además, MODIS cuenta con un sistema de calibración que le permite realizar ensayos de autocalibración en órbita y así garantizar la precisión y consistencia de los datos recopilados a lo largo del tiempo.

Para este trabajo, es de interés el producto *Cloud Top Properties* de MODIS. En particular, se eligió trabajar con la Colección 6 del producto derivado de las observaciones MODIS TERRA (identificado como MOD06¹). El reporte técnico de [Menzel et al. \(2015\)](#) detalla que este producto ofrece estimaciones de CTT, CTP y CTH con resolución horizontal de 1 km. En su trabajo [Menzel et al.](#) presenta una descripción detallada del algoritmo de base física Cloud Top Properties and Cloud Phase Algorithm de MODIS (CTPCPA) que produce estas estimaciones. Se aclara que CTPCPA trabaja sobre píxeles previamente identificados como nubosos por el Discriminating Clear-Sky from Cloud With Modis Algorithm (DCSFCMA). [Menzel et al.](#) menciona que CTPCPA utiliza las bandas 31 (11.03 μm), 33 (13.335 μm), 34 (13.635 μm), 35 (13.935 μm) y 36 (14.235 μm) de MODIS combinando la técnica *CO2-Slicing* para las nubes medias o altas y la técnica *IR-Window* para la nubosidad baja, en un enfoque que denominan de arriba hacia abajo (*top-down*). El desarrollo detallado de esta técnica y sus refinamientos excede el espíritu de esta introducción, de todas formas se desarrollará parte de la descripción disponible en [Menzel et al. \(2015\)](#). En particular, el enfoque *top-down* consiste en la evaluación selectiva de la técnica *CO2-Slicing*. En primer lugar, se resuelve P_c de la Ec. (2.3) para los pares de bandas 36/35, 35/34 y 34/33. Cada uno de estos tres valores P_c será candidato a CTP. Sin embargo, estos $P_c(36/35)$, $P_c(35/34)$ y $P_c(34/33)$ serán considerados según la importancia de cada par de bandas en los niveles altos de la tropósfera. En primer lugar se evaluará si $P_c(36/35)$ es menor a 450 hPa. Si esto fuera así, entonces $P_c(36/35)$ será CTP. En caso contrario, se considerará si $P_c(35/34)$ es menor a 550 hPa, en caso afirmativo $P_c(35/34)$ será CTP. En caso negativo, se evaluará si $P_c(34/33)$ es menor a 650 hPa, de ser así $P_c(34/33)$ será CTP. En caso de no tener éxito con ninguna de estas evaluaciones de *CO2-Slicing*, se considerará la TB-IR de

¹MOD06 refiere al producto derivado de MODIS TERRA, mientras que MYD06 refiere al producto derivado de MODIS AQUA

Banda MODIS	Longitud de Onda Central (μm)	Ancho Espectral (μm)	Nombre Descriptivo
1	0.645	0.620 - 0.670	Rojo
2	0.865	0.841 - 0.876	Infrarrojo cercano
3	0.469	0.459 - 0.479	Azul
4	0.555	0.545 - 0.565	Verde
5	1.240	1.230 - 1.250	Infrarrojo de onda corta
6	1.640	1.628 - 1.652	Infrarrojo de onda corta
7	2.130	2.105 - 2.155	Infrarrojo de onda corta
8	0.413	0.405 - 0.420	Azul oscuro
9	0.443	0.438 - 0.448	Azul claro
10	0.488	0.483 - 0.493	Azul
11	0.531	0.526 - 0.536	Verde
12	0.551	0.546 - 0.556	Verde
13	0.667	0.662 - 0.672	Rojo
14	0.678	0.673 - 0.683	Rojo
15	0.748	0.743 - 0.753	Infrarrojo cercano
16	0.865	0.862 - 0.877	Infrarrojo cercano
17	0.905	0.890 - 0.920	Infrarrojo cercano
18	0.936	0.931 - 0.941	Vapor de agua
19	0.940	0.915 - 0.965	Vapor de agua
20	3.750	3.660 - 3.840	Infrarrojo de onda media
21	3.959	3.929 - 3.989	Infrarrojo de onda media
22	3.959	3.929 - 3.989	Infrarrojo de onda media
23	4.050	4.020 - 4.080	Infrarrojo de onda media
24	4.465	4.433 - 4.498	Infrarrojo de onda media
25	4.520	4.482 - 4.549	Infrarrojo de onda media
26	1.375	1.360 - 1.390	Infrarrojo de onda corta
27	6.750	6.535 - 6.895	Vapor de agua
28	7.325	7.175 - 7.475	Vapor de agua
29	8.550	8.400 - 8.700	Infrarrojo de onda larga
30	9.730	9.580 - 9.880	Ozono
31	11.030	10.780 - 11.280	Infrarrojo de onda larga
32	12.020	11.770 - 12.270	Infrarrojo de onda larga
33	13.335	13.185 - 13.485	Infrarrojo de onda larga
34	13.635	13.485 - 13.785	Infrarrojo de onda larga
35	13.935	13.785 - 14.085	Infrarrojo de onda larga
36	14.235	14.085 - 14.385	Infrarrojo de onda larga

Tabla 2.2: Bandas del espectroradiómetro MODIS

la banda 31 como CTT. Luego de haber definido CTP por *CO2-Slicing* o CTT por *IR-Window*, se encontrarán CTH y CTT o CTP y CTH con el apoyo de perfiles verticales obtenidos de NWP. Sin embargo, el trabajo de Baum et al. (2012) detalla como el algoritmo CTPCPA considera, para la Colección 6, el caso particular de las inversiones térmicas en los niveles bajos de la atmósfera. Esta consideración se debe a que el trabajo de Holz et al. (2008) comprobó sesgos mayores a 2 km entre la Colección 5 de CTH de MODIS y respectivos valores de CTH derivados del LIDAR de Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) (de ahora en adelante CTH CALIPSO por simplicidad). Estas diferencias son atribuidas a la aplicación de la técnica *IR-Window* en condiciones de inversiones térmicas marítimas no resueltas en los perfiles auxiliares obtenidos mediante NWP. Para esto, Baum et al. (2012) propone como solución la construcción mensual de gradientes verticales medios zonales de TB-IR en 11.03 μm aparentes (de ahora en adelante LRC31 por simplicidad). Estos gradientes son actualizados considerando observaciones concurrentes de CTH CALIPSO, TB-IR en 11.03 μm de MODIS y temperaturas de superficie de mar obtenidas por NWP ajustadas. De esta forma, cuando la técnica *IR-Window* resulta en valores CTP mayores a 600 hPa, se utiliza LRC31 para determinar CTH de la siguiente forma:

$$CTH = \frac{SST - BT_{31}}{LRC31} \quad (2.4)$$

Donde SST es la temperatura de superficie de mar ajustada, (BT_{31}) la TB-IR de la banda 31 ($11.03 \mu\text{m}$) y LRC31 el gradiente vertical medio zonal de TB-IR en $11.03 \mu\text{m}$.

Con respecto al desempeño del producto MOD06/MYD06, el trabajo de [Baum et al. \(2012\)](#) presenta una mirada general sobre la mejora alcanzada por la Colección 6 frente a la Colección 5 previa. Para esto recurre a evaluar la diferencia entre estimaciones CTH MODIS y CTH CALIPSO. En este contexto, [Baum et al.](#) presenta la mejora lograda en la Colección 6 por la inclusión del refinamiento LRC31 (entre otros), para el caso de nubosidad *single layer* y espesor óptico ≥ 0.5 . La Figura 2.2 es tomada del trabajo de [Baum et al.](#) y presenta la mejora de desempeño notoria en las regiones costeras occidentales de América del Norte, América del Sur, Europa y África, donde domina la persistencia de nubosidad del tipo Stratocúmulus. En el mismo sentido, el trabajo de [Mitra et al. \(2021\)](#) presenta un estudio pormenorizado de la Colección 6, específicamente para el producto MOD06 (derivado de MODIS TERRA). En su análisis, [Mitra et al.](#) compara los valores CTH MODIS con los CTH del International Space Station Cloud-Aerosol Transport System (CATS) en forma global y particular. El análisis particular presenta resultados desglosados según la altura de la nubosidad y su espesor óptico. Para evaluar el desempeño, [Mitra et al.](#) considera la moda del sesgo (*bias*) MODIS - CATS. Para definir la precisión el trabajo recurre al Full Width Half Maximum (FWHM), utilizando la relación gaussiana:

$$\sigma = \frac{FWHM}{2\sqrt{2\ln(2)}} \quad (2.5)$$

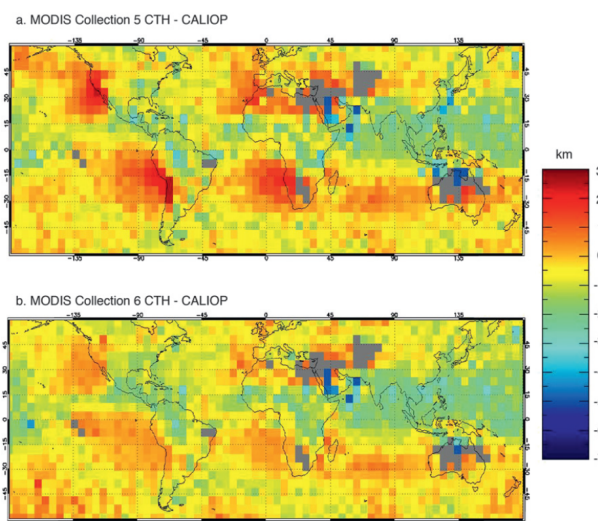


Figura 2.2: Mejoras de la Colección 6 del producto MOD06/MYD06 frente a la Colección 5. Se considera la diferencia entre CTH MODIS y CTH CALIPSO. Arriba se tiene Colección 5, abajo la Colección 6. Es notoria la mejora del producto en las regiones de persistencia de Stratocúmulus. Figura tomada de ([Baum et al., 2012](#)).

Una vez definidas las métricas de la evaluación, [Mitra et al.](#) explora el sesgo de CTH MODIS en distintas condiciones de altura de nubosidad. En primer lugar, la Tabla 2.3 presenta estas métricas para el caso particular de diferencias absolutas menores o iguales a 2.5 km y porcentaje de opacidad $CATS = 1$. De la tabla se resalta el comportamiento diferencial CTH MODIS con la altura de la nubosidad. Según [Mitra et al.](#), el producto ofrece resultados más exactos en nubosidades bajas (moda de sesgo 60 m) que en nubosidades medias y altas (-350 m y -950 m, respectivamente). En todos los casos, la precisión del producto se mantiene en valores casi constantes. Para el caso general, sin discriminar por niveles de nubosidad, la exactitud de CTH MODIS (en la plataforma TERRA) es de -540 m con una precisión de 740 m. Como se mencionó, el trabajo ([Mitra et al., 2021](#)) presenta una validación desglosada de CTH MODIS obtenido desde la plataforma TERRA. Esta validación excede el espíritu de esta introducción, de todas formas mencionaremos resultados que permiten

comprender mejor el desempeño de CTH MODIS. En primer lugar, los autores mencionan que el suyo es el primer estudio global de este tipo para este producto y en esa plataforma. El desglose según altitud de nubosidad, espesor óptico y nubosidad multi-capas es de gran valor. Este estudio pone los resultados en contexto del enfoque *top-down* llevado adelante por el algoritmo CTPCPA, aplicando sucesivamente las técnicas de *CO2-Slicing*, *IR-Window* y *LRC31-IR-Window* presentadas. De esta forma, [Mitra et al.](#) concluye que para la nubosidad de tipo bajo, donde domina la técnica *IR-Window*, el resultado es altamente dependiente del espesor óptico, teniendo sesgos de -440 m con precisiones de 600 m para el caso de nubosidad baja con alto espesor óptico. Mientras que para el caso de nubes bajas de bajo espesor óptico se tienen sesgos de 500 m con precisiones de 430 m. En su análisis, [Mitra et al.](#) explica este comportamiento dispar partiendo de que para la nubosidad más transmisiva (espesor óptico bajo) la técnica se ve contaminada por la consideración de temperaturas más representativas de la superficie. Para el caso de nubes con espesor óptico alto, más emisivas, proponen que existe una desviación mayor entre el gradiente vertical de temperatura y el LRC31 supuesto. En contrapartida, se presenta el caso de la nubosidad de tipo alto, donde domina la técnica *CO2-Slicing*. Para el caso de las nubes altas de espesor óptico alto, se tienen sesgos de -280 m con precisiones de 730 m. Mientras que para el caso de nubes altas con bajo espesor óptico, se tienen sesgos de -1200 m con precisiones de 1020 m. El trabajo justifica este comportamiento diferencial diciendo que las nubes del tipo Cirrus que presentan un espesor geométrico considerable, pero un bajo espesor óptico, escapan a la hipótesis de la técnica *CO2-Slicing*.

	Moda del sesgo (m)	Precisión (m)
Toda nubosidad	-540	690
Nubosidad alta (CATS CTH >10 km)	-950	740
Nubosidad media (10 km >CTH >5 km)	-350	690
Nubosidad baja (CATS CTH <5 km)	60	660

Tabla 2.3: Resultados de ([Mitra et al., 2021](#)) según tipo de nubosidad, para el caso diferencias absolutas menores o iguales 2.5 km y porcentaje de opacidad CATS = 1. Se observa que a medida que la altura de la nubosidad aumenta, aumenta el sesgo de CTH MODIS respecto a CATS, mientras que la precisión se mantiene relativamente constante

Concluyendo la presentación MODIS, es oportuno señalar que este instrumento ofrece más de 20 años ininterrumpidos de observación del sistema terrestre, en particular de la nubosidad. Los productos MODIS son ampliamente utilizados y sirven de referencia en el campo de la observación satelital de nubosidad con medios pasivos. Diversas campañas de validación se han llevado adelante y el refinado de las colecciones de productos permite actualizar el desempeño de estos a medida que avanza el conocimiento en el campo. Por esto, el producto CTH de MODIS tuvo un rol fundamental en este trabajo, al ser la referencia para el entrenamiento y validación de los clasificadores desarrollados que se presentan. Sin embargo, las herramientas aquí desarrolladas son satelitales puras, es decir, no utilizan como entrada otra información que no sean las imágenes de satélite de GOES16. Por esto, al trabajar sobre imágenes GOES16, los clasificadores desarrollados pueden actualizarse cada 10 minutos. De esta forma, se obtiene un producto con la misma frecuencia temporal y resolución espacial que las imágenes GOES16. Al contrario, el producto CTH de MODIS utiliza información auxiliar de NWP y RTM, además de las observaciones realizadas por el instrumento. Además, para el caso del territorio uruguayo, TERRA lo sobrevuela no más de 3 veces al día. Por otra parte, el producto CTH de MODIS presenta una demora de 24 horas entre el pasaje del satélite y su disponibilización al público.

Advanced Baseline Imager

El instrumento ABI es uno de los instrumentos principales a bordo del satélite geoestacionario GOES16 de la NOAA ([NOAA, 2018](#)). Se encuentra posicionado en la órbita geosincrónica en la longitud -75°W, observando de forma ininterrumpida el disco planetario centrado en las Américas. ABI consiste en un radiómetro de 16 canales espectrales y ofrece una resolución horizontal bajo el nadir del satélite de 500 metros (en el canal 2), 1000 metros (en los canales 1, 3-5) y 2000 metros (en los canales 6 a 16) ([Schmit et al., 2017](#)). Los detalles de estos canales se encuentran en la Tabla 2.4. Existe una variedad de productos derivados de las observaciones realizadas por ABI. En particular para el presente trabajo son de interés tres productos de tope de nube obtenidos tras la

aplicación del algoritmo ABI Cloud Height Algorithm (ACHA): CTP, CTT y CTH. Según detalla [A. K. Heidinger et al. \(2020\)](#), los productos de nubosidad derivados de ABI mediante ACHA son insumo para otros desarrollos de la NOAA. Entre los productos que aprovechan la salida de ACHA se cuentan productos enfocados a la microfísica y óptica de nubes, estimación de viento por satélite e inicialización de nubosidad de los NWP que corre NOAA. El reporte técnico de [A. Heidinger \(2012\)](#) presenta una descripción detallada del algoritmo ACHA. En particular, [A. Heidinger](#) resalta que el algoritmo utiliza los canales 14 (11.21 μm), 15 (12.29 μm) y 16 (13.28 μm) de ABI y utiliza un enfoque basado en la combinación de las técnicas *CO2-Slicing* y *Split-Window*. Esta técnica combinada incluye además un proceso de estimación óptima, también conocido como *1DVAR*, ampliamente utilizada en los procesos de *retrievals*² satelitales y asimilación de datos para NWP. El desarrollo de esta técnica combinada excede el alcance de esta introducción, sin embargo, se comentará brevemente el proceso para presentar el vector objetivo de la optimización que realiza ACHA. Para un desarrollo detallado se puede consultar el reporte técnico de [A. Heidinger \(2012\)](#) en su sección 1.11.2 y el trabajo de [Maahn et al. \(2020\)](#) sobre estimación óptima. En particular, la técnica *1DVAR* busca minimizar una función de costo Φ que integra las observaciones realizadas por ABI con los *retrievals* de CTP, CTT y CTH. Además, la función de costo Φ integra una primera aproximación a estos *retrievals* en un proceso iterativo que busca la minimización de Φ . Dicho esto, el vector y de observaciones tiene la forma:

$$y = (\text{TBC14}, \text{DTB}(\text{C14-C15}), \text{DTB}(\text{C14-C16})), \quad (2.6)$$

donde TBC14 es la TB-IR del canal 14 y las diferencias DTB(C14-C15) y DTB(C14-C16) corresponden a la resta de las TB-IR del canal 14 con los canales 15 y 16 respectivamente. Como se dijo, el objetivo de la optimización *1DVAR* es encontrar el vector x que minimice la función de costo Φ :

$$x = (T_c, E(11.2\mu\text{m}), \beta \left(\frac{12\mu\text{m}}{11\mu\text{m}} \right)), \quad (2.7)$$

donde T_c corresponde a la CTT, $E(11.2\mu\text{m})$ corresponde a la emisividad de la nube en los 11.21 μm y $\beta \left(\frac{12\mu\text{m}}{11\mu\text{m}} \right)$ es el cociente entre las emisividades de la nube en los 12.29 μm y 11.21 μm . La justificación de la elección de los *retrievals* reunidos en el vector x surge del desarrollo de la ecuación de transferencia radiativa descrito en el reporte de [A. Heidinger](#). Este desarrollo es similar al utilizado para describir la técnica de *CO2-Slicing* que puede encontrarse en la introducción de la presente sección 2.3.1, porque en esencia es la aplicación de esta técnica en el contexto del enfoque combinado con *Split-Window* que emplea ACHA. Continuando con la descripción, los *retrievals* serán determinados en la optimización de Φ , mediante un proceso iterativo limitado a no más de 10 iteraciones. Terminada la iteración, se tendrá CTT y las emisividades características del píxel bajo estudio. Luego, y mediante el uso de información auxiliar obtenida de salidas de modelos NWP y modelos RTM, ACHA obtiene CTP, CTT y CTH.

El reporte técnico de [A. Heidinger \(2012\)](#) detalla la validación del algoritmo ACHA. Para esto, utiliza como *proxy* de ABI al instrumento Spinning Enhanced Visible and Infrared Imager (SEVIRI) embarcado en el satélite geoestacionario Meteosat Second Generation (MSG). Antes de pasar a la validación, el reporte se enfoca en comparar ACHA con el producto *Cloud Top Properties* MYD06 de MODIS AQUA. Para esto, [A. Heidinger](#) describe que fue necesario resolver la concurrencia espacio temporal de las dos plataformas de observación. El resultado de esta comparación se desglosa para CTP y CTT. Para el caso de CTP se tiene un sesgo (*bias*) ACHA-MODIS de -23.48 hPa, con una desviación estándar de 80 hPa. Además los productos muestran una correlación de 0.94. Para el caso de CTT se tiene un sesgo ACHA-MODIS de -3.6 K, con una desviación estándar de 8.9 K. En este caso los productos muestran una correlación de 0.91. En esta comparación se observa que en promedio MODIS resuelve la posición de los topes nubosos en niveles inferiores de la atmósfera con respecto a ACHA. Al momento de evaluar estos resultados, [A. Heidinger](#) comenta que, por más que una comparación entre dos instrumentos pasivos no puede ser considerada una validación, el más que aceptable acuerdo entre los dos algoritmos es un indicador de la habilidad de ACHA. Dicho esto, la validación del algoritmo ACHA se lleva adelante considerando las medidas CTH, CTT y CTP obtenidas por CALIPSO. Nuevamente SEVIRI es elegido como *proxy* de ABI. Para el

²En el contexto de la observación satelital, se denomina *retrieval* al proceso de derivar o calcular valores de variables geofísicas o atmosféricas a partir de datos de sensores satelitales.

Canal ABI	Longitud de Onda (μm)	Ancho espectral (μm)	Nombre Descriptivo
1	0.47	0.45 - 0.49	Azul
2	0.64	0.60 - 0.68	Rojo
3	0.864	0.847 - 0.882	Vegetación
4	1.373	1.366 - 1.380	Cirros
5	1.61	1.59 - 1.63	Nieve/Hielo
6	2.24	2.22 - 2.27	Tamaño de partículas de nubes
7	3.90	3.80 - 3.99	Ventana de onda corta
8	6.19	5.79 - 6.59	Vapor de agua de nivel superior
9	6.93	6.72 - 7.14	Vapor de agua de nivel medio
10	7.34	7.24 - 7.43	Vapor de agua de nivel bajo/medio
11	8.44	8.23 - 8.66	Fase de la cima de las nubes
12	9.61	9.42 - 9.80	Ozono
13	10.33	10.18 - 10.48	Ventana de onda larga limpia
14	11.21	10.82 - 11.60	Ventana de onda larga
15	12.29	11.83 - 12.75	Ventana de onda larga sucia
16	13.28	12.99 - 13.56	Dióxido de carbono

Tabla 2.4: Descripción de los canales ABI

estudio se consideraron nuevamente el sesgo (ACHA-CALIPSO) y la desviación estándar del error. La validación se llevó adelante en todas las estaciones del año y en todo el dominio de observación de SEVIRI. El estudio del sesgo y la desviación estándar se presenta en función de la CTH y emisividad determinadas por CALIPSO, esto permite entender el comportamiento de ACHA frente a distintos tipos de nubosidad. El análisis de resultados considera el sesgo y la desviación estándar como exactitud y precisión de ACHA. De esta forma, la validación de ACHA ofrece una exactitud de -0.0002 km y una precisión 0.94 km en CTH. Para el caso de CTT se tiene una exactitud de -0.2 K y una precisión de 4.8 K. Para el caso de CTP, se tiene una exactitud de -0.02 hPa y una precisión de 94 hPa.

Producto ACHA	Exactitud	Precisión
CTH	-0.0002 km	0.94 km
CTT	-0.22 K	4.75 K
CTP	-0.02 hPa	94 hPa

Tabla 2.5: Resultados de la validación de ACHA contra CALIPSO (A. Heidinger, 2012).

El producto CTH de ABI es también especialmente relevante para este trabajo, dado que utiliza imágenes ABI como insumo igual que el presente trabajo. No obstante y como se mencionó anteriormente, este producto precisa de información auxiliar de NWP y RTM para ser calculado. En cualquier caso, al tratarse de un producto GOES16 operativo, que obtiene estimaciones de CTH, es de interés conocer su desempeño y compararlo contra la herramienta desarrollada en este trabajo. En este sentido, la evaluación de este producto provee una primera referencia de desempeño.

2.3.2. Sistemas de observación activos

Los instrumentos activos se caracterizan por emitir pulsos de radiación electromagnética y detectar los cambios en la señal reflejada en función del tiempo. Como ejemplo de este tipo de instrumentos se tienen los sistemas RADAR y LIDAR. Las características generales de estos instrumentos se presentaron en la sección 2.2, en el marco de la observación terrestre de la nubosidad. La misma descripción física aplica para la observación con sistemas RADAR y LIDAR embarcados en satélites y observando hacia el nadir (o casi nadir) de estos. Sin embargo, deben considerarse las restricciones energéticas y de espacio, entre otras, que condicionan el diseño de un instrumento de observación satelital. Además, los sistemas RADAR y LIDAR comparten varias limitaciones ya que la base física de funcionamiento de estos equipos es esencialmente la misma. En particular se destaca la atenuación de las señales emitidas al interactuar con objetivos altamente reflectivos y de esta forma se limita la observación más allá de esta interacción. También se tiene el fenómeno conocido como *clutter*, que

consiste en ecos no deseados producto de la interacción con objetos de variados orígenes (montañas, bandadas de pájaros o grupos de insectos, construcciones, aerogeneradores, etc).

A continuación se presentan dos ejemplos destacados de instrumentos activos a bordo de misiones satelitales de los cuales se obtienen medidas que permiten caracterizar la nubosidad y en particular obtener medidas de CTH. Como se mencionó anteriormente, existen estaciones terrestres equipadas con este tipo de instrumentos activos, de las cuales se obtienen productos de nubosidad u otros componentes atmosféricos, como los aerosoles. No se detallará en este documento casos particulares de este tipo de estaciones, sino que se pondrá en consideración dos instrumentos a bordo de misiones satelitales que sirven de ejemplo para este trabajo.

Cloud Aerosol Lidar with Orthogonal Polarization

En abril del año 2006 fue puesto en órbita el instrumento Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP). Este equipo consiste en un LIDAR embarcado en el satélite heliosincrónico de órbita baja CALIPSO, el cual ofrece un tiempo de revisita de 16 días. Las observaciones de CALIOP se centran en la detección de aerosoles y nubosidad. En el trabajo de [Winker et al. \(2009\)](#) se describen las características técnicas de CALIOP, donde se destaca que los pulsos láser prospectivos son generados en las longitudes de onda de 1064 nm y 532 nm. Además, [Winker et al.](#) detalla que el pulso de 532 nm está polarizado ortogonalmente, rasgo distintivo que permite a las observaciones de CALIOP conocer características de las partículas que reflejan el pulso láser, tales como: forma, estado u orientación, entre otras. La resolución vertical de CALIOP es de 30 metros desde la superficie hasta los 8.2 km de la tropopausa, con una resolución horizontal de 0.33 km. A partir de los 8.2 km y hasta los 20.2 km ofrece una resolución vertical de 60 metros y una resolución horizontal de 1 km. En [Vaughan et al. \(2004\)](#) se tiene el desarrollo detallado de la estrategia observacional CALIOP. Además, [Vaughan et al.](#) presenta la complejidad de la observación satelital con sistemas LIDAR, donde la relación entre señal y ruido en objetivos altamente transmisivos (como las nubes del género Cirrus) obliga a realizar varias medidas para establecer la ubicación de las partículas. El trabajo [Winker et al. \(2007\)](#) destaca el amplio rango dinámico de detección de señales que logra CALIOP. Esto es un elemento diferenciador de CALIOP con respecto a experiencias previas en el campo, ya que le permite detectar la gran variedad de señales reflejadas por los objetivos circunstanciales. Debe considerarse que las señales reflejadas pueden ser tan potentes como la emitida para el caso objetivos tales como la superficie o la nubosidad de gran espesor óptico; en cambio, los aerosoles y la nubosidad de bajo espesor óptico devuelven señales débiles que complejizan el proceso de detección. En su trabajo, [Vaughan et al.](#) también detalla los productos de nubosidad y aerosoles derivados de CALIOP. En particular, el producto *Vertical Feature Mask (VFM)* es de interés para el presente trabajo. A continuación se presenta una breve descripción de este producto.

El producto *VFM* permite conocer la ubicación y tipo de las partículas detectados por CALIOP ([Winker et al., 2009](#)). El trabajo de [Vaughan et al.](#) detalla el producto, describiéndolo como "*un archivo booleano de la atmósfera*", del cual se puede determinar la distribución espacial de la nubosidad y los aerosoles. En el trabajo de [Lu et al. \(2021\)](#) se presenta el uso del producto *VFM*, del cual se deriva la Altura de Base de Nube (CBH). En condiciones de nubosidad de espesor óptico alto, [Lu et al.](#) menciona que, dada la atenuación que sufre la señal LIDAR de CALIOP, es posible recuperar CTH pero no CBH del producto *VFM*. Por esto, [Lu et al.](#) propone una estrategia de asignación de CBH basada en el análisis del entorno. De esta forma, se exploran los perfiles cercanos en búsqueda de nubosidad cercana de menor espesor óptico que permita recuperar CBH y completar el perfil requerido. Esta estrategia es contrastada contra observaciones coincidentes de ceilómetros en superficie. De esta forma, completan el faltante de datos CBH, logrando un conjunto de datos del espesor geométrico de la nubosidad observada por CALIOP que analizan.

Cloud Profiling Radar

El instrumento Cloud Profiling Radar (CPR) fue un RADAR banda W embarcado en el satélite heliosincrónico y de órbita baja CloudSat. Fue el primer RADAR satelital de su tipo en ser desplegado específicamente para el estudio de la nubosidad a escala global [Stephens et al. \(2008\)](#). La frecuencia del pulso emitido por CPR era de 94 GHz, en la región del espectro electromagnético correspondiente a las microondas. El satélite CloudSat (transportando al CPR) fue puesto en órbita en abril del año

2006. Con una resolución vertical de 500 metros y una resolución horizontal de 1.4 km, permitió generar registros continuos de reflectividad vertical en la atmósfera, hasta el fin de su misión en diciembre de 2023 por problemas con su fuente energética. De las observaciones de CPR se han derivado una variedad de productos, para este trabajo son de interés los productos 2B-CLDCLASS-R05 y 2B-GEOPROF-R05 que se detallan a continuación.

En primer lugar, el producto 2B-CLDCLASS-R05 proporciona una clasificación de nubosidad derivada de las observaciones de CPR. Esta clasificación sigue la clasificación del AIN y categoriza las porciones de atmósfera observadas en 8 categorías: Stratus, Stratocumulos, Cumulus, Nimbostratus, Altocumulus, Altostratus, Cumulonimbus, y nubes altas (Cirrus y Cirrostratus). Para lograr esta clasificación, el algoritmo base de 2B-CLDCLASS utiliza además de la reflectividad obtenida de CPR, registros de precipitación y perfiles verticales de temperatura obtenidos del modelo European Centre for Medium-Range Weather Forecasts (ECMWF), además de un modelo topográfico de la superficie. En [Sassen y Wang \(2008\)](#) se tiene un estudio del producto 2B-CLDCLASS-R05 y una comparación del primer año de datos del producto contra dos climatologías de clasificación de nubosidad existentes, una derivada de observaciones de estaciones meteorológicas de superficie y otra la climatología derivada del International Satellite Cloud Climatology Project (ISCCP)³. El trabajo de [Sassen y Wang](#) permite caracterizar la clasificación del producto 2B-CLDCLASS-R05, reconociendo limitaciones en el reconocimiento de ciertos tipos de nubosidad. Específicamente se detalla que la detección de la nubosidad de las categorías Cumulus, Altocumulus y Cirrus de bajo espesor óptico se ve afectada por las limitaciones de sensibilidad y resolución vertical de CPR. Sin embargo, se resalta que la comparación considera un año de datos contra climatologías, y que los resultados podrían mejorar considerando futuros ciclos anuales.

Por otra parte, el producto 2B-GEOPROF-R04 identifica los niveles verticales de alta reflectividad asociados a hidrometeoros. De esta forma identifica la estructura vertical de las nubes. El trabajo de [Marchand et al. \(2008\)](#) describe el algoritmo base del producto 2B-GEOPROF-R04 y compara este producto contra observaciones respectivas de CALIOP. La conclusión general de [Marchand et al.](#) subraya la complementariedad de la estrategia observacional basada en la combinación de CPR y CALIOP. También se resaltan la reducción de la sensibilidad de CPR en los volúmenes cercanos a la superficie, debido al fenómeno *clutter*.

El despliegue del CPR sirvió para comprender las fortalezas y debilidades de esta estrategia observacional. Al día de hoy se encuentra activa la misión *Earth Clouds, Aerosols and Radiation Explorer (EarthCARE)* que aprovecha el aprendizaje logrado con CPR. De esta forma, *EarthCARE* combina un LIDAR y un RADAR Doppler (además de un radiómetro de amplio espectro y otro radiómetro multiespectral).

2.3.3. Comparación entre sistemas activos y sistemas pasivos

La combinación de observaciones de sistemas activos y pasivos satelitales ha sido fundamental para mejorar la comprensión de la nubosidad. Los sistemas activos ofrecen alta precisión y capacidad para detectar nubes, especialmente las de bajo espesor óptico, al tener mayor capacidad de detección a través de la interacción con el objeto. Estos instrumentos, por motivos de costo y constructivos, focalizan su detección en algunas longitudes de onda, con misiones y formas de funcionamiento especializadas a la propiedad o característica a detectar. En cambio los sistemas pasivos cubren una mayor región de espectro pudiendo observar más fenómenos y características, pero no siempre con la misma precisión, y, especialmente los embarcados en satélites geostacionarios, permiten una cobertura global y continua en el tiempo. La estrategia observacional llevada adelante en lo que se conoce como el *A-Train* ([Stephens et al., 2002](#)) integra datos de los instrumentos CALIOP, CPR y MODIS, permitiendo estudios comparativos. Así, fue posible investigar las diferencias entre los distintos sistemas de observación enfocados en nubosidad y medidas de referencia en tierra, en el contexto de los diferentes principios físicos de cada instrumento.

El trabajo de [Minnis et al. \(2008\)](#) compara los CTH derivados de MODIS y de CALIOP. Por más que estos productos se denominen de forma similar, se remarca que refieren a cantidades físicas diferentes, es decir, las alturas que estiman no se corresponden directamente. En el caso del instrumento CALIOP, este es un LIDAR altamente sensible a los cambios en la reflexión del haz láser que

³Primer proyecto internacional del World Climate Research Program (WCRP) iniciado en el año 1983 y enfocado en la climatología satelital de nubes, para esto combina información de diversos satélites.

genera el borde de una nube, identificando esta interfase con gran precisión. Por otra parte MODIS, al ser un instrumento pasivo, obtiene las alturas del centro radiativo efectivo de la nube, que no coincide exactamente con el tope de la nube. En el trabajo de [Minnis et al. \(2008\)](#) se destaca que típicamente este centro radiativo efectivo de la nube, reportado por MODIS, se ubica 1.6 km por debajo de la altura reportada por CALIOP. Además, estas discrepancias se hacen notables cuando las diferencias de sensibilidad de los instrumentos se intensifican, por ejemplo, MODIS tiene una baja sensibilidad a nubes del género Cirrus, que presentan espesores ópticos típicamente menores a 0.4 ([Ackerman et al., 2008](#)), mientras CALIOP es capaz de detectar nubes con espesores ópticos menores a 0.01 ([Winker et al., 2007](#)). En general, se tiene que MODIS puede subestimar la altura del tope de nube en aproximadamente 2.5 km con respecto a CALIOP ([Holz et al., 2008](#)).

2.4. Weather Research and Forecasting

El modelo WRF es una herramienta de NWP ampliamente utilizada, tanto en investigación como en predicciones meteorológicas operativas ([Powers et al., 2017](#)). Siendo en esencia un modelo de mesoescala, se ha utilizado típicamente en el estudio de fenómenos mesoescalares y sinópticos, tales como ciclones extratropicales, sistemas frontales, sistemas convectivos de mesoescala y huracanes. De esta forma, ha permitido aumentar la resolución espacial y temporal en estos eventos, resolviendo la física en escalas más pequeñas y logrando una simulación detallada a un costo computacional posible. El modelo y su plataforma de libre distribución fue originariamente desarrollado por un conjunto de instituciones estadounidenses, entre las que se destacan el National Center for Atmospheric Research (NCAR) y la NOAA. Aunque inicialmente fue pensado como una herramienta de apoyo a la investigación, actualmente el WRF también es utilizado ampliamente en forma operativa. Por esto, NCAR contempla ambos enfoques en los trabajos de coordinación del desarrollo del modelo que lleva adelante. Para esto, existen los grupos de expertos *WRF Developers' Committee* y *WRF Release Committee*. También existe el panel *WRF Physics Review Panel* que estudia nuevos esquemas o paquetes físicos propuestos por desarrolladores y recomienda su inclusión (o no) al *WRF Developers' Committee*. Desde sus inicios, el WRF acepta contribuciones de la comunidad, no restringiendo la colaboración a grupos específicos. Para que una contribución sea aceptada, deberá remitirse el código testeado de la contribución al *WRF Developers' Committee*. Esta apertura a la colaboración ha fortalecido el uso global del modelo.

Las simulaciones con WRF se realizan en dos etapas. En primer lugar se ejecuta el programa WRF Preprocessing System (WPS), donde se establecen los dominios espaciales de la simulación. Se establecen cuantos dominios se utilizarán, y si será solo uno o varios dominios anidados. También se definen los niveles verticales con los que se realizará la simulación, el periodo de tiempo a simular y, además, se incorpora información geográfica relevante del dominio, como la topografía y los usos del suelo. En esta etapa se elige también el núcleo dinámico de la simulación, siendo el núcleo dinámico Advance Research WRF (ARW) el más utilizado. Este programa también adapta las condiciones iniciales (a partir de análisis que deben ser proporcionadas previamente) al núcleo dinámico y dominios elegidos por el usuario; y prepara las condiciones de borde necesarias para la simulación. Como resultado se obtienen archivos que serán aprovechados en la segunda etapa, donde se ejecuta el programa WRF para resolver las ecuaciones dinámicas y lograr la simulación. Es aquí donde se definen las parametrizaciones de procesos físicos a utilizar, es decir, los esquemas elegidos para resolver los intercambios originados por procesos de escala menor a la escala de la simulación. El WRF incluye una amplia variedad de esquemas físicos para la parametrización de procesos ([Hong y Lim, 2006](#); [Lin et al., 1983](#)), siendo una característica destacable del modelo, ya que permite experimentar con varias parametrizaciones, habilitando el análisis de impacto de estas en el pronóstico ([Y. Zhang et al., 2012](#)). Particularmente, entre los procesos físicos de interés que se parametrizan en un estudio de pronóstico de nubosidad, se tienen las transferencias radiativas (tanto de onda corta como de onda larga), la microfísica de nubes, los intercambios convectivos, los intercambios de PBL y los intercambios de superficie. La Figura 2.3 muestra la interrelación de estos procesos físicos parametrizados, para el caso de WRF utilizando su núcleo dinámico ARW. Más adelante en este trabajo se presentarán las parametrizaciones elegidas para las simulaciones realizadas.

El trabajo de [Powers et al. \(2017\)](#) resume las versiones de WRF que se tiene en la actualidad. Entre estas se tiene: *WRF-Chem*, enfocado en simulaciones de química atmosférica, *WRF-Hydro* modelo hidrológico acoplado al WRF, *WRF-Fire* enfocado en la simulación de incendios forestales,

Direct Interactions of Parameterizations

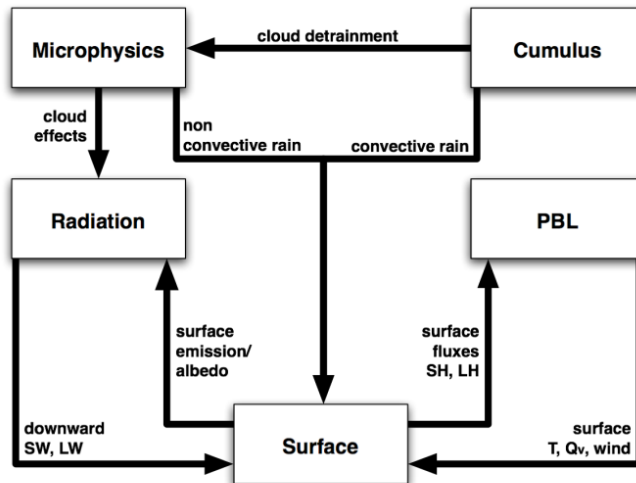


Figura 2.3: Interacción entre parametrizaciones para el WRF-ARW, tomada de [Dudhia \(2015\)](#)

HWRF ajustado específicamente para el pronóstico operativo de huracanes, *WRF-Urban* enfocado en predicciones meteorológicas urbanas, *Polar WRF* enfocado en el pronóstico en regiones polares y *WRF-Solar* aplicado al pronóstico de radiación solar para emprendimientos de energía solar. Además se tiene la adaptación *WRF-LES*, enfocada en la resolución de la turbulencia ("Large Eddy Simulation") y que permite trabajar con dominios cuya resolución horizontal es del orden de los 50 metros ([Liu et al., 2020](#)). Recientemente se tiene el desarrollo *Global-WRF*, una versión de WRF que funciona como un modelo global, permitiendo unificar el modelado de los procesos atmosféricos y sus interacciones desde la escala global a la regional, habilitando incluso a acoplar estos intercambios. En particular, [Y. Zhang et al. \(2012\)](#) comenta que puede utilizarse *Global-WRF* para proporcionar las condiciones iniciales y de borde al WRF regional, anidando las simulaciones. De esta forma se reducen imprecisiones que se originan al utilizar modelos globales y regionales cuyos enfoques dinámicos y físicos son inconsistentes.

Como se mencionó en el Capítulo 1, WRF es utilizado en Uruguay como herramienta de NWP operativa. En particular, es utilizado en el pronóstico de generación de parques eólicos y plantas fotovoltaicas que genera el Instituto de Mecánica de Fluidos e Ingeniería Ambiental (IMFIA) de la Facultad de Ingeniería de la UDELAR. Los trabajos de [Orteli \(2023\)](#); [Teixeira-Branco et al. \(2023\)](#); [Cazes Boezio y Ortelli \(2019\)](#); [De Almeida \(2019\)](#); [Porrini \(2017\)](#); [de Mello et al. \(2015\)](#) son ejemplo de esta línea de investigación y desarrollo.

2.5. Clasificación de nubosidad por aprendizaje automático

Los algoritmos de aprendizaje automático están diseñados para encontrar relaciones o patrones complejos dentro de un conjunto de datos previamente anotados ([Wang et al., 2020](#)). Esta anotación toma la forma de relación entre un conjunto de entradas y un conjunto de salidas. En el problema de clasificación, el conjunto anotado es una serie de categorías de salida asociadas a los datos de entrada. Utilizando estos datos ya clasificados es posible entrenar algoritmos de clasificación que puedan ser aplicados a nuevos datos aún sin clasificar. Esto los convierte en una herramienta destacada en el campo de la detección y clasificación. Los avances en la optimización de estos algoritmos y las mejoras en las capacidades de cómputo han permitido cada vez más la implementación de aplicaciones que buscan aprovechar estas herramientas en el campo de la meteorología y el monitoreo medioambiental y, en particular, en el estudio de nubosidad. Esta articulación entre áreas del conocimiento tiene el potencial de ser virtuosa cuando su desarrollo integra a profesionales de ciencias de la atmósfera y de ciencia de datos, o profesionales en la interfaz con formación adecuada en ambas disciplinas. A continuación se presentan ejemplos de experiencias internacionales y regionales en el área de la

clasificación de nubosidad por aprendizaje automático.

Para comenzar, el trabajo de [Min et al. \(2020\)](#) presenta la implementación de cuatro algoritmos de aprendizaje automático para la estimación de propiedades de nubes. Los datos de entrada utilizados para el entrenamiento de los algoritmos proceden de dos fuentes distintas. Por una parte recurren a los datos observados por el satélite geoestacionario Advanced Himawari Imager (AHI). La otra fuente de información que consideran son las salidas del modelo NWP Global Forecast System (GFS). El etiquetado del conjunto de datos es en base al producto CTH derivado de CALIOP. Los cuatro algoritmos implementados por [Min et al. \(2020\)](#) son: *K-Nearest-Neighbor (KNN)*, *Support Vector Machine (SVM)*, *Random Forest (RF)* y *Gradient Boosting Decision Tree (GBDT)*. El objetivo de ([Min et al., 2020](#)) es lograr un producto CTH, por lo que el trabajo no busca desarrollar clasificadores, sino ajustar una regresión paramétrica. El resultado de esta experiencia muestra un rendimiento superior a la estimación de CTH basada en métodos físicos. Además, comprueban que la combinación de los algoritmos de base física con herramientas de *Machine Learning* ofrecen un mejor desempeño que su aplicación individual.

El trabajo de [Wang et al. \(2020\)](#) presenta el entrenamiento de clasificadores de nubosidad basados en el algoritmo de aprendizaje automático *Random Forest*. El objetivo de la clasificación es detectar la nubosidad y clasificar su fase física (nubes líquidas, nubes de hielo, nubes mixtas). Para esto utiliza un conjunto de datos del Visible Infrared Imaging Radiometer Suite (VIIRS), tomando la etiqueta de clasificación a partir de las observaciones de CALIOP. La evaluación de desempeño del clasificador obtenido fue comparable al desempeño de la clasificación derivada de VIIRS y MODIS, que emplean algoritmos de base física.

Un trabajo de especial relevancia regional es el desarrollado por el Centro de Previsão de Tempo e Estudos Climáticos (CPTEC) del Instituto Nacional de Pesquisas Espaciais (INPE). El trabajo de ([Ceballos y Negri, 2022](#)) resume esta experiencia, cuyo producto puede encontrarse en el sitio web de la División de Sistemas Ambientales (DSA) (<https://satelite.cptec.inpe.br/>). Los trabajos de [Bottino y Ceballos \(2006\)](#); [Oliveira et al. \(2008\)](#) presentan una descripción de la técnica empleada. Originalmente, esta herramienta se basaba en una clasificación automática no supervisada (clustering) con $K = 30$ clusters, aprovechando la RP del canal visible y la TB-IR del canal 4 (10.70 μm) del satélite GOES12. Luego de realizada la clasificación automática, a cada cluster reconocido se le es asignada una categoría de nubosidad. Para la asignación de categorías de nubosidad se recurre a especialistas entrenados en el reconocimiento de nubosidad por satélite. En la actualidad, y debido a la salida de servicio del satélite GOES12, este clasificador ha sido rediseñado para su uso con las imágenes del instrumento ABI del GOES16. En [Ceballos y Negri \(2022\)](#) se presenta la metodología actualizada y se menciona que la base del algoritmo de clasificación continúa siendo similar al enfoque original. En particular, [Ceballos y Negri](#) concluye que el desempeño de esta técnica en la detección de nubosidad es similar al producto CM de NOAA. Este producto ha recibido sucesivas mejoras y, si bien es operativo y ampliamente utilizado en nuestra región, no ha sido a la fecha validado por observaciones terrestres sistemáticas u otros productos satelitales independientes a GOES16.

Otra experiencia regional en el área se presenta en el trabajo argentino de [Yuchechen et al. \(2020\)](#). En este estudio, se implementa un método de *clustering* basado en la combinación de los algoritmos *k-means* y *k-means++* sobre imágenes diarias de anomalías estandarizadas de la TB-IR del canal 4 (10.70 μm) obtenidas por el satélite GOES-13. El trabajo se centra en el período comprendido entre el 1 de diciembre de 2010 y el 30 de noviembre de 2016, tomando como área de interés el continente sudamericano y los océanos adyacentes. En particular, presentan una regresión lineal entre la TB-IR y CTH (obtenido de reanalysis), para así estimar CTH utilizando TB-IR como predictor. En el trabajo se estima un número óptimo de cuatro clusters ($K = 4$), cada uno asociado a diferentes tipos de nubes. Generalmente, cuanto menor es el orden del cluster, mayor es la altura de las nubes que representa. El cluster 4 representó el tipo de nubes bajas, con valores de CTH hasta los 2 km. Este fue el grupo más dominante, representando en promedio, el 32% de los píxeles. En cuanto a los clusters 2 y 3, se resalta que se confunden, en el sentido de que ambos tienen valores de CTH que abarcan la tropósfera media. En conjunto, los clusters 2 y 3 representaron, en promedio, el 62% de los píxeles. Finalmente, el cluster 1 fue el menos dominante (6% de los píxeles en promedio), representando nubes con los CTH más altos y mostrando las mayores variabilidades. Los autores mencionan que el cluster 1 agrupa nubes de diferentes propiedades microfísicas y que la metodología de clustering por sí sola no pudo resolver en forma adecuada estos tipos de nubes, señalando que la inclusión de parámetros adicionales en el análisis (por ejemplo, espesores ópticos) podría ayudar a refinar el método de clasificación. También realizan un análisis detallado vinculando los valores

CTH estimados, con una gran variedad de índices de circulación atmosférica.

Esta serie de trabajos representan los antecedentes más cercanos al presente trabajo. Se destaca que los antecedentes regionales son escasos. También se tiene que los métodos, información de entrada, y objetivos de estimación son bien distintos. Se destaca que la validación de los métodos con medidas terrestres es, sorprendentemente, bastante más escasa, favoreciéndose en general la intercomparación entre plataformas satelitales.

Página intencionalmente dejada en blanco.

Capítulo 3

Construcción de la base de datos

Este capítulo presenta la confección de la base de datos que es insumo de este trabajo y su análisis. Se detallan las fuentes de información aprovechadas y se justifica su uso. El desarrollo de clasificadores basados en algoritmos de aprendizaje automático implica la consolidación de una base de datos correctamente etiquetada que permita el entrenamiento y validación de estos algoritmos. Esto requirió consistir bases de datos de distinto origen para que tengan correspondencia espacial y temporal. Se presentan las características de la base de datos lograda y una exploración primaria.

3.1. Fuentes de información

La fuente primaria de datos de este trabajo son las imágenes multiespectrales generadas por el ABI embarcado en GOES16. Estas imágenes y sus derivados componen el conjunto principal de características de la BDT (en el campo del aprendizaje automático, las variables de entrada que se ingresan a un algoritmo son denominadas “características”). Para esto, se hizo uso del archivo de imágenes satelitales GOES16 del LES. Además se utiliza el producto CTH de MODIS como verdad de referencia (ground truth) para el etiquetado de la BDT, y así conseguir un conjunto etiquetado entrada-salida para el entrenamiento y evaluación de los clasificadores. El producto CTH de ABI es incorporado como base de desempeño, es decir como una referencia de desempeño para este tipo de productos. En una segunda etapa este producto es incorporado como característica de entrada, por lo que la función del clasificador es hacer un refinamiento (postproceso) de dicho producto.

3.2. Características primarias: imágenes del ABI GOES16

La Tabla 3.1 resume los canales ABI elegidos para integrar la BDT. Dadas las características físicas del problema y el estudio de antecedentes, se decidió el uso de los canales visible e infrarrojos indicados en la Tabla.

Código de Característica	Descripción
C02-RP	Reflectividad Planetaria derivada del Canal 02
C13-TB	Temperatura de Brillo derivada del Canal 13
C14-TB	Temperatura de Brillo derivada del Canal 14
C15-TB	Temperatura de Brillo derivada del Canal 15
C16-TB	Temperatura de Brillo derivada del Canal 16

Tabla 3.1: Canales de ABI integrados a la BDT como características principales.

El GOES16 tiene dos canales en el espectro visible (ver Tabla 2.4), uno más hacia el azul (C01) y otro más hacia el rojo (C02), de los cuáles este segundo es el más utilizado dado que tiene mayor resolución espacial. EL canal C02 es el utilizado en este trabajo, en parte por lo anterior, pero principalmente porque el LES no cuenta con un histórico del canal C01. La potencialidad de incorporar

el canal visible C01 a este problema no ha sido explorada en este trabajo, y queda para trabajo futuro. La radiancia observada por el ABI en los canales dentro del espectro solar, como es el caso de los canales en el visible, es transformada en reflectividad del sistema Tierra-Atmósfera (denominado RP, reflectividad planetaria). Este proceso se realiza conociendo la distribución sobre el tope de la atmósfera de la irradiancia solar (en plano horizontal) que llega desde el Sol en cada instante y la respuesta espectral del canal del ABI, y se asume una reflexión isotrópica de la radiancia solar en el sistema Tierra-Atmósfera (Alonso-Suárez, 2017). La reflectividad es mayor en condiciones de nubosidad, siendo la mayor parte del valor en comparación a la retro-dispersión en el resto de los constituyentes atmosféricos, y en condiciones de cielo claro es menor, conteniendo información de reflexión en el fondo (suelo, masas de agua, nieve, etc.) y de reflexión por retro-dispersión en la atmósfera. En los canales infrarrojos donde solo se observa en forma significativa emisión terrestre, como es el caso de los canales utilizados, se transforma la radiancia observada por el ABI en TB-IR. Esto se realiza asumiendo un comportamiento de cuerpo negro e invirtiendo la ley de Stefan-Boltzmann. La información de estos canales rara vez se distribuye en forma de radiancia, sino en TB-IR ya calculada y típicamente expresada en Kelvin (K).

La Figura 3.1 muestra la extensión latitudinal y longitudinal de la región considerada en este trabajo, ajustada al territorio uruguayo. También ilustra la grilla de trabajo elegida, con un factor de diezado de 10 (solo se muestran 1 de cada 10 muestras) por motivos de legibilidad. La resolución de las imágenes consideradas es de 0.035° en latitud y longitud, unos 4 km, que es el mínimo valor de resolución espacial común a los 5 canales ABI utilizados. La Figura 3.1 muestra por tanto una grilla espaciada por 0.35° . De esta forma la grilla de trabajo se compone de un total de 24795 puntos o píxeles en la región considerada, conformados por las 145 latitudes y las 171 longitudes que hacen a la grilla densa de 0.035° . El soporte espacial (la grilla) es un elemento importante en este trabajo, dado que hubo que construir una base de datos con información de distintas fuentes satelitales a una misma base espacial.

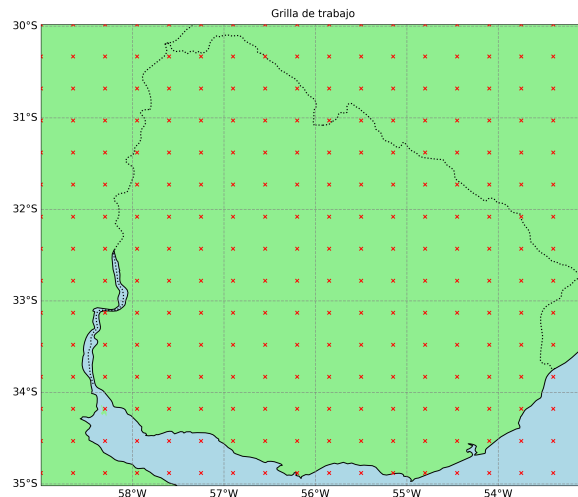
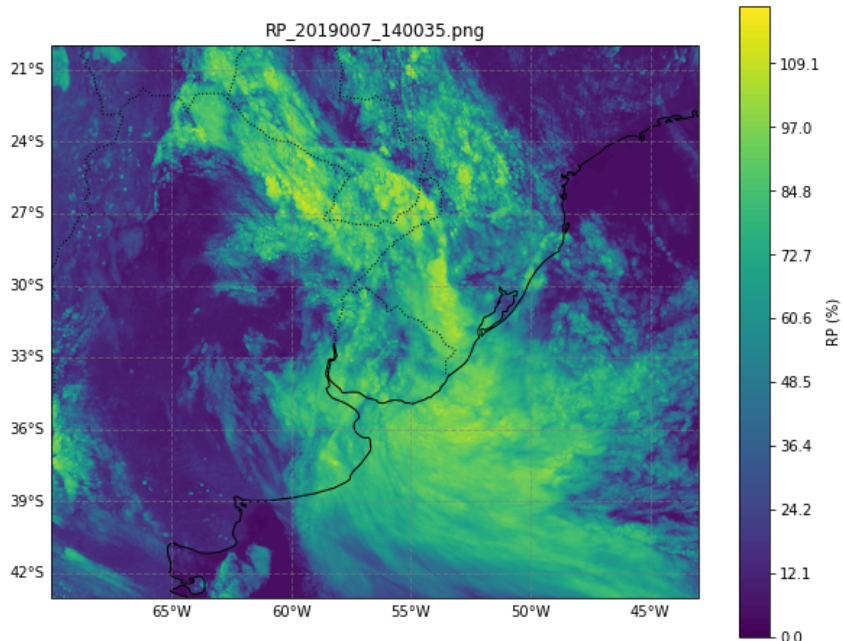


Figura 3.1: Grilla de trabajo ABI utilizada, con una resolución horizontal de 0.035° . Latitud máxima: -35.02° . Latitud mínima: -29.98° . Longitud máxima: -59.00° . Longitud mínima: -53.05° . Los puntos de grilla están subrepresentados con factor 10 a modo ilustrativo, por lo que están ubicados a cada 0.35° , y no cada 0.035° que es la grilla de trabajo.

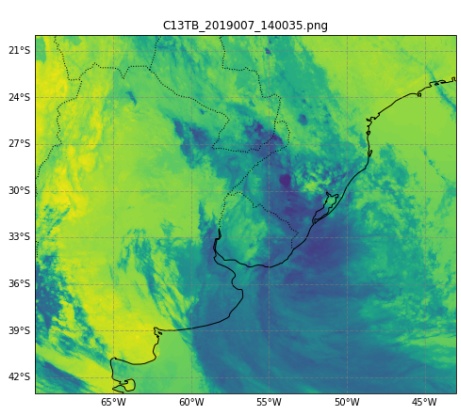
Las Figuras 3.2 (a), (b), (c), (d) y (e) muestran ejemplos de las imágenes de las características principales para el día 7 de enero de 2019 a las 14:00 hs UTC. En la Figura 3.2 (a) se muestra la imagen de reflectancia planetaria del canal C02. Como se puede observar, la presencia de nubosidad es fácilmente distinguible del fondo. Los píxeles de cielo despejado tienen típicamente reflectividades entre $\simeq 7-12\%$, dependiendo el tipo de fondo y sus características ópticas (tipo de suelo, río, océano)¹. Los valores más altos de reflectividad representan nubosidad. Notar que estas imágenes a veces muestran algunos pocos píxeles con valores por encima del 100% en nubes de gran espesor óptico, como es el caso de la figura mostrada. Esto puede deberse a la validez de la hipótesis isotrópica o pequeñas diferencias en la calibración del ABI. Típicamente estos pocos píxeles son limitados a

¹Este valor presenta también leves variaciones durante el día dependiendo de la validez de la hipótesis de reflexión isotrópica y las características del fondo. También puede tener variaciones en función de la presencia de aerosoles, que aumentan la retro-dispersión en el espectro visible.

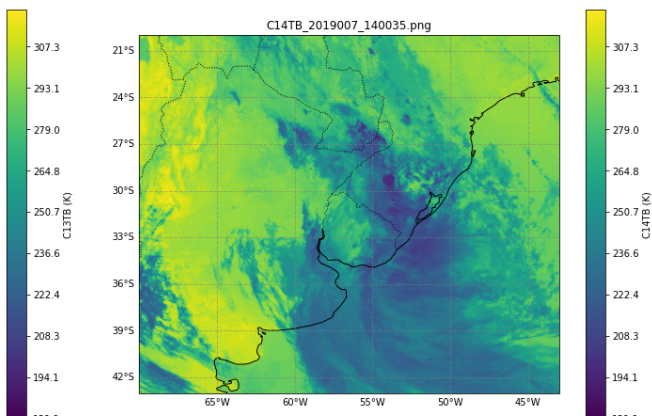
100%. Este topeo introduce una anomalía no física y visible en la estadística de píxeles que puede afectar el funcionamiento de algoritmos de aprendizaje automático. Debido a esto, se decidió en este trabajo no aplicar la limitación. El resto de las figuras (b, c, d, e) muestran las imágenes de los canales infrarrojos para la misma fecha y hora.



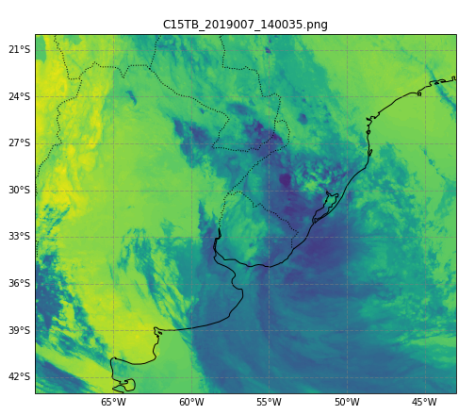
(a) Imagen C02-RP.



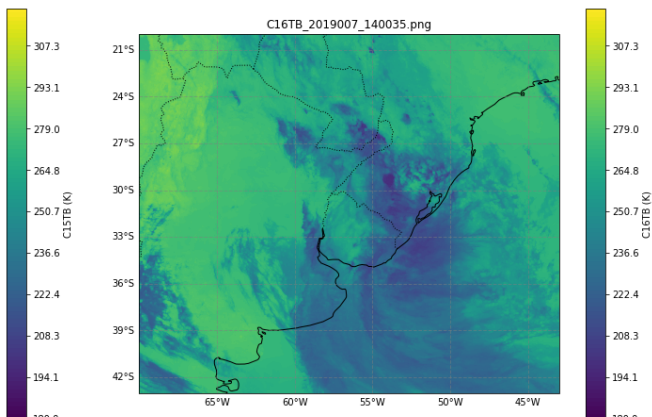
(b) Imagen C13-TB.



(c) Imagen C14-TB.



(d) Imagen C15-TB.



(e) Imagen C16-TB.

Figura 3.2: Ejemplo de imágenes ABI para el día 7 de enero de 2019 a las 14:00hs.

3.3. Características secundarias: derivadas del ABI GOES16

Las imágenes anteriores no fueron la única información que se utilizó como característica de entrada. A continuación se presentan una serie de características adicionales que fueron derivadas de las imágenes ABI. El objetivo de estas características fue cuantificar información sobre texturas o movimientos de la escena, que pudieran ser de utilidad al momento de la clasificación de altura de nubosidad. La utilidad de estas características fue una hipótesis que se testeó en este trabajo, sin antecedentes en la literatura.

3.3.1. Movimiento horizontal

Una característica interesante a explorar es la velocidad de movimiento horizontal de la nubosidad en las escenas (en rigor, una estimación de la velocidad $x-y$ de los píxeles en la imagen). Una forma clásica de hacer esta estimación es mediante la técnica de "flujo óptico". Esta técnica se basa en la hipótesis de conservación de brillo, es decir, que en un intervalo de tiempo infinitesimal, el "brillo" de los píxeles se mantiene aproximadamente constante. Esta hipótesis resulta en la Ecuación de Continuidad, a la cuál se le deben adicionar vínculos o restricciones para obtener un problema matemáticamente determinado. Existen varias formas de hacer esto (Horn y Schunck, 1981; Lucas y Kanade, 1981). En Aicardi et al. (2022) se analizan y adaptan a nuestra región varias alternativas para imágenes ABI, resultando que la técnica de Farnebäck (2003) con parámetros ajustados localmente es un buen compromiso entre desempeño y facilidad de implementación, obteniendo resultados de estimación de movimiento que son de los mejores. Esta técnica permite estimar el movimiento de la escena para cada píxel considerando la imagen actual y la inmediata anterior y el paso de tiempo entre estas imágenes. De esta forma, el flujo óptico, al igual que el viento, se presenta como una magnitud vectorial, cuyo módulo se expresa en unidades de píxeles de desplazamiento horizontal por segundo (en este caso se escoge la unidad de píxeles por minuto, px/min). El campo vectorial resultante se conoce como Cloud Motion Vectors (CMV). En la Figura 3.3 se muestra un ejemplo de esta característica en su forma vectorial para la imagen de la Figura 3.2 (a), mostrando sólo 1 de cada 10 vectores por legibilidad. Se observa que el movimiento de la nubosidad es consistente con la escena de la figura anterior. Además, en los píxeles despejados, el movimiento es nulo.

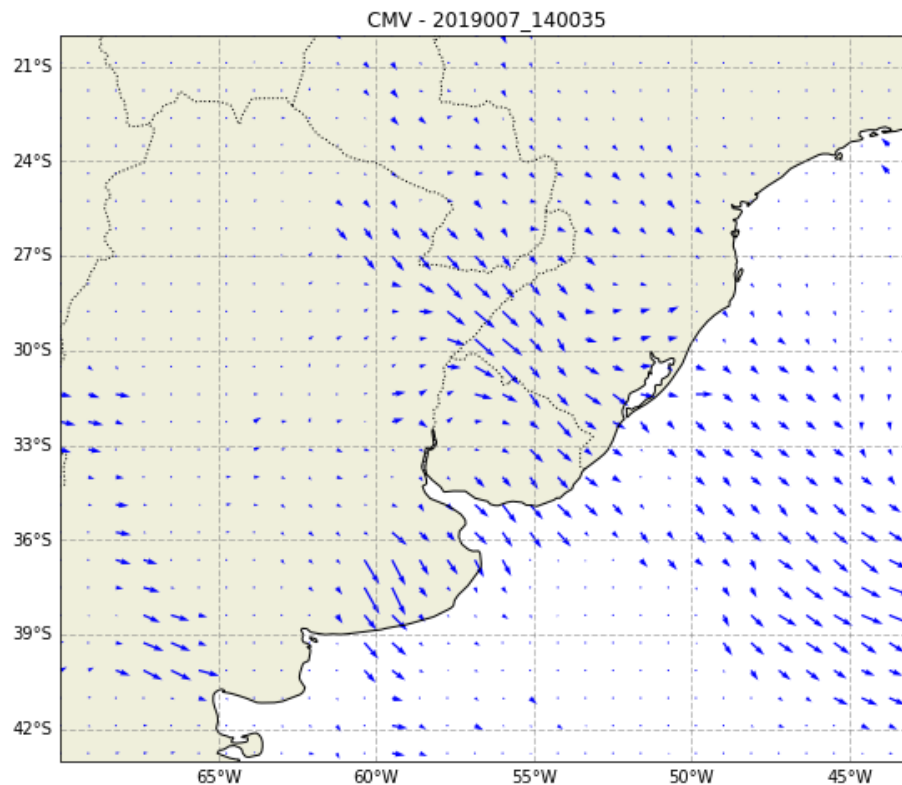


Figura 3.3: Flujo Óptico calculado para el día 7 de enero de 2019 a las 14:00hs.

La Tabla 3.2 presenta las características integradas a la BDT derivadas de esta técnica, que son el módulo y la dirección del campo vectorial en cada píxel.

Nombre de Característica	Descripción
OFORMOD	Módulo del Vector de Flujo Óptico
OFORMDIR	Dirección del Vector de Flujo Óptico

Tabla 3.2: Características de Flujo Óptico integradas a la BDT.

3.3.2. Promedios espaciales

Este grupo de características busca integrar a un píxel dado la información de píxeles cercanos. Esto se logra calculando el promedio de los valores en una ventana de 11×11 píxeles, asignando el resultado al píxel central de dicha ventana. Esto además permite reducir problemas de geo-referenciamiento preciso o diferencias que surjan del procesos de consistir la información a una única base espacial. Dada la resolución de la grilla de trabajo, la ventana móvil de promediado equivale a un área de aproximadamente unos 50×50 km ($\simeq 250$ km²). La Tabla 3.3 muestra el nombre escogido para cada variable aplicando este procedimiento a cada canal del ABI. Este cálculo funciona como un suavizado de la imagen original. La Figura 3.4 (a) presenta una imagen de ejemplo del resultado que se obtiene al aplicar el procedimiento a la imagen del canal visible de la Figura 3.2 (a), lo que es la característica AvgC02RP.

Nombre de Característica	Descripción
AvgC02RP	Promedio de C02RP en una ventana de 11×11 píxeles
AvgC13TB	Promedio de C13TB en una ventana de 11×11 píxeles
AvgOFORMOD	Promedio de OFORMOD en una ventana de 11×11 píxeles
AvgOFORMDIR	Promedio de OFORMDIR en una ventana de 11×11 píxeles

Tabla 3.3: Promedios espaciales integrados a la BDT.

3.3.3. Variabilidad espacial

Este grupo de características busca representar la variabilidad espacial de reflectividad en un entorno de cada píxel. Esto se logra calculando la desviación estándar de los valores en una ventana de 11×11 píxeles, asignando el resultado al píxel central de dicha ventana. La Tabla 3.4 presenta el nombre de estas características para cada canal y la Figura 3.4 (b) presenta una imagen de ejemplo de su aplicación al canal C02, lo que es la característica StdC02RP. Desde el punto de vista de tratamiento de imágenes, esta característica da información de textura en un entorno del píxel. Observando la figura, tiende a ser mayor en los bordes de la nubosidad, donde hay interfaces nube-fondo, y tiene valores más bajos en regiones despejadas y en regiones donde la reflectividad de la nubosidad es uniforme.

Nombre de Característica	Descripción
StdC02RP	Desviación estándar de C02RP en una ventana de 11×11 píxeles
StdC13TB	Desviación estándar de C13TB en una ventana de 11×11 píxeles
StdOFORMOD	Desviación estándar de OFORMOD en una ventana de 11×11 píxeles
StdOFORMDIR	Desviación estándar de OFORMDIR en una ventana de 11×11 píxeles

Tabla 3.4: Características de variabilidad espacial integradas a la BDT.

3.3.4. Gradientes

Este grupo de características se construyó a partir del calculo del gradiente asociado a un píxel en particular. Se calcularon las derivadas parciales en latitud y longitud (en y y en x , respectivamente) utilizando el esquema de diferencias centradas. Para esto, se consideraron los quintos píxeles

posteriores y anteriores al píxel central. Luego, se calculó el módulo del vector gradiente como es usual, utilizando el Teorema de Pitágoras. La Figura 3.4 (c) y (d) presentan imágenes de ejemplo del módulo del gradiente utilizando las imágenes en su resolución original y promediadas en el espacio, lo que son las características GradC02RP y GradAvgC02RP, respectivamente. Estas características tienden a resaltar los bordes de la nubosidad, y la diferencia es más notoria al utilizar las imágenes promediadas (suavizadas) como insumo. Este procedimiento se aplicó a varias de las características anteriores, siendo las integradas a la BDT las que se muestran en la Tabla 3.5.

Nombre de Característica	Descripción
GradC02RP	Módulo del gradiente de C02RP
GradC13TB	Módulo del gradiente de C13TB
GradOFMOD	Módulo del gradiente de OFMOD
GradAvgC02RP	Módulo del gradiente de AvgC02RP
GradAvgC13TB	Módulo del gradiente de AvgC13TB
GradAvgOFMOD	Módulo del gradiente de AvgOFMOD

Tabla 3.5: Características gradientes integradas a la BDT.

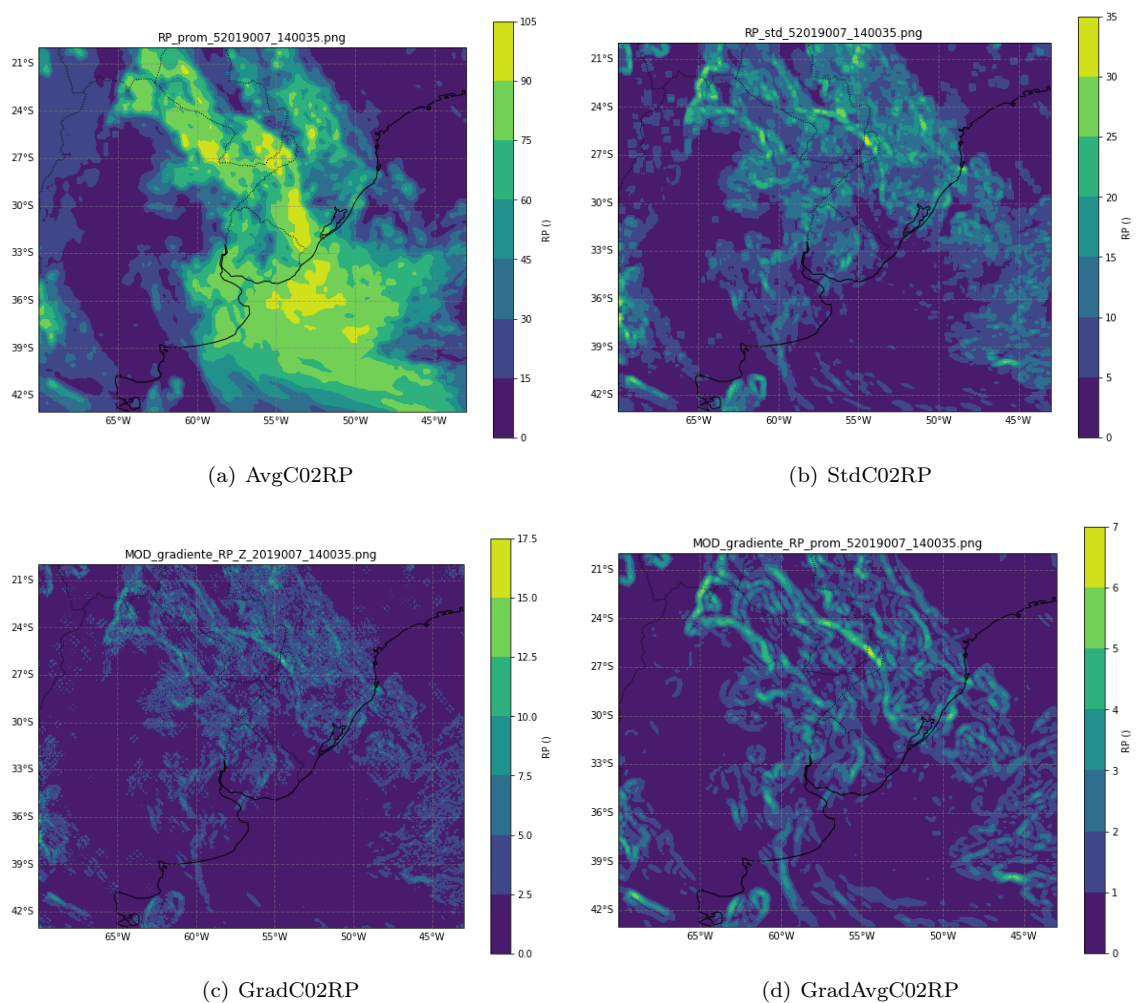


Figura 3.4: Ejemplos de características derivadas del ABI utilizando las imágenes de la Figura 3.2.

3.4. Características astronómicas

Por completitud, y en forma exploratoria, se incluyeron características astronómicas en la BDT. Estas características buscan aportar una referencia temporal y de posición relativa entre la Tierra y el Sol a la BDT, que podría tener información para el problema de clasificación, ya sea por la referencia

Este producto tiene una resolución espacial de 1 km, lo que es mayor que la resolución de 4 km de la información satelital anterior (0.035°). También se incorporó la estimación de CTH de ABI, que tiene una resolución espacial de 0.150° , lo que equivale aproximadamente de 18 km. Al contrario que el anterior, este producto tiene una resolución espacial menor que la información satelital anterior. Como ya se ha explicado, el producto CTH ABI da en este trabajo una base de desempeño para evaluar el resultado de la clasificación desarrollada en base a las características anteriores y también, en una segunda etapa, se utiliza como información de entrada.

La Figura 3.6 muestra la comparación entre los productos CTH de MODIS y ABI para la misma fecha y hora que en las imágenes anteriores. Como se observa en la figura, la información presentada es similar, pero con una cobertura y forma de presentación espacial distinta. En particular, es fácilmente apreciable que en la zona sin nubosidad (a la izquierda de ambas figuras) los productos coinciden. El rango de variación es también aproximadamente el mismo.

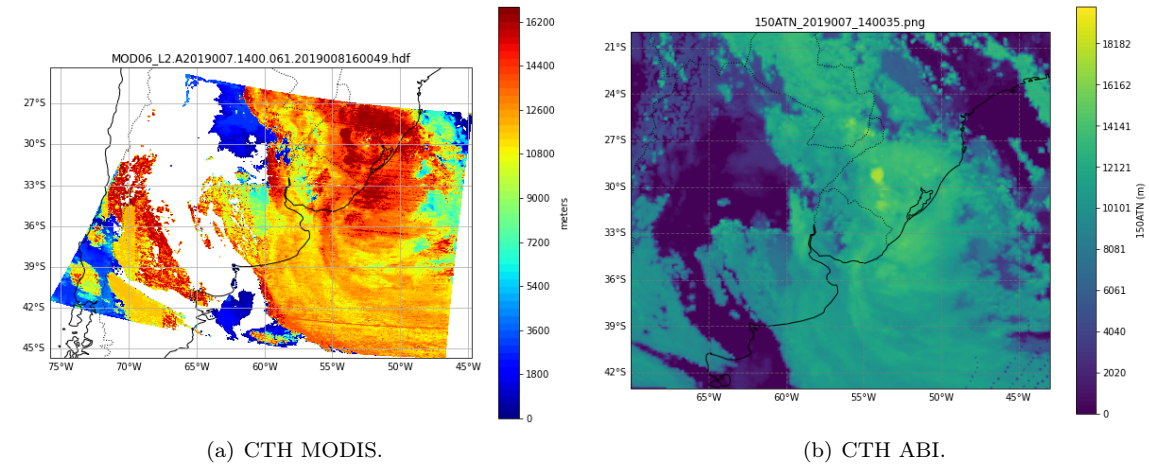


Figura 3.6: Imagen de los productos CTH (en metros) para la misma fecha y hora de la Figura 3.2.

3.6. Concurrencia espacio temporal

Una parte importante de la construcción de la BDT fue la resolución de la concurrencia espacial y temporal de las dos plataformas de observación que se consideraron y las diferencias entre sus productos. Incluso la información ABI tiene distinta base espacial, estando el CTH ABI con resolución de 0.150° y las imágenes y derivados en resolución de 0.035° . Esto obligó a compatibilizar los productos MODIS y ABI con la grilla de la BDT anterior (ver Figura 3.1). Esta diferencia implicó resolver la sincronía de las observaciones en el espacio y en el tiempo, imponiendo restricciones y aplicando los procedimientos que se detallan a continuación.

Para resolver la sincronía temporal, se buscó encontrar la imagen de ABI que más representara el pasaje de MODIS TERRA sobre el territorio uruguayo. Para esto, hay que tener en cuenta que la etiqueta temporal de las imágenes ABI se definen al comienzo del proceso de adquisición de datos por parte del radiómetro en órbita. Este proceso demora unos 10 minutos para lograr la imagen del disco planetario por completo. Por este motivo se definió elegir la imagen de ABI con etiqueta temporal más cercana, pero anterior, a la etiqueta temporal de MODIS.

Luego, se decidió llevar ambos productos a la grilla de 0.035° de la Figura 3.1, dado que es la grilla de las imágenes que serán consideradas por los clasificadores. Como las imágenes MODIS cuentan con mayor resolución espacial, habilita a realizar un promedio en regiones de la imagen para ir de la grilla origen a la grilla destino. Esto implicó bajar la resolución de las imágenes MODIS de 1km a 4km, y elegir el punto de grilla MODIS de baja resolución más cercano a cada punto de la grilla destino. En cambio, el producto CTH ABI debe interpolarse a la grilla destino para hacer el aumento de resolución. Existen, a priori, varias formas de hacer este procedimiento. En este trabajo se optó por la técnica de interpolación bilineal simple, dado que no es objetivo de este trabajo explorar la mejor forma de downscaling espacial.

Otra restricción importante que se debió considerar fue que las imágenes CTH MODIS cubrieran totalmente al territorio uruguayo, que es la zona objetivo. En varios casos la concurrencia temporal MODIS y ABI ocurre cuando la imagen MODIS no cubre todo el territorio nacional. En la figura 3.6 se presentó un ejemplo de imagen MODIS y ABI donde MODIS cubre totalmente el territorio de Uruguay, por lo que sería aceptada para integrar la BDT.

Luego de considerar las restricciones de concurrencia, solamente 125 escenas fueron aceptadas en el período comprendido entre el 1° de noviembre del 2019 y el 31 de diciembre del 2020. La BDT se compone entonces de estos 125 instantes de concurrencia MODIS y ABI sobre todo el territorio de Uruguay, y las características presentadas anteriormente, todo en una resolución espacial regular de 0.035° en latitud y longitud.

3.7. Asignación de categorías de nubosidad

La asignación de categorías de nubosidad a la BDT se consiguió a través de un proceso de dos pasos. En primer lugar se define si el píxel observado es nuboso o no en base a un filtro de Brillo de Fondo. Para esto, se evalúa el valor de RP de ese píxel con respecto a su correspondiente Brillo de Fondo (RP0), que representa la RP típica del píxel en condiciones de cielo despejado, calculada para ese momento exacto del año. La estimación de RP0 se realiza en base a una parametrización descrita en [Alonso-Suárez et al. \(2014\)](#), y requiere considerar para su ajuste muestras de RP en el píxel en ausencia de nubosidad durante largos periodos de tiempo.

Para este trabajo fue necesario recalcular el ajuste de la parametrización de RP0 considerando las imágenes del canal visible C02 de ABI GOES16 disponibles en el archivo del LES, ya que el ajuste disponible había sido realizado para satélites GOES-East anteriores al GOES16. Esto fue realizado con 4 años de imágenes, y del proceso de ajuste se extraen cuatro campos escalares $A(x, y)$, $B(x, y)$, $C(x, y)$ y $D(x, y)$ que permiten estimar la reflectividad del fondo de cada píxel según:

$$RP0(x, y) = \frac{A(x, y) + B(x, y) \cos \theta_z + C(x, y) \sin \theta_z \cos \gamma_{rs} + D(x, y) \sin \theta_z \cos^2 \gamma_{rs}}{\cos \theta_z}, \quad (3.1)$$

donde γ_{rs} es el ángulo azimutal entre la dirección del Sol y la dirección del satélite. Esta parametrización es una actualización de la propuesta original de [Justus et al. \(1986\)](#), y su ajuste se realiza a partir de un proceso de filtrado iterativo de muestras de cielo cubierto, llegando a un ajuste final exclusivamente con muestras de cielo despejado ([Alonso-Suárez et al., 2012](#)). Realizado el ajuste, si el valor de RP del píxel en cuestión supera a su correspondiente RP0, es probable que dicho píxel contenga nubosidad.

El filtro de brillo de fondo se aplica en conjunto con la evaluación del umbral inferior de la categoría 1 (ver Tabla 3.7) tanto con el CTH MODIS como con el CTH ABI. De esta forma, se logra que el total de píxeles clasificados como nubosos tenga la misma cantidad de elementos, ya sea en la clasificación derivada de CTH MODIS como en la clasificación derivada de CTH ABI. Dado que podría ocurrir que un píxel fuera considerado nuboso al evaluar la restricción de RP0, pero luego su valor CTH MODIS fuera menor a este umbral mínimo de la categoría 1, lo que lo categorizaría como no nuboso. De la misma forma, podría suceder que su valor CTH ABI sí le permitiera continuar en una categoría nubosa. Esto dificultaría la validación de la clasificación, al generarse conjuntos de píxeles nubosos de distintos tamaños, dependiendo de cual sea la estimación de CTH que se esté usando para clasificar.

A los efectos de este trabajo, la condición expresada en la ecuación 3.2 resume las restricciones definidas para considerar nuboso a un píxel. Cabe destacar que se eligió considerar un factor de cobertura de 40 % para la evaluación de RP0. Esto es porque el ajuste de la parametrización implica que al usar la Ec. (3.1) para estimar RP0 se reproducen condiciones medias, por lo que hay que dejar un margen. Por ejemplo, para un valor típico de reflectividad de fondo de 10 %, el factor de cobertura de 40 % eleva el valor anterior a 14 %, lo que es también un valor posible de cota para brillo de fondo. Este valor de 40 % se fijó en forma heurística, considerando el caso típico comentado en la oración anterior y la variación normal de la reflectividad de fondo en el territorio de Uruguay.

$$\text{El píxel es considerado nuboso si: } \begin{cases} \text{RP} \geq 1.4 \cdot \text{RP0} \\ \text{CTH ABI} \geq 500 \text{ m} \\ \text{CTH MODIS} \geq 500 \text{ m} \end{cases} \quad (3.2)$$

Luego de que un píxel es clasificado como nuboso, solamente el valor de CTH MODIS se considera decisor para asignar la categoría nubosa para el etiquetado, dado que es la verdad de referencia para el entrenamiento y validación de los clasificadores. De igual forma, solamente el valor de CTH ABI será considerado para asignar la categoría nubosa CTH ABI, la base de desempeño de esta experiencia. La Tabla 3.7 resume los criterios de asignación de las 4 categorías definidas en la BDT, a partir de la estimación de CTH MODIS. Los umbrales de las categorías nubosas se eligieron buscando representar la clasificación usual de la nubosidad vista en la sección 2.1.

Categoría	Descripción
0 - Sin nubosidad	El píxel no supera en, por lo menos, un 40 % al RP0 característico o la altitud del tope es menor a 500 metros, tanto en CTH ABI como CTH MODIS.
1 - Nubes Bajas	Valores CTH MODIS mayores a 500 metros, y menores a 2000 metros.
2 - Nubes Medias	Valores CTH MODIS mayores a 2000 metros, y menores a 6000 metros.
3 - Nubes Altas	Valores CTH MODIS mayores a 6000 metros, sin límite superior.

Tabla 3.7: Criterio para el etiquetado de categorías nubosas considerado en este trabajo

3.8. Resumen general de la BDT

La BDT considera 23 características descriptoras de cada píxel, además de los valores de los productos CTH ABI y CTH MODIS, y los etiquetados que surgen de ambos productos, siendo el basado en MODIS el etiquetado de referencia. La Tabla 3.8 resume las características presentadas en las secciones 3.2, 3.3 y 3.4. Las restricciones impuestas para la resolución de la concurrencia espacio temporal de las distintas plataformas y la elección de trabajar solamente con imágenes diurnas, llevó a que 125 escenas del periodo comprendido entre el 1 de noviembre de 2019 y el 31 de diciembre de 2020 fueran consideradas para conformar la BDT. De esta forma, y considerando que la grilla de trabajo (a 0.035°) está conformada por 24795 píxeles, se logró reunir información correlativa de 3099375 píxeles etiquetados en este proceso. Esta cantidad de datos es adecuada para trabajar con algoritmos de aprendizaje automático clásicos, que son los que se abordan para el desarrollo del clasificador satelital.

La Tabla 3.9 resume la composición de la BDT según la categorización derivada de CTH MODIS. Se observa que el 59.7% de los píxeles seleccionados corresponden a situaciones de cielo despejado (categoría 0). De esta forma, el restante 40.3% corresponden a píxeles considerados nubosos según los criterios desarrollados en 3.7. De este conjunto, el 8.2% de los píxeles de la BDT corresponden a nubes bajas (categoría 1), el 10.4% a nubes de media altura (categoría 2) y el 21.7% a nubes altas (categoría 3).

3.9. Exploración de la BDT

A continuación se presenta un análisis exploratorio de la BDT considerando las categorías nubosas derivadas del producto CTH de MODIS. Para comenzar, se estudia el comportamiento de la diferencia entre los valores CTH ABI y CTH MODIS. De esta forma es posible evaluar la calidad del procedimiento llevado a cabo para compatibilizar las diferencias entre las plataformas de observación satelital. La Figura 3.7 presenta los histogramas de la diferencia entre ambos productos CTH según la categorización de nubosidad derivada del producto CTH MODIS. Se observa que los

Ordinal	Código de Característica	Ordinal	Código de Característica
1	C13TB	14	AvgOFMOD
2	C14TB	15	StdOFMOD
3	C15TB	16	AvgOFDIR
4	C16TB	17	StdOFDIR
5	C02RP	18	GradC02RP
6	ASTCZ	19	GradC13TB
7	ASTWN	20	GradOFMOD
8	OFMOD	21	GradAvgC02RP
9	OFDIR	22	GradAvgC13TB
10	AvgC02RP	23	GradAvgOFMOD
11	StdC02RP	24	CTHABI
12	AvgC13TB	25	CTHMODIS
13	StdC13TB	26	etiqueta CTHABI
-	-	27	etiqueta CTHMODIS

Tabla 3.8: Resumen de características que conforman BDT. Estas características son obtenidas de 125 imágenes diurnas del periodo comprendido entre 1/11/2019 al 31/12/2020.

Categoría	Participación en BDT (%)
0 - Sin nubosidad	59.7
1 - Nubes Bajas	8.2
2 - Nubes Medias	10.4
3 - Nubes Altas	21.7

Tabla 3.9: Representación de cada categoría en el total de las 3099375 instancias de la BDT.

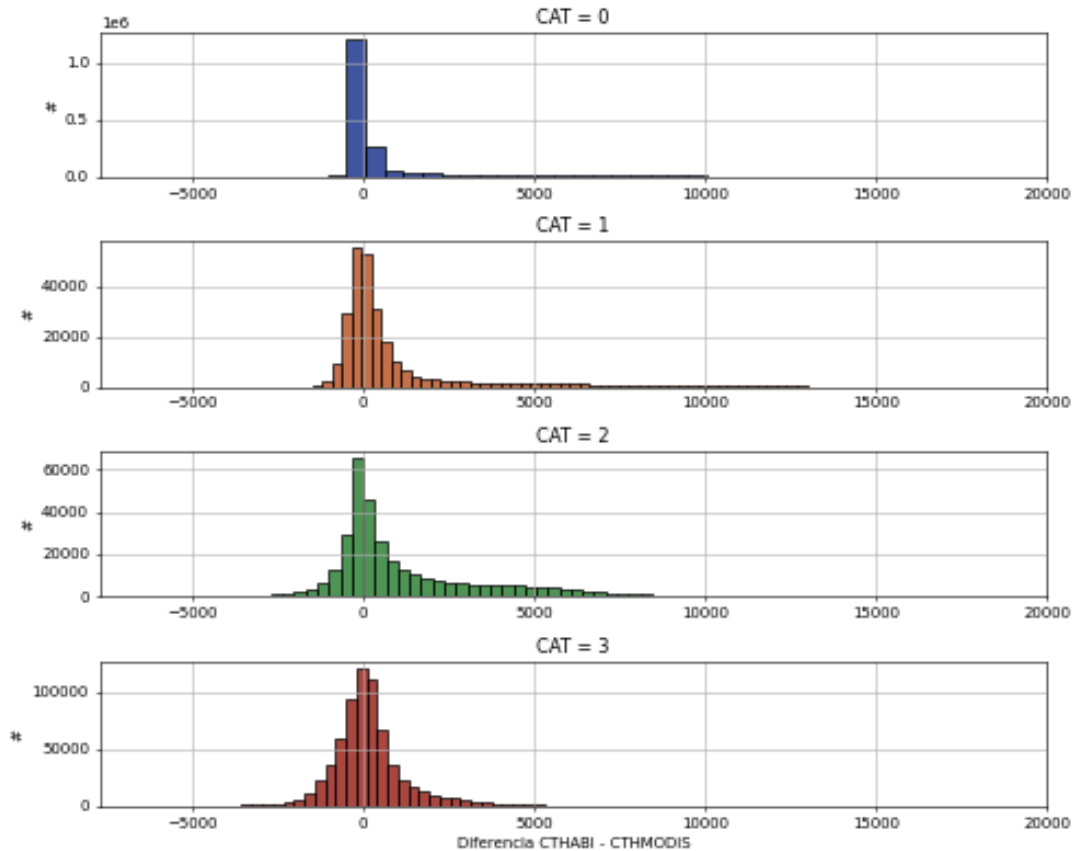


Figura 3.7: Diferencias entre CTHABI y CTHMODIS para varias categorías.

histogramas muestran la mayor cantidad de ocurrencias en torno a la diferencia nula, sin embargo, existen colas, principalmente hacia la derecha, indicando valores de CTH ABI superiores a aquellos de CTH MODIS. Es notorio como estas colas hacia la derecha disminuyen su extensión a medida que se avanza en las categorías de 1 a 3.

Las Figura 3.8 presenta los diagramas de dispersión entre CTH ABI y CTH MODIS según cada categoría nubosa. Se presentan tanto el diagrama de dispersión sin limitar los valores de CTH ABI a la categoría en cuestión ((a), (c) y (e)), donde las colas dificultan la visualización, como el diagrama de dispersión con los valores de CTH ABI limitados a la categoría en cuestión, ((b), (d) y (f)). En estos diagramas de dispersión se incorpora la densidad de puntos como barra de colores de forma tal que se destacan las regiones con mayor densidad en el diagrama y se observa como ambos productos se ajustan aceptablemente a la recta identidad. Esto es compatible con lo visto en los histogramas de la diferencia entre los productos (Figura 3.7). De esta exploración, es notable como en la categoría 1 la densidad de puntos está apreciablemente distribuida a lo largo de la recta identidad dentro del rango de la categoría. Esta es la categoría donde la estimación de alturas presenta colas más largas. Por el contrario, las categorías 2 y 3 presentan regiones del diagrama con apreciable mayor densidad de puntos que otras, centradas en torno a la recta identidad. En el caso de la categoría 2, la mayor densidad se da en el rango de CTH de 2000 a 2800 metros. En el caso de la categoría 3, la mayor densidad se observa en el rango de valores CTH entre 8200 a 14400 metros. Este comportamiento es esperable por lo visto en la comparación de los productos CTP ABI y CTP MODIS realizada en [A. K. Heidinger et al. \(2010\)](#) y respalda el trabajo realizado en la sección 3.6.

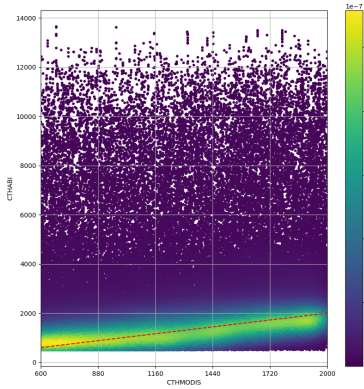
La Tabla 3.10 presenta los valores promedio y desviación estándar de la diferencia entre CTH ABI y CTH MODIS para las tres categorías de nubosidad. Por más que los picos de las distribuciones están centrados en la diferencia nula, los valores de diferencia promedio son positivos debido a la asimetría de las distribuciones. Se observa que estos valores obtenidos son mayores a los que se esperarían dada la diferencia reportada en [A. K. Heidinger et al. \(2020\)](#) entre los productos CTP ABI y CTP MODIS, presentada en la subsección 2.3.1. De todas formas, se cumple con que el CTH MODIS ubica a los topes nubosos en niveles más bajos de la atmósfera con respecto a CTH ABI. Cabe la posibilidad de que el procedimiento de pasaje de las grillas CTH ABI y MODIS a la grilla de trabajo haya afectado la diferencia entre los productos. En el caso de este trabajo se requirieron dos conversiones de grillas y no sólo una, debido a que la base de imágenes de satélite para el clasificador es en una resolución intermedia a ambos productos CTH, de 0.035°. Se observa además que la categoría intermedia, de nubes con topes nubosos de altura media, es la que presenta mayor error promedio y mayor desvío estándar. La categoría 1, a pesar de tener la distribución de diferencias con colas más largas, éstas también son de menor probabilidad, por lo que no se obtiene un desvío estándar mayor que la categoría 2. La categoría 3 es, por tanto, en la que ambos productos tienen más coincidencia en la estimación de la CTH.

Categoría	Promedio (m)	Desviación Estándar (m)
1 - Nubes Bajas	744.35	1917.91
2 - Nubes Medias	1137.98	2152.20
3 - Nubes Altas	173.97	1107.31

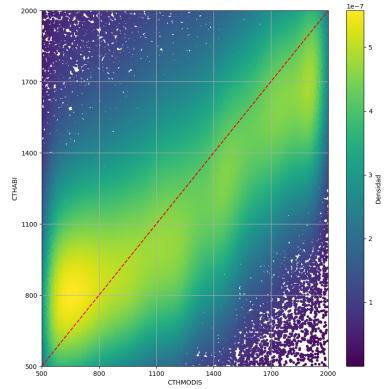
Tabla 3.10: Media y desviación estándar de la diferencia CTHABI - CTHMODIS.

A continuación, se presenta una exploración de las características de la BDT con respecto a la categorización derivada de CTH MODIS. Este análisis permite observar comportamientos esperados dada la física del problema. También se puede observar las diferentes habilidades de las características para representar a las distintas categorías nubosas. Esto será aprovechado por los algoritmos de aprendizaje automático para lograr la clasificación.

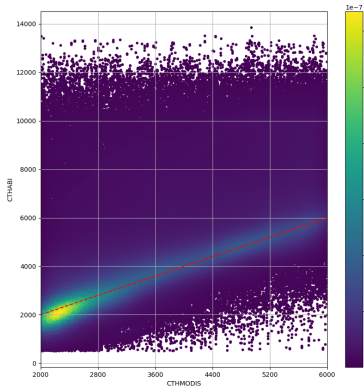
Las Figuras 3.9 y 3.10 presentan los histogramas de las características infrarrojas: C13TB, C14TB, C15TB y C16TB. Se observa que a medida que se incrementa la categoría de nubosidad, las temperaturas de brillo son en general más frías, cómo es característico. También se destaca el importante ensanchamiento del histograma en la categoría 3, para todas estas características y especialmente para la característica C13TB, algo que muestra que la clasificación de altura exclusivamente en base a la temperatura de brillo es limitada, al menos en comparación a la clasificación que proviene del CTH MODIS.



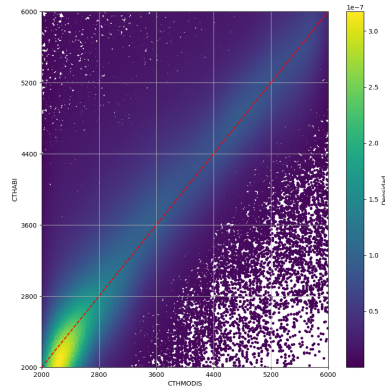
(a) Categoría 1. Sin restringir CTH ABI.



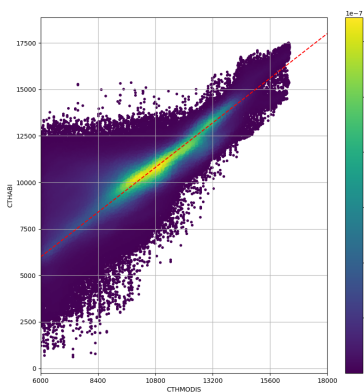
(b) Categoría 1. Restringiendo CTH ABI a rango de categoría 1.



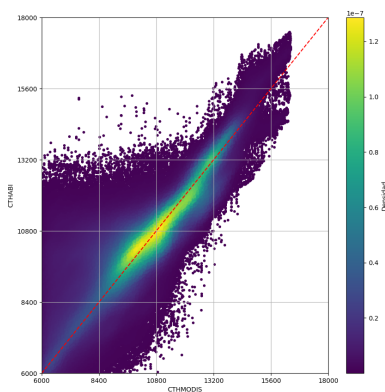
(c) Categoría 2. Sin restringir CTH ABI.



(d) Categoría 2. Restringiendo CTH ABI a rango de categoría 2.



(e) Categoría 3. Sin restringir CTH ABI.



(f) Categoría 3. Restringiendo CTH ABI a rango de categoría 3.

Figura 3.8: Diagramas de dispersión CTH ABI vs CTH MODIS, con densidad de puntos.

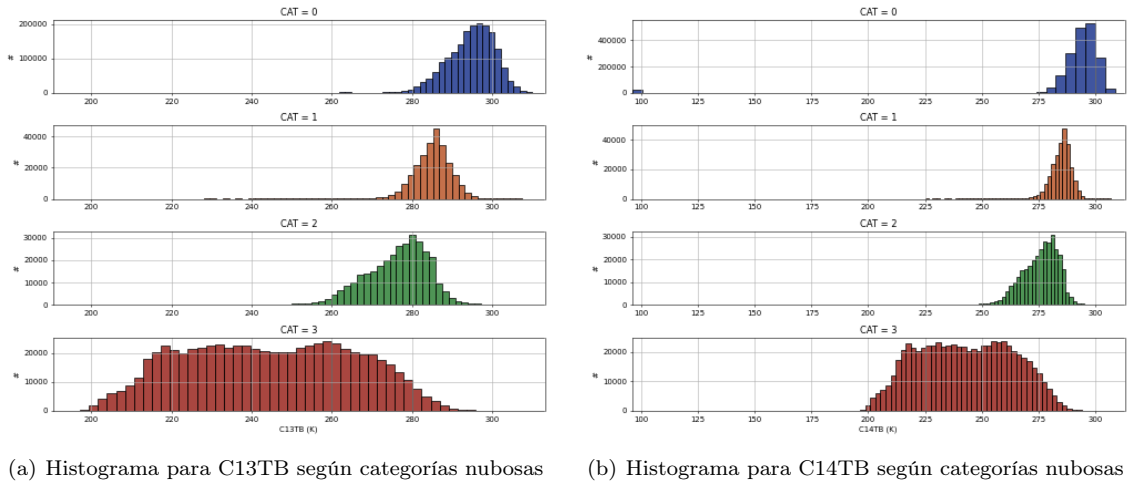


Figura 3.9: Histogramas para las características C13TB y C14TB según categorías nubosas.

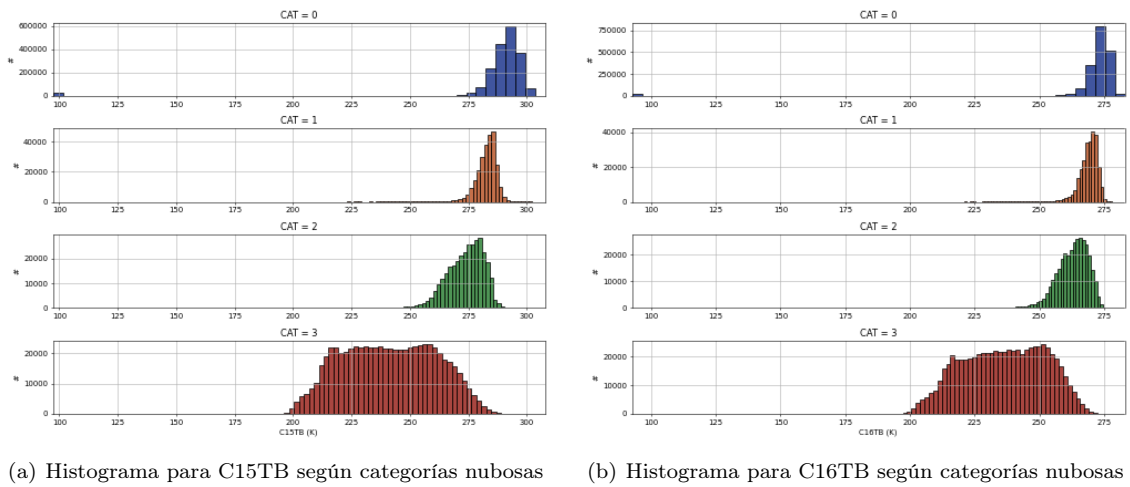


Figura 3.10: Histogramas para las características C15TB y C16TB según categorías nubosas.

La Figura 3.11 (a) presenta los histogramas por categoría para la característica C02RP. Estos histogramas muestran que la frecuencia de los valores altos de RP se incrementa con la altura de topes nubosos, aunque se pueden tener valores altos en todas las categorías, como muestra lo ancho de los histogramas para las categorías 1, 2 y 3. También se observa el corte abrupto en la cola izquierda de los gráficos, relacionado al filtro de brillo de fondo realizado. También se puede observar que los valores bajos de C02RP representan píxeles de cielo despejado. Los valores por encima del $\gtrsim 15\%$ de RP representan muestras cercanas al inicio y principio del día, donde RP puede mostrar valores un poco más altos debido a una menor validez de la hipótesis de reflexión isotrópica. Por otro lado, la Figura 3.11 (b) presenta los histogramas por categoría para la característica AvgC02RP. Como se mencionó en la descripción de las características de este tipo, este promediado espacial suaviza las imágenes. En la figura se observa que AvgC02RP suaviza las colas izquierdas de los histogramas en todas las categorías, respecto a los histogramas vistos para C02RP. Particularmente en las categorías 1 y 2 se observan valores de AvgC02RP más diferenciados de aquellos valores AvgC02RP característicos de la categoría 0.

La Figura 3.12 (a) presenta los histogramas por categoría para la característica ASTCZ (coseno del ángulo cenital). Por construcción, todos los valores de ASTCZ son positivos, debido a que en este trabajo se consideran solamente las horas diurnas. También, se destaca la alta frecuencia de los valores superiores a 0.8 en todas las categorías, esto se debe a la hora de pasaje del satélite TERRA sobre la región, cercana al mediodía solar. Por otro lado, la Figura 3.12 (a) presenta los histogramas por categoría para la característica ASTWN. En estos histogramas se observan nuevamente valores relacionados al horario de pasaje del satélite TERRA sobre la región. Se tiene que para todas las categorías los valores de ASTWN son negativos y tienen su frecuencia máxima alrededor de -0.30,

es decir las imágenes corresponden a horas de la mañana, en un horario relativamente cercano al mediodía solar.

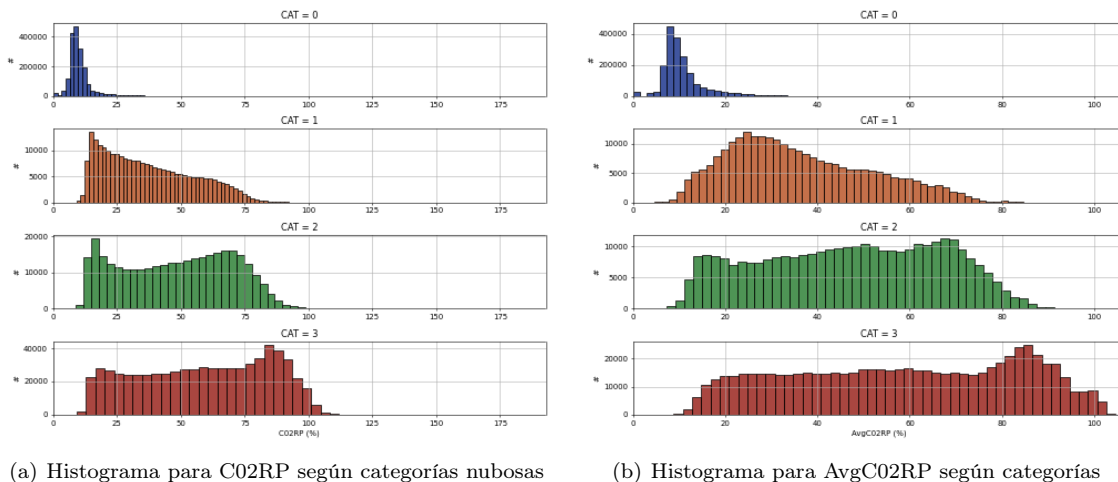


Figura 3.11: Histogramas para las características C02RP y AvgC02RP según categorías nubosas.

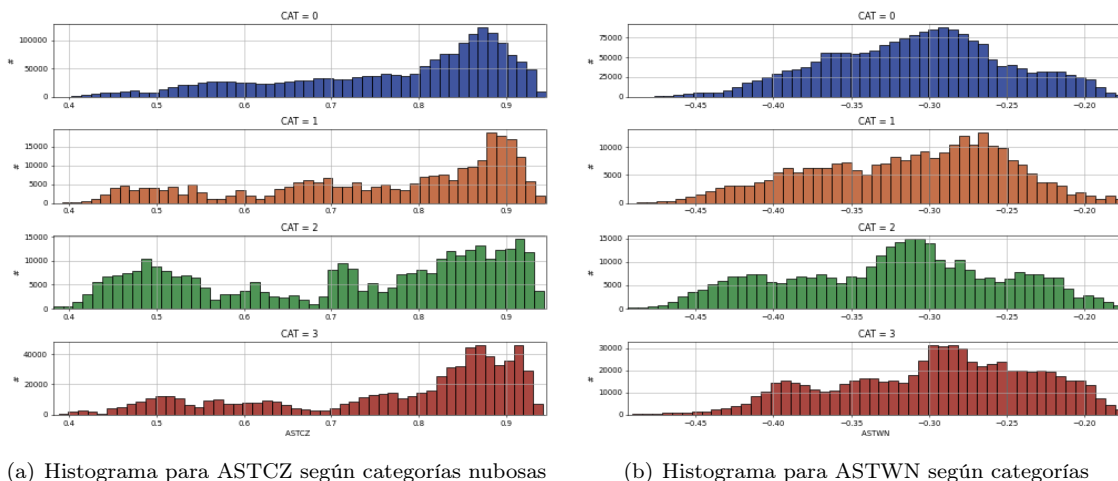


Figura 3.12: Histogramas para las características ASTCZ y ASTWN según categorías nubosas.

La Figura 3.13 (a) presenta los histogramas por categoría para la característica OFMOD. Como se mencionó en su descripción, esta característica es *proxy* de la intensidad del viento. Dicho esto, se destaca como al aumentar la categoría nubosa, se vuelven más frecuentes las intensidades altas de OFMOD. Esto está en acuerdo con el comportamiento del viento a medida que se incrementa la altura en la tropósfera. Las ocurrencias de OFMOD nulo se concentran, como era de esperar, en la categoría sin nubosidad. Se observan también en la categoría de nubes altas, pero esto podría tratarse de limitaciones del algoritmo de flujo óptico en captar el movimiento en algunas de esas situaciones. Por otro lado, la Figura 3.13 (b) presenta los histogramas por categoría para la característica OFDIR. Como se mencionó en su descripción, esta característica es *proxy* de la dirección del viento. En este caso se destaca como a medida que se incrementa la categoría de nubosidad se vuelven dominantes los valores de OFDIR que indican vientos del Oeste. Nuevamente esto está de acuerdo con los vientos que se observan a medida que se asciende en la troposfera. En cuanto a la categoría 1, se puede observar la relativa baja frecuencia de vientos con componente Sur en niveles bajos, algo que está de acuerdo con la climatología del Uruguay.

La Figura 3.14 (a) presenta los histogramas por categoría de la característica StdC02RP. Se observa como la característica StdC02RP en las categorías 0 y 3, categorías extremas de la clasificación, indica menor variabilidad espacial de C02RP que en los casos de las categorías 2 y 3. La característica StdC13TB se muestra en la Figura 3.14 (b), en donde las distribuciones se encuentran más concentradas hacia variabilidades menores que en el caso anterior, mostrando uniformidad espacial

de los valores de C13TB en la ventana de cálculo elegida.

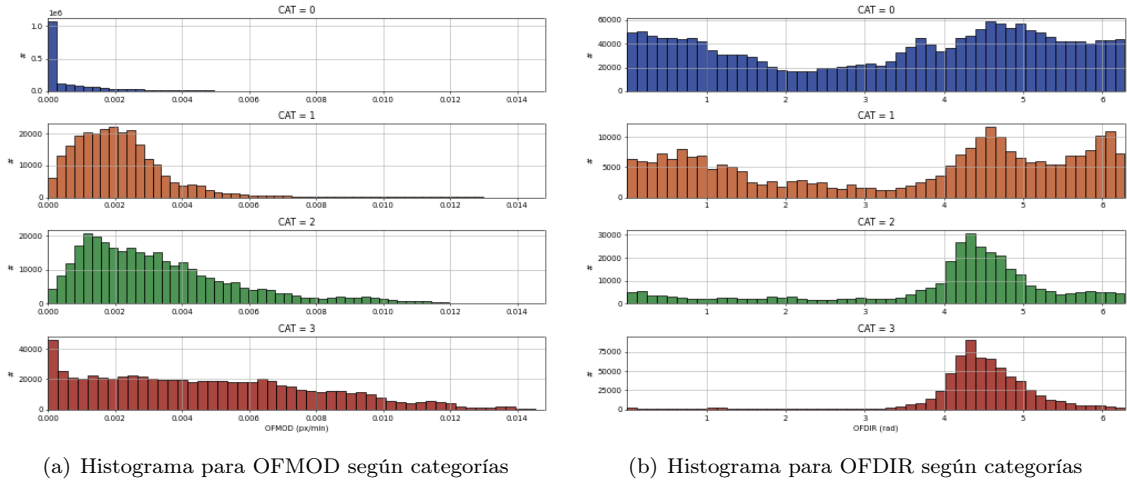


Figura 3.13: Histogramas para las características OFMOD y OFDIR según categorías nubosas.

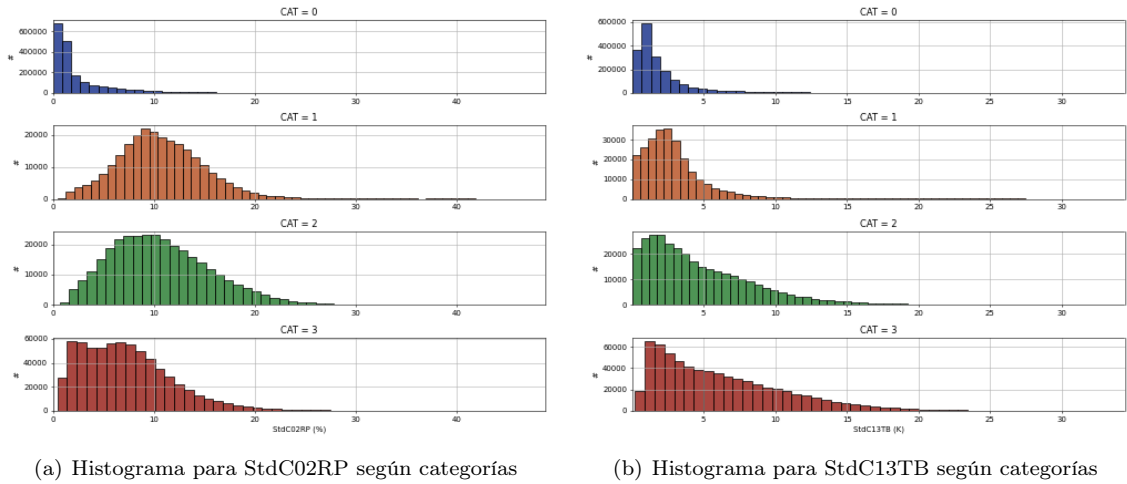


Figura 3.14: Histogramas para las características StdC02RP y StdC13TB según categorías nubosas.

La Figura 3.15 (a) presenta los histogramas por categoría de la característica AvgOFMOD. No se observan cambios destacables con respecto a su característica original, OFMOD. En este sentido, también se puede observar en la figura 3.16 (a), donde se presentan los histogramas para StdOFMOD, la aparente uniformidad espacial de la característica OFMOD independientemente de la categoría nubosa. Los histogramas de StdOFMOD presentan patrones similares a lo largo de las categorías, además presentan frecuencias máximas un orden de magnitud menor que los valores de OFMODs. De igual forma, los histogramas para la característica AvgOFDIR presentados en la figura 3.15 (b) no muestran cambios significativos respecto a la característica OFDIR. Mientras que en la figura 3.16 (b), donde se presentan los histogramas por categoría de la característica StdOFDIR, se observa una disminución de variabilidad espacial de OFDIR con el aumento de la categoría nubosa. Esto nuevamente está en acuerdo con la naturaleza de los niveles superiores de la tropósfera, donde predominan los vientos del Oeste, lo que justifica una disminución de la variabilidad de OFDIR con la altura.

En la Figura 3.17 (a) se tienen los histogramas por categoría nubosa para la característica GradC02RP. Se observa que los gradientes para las categorías 0 y 3 son menos intensos que aquellos de las categorías 1 y 2. De esta manera, se tiene que una mayor uniformidad espacial en las categorías 0 y 3 de esta característica, mientras que las categorías 1 y 2 presentan gradientes más intensos. Por otra parte, en la figura 3.17 (b) se tienen los histogramas de la característica GradC13TB, donde no existe una diferencia notable en la forma de los histogramas entre las categorías nubosas, solamente se destaca la extensión de la cola hacia la derecha derecha en el caso del histograma de GradC13TB

para la categoría 1. Esto indica que no se tienen gradientes abruptos de C13TB, para la ventana de cálculo que se considera en este trabajo, confirmando lo visto en la Figura 3.14 (b).

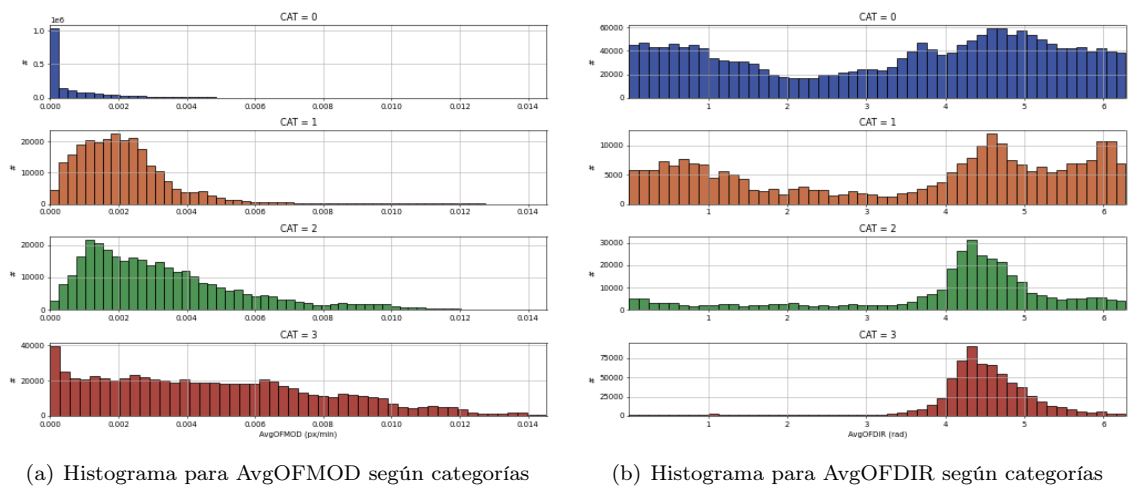


Figura 3.15: Histogramas para las características AvgOFMOD y AvgOFDIR según categorías.

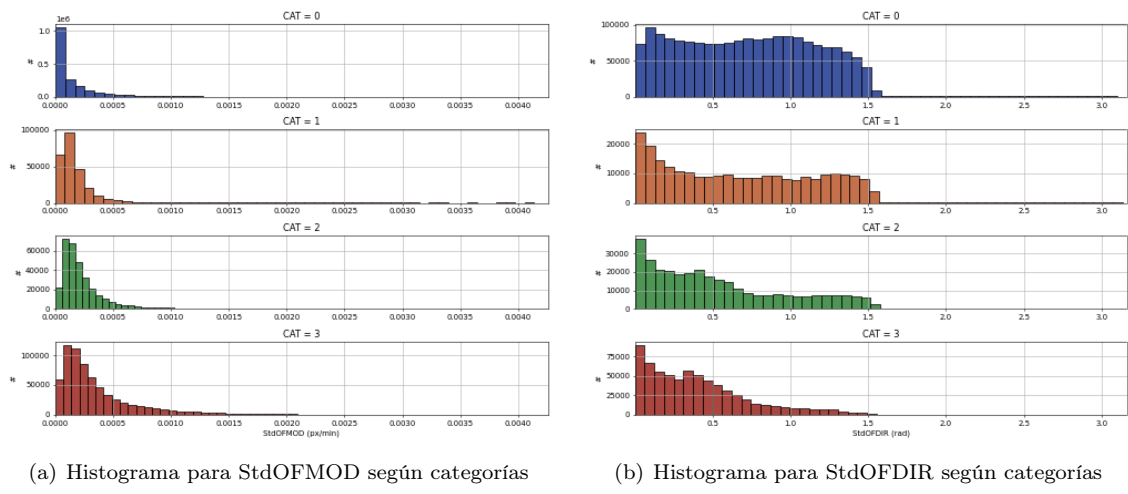


Figura 3.16: Histogramas para las características StdOFMOD y StdOFDIR según categorías.

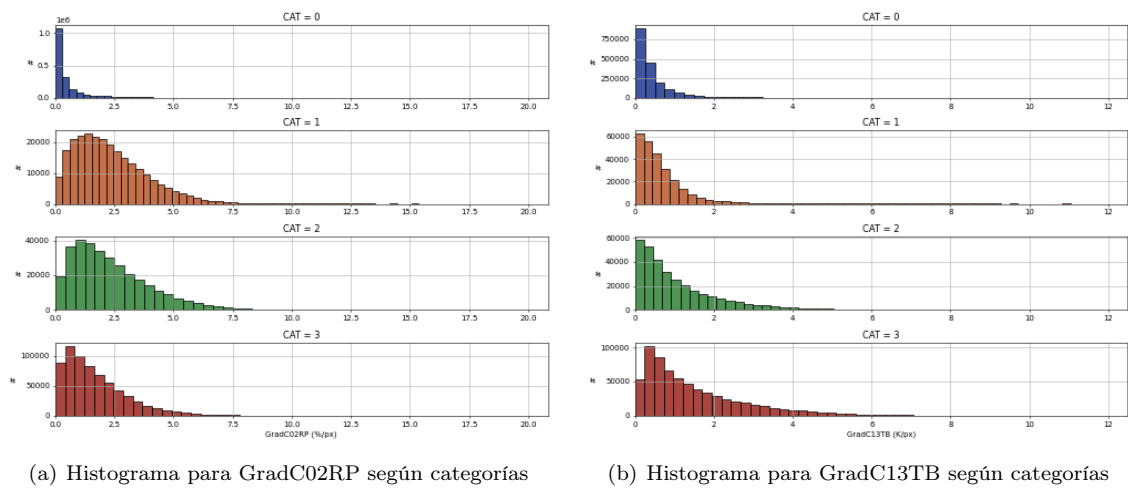


Figura 3.17: Histogramas para las características GradC02RP y GradC13TB según categorías.

En el caso de la característica GradAvgC02RP que se presenta en la figura 3.18 (a) se vuelve a observar dos comportamientos entre categorías. Por un lado la alta frecuencia de valores de GradAvgC02RP bajos en las categorías 0 y 3. A la vez que los valores GradAvgC02RP de las categorías 1 y 2, son en general mayores a las de las categorías extremas. Esto indica gradientes más intensos de AvgC02RP para las categorías nubladas 1 y 2, frente los gradientes menos intensos que se observan en las categorías 0 y 3. Por otra parte, la Figura 3.18 (b) presenta el caso de la característica GradAvgC13TB, no se observa un compartamiento que permita entender la variabilidad de AvgC13TB en las distintas alturas, solo se destaca que las categorías 1 y 3 presentan histogramas con colas derechas considerablemente alargadas, indicando la presencia (aunque poco frecuente) de gradientes intensos en la característica AvgC13TB, para estas categorías.

En la Figura 3.19 (a) se tienen los histogramas por categoría nublada para la característica GradOFMOD. En este caso, se observa que el comportamiento entre todas las categorías nubladas es similar, se tienen gradientes poco intensos de OFMOD, indicando cierta uniformidad espacial de la característica más allá de la categoría nublada que se considere. De igual forma sucede con la característica GradAvgOFMOD. En la Figura 3.19 (b) se presentan los histogramas para esta característica. Se observa que los histogramas son similares, independientemente de la categoría nublada que se considere. Es destacable para el caso de la categoría 1, la notable extensión de la cola derecha del histograma, nuevamente indicando casos poco frecuentes en los que se tienen gradientes intensos de AvgOFMOD.

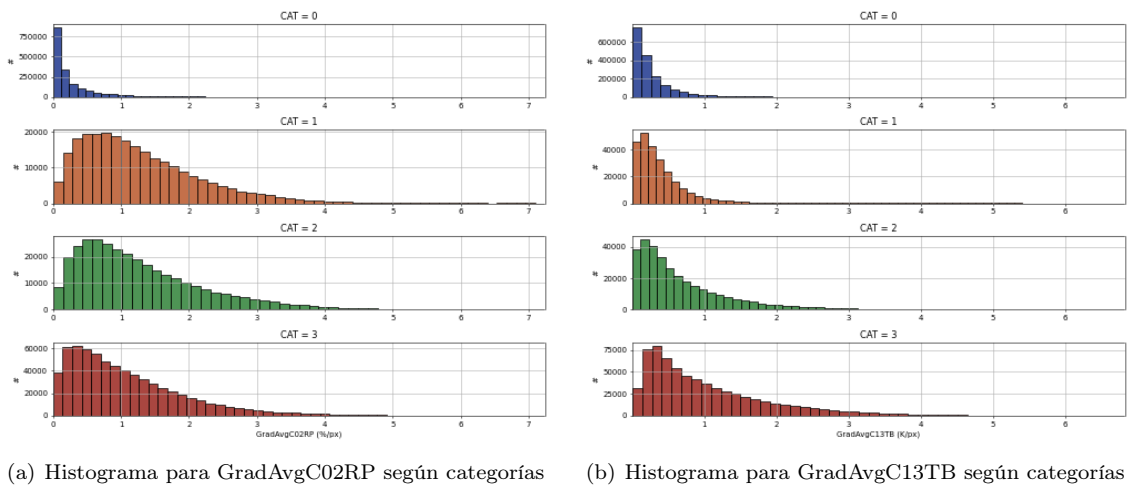


Figura 3.18: Histogramas de las características GradAvgC02RP y GradAvgC13TB según categorías.

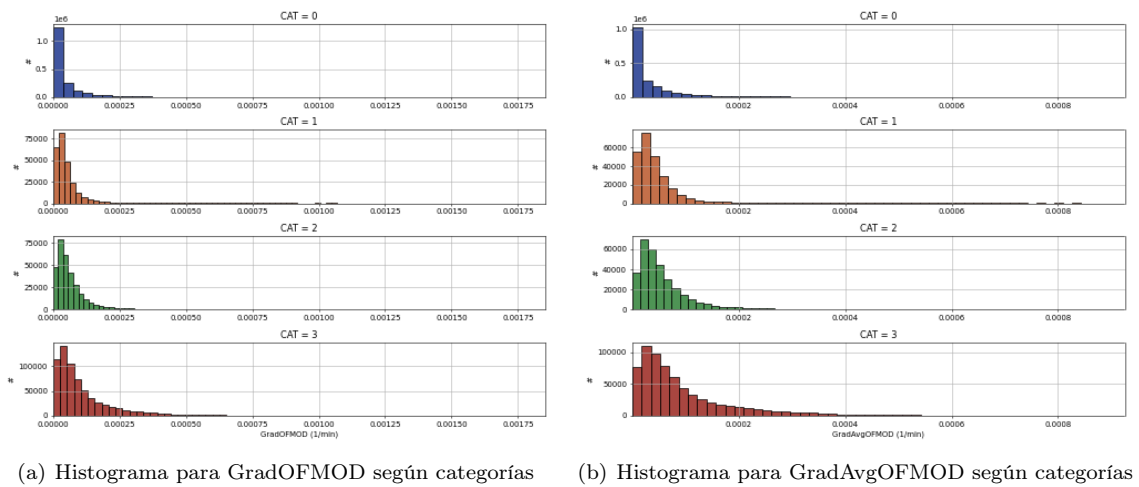


Figura 3.19: Histogramas para las características GradC13TB y GradOFMOD según categorías.

Capítulo 4

Clasificación satelital de altura de topes nubosos

Este capítulo presenta la evaluación de desempeño de tres tecnologías de aprendizaje automático desarrolladas para el problema de clasificación de altura de tope de nubes en base a la información satelital ABI y las variables derivadas. Esta evaluación de desempeño es comparada con la que se obtiene de utilizar el CTH ABI. Las mismas técnicas de aprendizaje automático se utilizan luego tomando al producto CTH ABI como característica de entrada. De esta forma, el nuevo enfoque propone el uso de los algoritmos de aprendizaje automático como herramientas de postproceso de este producto, lo que permitió mejorar el desempeño. El capítulo presenta además al inicio los elementos más importantes de las técnicas de aprendizaje automático y la forma de evaluación de los clasificadores.

4.1. Herramientas de aprendizaje automático utilizadas

Los clasificadores implementados están basados en tres tecnologías de aprendizaje automático distintas: un Árbol de Decisión simple (DTC), un Bósque aleatorio (RNDF) y una Red Neuronal Multicapa (MLP). Cada subsección debajo describe muy brevemente los aspectos más relevantes de estas herramientas. Durante este trabajo se tuvo que aprender a trabajar con estas metodologías y se realizaron varias pruebas. La implementación final de los clasificadores, en particular, la fijación de hiperparámetros (definiciones que ajustan la estrategia de trabajo del clasificador) se hizo por exploración manual, buscando no crecer en complejidad innecesaria o excesiva. No se presentan aquí el conjunto total de pruebas realizadas, sino los mejores resultados obtenidos de la exploración heurística. Como se mencionó, la determinación de estos hiperparámetros surge de ejercicios de exploración manual y no se puede asegurar que sea la configuración óptima de cada algoritmo para el problema que se presenta. Futuros trabajos podrían contemplar la optimización de los clasificadores en este sentido. Para la implementación de los algoritmos se utilizó el paquete python *scikit-learn* en su versión 1.5.

4.1.1. Árbol de decisión simple

Los *Árboles de Decisión*, DTC, son una técnica de aprendizaje automático construida en base a la partición recursiva de un conjunto de datos en base a criterios de decisión que son evaluados mediante algún modelo o condición de umbral. Su nombre deriva de la práctica de describir el proceso de partición comparándolo con un árbol con hojas y ramas. La Figura 4.1 presenta un esquema general de un algoritmo DTC. El proceso de aprendizaje conforma un árbol, compuesto por nodos internos, ramas y hojas. Los nodos surgen de las características de entrada del modelo, cada rama representa una decisión basada en esa característica y cada hoja representa el resultado de la clasificación (Loh, 2011). El objetivo es crear un modelo de clasificación a partir de la división de los datos de entrada en subconjuntos cada vez más puros con respecto a la clasificación base del

entrenamiento.

En particular, este trabajo implementó un DTC utilizando la clase python *DecisionTreeClassifier*. Se definió al índice de Gini como criterio para medir la calidad de las divisiones, junto con la estrategia de división denominada "best"(dentro del entorno *scikit-learn*). Esta estrategia selecciona la mejor división posible en cada nodo. Tras evaluar todas las características disponibles, asegura que se elijan las divisiones óptimas según el criterio escogido, mejorando así la precisión del algoritmo. La profundidad del DTC se limitó a cinco niveles, lo cual limita la complejidad y previene el sobreajuste. De esta forma, se logra que el clasificador generalice mejor cuando se enfrente a datos nuevos. También se estableció que un nodo solo pueda dividirse si contiene al menos quince muestras. Esto previene que el árbol se ajuste demasiado a los datos de entrenamiento, al evitar divisiones basadas en un número reducido de muestras.

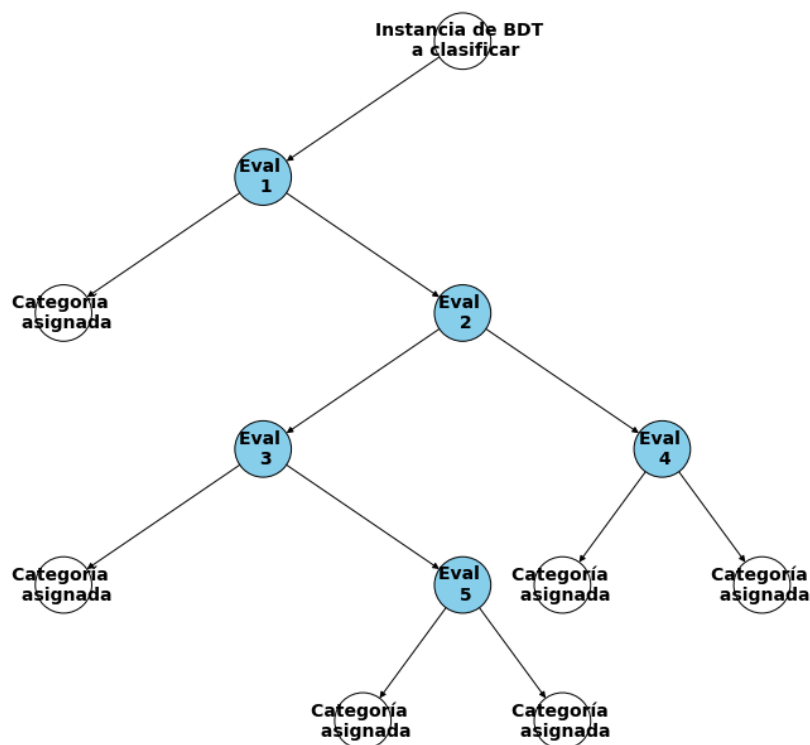


Figura 4.1: Esquema de funcionamiento del algoritmo DTC.

4.1.2. Bosques aleatorios

Los *Bosques Aleatorios*, RNDF, son una técnica de aprendizaje automático que implementa una cantidad predefinida de DTC. La aleatoriedad es introducida en cada nodo, donde la decisión de qué característica utilizar para hacer la división en base a la métrica escogida se realiza con una muestra aleatoria del conjunto de datos de entrenamiento, no con la totalidad. La clasificación de un RNDF se obtiene por consenso, votando todos los árboles que componen el RNDF. Esto se conoce como *bagging* y permite reducir la variación de la clasificación, promediando muchos modelos ruidosos (Breiman y Cutler, 2001). La Figura 4.2 muestra esquemáticamente el funcionamiento de un RNDF compuesto por 5 DTC.

Como se mencionó, la ventaja de los RNDF es el *bagging*, que permite reducir la variabilidad de la clasificación al promediar muchos modelos ruidosos pero aproximadamente imparciales. Los DTC son candidatos ideales para el *bagging*, dado que pueden encontrar estructuras de interacción compleja en los datos, y si crecen suficiente en profundidad, tienen relativamente baja parcialidad. Dado que los DTC son notoriamente ruidosos, se benefician de la decisión por consenso. RNDF es especialmente útil en casos donde las características tengan comportamiento ruidoso o incluso existan características irrelevantes. Al promediar los resultados de múltiples árboles de decisión,

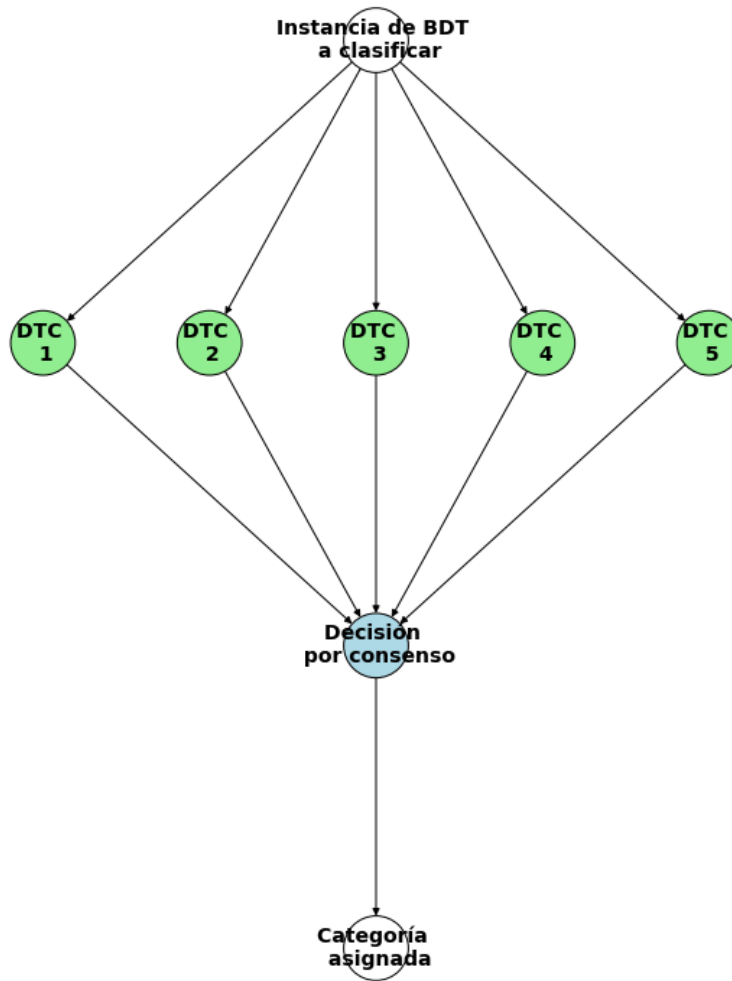


Figura 4.2: Esquema de funcionamiento de un algoritmo RNDf compuesto por 5 DTC.

RNDf reduce el riesgo de sobreajuste que podría ocurrir con un DTC. Cada DTC en el bosque se entrena con un subconjunto diferente de los datos de entrenamiento y a menudo, con un subconjunto diferente de características, esto mejora la capacidad del modelo para generalizar la clasificación.

Para este trabajo, se implementó un RNDf recurriendo a la clase python *RandomForestClassifier*. Se eligió que el clasificador contara con 305 estimadores (es decir, 305 DTC conforman el *Bosque Aleatorio*). De esta forma, el clasificador puede capturar una amplia gama de patrones en los datos de entrada de la BDT. Esto generalmente resulta en una mejora en la capacidad de generalización del clasificador, esto es, cuando el clasificador se enfrenta a datos no vistos en el entrenamiento.

4.1.3. Redes neuronales

Por último, las *Redes Neuronales*, MLP, son modelos computacionales inspirados en el cerebro humano que consisten en capas de neuronas interconectadas. En el caso de este trabajo se implementó una MLP con 3 capas ocultas, la primera capa de 15 neuronas, la segunda capa de 10 neuronas, y la tercera capa de 5 neuronas. De esta forma, la red captura y aprende patrones complejos en los datos, ya que cada capa oculta agrega un nivel de abstracción (Haykin, 1994). La Figura 4.3 muestra un ejemplo de red neuronal con 3 capas.

En este caso, se implementó un MLP utilizando la clase python *MLPClassifier*. En particular, se definió el número máximo de iteraciones 400. De esta forma se establece el número máximo de veces que el algoritmo puede recorrer el conjunto de datos de entrenamiento, para ajustar los pesos de la

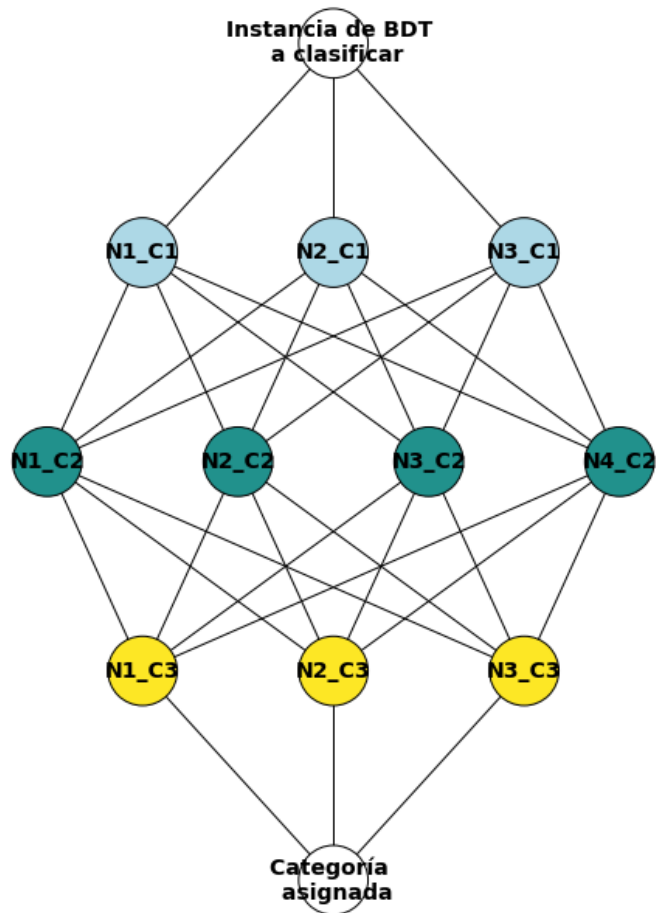


Figura 4.3: Esquema de funcionamiento del algoritmo MLP, con 3 capas ocultas de 3, 4 y 3 neuronas respectivamente.

red neuronal. Un número adecuado de iteraciones asegura que el modelo tenga suficiente oportunidad de converger a una solución óptima, sin sobreentrenarse.

4.2. Evaluación de desempeño de la clasificación

Se presenta en esta sección el procedimiento de evaluación de los clasificadores y las métricas de desempeño a utilizar.

4.2.1. Validación cruzada

Para la evaluación de desempeño de los clasificadores se utilizó la técnica de Validación Cruzada (VC) con 100 particiones aleatorias en conjuntos 90/10 . Esto significa que en cada iteración, el 90% de las imágenes (112 imágenes) son utilizadas para entrenamiento y el 10% son utilizadas para evaluación (13 imágenes). Las particiones del conjunto de datos se realizan por imágenes, no por píxeles. Este punto es importante, dado que los píxeles cercanos (o contiguos) están altamente correlacionados. Debido a esta correlación, se debe asegurar que píxeles de una misma imagen no son usados para entrenar y validar el algoritmo en ninguna iteración. Esto llevaría a métricas de desempeño artificialmente altas, debido a una capacidad predictiva del conjunto de entrenamiento irrealmente alta sobre el conjunto de evaluación. La no representatividad de esta situación de un caso real viene dada porque se busca que la herramienta sea capaz de clasificar píxeles en una imagen conocida la historia previa (su entrenamiento previo con otras imágenes), no clasificar algunos píxeles de la imagen conocidas las características de otros píxeles de la misma imagen (que es otro problema). En cada iteración se obtienen las métricas de desempeño sobre el conjunto de evaluación, que luego

son promediadas para obtener el desempeño final. La realización de 100 iteraciones asegura repetibilidad del experimento, y el procedimiento asegura que ninguna imagen o alguno de sus píxeles es utilizado en ningún momento para entrenamiento y evaluación al mismo tiempo.

4.2.2. Matriz de confusión y métricas de desempeño

La evaluación de un algoritmo de clasificación implica el uso de las métricas de desempeño que se detallan en esta subsección. En particular, la clasificación desarrollada en este trabajo es de clase múltiple ("*multi-class classification*") por lo que se consideran métricas de desempeño usuales en este tipo de experimentos. El trabajo de [Grandini et al. \(2020\)](#) presenta un abordaje detallado al respecto. En particular, en un problema de clasificación es útil desarrollar el concepto de Matriz de Confusión. Esta matriz permite resumir el resultado de la clasificación de forma sencilla y presentar, a partir de ella, la definición de métricas de desempeño usuales en este tipo de problemas. Para el caso de un problema de clasificación con tres categorías (A, B y C), la matriz de confusión se presenta de la siguiente manera:

	Predicción A	Predicción B	Predicción C
Real A	AP_A	$EP_{A,B}$	$EP_{A,C}$
Real B	$EP_{B,A}$	AP_B	$EP_{B,C}$
Real C	$EP_{C,A}$	$EP_{C,B}$	AP_C

Donde:

- AP_i son los aciertos de predicción para la categoría i .
- $EP_{i,j}$ representa los errores por predecir la categoría i como categoría j .

A partir de la matriz de confusión se pueden construir varias métricas que son de importancia para el análisis de desempeño de un clasificador. Las más comunes son la exactitud, la precisión, el recall, la especificidad y el F1 Score, que se describen a continuación:

Exactitud: La exactitud se define como la proporción de predicciones correctas realizadas por el clasificador. Esta métrica indica qué tan bien resulta la clasificación en general, considerando la proporción de predicciones correctas sobre el total de instancias a clasificar. Se debe notar que la exactitud es una métrica global del desempeño del modelo que no está desglosada por clase.

$$\text{Exactitud} = \frac{AP_A + AP_B + AP_C}{\text{Total de instancias}} \quad (4.1)$$

- Total de instancias = $AP_A + AP_B + AP_C + EP_{A,B} + EP_{A,C} + EP_{B,A} + EP_{B,C} + EP_{C,A} + EP_{C,B}$.

Precisión: La precisión mide la proporción de verdaderos positivos sobre el total de predicciones positivas. Es decir, considera cuántas de las predicciones positivas son realmente positivas. Es útil en contextos donde el costo de los falsos positivos es alto. Se calcula de manera independiente para cada categoría.

$$\text{Precisión}_i = \frac{AP_i}{AP_i + EP_{j,i} + EP_{k,i}} \quad (4.2)$$

Recall: El recall mide la proporción de verdaderos positivos sobre el total de instancias reales de esa categoría, comprobando cuántas de las instancias reales positivas fueron correctamente clasificadas. Es importante cuando el costo de los falsos negativos es alto.

$$\text{Recall}_i = \frac{AP_i}{AP_i + EP_{i,j} + EP_{i,k}} \quad (4.3)$$

Especificidad: La especificidad mide la proporción de verdaderos negativos sobre el total de verdaderos negativos y falsos positivos, comprobando cuántas de las instancias realmente negativas fueron correctamente clasificadas. Es útil cuando se necesita identificar correctamente las instancias negativas en la clasificación.

$$\text{Especificidad}_i = \frac{\text{TN}_i}{\text{TN}_i + \text{FP}_i} \quad (4.4)$$

Donde (por ejemplo, para $i = A$):

- $\text{TN}_A = \text{AP}_B + \text{AP}_C$.
- $\text{FP}_A = \text{EP}_{A,B} + \text{EP}_{A,C}$.

F1 Score: El F1 Score es la media armónica entre precisión y recall, combinándolas en una sola métrica. Es útil cuando se necesita un equilibrio entre precisión y recall:

$$\text{F1 Score}_i = 2 \times \frac{\text{Precisión}_i \times \text{Recall}_i}{\text{Precisión}_i + \text{Recall}_i} \quad (4.5)$$

Métricas ponderadas: Como se mencionó, la Exactitud es la única métrica global de desempeño de un clasificador. Esto quiere decir que no está definida para una categoría o clase en particular. Sin embargo, para el caso de categorías desbalanceadas (cómo es el caso de este trabajo) la Exactitud puede brindar una noción equivocada del desempeño global del clasificador. En estos casos es de utilidad considerar la Exactitud Balanceada, definida como la media aritmética de los Recall de cada clase. Para el caso de las demás métricas es necesario generalizarlas, transformándolas en globales. Para esto se deben calcular los valores medios ponderadas por categoría a través de la consideración de la cantidad de elementos (N_i con $i = \{A, B, C\}$) de cada clase. A continuación se definen la Exactitud Balanceada y las medias ponderadas que se utilizan en este trabajo.

$$\text{Exactitud Balanceada} = \frac{\text{Recall}_A + \text{Recall}_B + \text{Recall}_C}{3} \quad (4.6)$$

$$\text{Precisión Ponderada} = \frac{N_A \cdot \text{Precisión}_A + N_B \cdot \text{Precisión}_B + N_C \cdot \text{Precisión}_C}{N_A + N_B + N_C} \quad (4.7)$$

$$\text{Recall Ponderado} = \frac{N_A \cdot \text{Recall}_A + N_B \cdot \text{Recall}_B + N_C \cdot \text{Recall}_C}{N_A + N_B + N_C} \quad (4.8)$$

$$\text{Especificidad Ponderada} = \frac{N_A \cdot \text{Especificidad}_A + N_B \cdot \text{Especificidad}_B + N_C \cdot \text{Especificidad}_C}{N_A + N_B + N_C} \quad (4.9)$$

$$\text{F1 Score Ponderado} = \frac{N_A \cdot \text{F1 Score}_A + N_B \cdot \text{F1 Score}_B + N_C \cdot \text{F1 Score}_C}{N_A + N_B + N_C} \quad (4.10)$$

4.3. Desempeño de la clasificación

Durante el procedimiento de VC se ajustan y validan los clasificadores en cada iteración. El resultado del procedimiento de VC es esencialmente una evaluación de desempeño en condiciones realistas. El clasificador final, es decir, el que se pone en utilización, en general se obtiene de hacer un único entrenamiento sobre todo el conjunto de datos, y su desempeño esperado es el que se obtiene por VC. La Tabla 4.1 muestra las métricas de desempeño halladas para los tres clasificadores. Se presentan valores promedio de la VC a partir de las 100 iteraciones realizadas. El CTH ABI, al ser un producto que no se entrena con la BDT, se evalúa sobre todo el conjunto de la BDT.

Se puede observar que el desempeño de los clasificadores es competitivo con el desempeño de la clasificación derivada de CTH ABI. La Exactitud ubica al CTH ABI como el de mejor desempeño, seguido de cerca por el clasificador RNDF que está 0.5 % por debajo. Sin embargo, esto se debe a que el clasificador basado en CTH ABI es particularmente bueno para clasificar la categoría de nubes altas, y como ésta es la clase mayoritaria (ver Tabla 3.9) el desempeño se desbalancea a su favor. Esto se puede ver en los gráficos de la Figura 4.4, donde se ilustra la matriz de confusión de cada clasificador. Allí se observa un acierto de predicción de 98.3 % para las nubes altas del CTH ABI, en comparación a los registros entre 91-94 % de los otros clasificadores. Sin embargo, al considerar la Exactitud Balanceada, el algoritmo RNDF emerge como el de mejor desempeño considerando el desbalance de clases, seguido muy de cerca por MLP. En particular, en la categoría de nubes bajas, que es un foco de este trabajo debido al conocimiento previo de los errores en la predicción numérica de nubes y radiación solar, el clasificador RNDF provee el mejor desempeño, superando en un $\simeq 3\%$ al CTH ABI. Es notable observar que considerar el desbalance de clase reduce los porcentajes de Exactitud en 4-7 %, dependiendo del clasificador. Se destaca además que RNDF supera en Precisión Ponderada, F1 Score Ponderado y Especificidad Ponderada a la base de desempeño del CTH ABI. La Precisión Ponderada, no obstante, es muy parecida para todos los clasificadores, salvo el DTC, que se encuentra un par de puntos porcentuales por debajo.

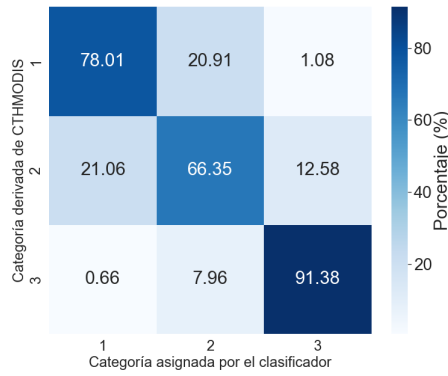
Se debe destacar en este punto que el producto CTH ABI utiliza insumos adicionales a las imágenes GOES16 para realizar su clasificación. Además de las imágenes de la observación satelital, se consideran salidas del modelo numérico global GFS y modelos de transferencia radiativa en la atmósfera, entre otros. De esta forma, CTH ABI utiliza más información que las imágenes operativas de ABI que son el insumo principal de los clasificadores de aprendizaje automático desarrollados en este trabajo. No obstante, el desempeño obtenido es competitivo con el CTH ABI, y en varios casos un poco superior, especialmente en la categoría de nubes de bajas, de especial interés para este trabajo.

Clasificador	Exactitud	Exactitud Balanceada	Precisión Ponderada	Recall Ponderado	F1 Score Ponderado	Especific. Ponderada
DTC	82.5 %	78.2 %	83.3 %	82.5 %	82.6 %	77.5 %
MLP	84.9 %	80.8 %	85.3 %	84.9 %	84.9 %	80.2 %
RNDF	85.2 %	80.9 %	85.5 %	85.2 %	85.2 %	80.4 %
CTHABI	85.7 %	78.9 %	85.4 %	85.7 %	85.1 %	78.1 %

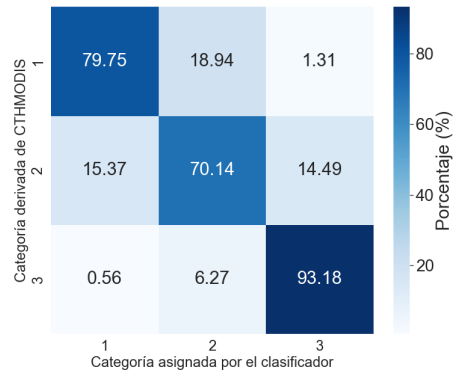
Tabla 4.1: Valores medios de las métricas obtenidas en las 100 iteraciones del proceso VC.

De la Figura 4.4 se pueden obtener más conclusiones, dado que permiten observar el comportamiento de los clasificadores para cada categoría. Se observa que todos los clasificadores desarrollados superan a la base de desempeño en las tasas de verdaderos positivos para las categorías 1 y 2. Esto indica una mayor habilidad de los clasificadores basados en aprendizaje automático para reconocer la nubosidad baja y media. También se destaca que los falsos positivos son máximos entre categorías contiguas, siendo la peor tasa de falsos positivos el caso de las instancias $EP_{2,3}$ (categoría 2 clasificados como categoría 3) por el CTH ABI. Por otra parte, los $EP_{3,1}$ y $EP_{3,2}$ son casi inexistentes en el CTH ABI, mientras que para los clasificadores desarrollados estas tasas empeoran un poco, manteniéndose en todos los casos por debajo de 8 %. Por otra parte, los $EP_{1,3}$ bajan considerablemente en las clasificaciones con aprendizaje automático respecto a la base de desempeño. El CTH ABI muestra una tasa de $EP_{3,1} = 5.3\%$, mientras que los clasificadores logran tasas de 1.1 % (DTC), 1.3 % (MLP) y 0.8 % (MLP).

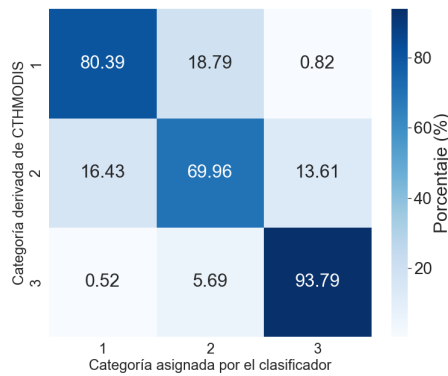
En general, se observa una apariencia más simétrica en las matrices de confusión de los clasificadores basados en aprendizaje automático, frente a la matriz de confusión del CTH ABI. Esto



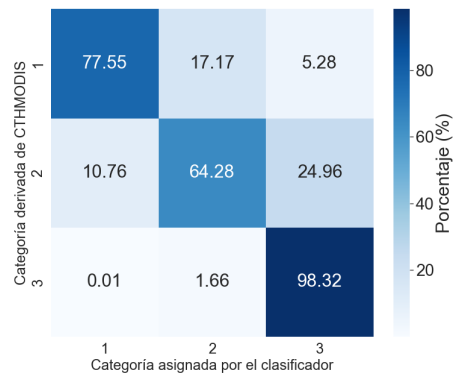
(a) Matriz de confusión para DTC



(b) Matriz de confusión para MLP



(c) Matriz de confusión para RNDF



(d) Matriz de confusión para CTH ABI

Figura 4.4: Matrices de confusión normalizadas para los clasificadores para los distintos clasificadores. Para los tres primeros se presenta el promedio de porcentajes de clasificación en cada iteración. Para el CTH ABI se presenta los porcentajes sobre toda la BDT.

indica un desempeño más parejo entre las categorías por parte de los primeros. Al contrario, como se mencionó en el párrafo anterior, la clasificación derivada de CTH ABI muestra un muy buen desempeño en la categoría 3, con una $AC_{3,3} = 98.3\%$ pero no tan bueno en las demás categorías: $AC_{1,1} = 77.6\%$ y $AC_{2,2} = 64.3\%$. En el caso de RNDF las tasas son $AC_{1,1} = 80.4\%$, $AC_{2,2} = 70.0\%$ y $AC_{3,3} = 93.8\%$, superando a CTH ABI en la predicción correcta de las categorías 1 y 2.

En la Figura 4.5 se presenta a modo de ejemplo el resultado de las 4 clasificaciones para un día en particular, el día 30 de enero del año 2020 a las 13:30 hs UTC sobre el territorio uruguayo. Las clasificaciones obtenidas son todas consistentes espacialmente y reflejan situaciones similares, con algunas diferencias. Como es de esperar luego de estudiar las métricas de desempeño de las clasificaciones, se tiene que en general todos los clasificadores desarrollados logran ofrecer una imagen de clasificación similar a la del producto CTH ABI. En particular, la clasificación lograda por RNDF en la figura 4.5 (c) es prácticamente indistinguible de la clasificación del producto CTH ABI en la figura 4.5 (d). Al mismo tiempo, se puede observar que las diferencias se observan en general en las interfaces de las categorías (en los bordes espaciales de la clasificación).

4.4. Desempeño de la clasificación incluyendo el CTH ABI

Dada la disponibilidad del CTH ABI, y su buen desempeño, tiene sentido utilizar este producto como característica de entrada a los clasificadores, dado que su disponibilidad es la misma que la de las imágenes satelitales. En este enfoque, además de las 23 características consideradas en el enfoque

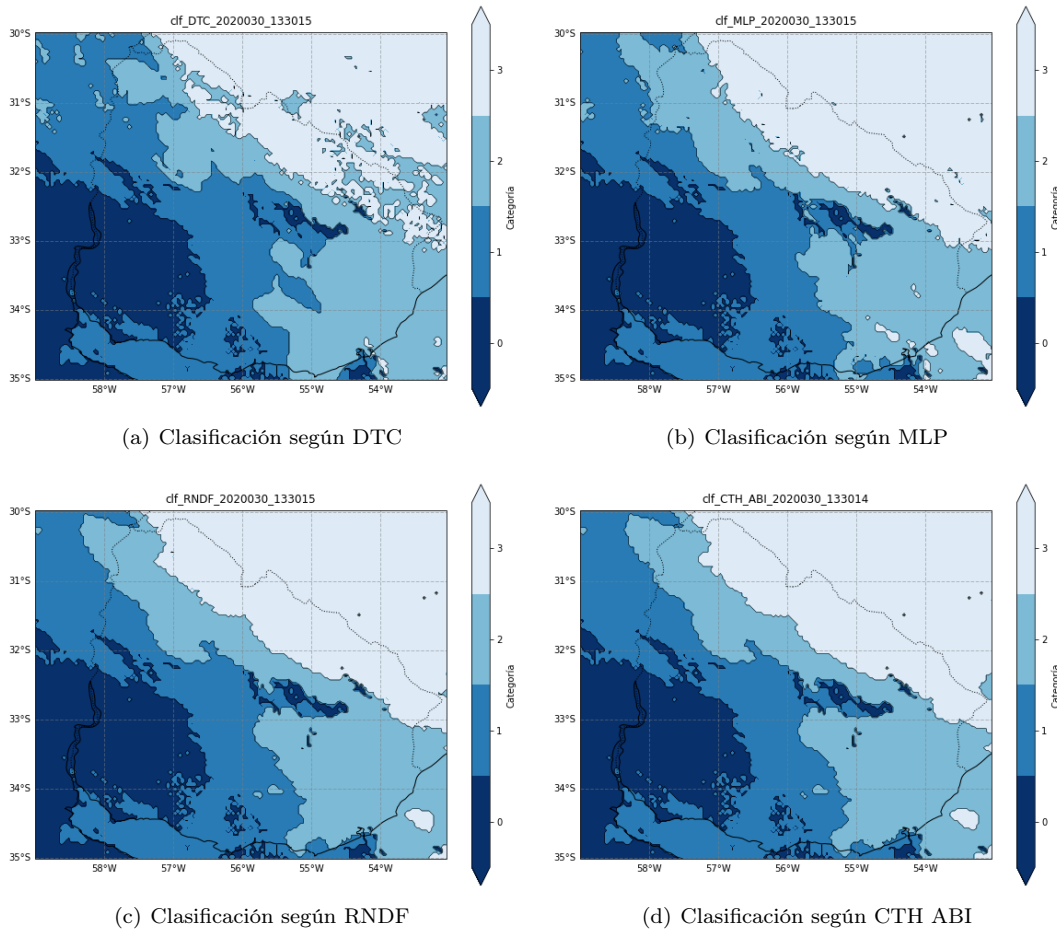


Figura 4.5: Clasificación para el día 30 de enero de 2020, 13:30hs.

original visto en la sección anterior, se suma como característica 24 el producto CTH ABI. De esta forma, ahora la estimación de CTH obtenida por ABI ingresa al espacio de las características del problema y el uso del aprendizaje automático de esta manera podría verse como un postproceso de la clasificación del CTH ABI. La aplicación de las tres metodologías de aprendizaje automático con esta modalidad se denominan aquí CTH-DTC, CTH-MLP y CTH-RNDF.

La Tabla 4.2 reúne las métricas de desempeño para cada clasificador de la misma forma que la Tabla 4.1 e incluyendo en la última fila la base de desempeño del CTH ABI. Se observa que todos los clasificadores presentan un mejor desempeño que el producto CTH ABI y también un mejor desempeño que los algoritmos de la sección anterior, cuando no se usaba el CTH ABI como característica de entrada. Esto ocurre incluso para la métrica de Exactitud sin considerar su adecuación por el desbalance de clases. Se observa nuevamente que el algoritmo basado en RNDF es el de mejor desempeño en la clasificación. Se observan mejoras notables de estos clasificadores con respecto a los anteriores, con un aumento de Exactitud Balanceada entre 4-6 %, de Precisión Ponderada de 4-5 %, de Recall Ponderado de 3-5 %, de F1 Score Ponderado de 3-5 %, y de Especificidad Ponderada de 4-7 %. Es decir, un aumento promedio de cinco puntos porcentuales en todas las métricas debido al uso del CTH ABI como entrada adicional a la clasificación. Esto muestra que el CTH ABI cuenta con valor adicional a la información en las imágenes de satélite, ya sea por la física modelada o por la información adicional que se utiliza para la confección del producto.

En la Figura 4.6 se presentan las matrices de confusión normalizadas. Se observa que el uso de los algoritmos de aprendizaje automático como herramientas de postproceso mejora las matrices de confusión en comparación con las de la sección anterior, volviéndolas más simétricas y minimizando los falsos positivos de categorías no contiguas ($EP_{3,1}$ y $EP_{1,3}$). Esto indica que el uso de CTH ABI como característica de entrada ofrece una mejora, reconociendo de forma más uniforme a las distintas categorías, no como en el caso de la base de desempeño que presenta una grandísima habilidad para reconocer a la nubosidad de categoría 3 pero no consigue igual desempeño en las restantes categorías

Clasificador	Exactitud	Exactitud Balanceada	Precisión Ponderada	Recall Ponderado	F1 Score Ponderado	Especific. Ponderada
CTH-DTC	87.9 %	84.4 %	88.4 %	87.9 %	88.0 %	84.1 %
CTH-MLP	88.3 %	84.6 %	88.6 %	88.3 %	88.3 %	84.2 %
CTH-RNDF	89.1 %	85.9 %	89.3 %	89.1 %	89.1 %	85.5 %
CTHABI	85.7 %	78.9 %	85.4 %	85.7 %	85.1 %	78.1 %

Tabla 4.2: Valores medios de las métricas obtenidas en las 100 iteraciones del proceso VC.

nubosas.

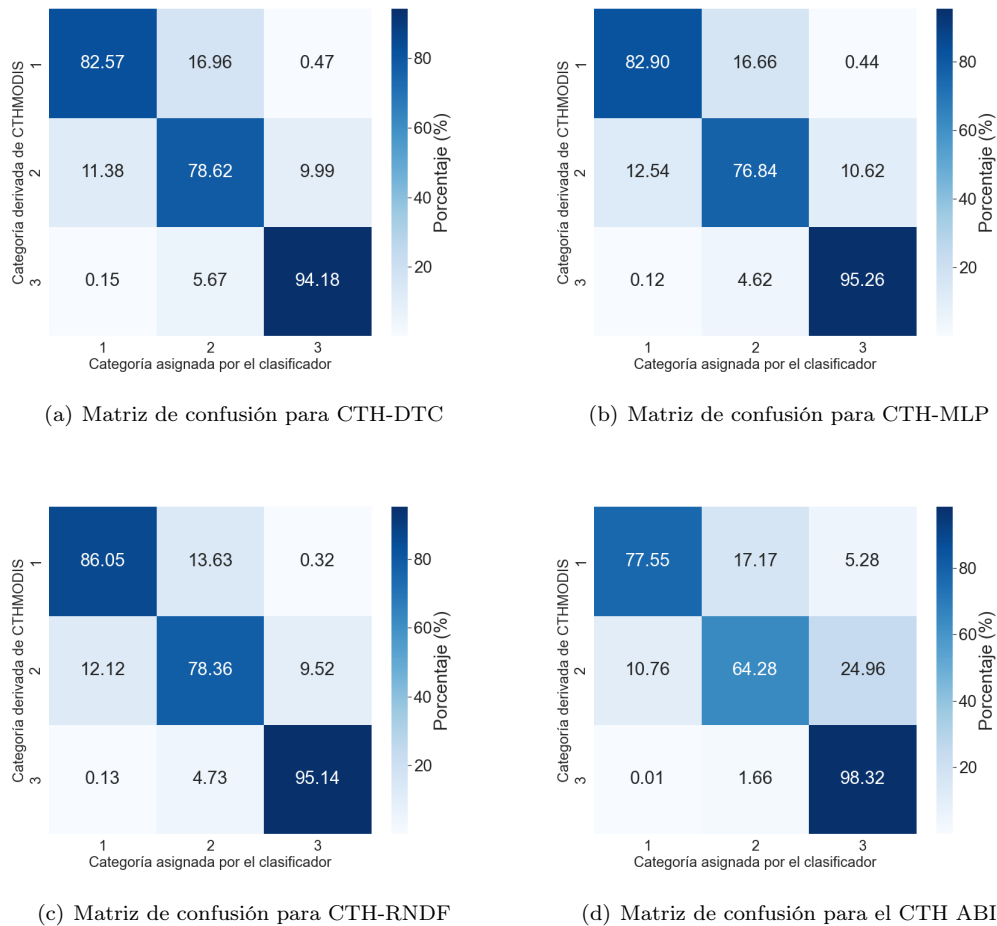


Figura 4.6: Matrices de confusión normalizadas para los clasificadores CTH-DTC, CTH-MLP y CTH-RNDF. Se agrega también la matriz de confusión de la base de desempeño de CTH ABI.

En la Figura 4.7 se presenta el resultado de las 4 clasificaciones para el día 30 de enero del año 2020 a las 13:30 hs UTC sobre el territorio uruguayo, la misma escena que en la Figura 4.5 de la sección anterior. Nuevamente, se tiene que en general todos los clasificadores desarrollados logran ofrecer una imagen de clasificación similar a la de la base de desempeño. Solo se observan discrepancias en la imagen del clasificador CTHMLP, donde una aparece una región de nubosidad de categoría 2 en la zona de los -31.0°S , -58.5°W , que no se observa en las demás clasificaciones. Dicho esto, nuevamente se observa que la clasificación lograda por CTHRNDF en la figura 4.7 (c) es prácticamente indistinguible de la clasificación del producto CTHABI en la figura 4.7 (d). Se observa en particular que los clasificadores basados en DTC y MLP mejoran los outliers de clasificación, acercándose a la clasificación de la escena provista por RNDF, CTHRNDF y CTHABI.

El análisis proporcionado en este capítulo permitió desarrollar y evaluar clasificadores de altura de tope de nube sobre el territorio de Uruguay en base a información operativa del satélite GOES16. El trabajo se realizó en forma incremental, partiendo de clasificadores que utilizan sólo imágenes

satelitales y variables calculadas de ellas, y llegando a un clasificador que además integra a su entrada el producto GOES16 de CTH ABI. En este proceso, se alcanzó un clasificador de escenas (el CTH-RNDF) de 89 % de Exactitud y 86 % de Exactitud Balanceada, y que logra clasificar nubes bajas con un 86 % de acierto. Asimismo logra clasificar nubes altas con un 96 % de acierto y nubes intermedias con un 78 % de acierto. Estos son los mejores resultados obtenidos en este trabajo, y por tanto este será el clasificador utilizado para evaluar eventos de predicción de nubosidad en el Uruguay y diagnosticar su comportamiento en presencia de nubes con distintas alturas de tope.

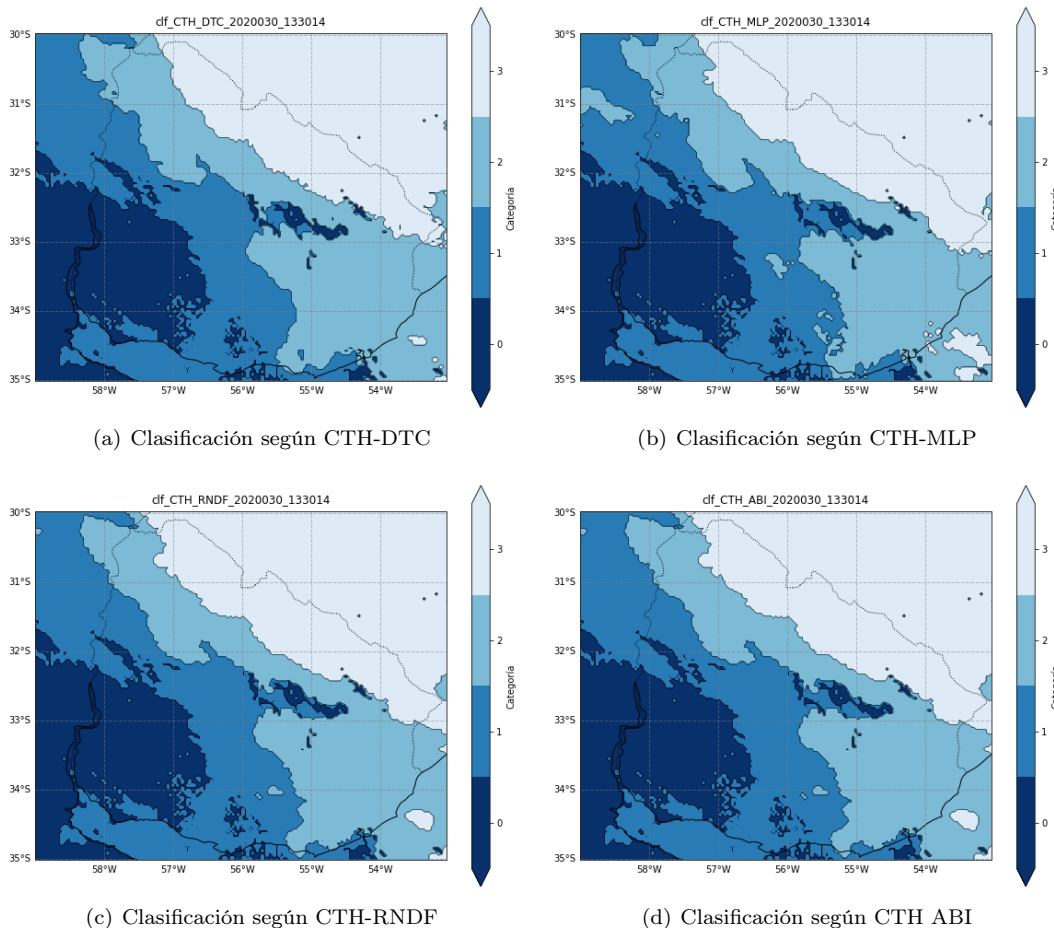


Figura 4.7: Imágenes clasificadas para el día 30 de enero de 2020, 13:30hs. Las figuras (a), (b) y (c) son el resultado de las técnicas de aprendizaje automático utilizando el CTH ABI como característica adicional. La figura (d) repite la clasificación de base del CTH ABI.

4.5. Análisis de sensibilidad

Para la estimación de intervalos de confianza de las métricas presentadas en las secciones 4.3 y 4.4, se eligió el método de *Bootstrap*. Este método de remuestreo es desarrollado en profundidad en el trabajo de Efron y Tibshirani (1994). Su ventaja consiste en que permite la estimación de intervalos de confianza independientemente de la distribución específica de la muestra. De esta forma, el método *Bootstrap* resulta apropiado en situaciones donde la hipótesis de normalidad de la distribución muestral no aplica o cuando se desconoce la distribución de los datos. Para ejemplificar el método, se define un conjunto de datos muestrales $\mathcal{X} = \{x_1, x_2, \dots, x_n\}$. De este conjunto se pretende estimar un estadístico θ , por ejemplo la media muestral. El *bootstrapping* consiste en los siguientes pasos:

1. Primero se generan m conjuntos \mathcal{X}_b^* de tamaño n aplicando Bootstrap. Esto se logra mediante muestreo con reemplazo a partir de la muestra original \mathcal{X} , lo que permite generar nuevas muestras con la misma estructura estadística.
2. Luego, se calcula el estimador $\hat{\theta}_b^*$ para cada conjunto b . En el caso de que se busque estimar la media muestral, se calculará la media en cada uno de los m conjuntos \mathcal{X}_b^* , de esta forma se obtiene una distribución empírica de la media.
3. Para finalizar, obtenida la distribución empírica del estimador $\hat{\theta}_b^*$ se llega a los límites del intervalo de confianza $1-\alpha$ a partir de los percentiles $\alpha/2$ y $1-\alpha/2$ de la distribución empírica. Así, se obtiene el Intervalo de Confianza con un nivel de confianza $1-\alpha$ para el estadístico calculado:

$$IC_{(1-\alpha)} = \left[\hat{\theta}_{\alpha/2}^*, \hat{\theta}_{1-\alpha/2}^* \right]. \quad (4.11)$$

De esta forma se resume la base del cálculo del Intervalo de Confianza para las métricas obtenidas que se presenta a continuación. Se definió $\alpha = 5\%$, determinando así Intervalos de Confianza con nivel de confianza del 95%. Para este cálculo se aprovechó el conjunto de 100 muestras de cada métrica obtenido en las 100 iteraciones del proceso de validación cruzada descrito en 4.2.1. A continuación, las Figuras 4.8, 4.9, 4.10, 4.11, 4.12 y 4.13 resumen los intervalos de confianza obtenidos mediante el método de Bootstrapping para las métricas calculadas y presentadas en 4.3 y 4.4.

Se observa que los intervalos de confianza respaldan el análisis de las secciones 4.3 y 4.4. En particular, la Figura 4.10 muestra la validez de afirmar el mejor desempeño en Exactitud Balanceada para los clasificadores RNDF y MLP frente a la base de desempeño CTH ABI destacada en 4.3. También, se destaca la superioridad general a todas las métricas de los clasificadores CTH-DTC, CTH-MLP y CTH-RNDF frente a la base de desempeño CTH ABI visto en 4.4.

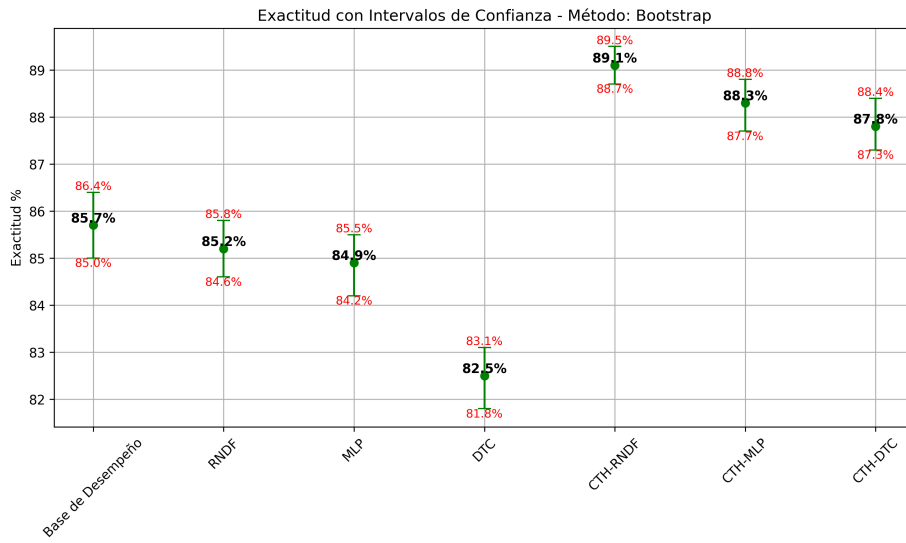


Figura 4.8: Métrica Exactitud obtenida para la Base de Desempeño y los clasificadores implementados, con los intervalos de confianza calculados por la técnica de Bootstrapping.

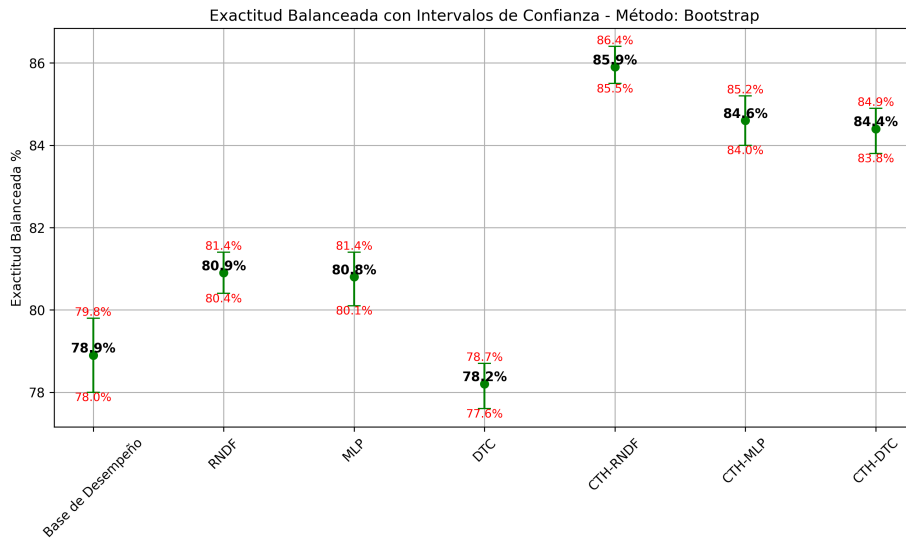


Figura 4.9: Métrica Exactitud Balanceada obtenida para la Base de Desempeño y los clasificadores implementados, con los intervalos de confianza calculados por la técnica de Bootstrapping.

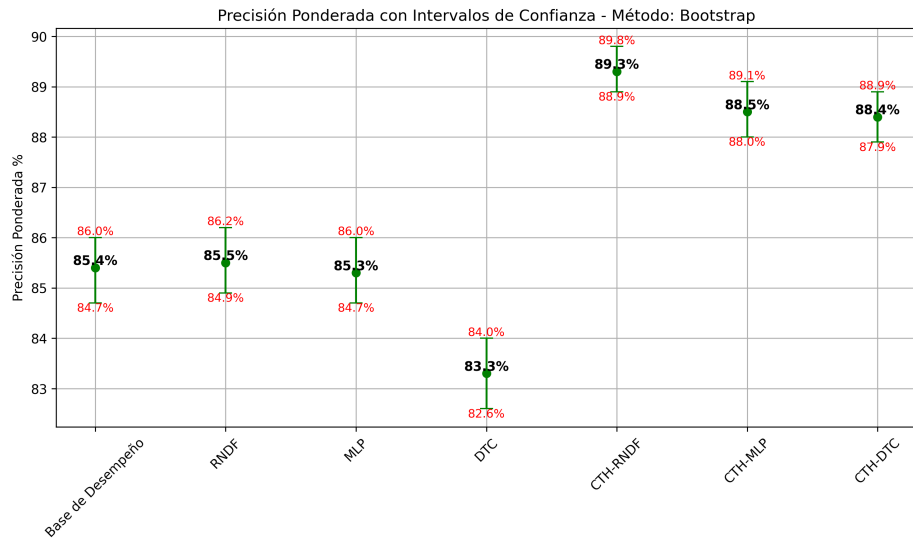


Figura 4.10: Métrica Precisión Ponderada obtenida para la Base de Desempeño y los clasificadores implementados, con los intervalos de confianza calculados por la técnica de Bootstrapping.

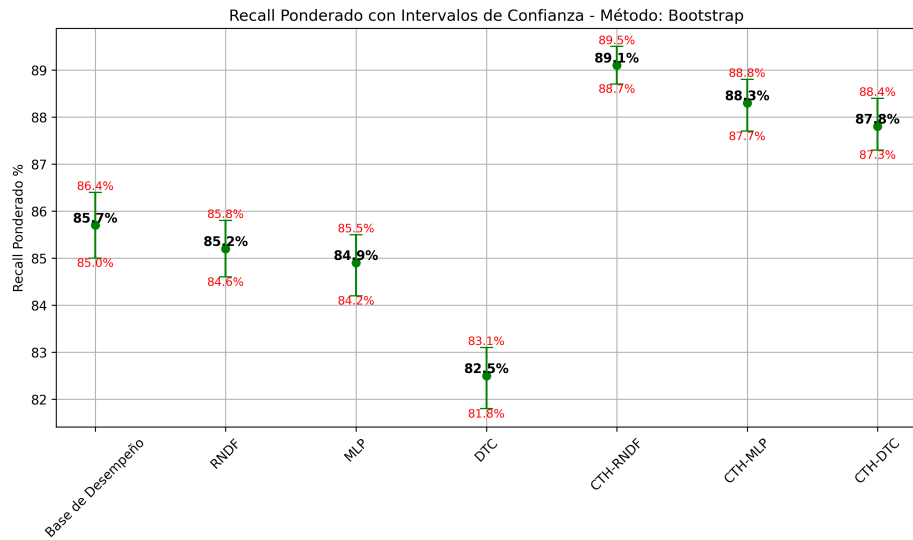


Figura 4.11: Métrica Recall Ponderado obtenido para la Base de Desempeño y los clasificadores implementados, con los intervalos de confianza calculados por la técnica de Bootstrapping.

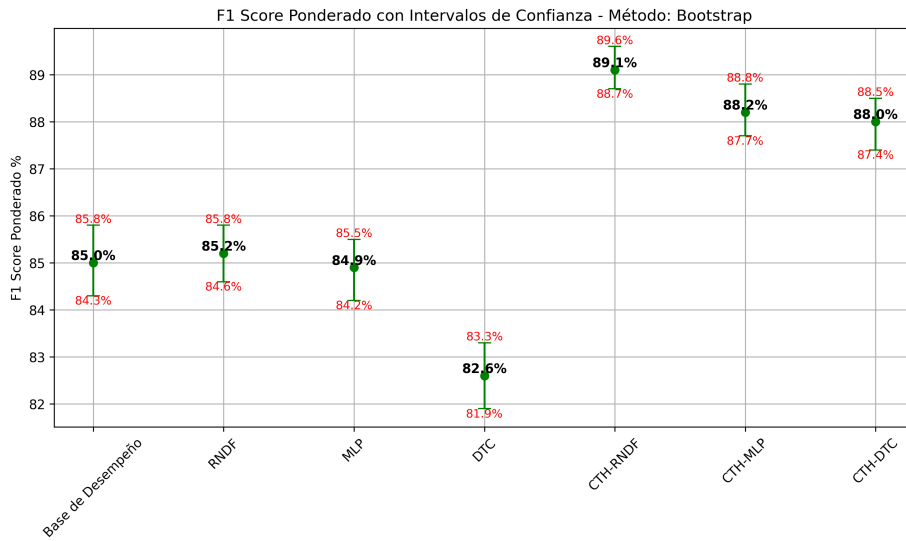


Figura 4.12: Métrica F1 Score Ponderado obtenido para la Base de Desempeño y los clasificadores implementados, con los intervalos de confianza calculados por la técnica de Bootstrapping.

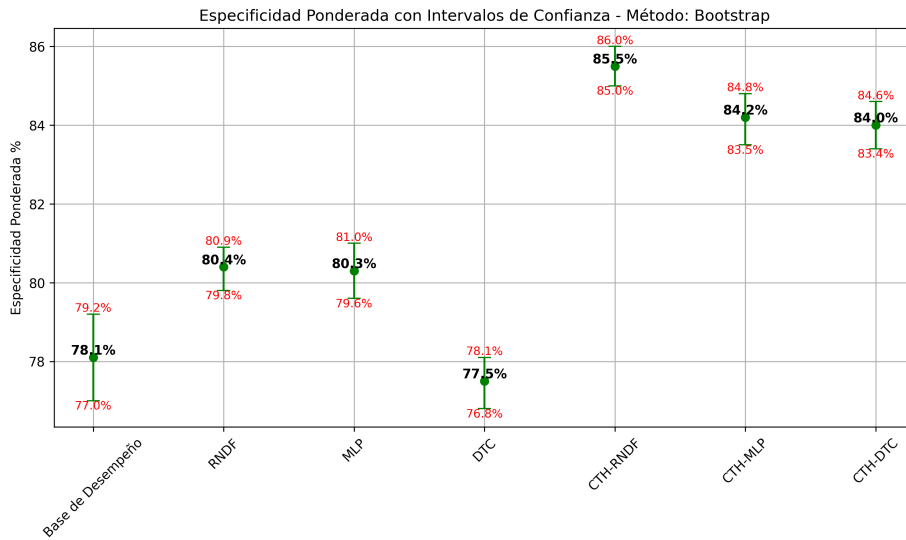


Figura 4.13: Métrica Especificidad Ponderada obtenida para la Base de Desempeño y los clasificadores implementados, con los intervalos de confianza calculados por la técnica de Bootstrapping.

Página intencionalmente dejada en blanco.

Capítulo 5

Selección y análisis de eventos

Este capítulo trata sobre el uso del clasificador desarrollado para seleccionar días de nubosidad significativa y casos de estudio de interés para la predicción numérica de nubosidad. Se define el método de selección utilizado, que surge de la aplicación de indicadores derivados del clasificador de mejor desempeño (CTH-RNDF). Luego, ciertos días seleccionados son estudiados tanto dinámicamente como termodinámicamente a partir de simulaciones numéricas obtenidas con el modelo WRF. El objetivo es evaluar primariamente la capacidad del modelo WRF de reproducir los patrones de nubosidad de los casos seleccionados, y sugerir hipótesis sobre procesos físicos asociados a la nubosidad en los casos en que el modelo es exitoso. Se detallan las características de la versión WRF elegida y la configuración definida.

5.1. Selección de días

Se busca identificar días de nubosidad significativa en el periodo comprendido entre el 1° de noviembre del 2019 y el 29 de febrero del año 2020. Para lograr esto, se construyó un índice de fracción de categoría nubosa para cada imagen clasificada. La fracción de cada categoría se calculó dividiendo el número de píxeles pertenecientes a esa categoría entre el número total de píxeles en la imagen (24795, ver sección 3.2). La Ec. (5.1) presenta la definición de la fracción de categoría nubosa que se deriva de la clasificación.

$$fr_{cat_i} = \frac{pxs_{cat_i}}{pxs_{tot}} \quad \text{donde } i = 0, 1, 2, 3 \quad (5.1)$$

Donde:

- fr_{cat_i} es la fracción de categoría nubosa para la categoría i .
- pxs_{cat_i} es la cantidad de píxeles con categoría i en la imagen.
- pxs_{tot} es la cantidad de píxeles totales en la imagen.

Así, se tiene que cada día reúne una cierta cantidad de imágenes 10-minutales diurnas de ABI y esta cantidad de imágenes depende del momento del año, debido al cambio natural en la duración del día. Entonces, para cada día, se determina el máximo valor de fracción de categoría nubosa en cada categoría de nubosidad. Luego, para cada categoría, se ordenan los días según sus valores de fracción máxima de mayor a menor. De este modo, los primeros diez días serán los que tengan los diez valores mayores de fracción de categoría nubosa máxima diaria dentro de cada categoría.

Para el estudio de eventos, se eligieron días con nubosidad de categoría 1 dominantes, dado que las nubes bajas tiene efectos radiativos importantes y está comprobada la dificultad en predecirlas de los NWP, en particular del modelo WRF. También, se consideró especialmente a la categoría 3 por ser la categoría nubosa más frecuente en la BDT (ver 3.9), de esta forma también se seleccionaron

días con nubosidad categoría 3 dominante. Para este trabajo no se consideran casos de eventos con nubosidad categoría 2, porque se decidió dedicar especial atención a las categorías 1 y 3. La Tabla 5.1 resume los días seleccionados en el periodo comprendido entre el 1° de noviembre del año 2019 y el 29 de febrero del año 2020.

Categoría 1 - Nubes bajas	Fracción Nubosa Máx.	Categoría 3 - Nubes Altas	Fracción Nubosa Máx.
07/01/2020	0.78	04/11/2019	1.00
01/01/2020	0.66	26/11/2019	1.00
15/11/2019	0.66	15/12/2019	1.00
08/01/2020	0.64	17/02/2020	1.00
02/01/2020	0.54	20/12/2019	0.99
21/12/2019	0.53	02/11/2019	0.99
17/01/2020	0.50	15/01/2020	0.99
10/11/2019	0.49	10/01/2020	0.98
05/11/2019	0.47	25/11/2019	0.94
15/02/2020	0.46	21/01/2020	0.92

Tabla 5.1: Fechas seleccionadas para la Categoría 1 - Nubes bajas y la Categoría 3 - Nubes altas. Se presentan además las fracciones de categorías nubosas máximas registradas en cada día.

Con los diez días seleccionados para cada categoría (nubes bajas y nubes altas), se elaboraron gráficos que representan la evolución diaria de las fracciones de categorías nubosas dentro de cada día. Estos gráficos permiten visualizar y analizar cómo varían las fracciones de categorías de nubes a lo largo de los días seleccionados. En la Figura 5.1 se muestra un ejemplo de un diagrama construido para visualizar el cambio de la nubosidad en un día (el 16/02/2020) sobre el territorio de Uruguay. Se observa que el día (diurno) inicia con un alto porcentaje de píxeles despejados, y paulatinamente aumenta la nubosidad en el país durante el día, especialmente las nubes altas a partir de la mitad del día. El día culmina con un $\simeq 60\%$ de nubes altas en el territorio y el restante $\simeq 40\%$ de píxeles despejados. Este tipo de visualización fue construida en este trabajo, y fue muy útil para poder seleccionar y entender casos a partir de previsualizar la evolución temporal de la nubosidad en el día, siendo más sencillo que visualizar las imágenes clasificadas en secuencia de ese día.

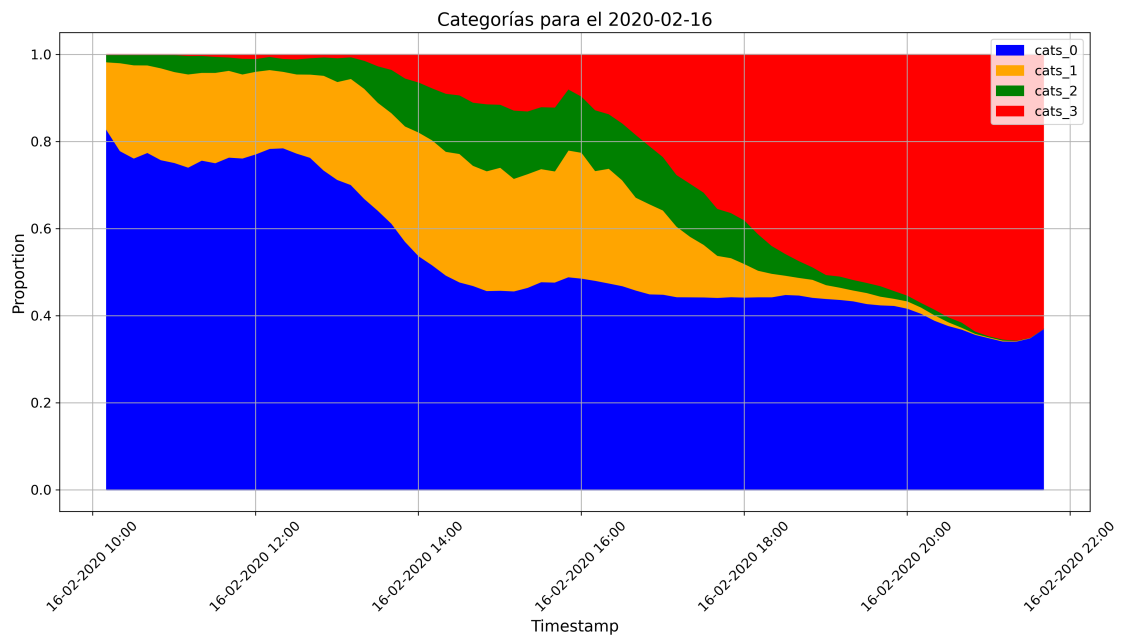


Figura 5.1: Ejemplo de evolución diaria de las fracciones de categorías nubosas para el día 16 de febrero de 2020. El diagrama permite visualizar la evolución de la nubosidad en el día.

5.2. Setup del WRF

Para las simulaciones se utilizó el modelo WRF versión 4.1.2. Se definió un paso temporal de 60 minutos. El dominio de simulación está centrado en Uruguay, tiene una resolución horizontal de 30 km y una resolución vertical de 33 niveles. La grilla de simulación quedó conformada por 110 puntos en la dirección este-oeste y 101 puntos en la dirección norte-sur, cubriendo una extensión latitudinal de aproximadamente 3030 km y longitudinal de aproximadamente 3330 km. La caja de simulación abarca aproximadamente desde 20°S a 45°S en latitud y desde 70°W a 40°W en longitud. El modelo se ejecutó en su versión no hidrostática, y las condiciones de borde e iniciales de las simulaciones fueron obtenidas del modelo GFS. La Tabla 5.2 muestra las parametrizaciones definidas para las 4 simulaciones.

Parámetro	Valor	Descripción
SW Radiation (ra_sw_physics)	1	Dudhia scheme
LW Radiation (ra_lw_physics)	1	RRTM scheme
Microphysics (mp_physics)	2	Purdue Lin scheme
Convección (cu_physics)	1	Kain-Fritsch (new Eta) scheme
Capa Límite (bl_pbl_physics)	2	YSU (Yonsei University) scheme
Surface Layer Physics (sf_sfclay_physics)	2	Eta similarity

Tabla 5.2: Parametrizaciones de interés definidas en la configuración del modelo WRF para la simulación de los casos seleccionados.

Para el análisis de casos, se eligieron las horas principales del día: 0, 3, 6, 9, 12, 15, 18 y 21 UTC. En la inspección de los perfiles verticales, se eligió una sección transversal con dirección SW a NE, según la figura 5.2 (b), donde se evaluó el pronóstico de nubosidad a lo largo de esta sección. Además, se definió el punto central de esta sección transversal, de coordenadas -32.82°S , -56.44°W , donde se evalúan los perfiles verticales de Humedad Relativa, Temperatura Potencial y se realiza el diagrama aerológico SkewT-Lp para inspeccionar el potencial de inestabilidad y la humedad de los perfiles. También en este punto se presenta el nivel de presión del tope de la PBL.

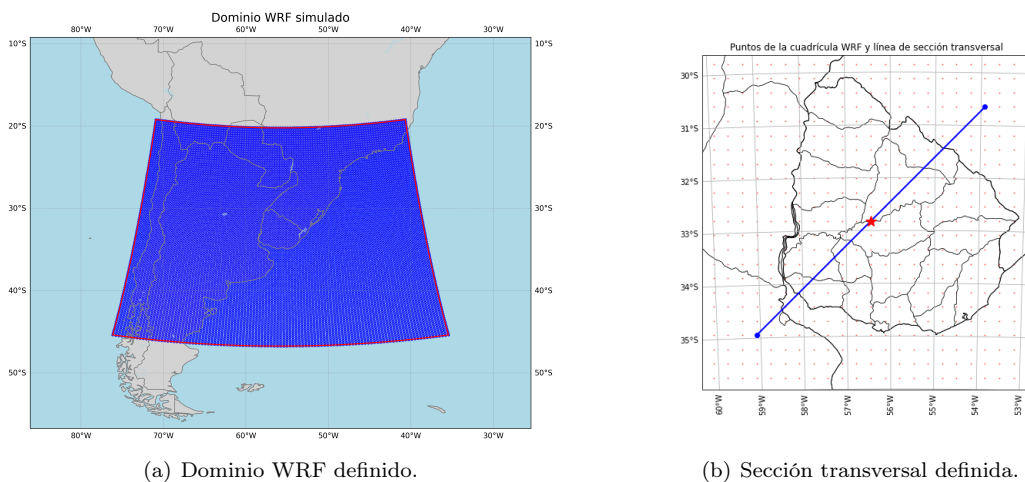


Figura 5.2: (a) Dominio y grilla base de las simulaciones realizadas en WRF. (b) Sección transversal para estudio de perfiles. El punto central de la sección transversal fue elegido para evaluar en detalle los perfiles verticales de la simulación.

También se eligió la variable *Cloud Fraction*, CLDFRA, del modelo WRF, que considera 3 categorías nubosas similares a las categorías determinadas por los clasificadores. Como se mencionó en la Sección 2.1, esta variable muestra la fracción de cielo cubierto por la nubosidad de cada categoría, para cada punto de grilla. De esta forma, con la variable CLDFRA se tiene una visión sinóptica de la nubosidad pronosticada por el modelo. Además, se presentan los vientos y el geopotencial en 200 hPa y 850 hPa para evaluar los patrones sinópticos de circulación dominantes.

5.3. Análisis de casos

5.3.1. Evento del 21 de diciembre de 2019

El día 21 de diciembre de 2019 presenta un valor de Fracción Nubosa máximo de 0.53 en la categoría 1 (nubes bajas), lo que lo posiciona en el sexto lugar dentro del conjunto de días seleccionados por ser dominante la nubosidad de esta categoría. La Figura 5.3 muestra la evolución de las fracciones nubosas a lo largo del día. Se puede observar como la nubosidad categoría 3 domina en las primeras horas de la mañana, dando paso luego al dominio de la nubosidad categoría 1 hasta el mediodía, acompañado también por un incremento de la nubosidad categoría 2. Luego del mediodía comienza a dominar la categoría 0 (se empieza a despejar el territorio).

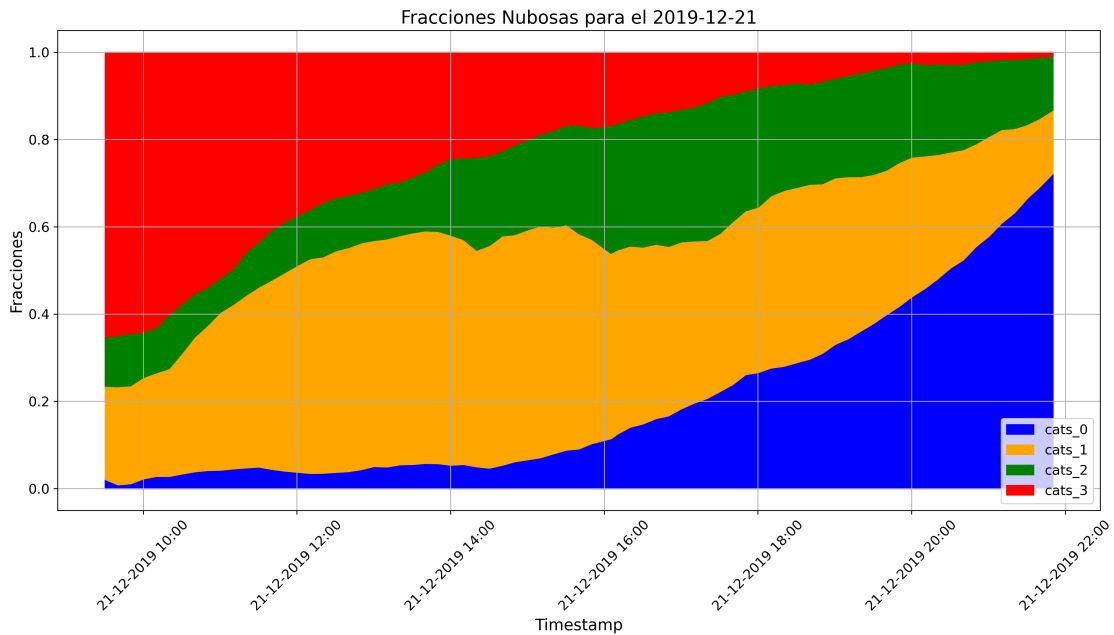


Figura 5.3: Evolución diaria de las fracciones nubosas para el día 21 de diciembre de 2019.

Las Figuras 5.4 y 5.5 muestran la circulación dominante. En el nivel de 200 hPa (Figura 5.4) se puede observar el rápido pasaje de una vaguada. Asociado a esta vaguada en altura se observa el desarrollo rápido de un ciclón extratropical en el nivel de 850 hPa (Figura 5.5).

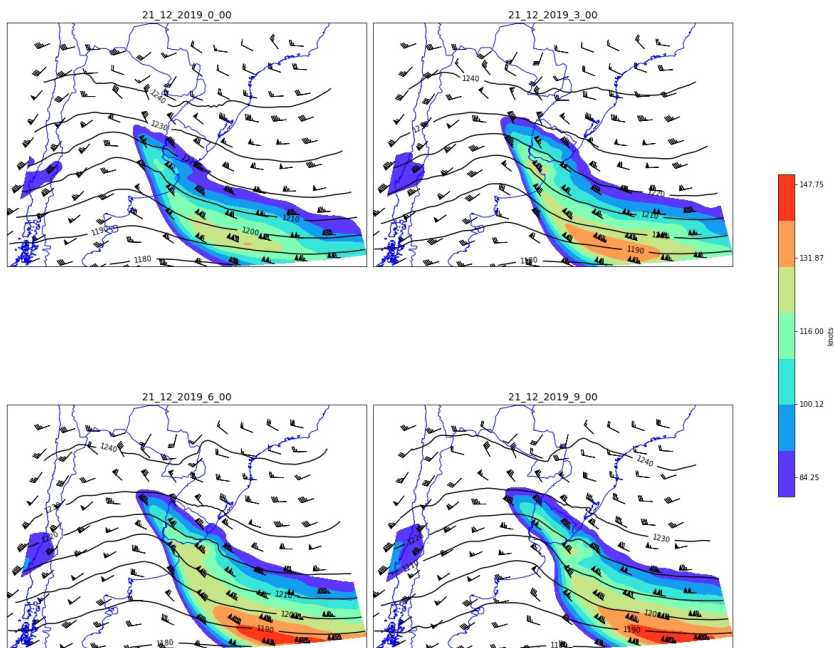
El corte transversal de nubosidad se tiene en la Figura 5.6 y muestra como el modelo reconoce la nubosidad en consonancia con la evolución vista en la Figura 5.3. En la madrugada se observa nubosidad en alturas superiores a los 6000 metros, correspondiente a la categoría 3, mientras que en las demás horas del día se observa nubosidad en alturas correspondientes a las categorías 1 y 2.

Los perfiles verticales de Humedad Relativa y Temperatura Potencial se observan en las Figura 5.7 y también son coherentes con la evolución vista en la Figura 5.3. Durante la madrugada se tienen valores altos de Humedad Relativa en la troposfera baja y media (800 a 500 hPa), llegando en las horas 3 y 6 a alcanzar el 100 % asociado a la presencia de nubosidad. A medida que avanza el día se observa la disminución de la Humedad Relativa en los perfiles, asociada al desarrollo de la PBL y al cambio de circulación dado por el ciclón extratropical. Esto se ve reflejado también en los diagramas SkewT-Lp que se presentan en la Figura 5.8 y además se destaca la baja inestabilidad potencial, otra vez asociada a la consolidación ciclón extratropical que afecta con su parte fría al territorio de Uruguay.

Por último, las Figuras 5.9, 5.10 y 5.11 presentan la evolución de la cobertura nubosa pronosticada por el modelo, según categorías de altura de nubosidad. Nuevamente, el acuerdo con la evolución diaria vista en la Figura 5.3 es destacable. Se puede observar como la categoría 3 domina en horas de la madrugada, hasta ser inexistente al final del día. Por el contrario, las categorías 1 y 2 dominan en las horas diurnas.

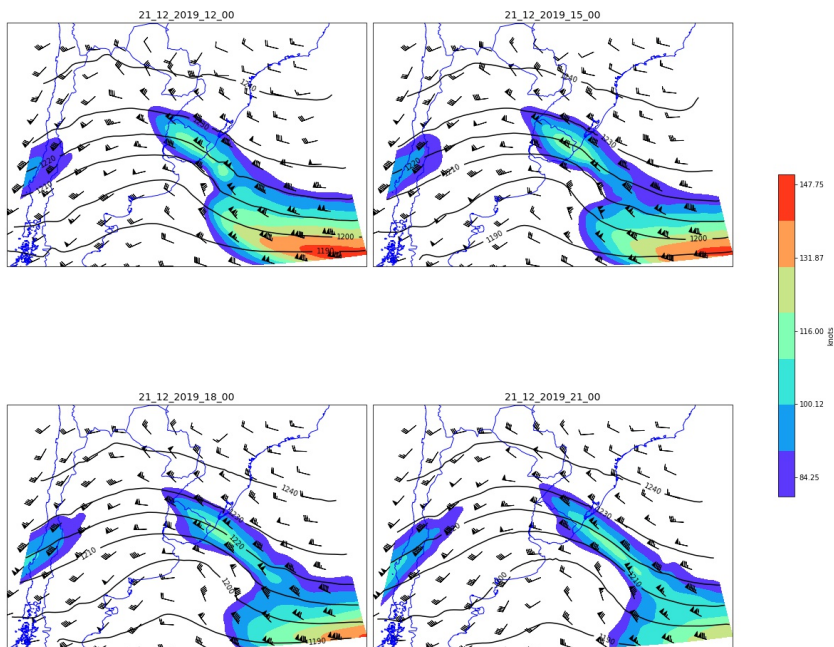
En líneas generales se puede concluir que este evento fue bien pronosticado por el modelo WRF.

Viento 200 hPa y Z200



(a) (Horas de la Mañana)

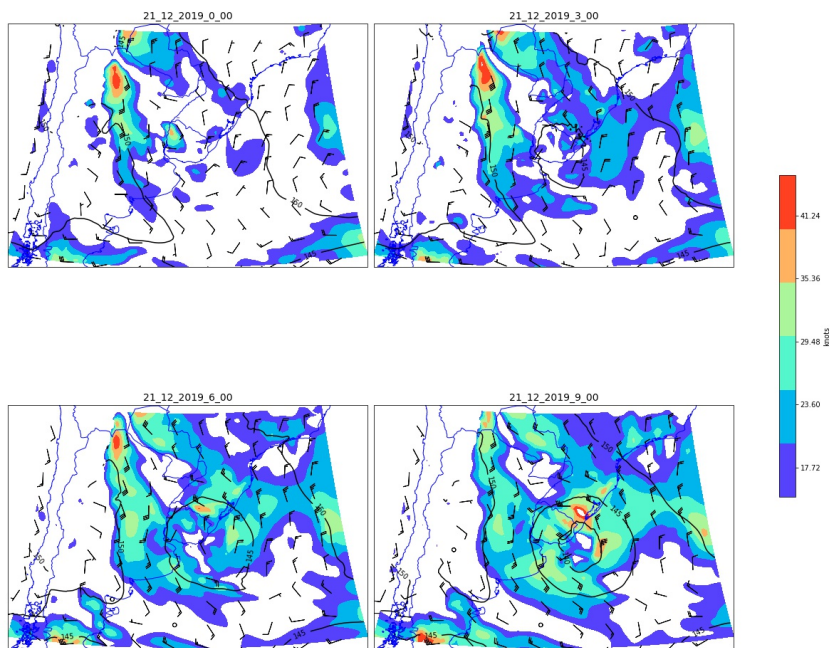
Viento 200 hPa y Z200



(b) (Horas de la Tarde)

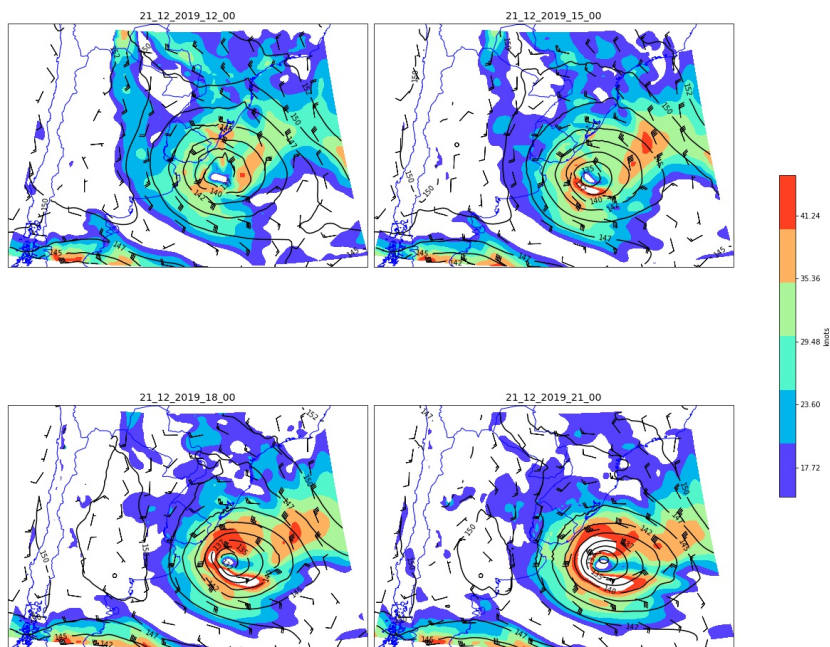
Figura 5.4: Viento y Geopotencial en 200 hPa para el día 21 de diciembre 2019.

Viento 850 hPa y Z850



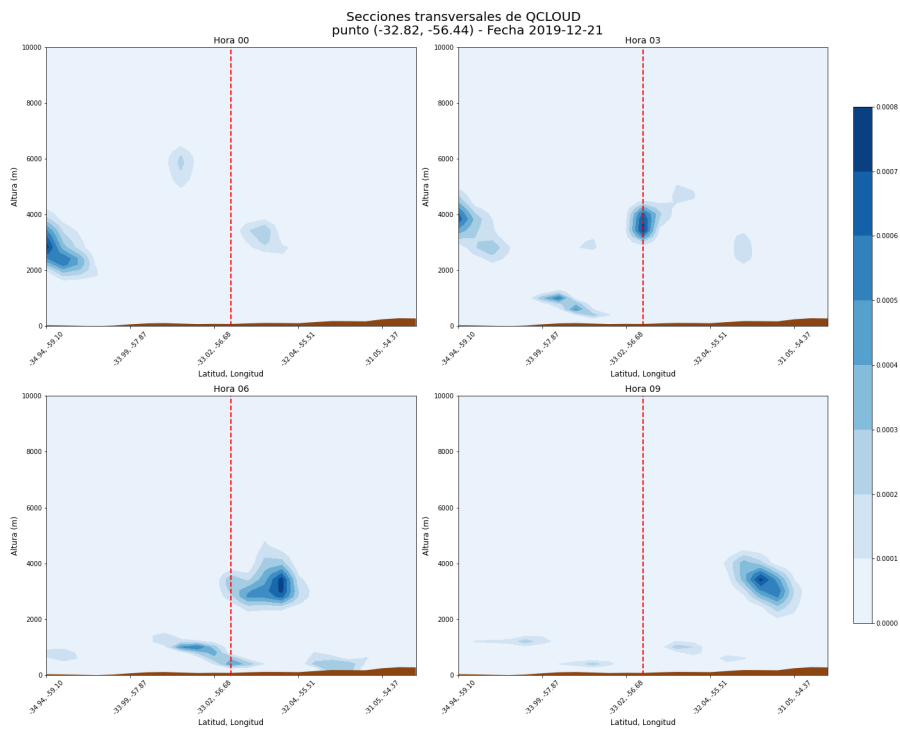
(a) (Horas de la Mañana)

Viento 850 hPa y Z850

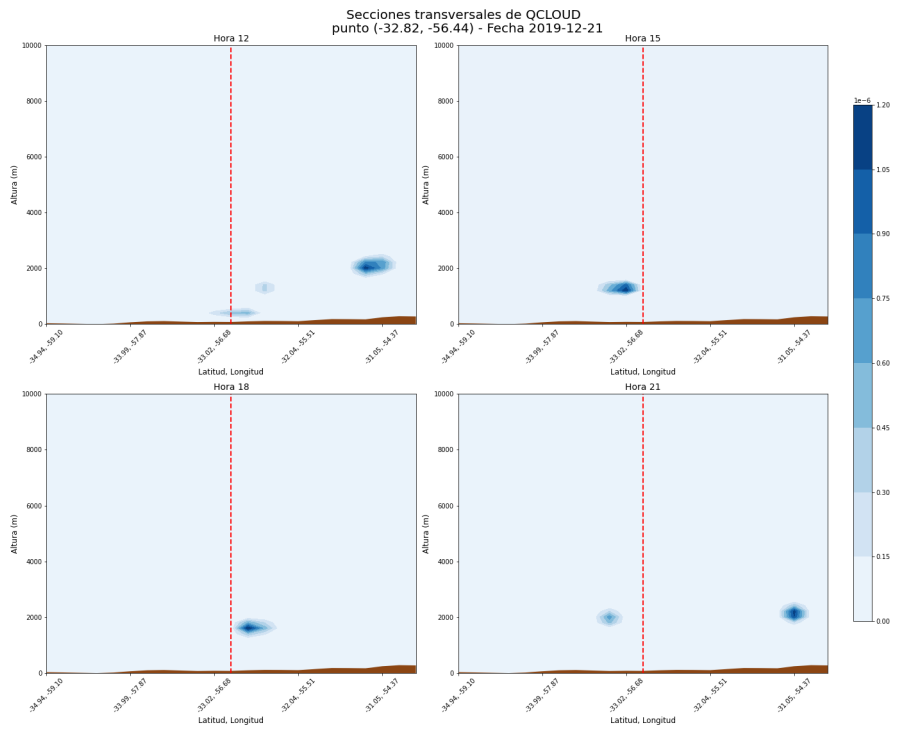


(b) (Horas de la Tarde)

Figura 5.5: Viento y Geopotencial en 850 hPa para día 21 de diciembre 2019.



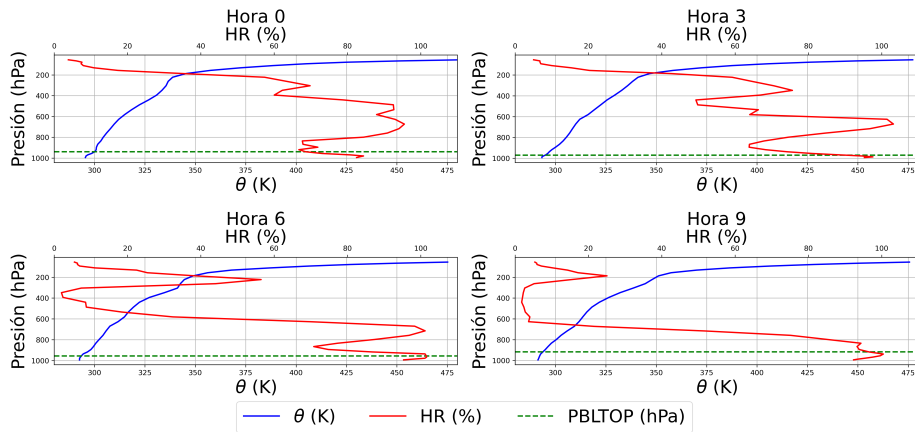
(a) (Horas de la Mañana)



(b) (Horas de la Tarde)

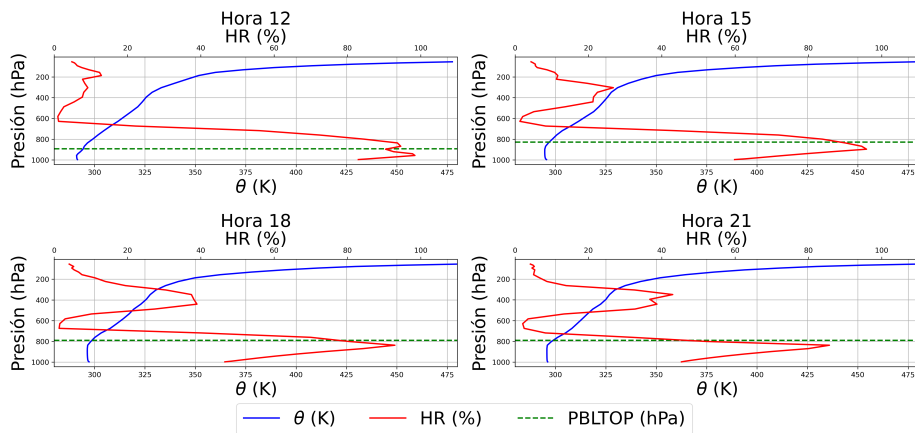
Figura 5.6: Sección transversal para el día 21 de diciembre 2019, la línea vertical roja muestra el punto central de coordenadas -32.82°S , -56.44°W .

Perfiles Verticales para el 21-12-2019 en lat=-32.82, lon=-56.44
 Humedad Relativa (HR) y Temperatura Potencial (θ)



(a) (Horas de la Mañana)

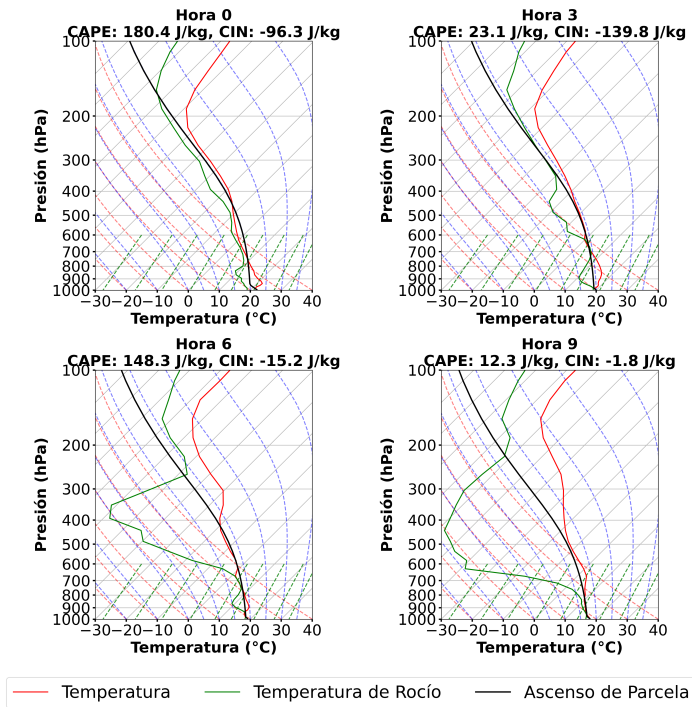
Perfiles Verticales para el 21-12-2019 en lat=-32.82, lon=-56.44
 Humedad Relativa (HR) y Temperatura Potencial (θ)



(b) (Horas de la Tarde)

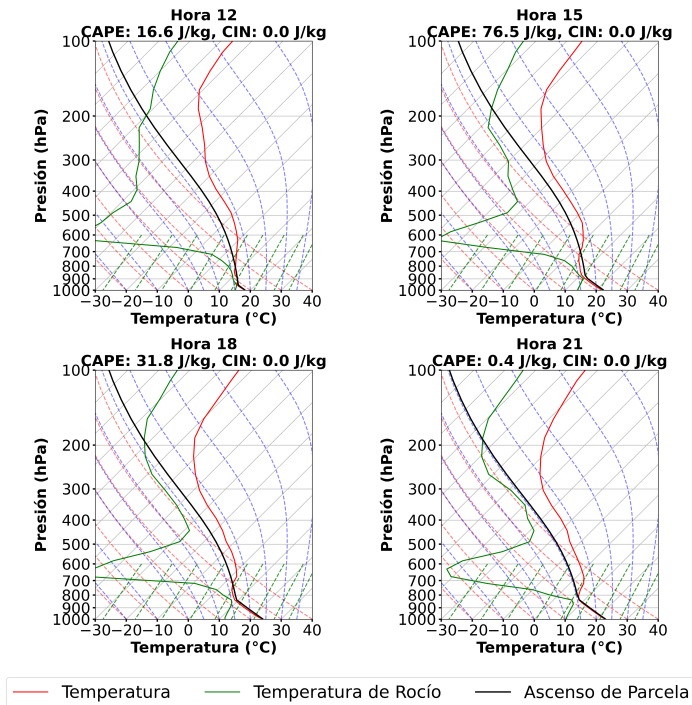
Figura 5.7: Perfiles HR, Temperatura Potencial y Tope de PBL para el día 21 de diciembre 2019.

SkewT-logP : 21-12-2019
Lat: -32.82, Lon: -56.44



(a) (Horas de la Mañana)

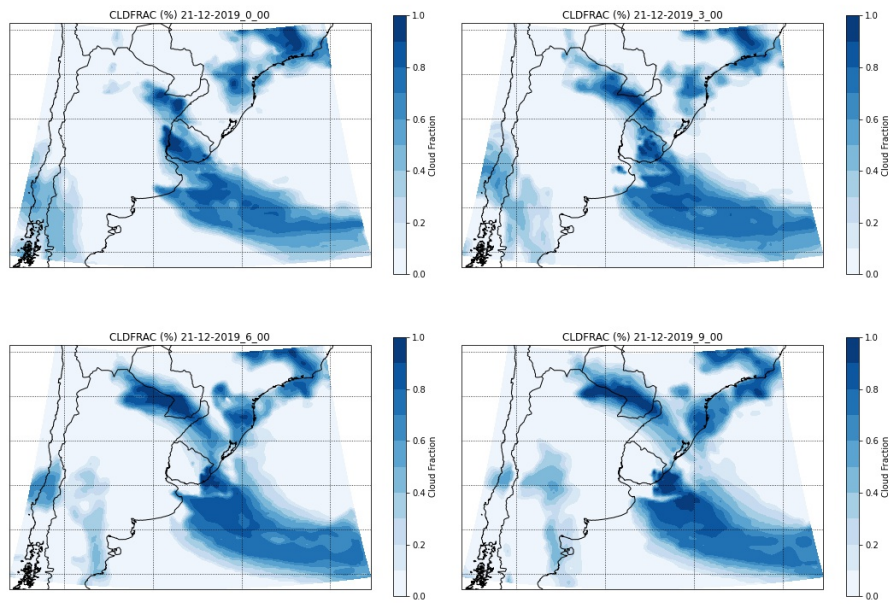
SkewT-logP : 21-12-2019
Lat: -32.82, Lon: -56.44



(b) (Horas de la Tarde)

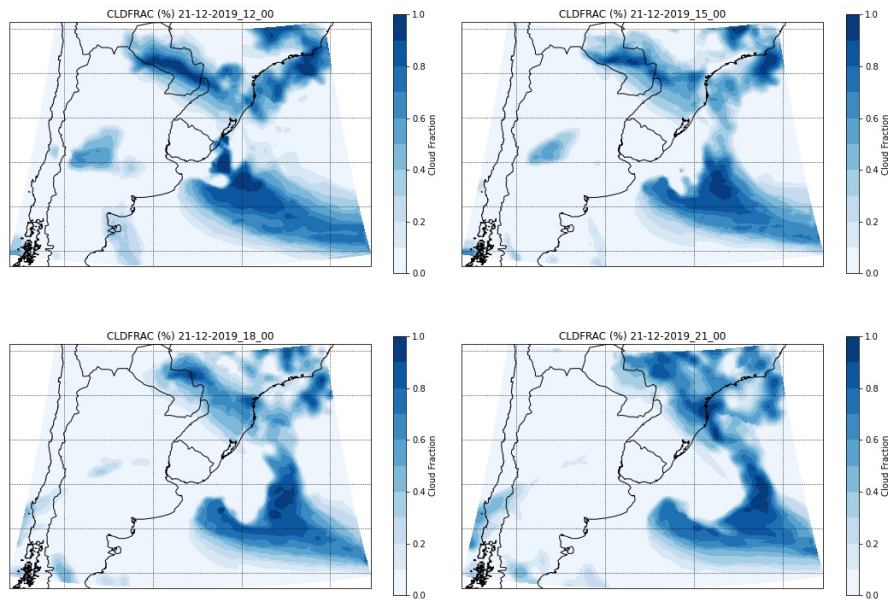
Figura 5.8: Diagramas SkewT-Lp para el día 21 de diciembre 2019.

CLDFRAC for h1 - High-Level Clouds



(a) (Horas de la Mañana)

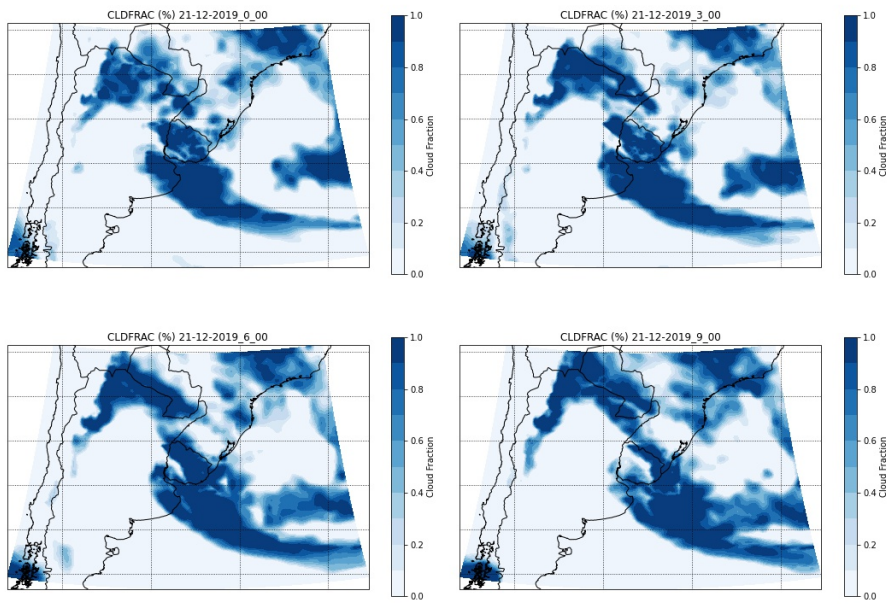
CLDFRAC for h2 - High-Level Clouds



(b) (Horas de la Tarde)

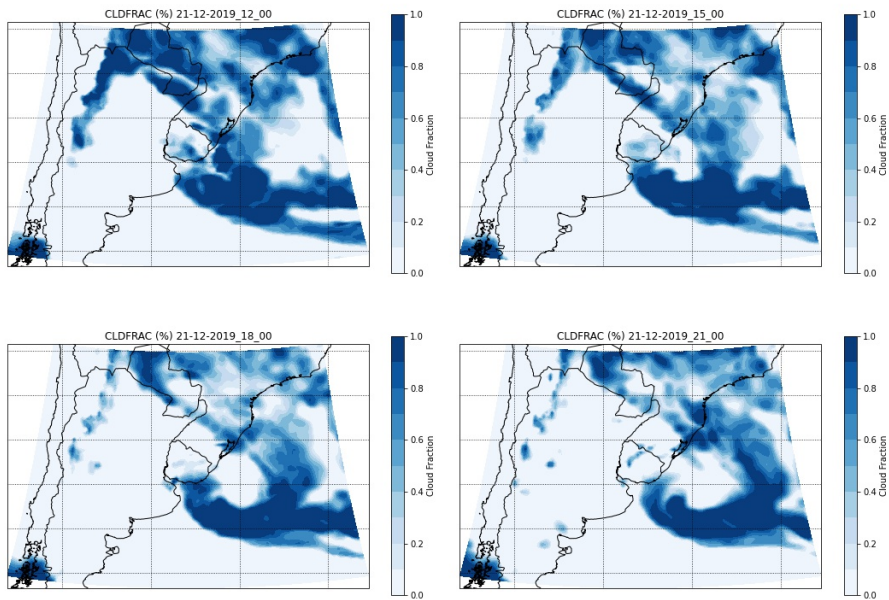
Figura 5.9: Fracción de nubes altas para el día 21 de diciembre 2019.

CLDFRAC for h1 - Mid-Level Clouds



(a) (Horas de la Mañana)

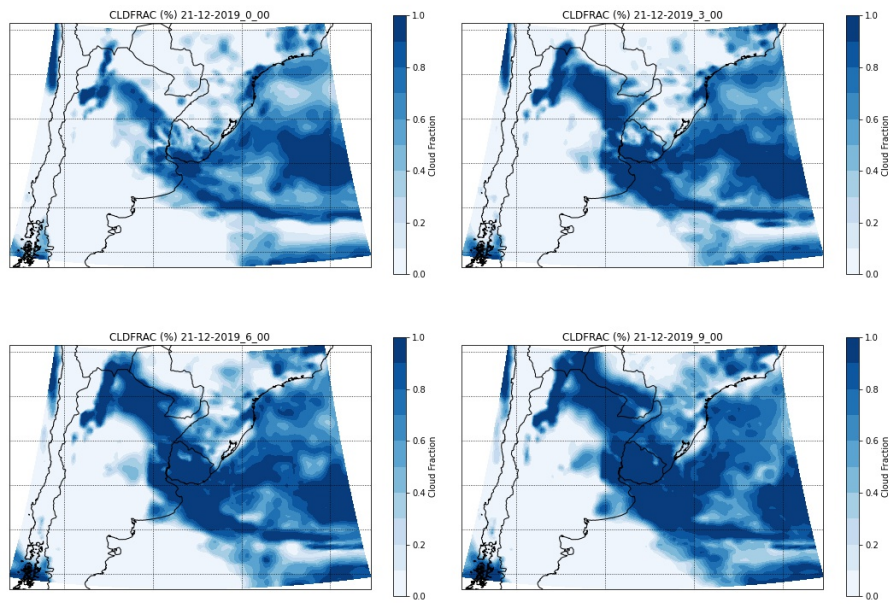
CLDFRAC for h2 - Mid-Level Clouds



(b) (Horas de la Tarde)

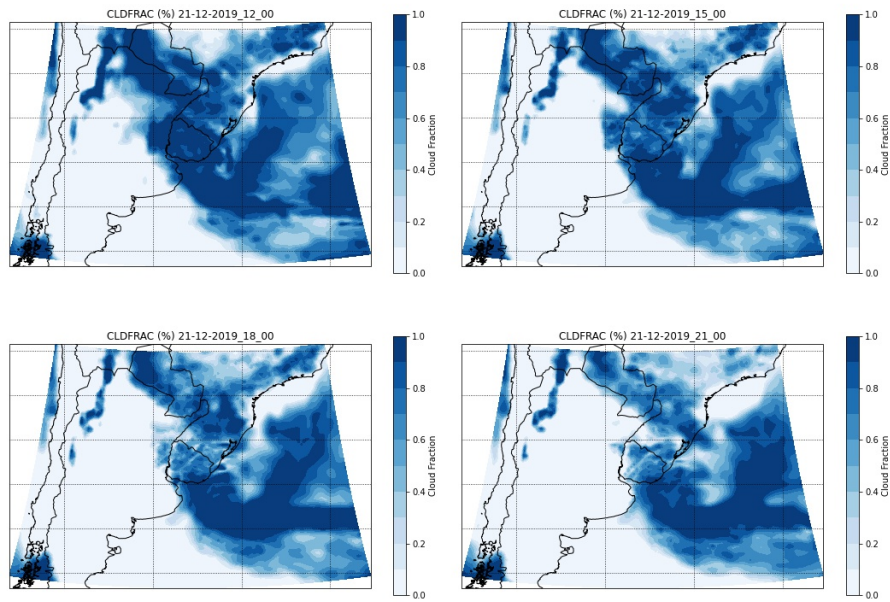
Figura 5.10: Fracción de nubes medias para el día 21 de diciembre 2019.

CLDFRAC for h1 - Low-Level Clouds



(a) (Horas de la Mañana)

CLDFRAC for h2 - Low-Level Clouds



(b) (Horas de la Tarde)

Figura 5.11: Fracción de nubes bajas para el día 21 de diciembre 2019.

5.3.2. Evento del 7 de enero de 2020

El día 7 de enero de 2020 presenta un valor de Fracción Nubosa máxima de 0.78 en la categoría 1, lo que lo posiciona en el primer lugar dentro del conjunto de días seleccionados por ser dominante la categoría de nubes bajas. La Figura 5.12 muestra la evolución de las fracciones nubosas a lo largo del día. Se puede observar como la nubosidad categoría 1 logra su fracción máxima a las 10:00 hs UTC para luego dar paso al dominio del cielo despejado (categoría 0).

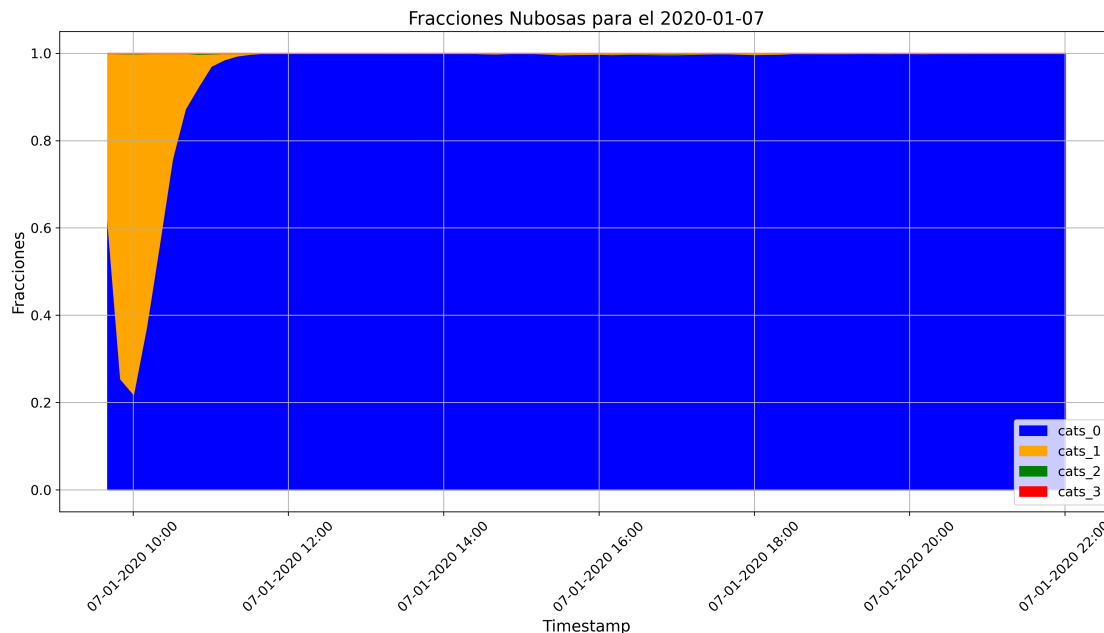


Figura 5.12: Evolución diaria de las fracciones nubosas para el día 7 de enero de 2020.

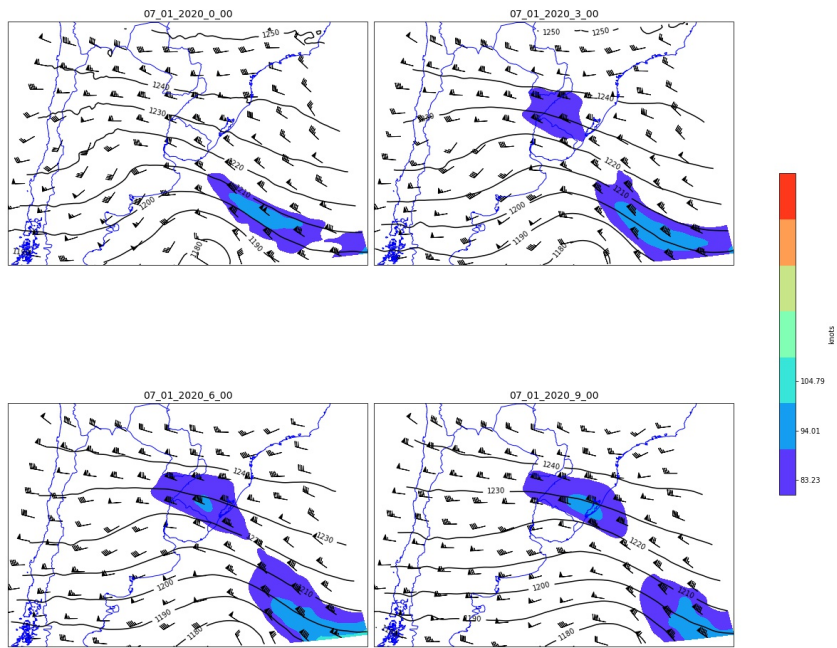
Las Figuras 5.13 y 5.14 muestran la circulación dominante. En el nivel de 200 hPa se puede observar una leve vaguada en altura con algunos máximos de vientos en altura que, por su ubicación, principalmente favorecen la estabilidad sobre el territorio de Uruguay. Asociado a esta vaguada en altura en el nivel de 850 hPa se observa un ciclón extratropical ya en etapa madura y que con el correr de las horas parecería disiparse.

El corte transversal de nubosidad, se puede observar en las figuras 5.15 y muestra la ausencia total de nubosidad durante todo el día. De la misma forma, el análisis de las salidas de *Cloud Fraction* en las Figuras 5.18, 5.19 y 5.20 muestra la ausencia generalizada de nubosidad en el pronóstico del modelo para el territorio de Uruguay.

Para el caso de los perfiles verticales de Humedad Relativa y Temperatura Potencial, Figura 5.16, se pueden observar los bajos valores de Humedad Relativa en todas las horas y la fuerte inversión en la tropósfera baja durante la madrugada y parte de la mañana. Dicha inversión desaparece con el desarrollo de la PBL. Esto se ve reflejado también en los diagramas SkewT-Lp que se presentan en la Figura 5.17 y además se observa la nula inestabilidad potencial, confirmando la estabilidad generalizada que se observa en el pronóstico.

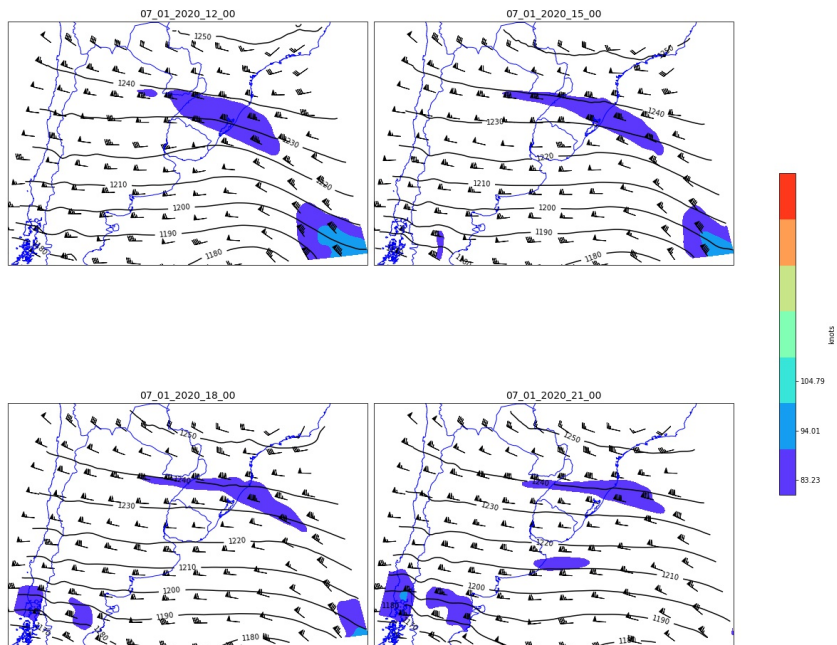
A modo de resumen, se destaca que este evento no es bien representado por el modelo, no pudiendo resolver las nubes bajas matinales que detecta el clasificador. Sin embargo, es notable la localización del pico de nubosidad de categoría 1 en horas de la mañana detectado por el clasificador, cuando el modelo muestra una inversión térmica en los niveles cercanos a la superficie, en conjunto con valores de Humedad Relativa bajos. Es probable que en el tope de la PBL haya un exceso de difusión de humedad, posiblemente ayudado por la resolución gruesa de la simulación. Al estar el tope de la PBL algo más caliente y seco que lo que debería, no se logra representar el Nivel de Condensación por Ascenso (LCL) en niveles inferiores al tope de la PBL, dificultando la generación de Stratocúmulus en la simulación. Tampoco se descarta como hipótesis que las condiciones de borde proporcionadas por el modelo GFS no representen correctamente la humedad y esto afecte la simulación del WRF en estos casos, o que la simulación WRF esté suponiendo un entrainment mayor al que realmente ocurra y por este motivo no se llegue a la condensación en niveles cercanos al tope de la PBL.

Viento 200 hPa y Z200



(a) Horas de la Mañana

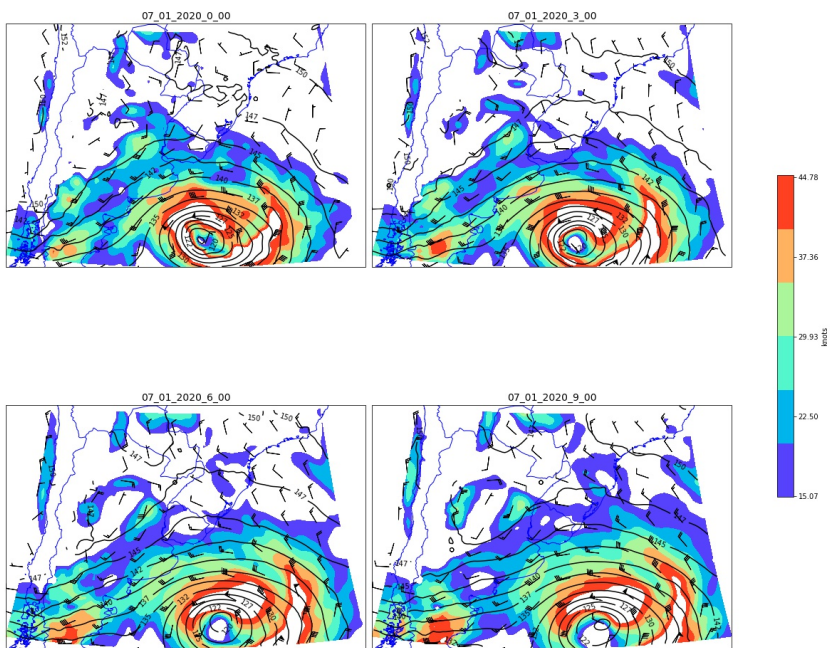
Viento 200 hPa y Z200



(b) Horas de la Tarde

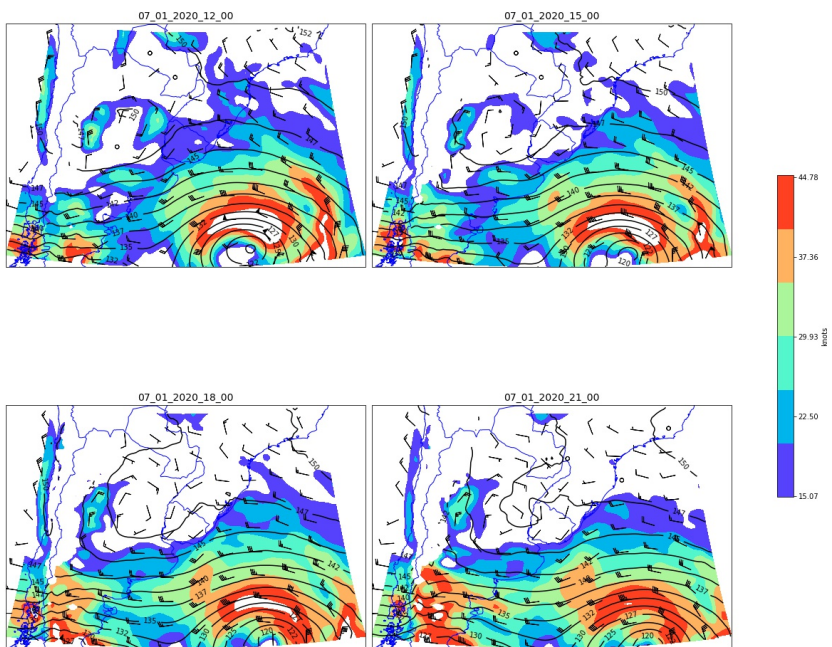
Figura 5.13: Viento y Geopotencial en 200 hPa para el día 7 de enero de 2020.

Viento 850 hPa y Z850



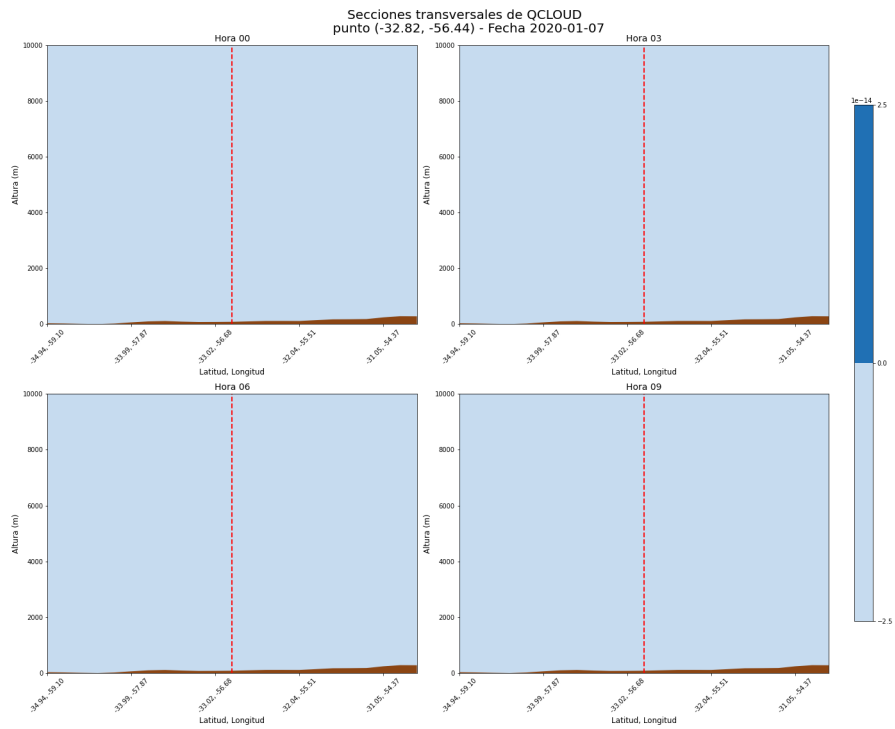
(a) Horas de la Mañana

Viento 850 hPa y Z850

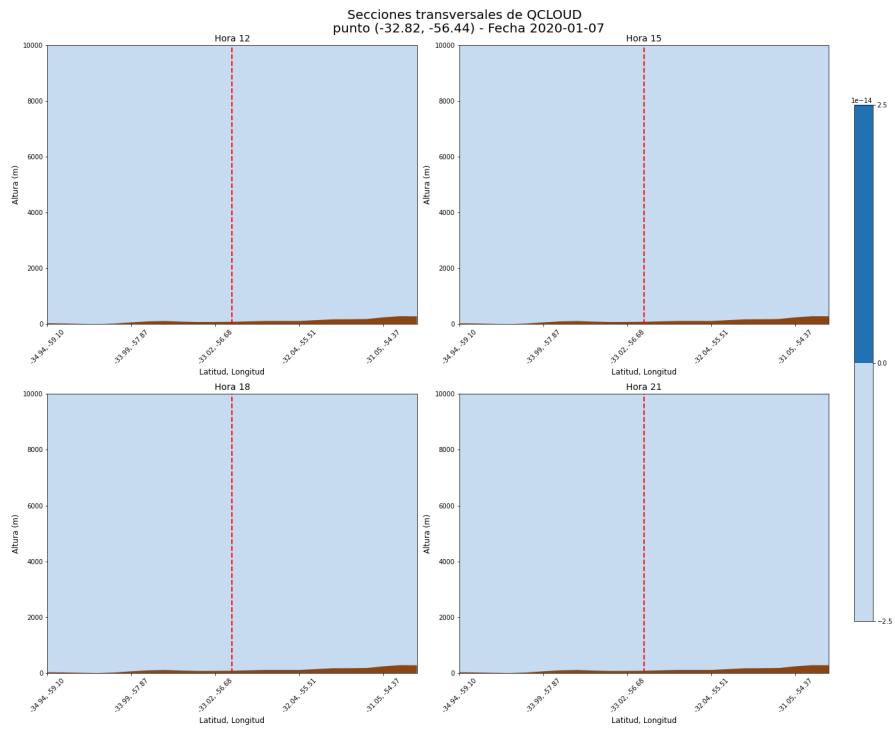


(b) Horas de la Tarde

Figura 5.14: Viento y Geopotencial en 850 hPa para el día 7 de enero de 2020.



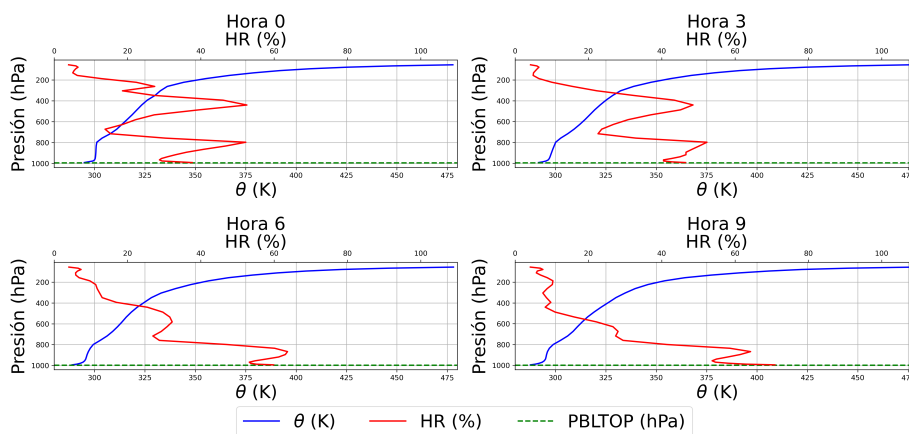
(a) Horas de la Mañana



(b) Horas de la Tarde

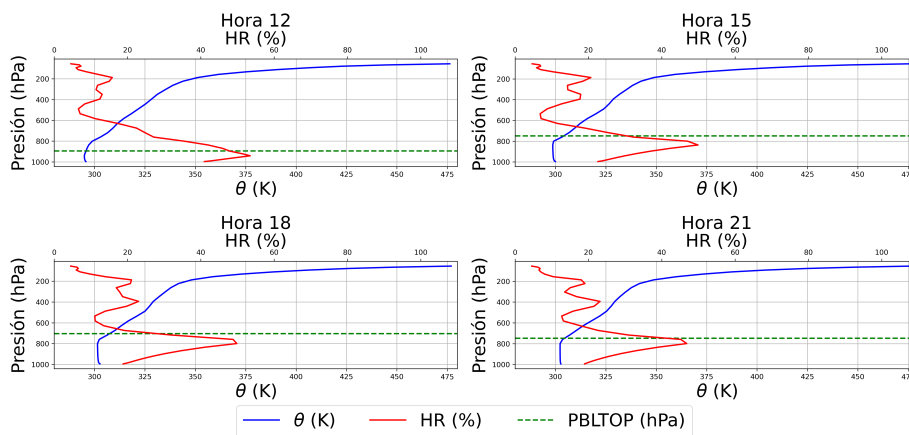
Figura 5.15: Sección transversal para el día 7 de enero de 2020, la línea vertical roja muestra el punto central de coordenadas -32.82°S , -56.44°W .

Perfiles Verticales para el 07-01-2020 en lat=-32.82, lon=-56.44
 Humedad Relativa (HR) y Temperatura Potencial (θ)



(a) Horas de la Mañana

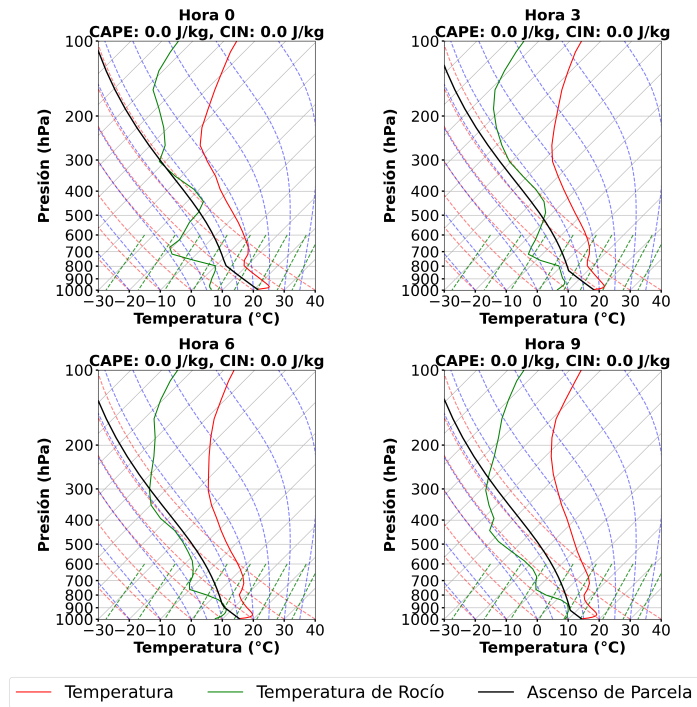
Perfiles Verticales para el 07-01-2020 en lat=-32.82, lon=-56.44
 Humedad Relativa (HR) y Temperatura Potencial (θ)



(b) Horas de la Tarde

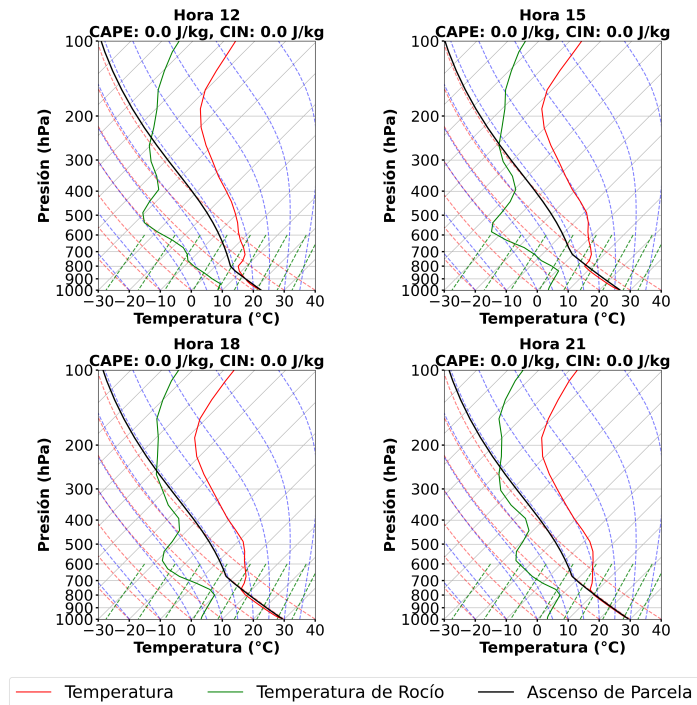
Figura 5.16: Perfiles HR, Temperatura Potencial y Tope de PBL para el día 7 de enero de 2020.

SkewT-logP : 07-01-2020
Lat: -32.82, Lon: -56.44



(a) Horas de la Mañana

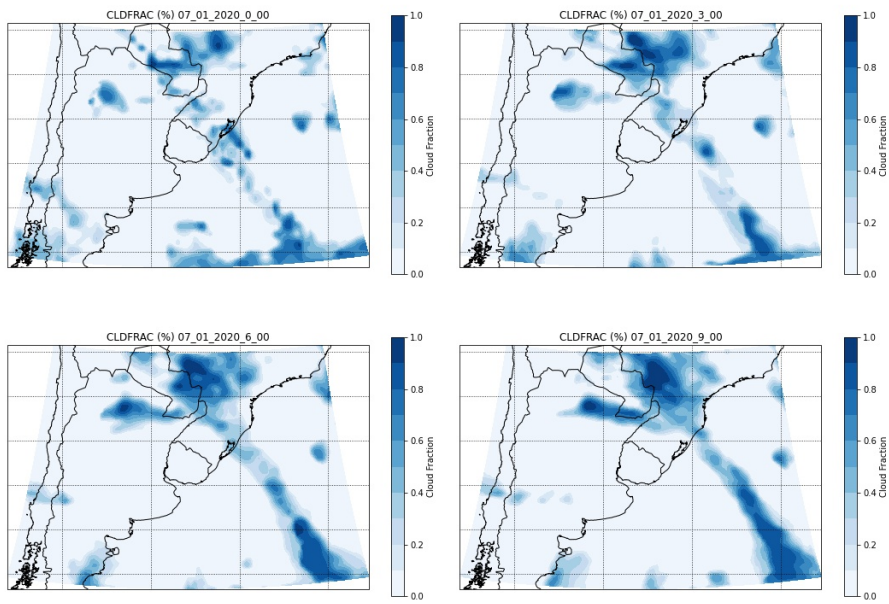
SkewT-logP : 07-01-2020
Lat: -32.82, Lon: -56.44



(b) Horas de la Tarde

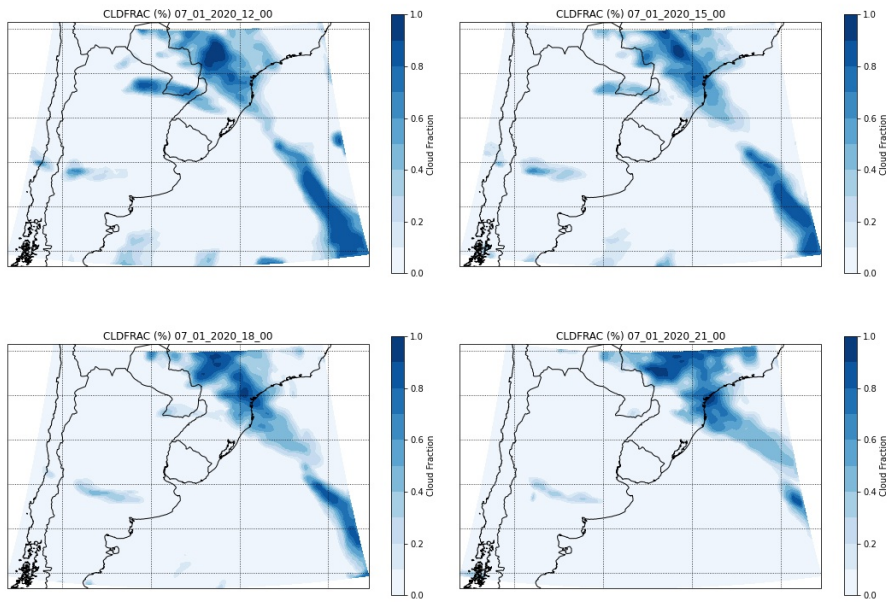
Figura 5.17: Diagramas SkewT-logP para el día 7 de enero de 2020.

CLDFRAC for h1 - High-Level Clouds



(a) Horas de la Mañana

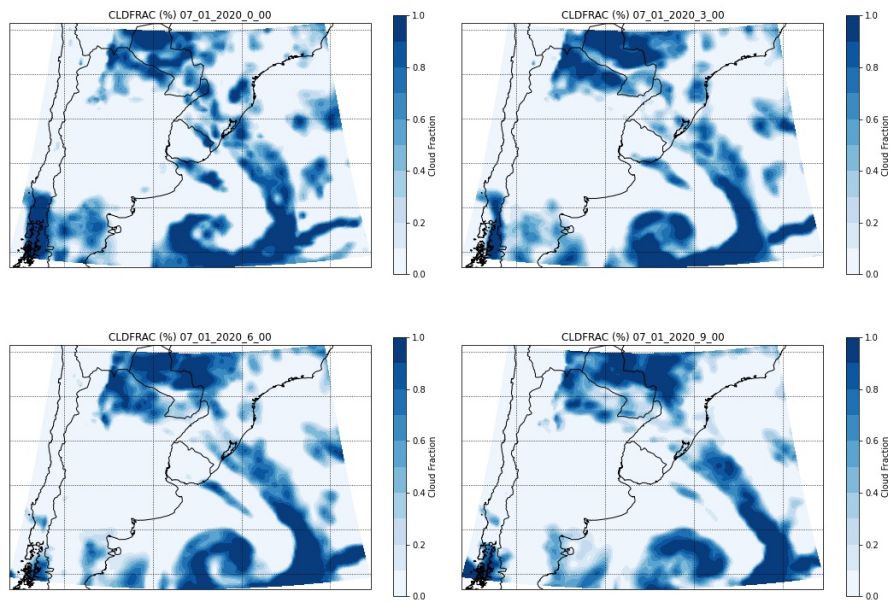
CLDFRAC for h2 - High-Level Clouds



(b) Horas de la Tarde

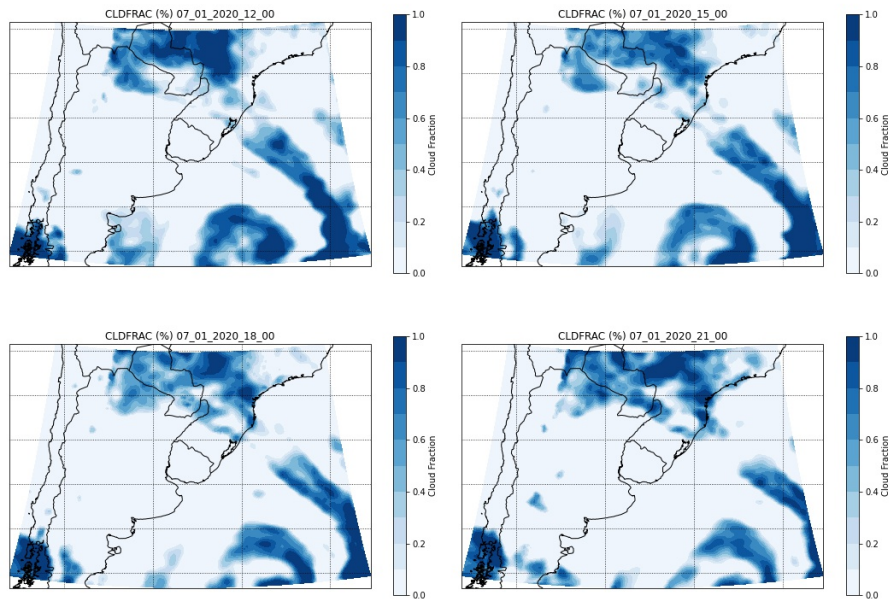
Figura 5.18: Fracción de nubes altas para el día 7 de enero de 2020.

CLDFRAC for h1 - Mid-Level Clouds



(a) Horas de la Mañana

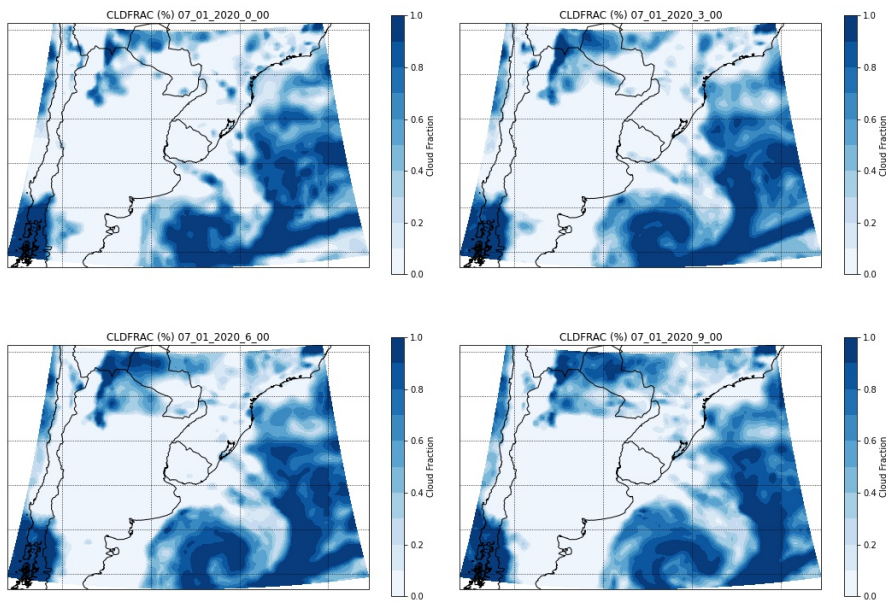
CLDFRAC for h2 - Mid-Level Clouds



(b) Horas de la Tarde

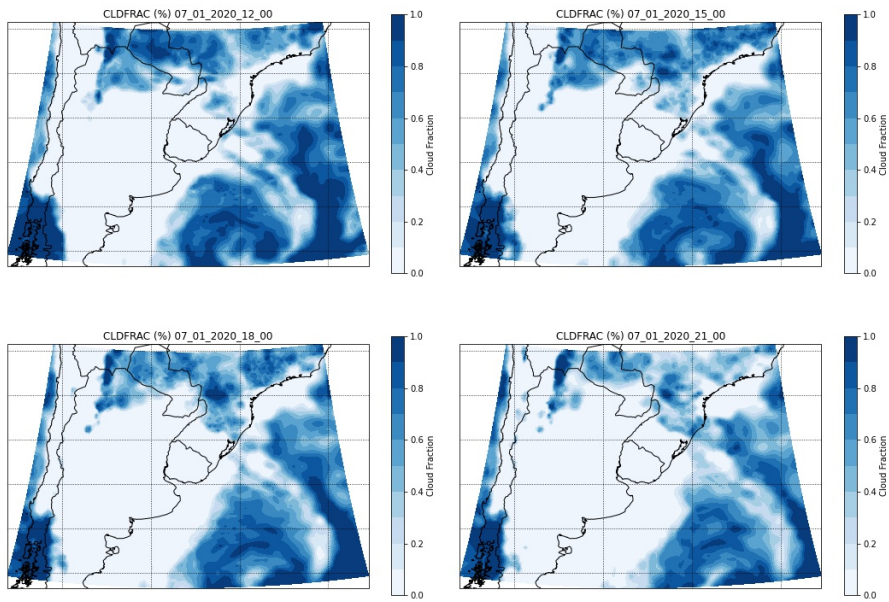
Figura 5.19: Fracción de nubes medias para el día 7 de enero de 2020.

CLDFRAC for h1 - Low-Level Clouds



(a) Horas de la Mañana

CLDFRAC for h2 - Low-Level Clouds



(b) Horas de la Tarde

Figura 5.20: Fracción de nubes bajas para el día 7 de enero de 2020.

5.3.3. Evento del 4 de noviembre de 2019

El día 4 de noviembre de 2019 presenta un valor de Fracción Nubosa máximo de 1.00 en la categoría 3, lo que lo posiciona en el primer lugar dentro del conjunto de días seleccionados por ser dominante la categoría de nubes altas. La Figura 5.21 muestra la evolución de la fracciones nubosas a lo largo del día. Se puede observar como la nubosidad categoría 3 domina durante todo el día, comenzando disminuir su fracción en las últimas horas del día a la vez que aumenta la fracción de las categorías 2 y 0.

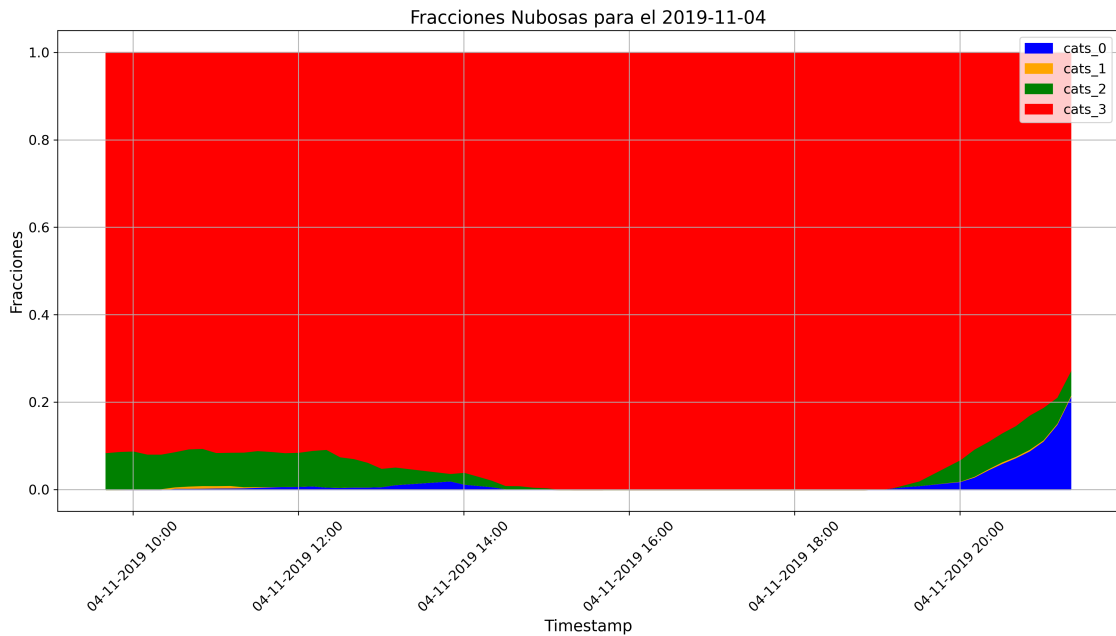


Figura 5.21: Evolución diaria de las fracciones nubosas para el día 4 de noviembre de 2019.

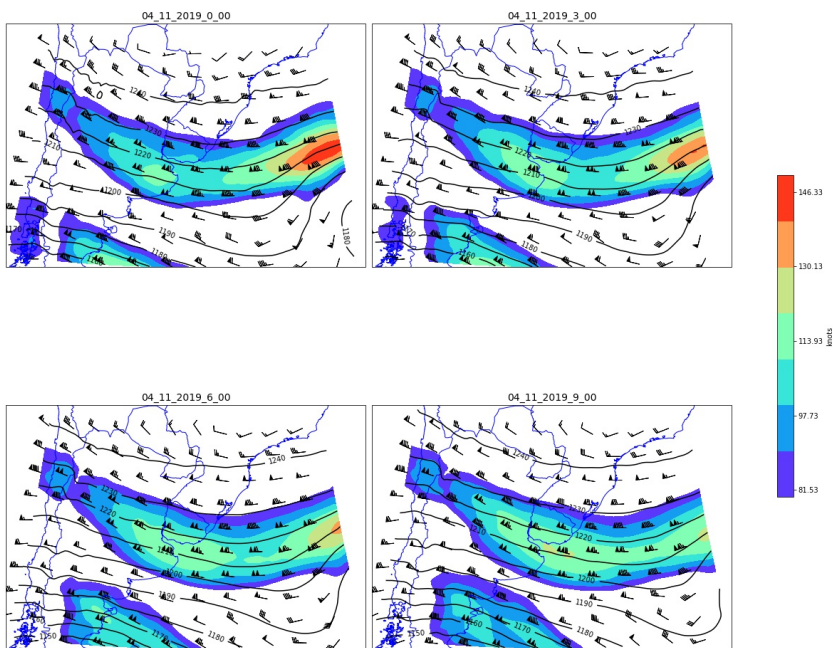
Las Figuras 5.22 y 5.23 muestran la circulación dominante. En el nivel de 200 hPa se puede observar la Corriente en Chorro con un flujo principalmente zonal instalada sobre el sur de Uruguay. También, a medida que pasan las horas, se observan máximos locales de viento en la Corriente en Chorro que favorecen la inestabilidad sobre Uruguay. En el nivel de 850 hPa se observa la Corriente en Chorro de Bajos Niveles afecta el norte de Uruguay donde converge su transporte de humedad y temperatura.

El corte transversal de nubosidad, Figura 5.15, muestra nubosidad de desarrollo vertical en horas de la tarde, en consonancia con el dominio de la nubosidad de categoría 3 visto previamente. De la misma forma, las Figuras 5.18, 5.19 y 5.20 muestran la presencia generalizada de nubosidad en el pronóstico del modelo para el territorio de Uruguay, en las tres categorías.

De los perfiles verticales de Humedad Relativa y Temperatura Potencial (Figura 5.7) es notorio el escaso desarrollo de la PBL a medida que transcurre el día. Esto puede deberse a la afectación del ciclo diario por la nubosidad generalizada y su impacto en la disminución del calentamiento radiativo de la superficie. En los diagramas SkewT-Lp de la Figura 5.17 se observa que el contenido de humedad en los distintos niveles verticales es cercano a la saturación y existen condiciones de inestabilidad potencial.

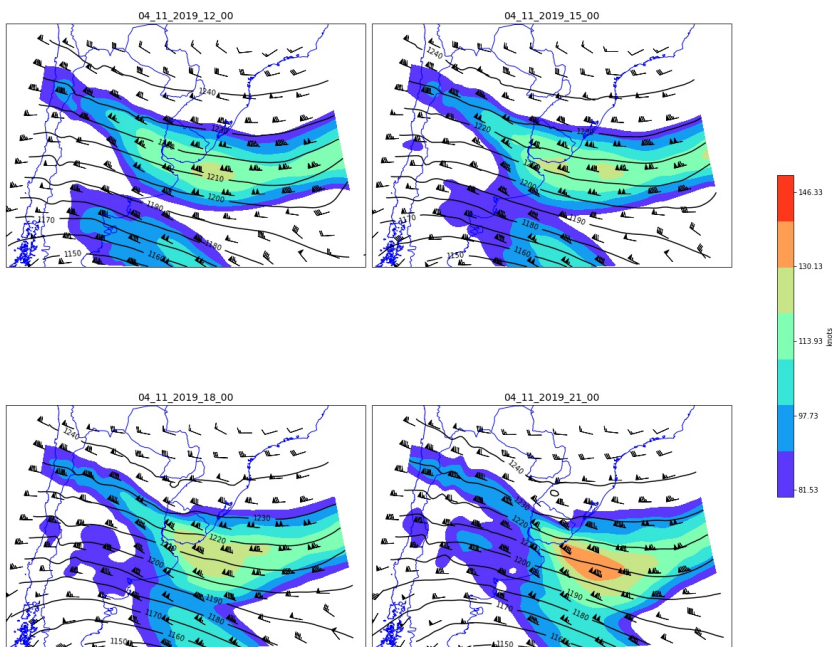
En conclusión, en este caso el modelo logra representar la nubosidad de desarrollo vertical de este día, encuadrada en la categoría 3 de la clasificación implementada.

Viento 200 hPa y Z200



(a) Horas de la Mañana

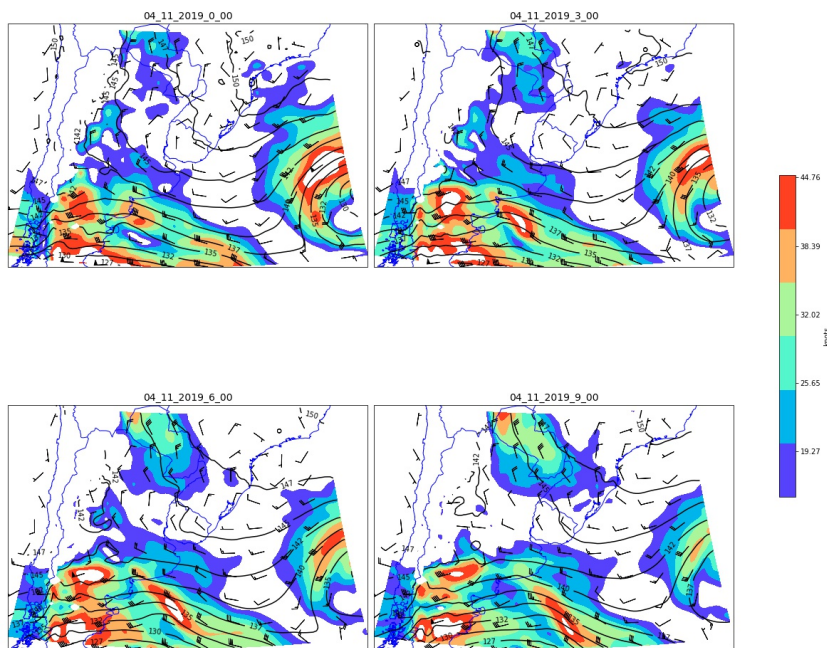
Viento 200 hPa y Z200



(b) Horas de la Tarde

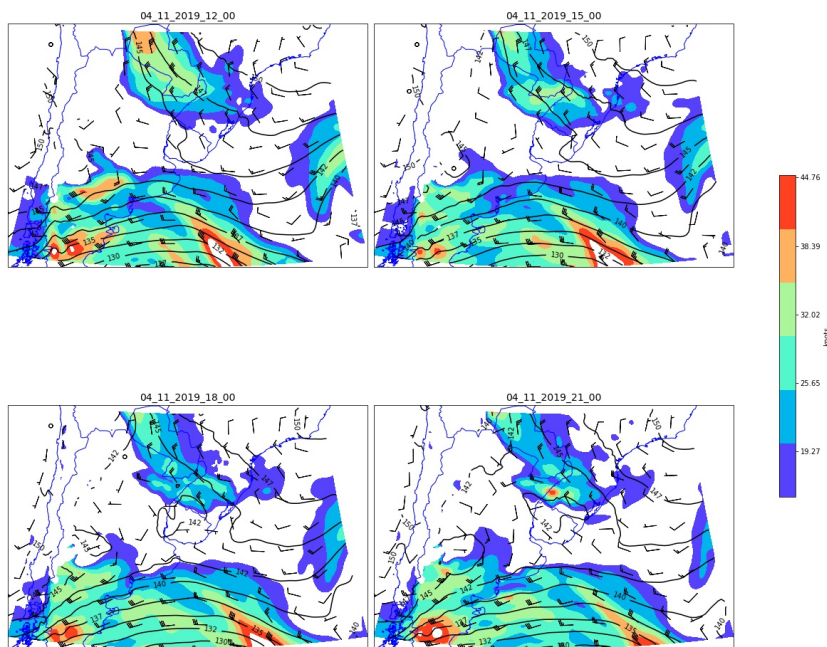
Figura 5.22: Viento y Geopotencial en 200 hPa para el día 4 de noviembre de 2019.

Viento 850 hPa y Z850



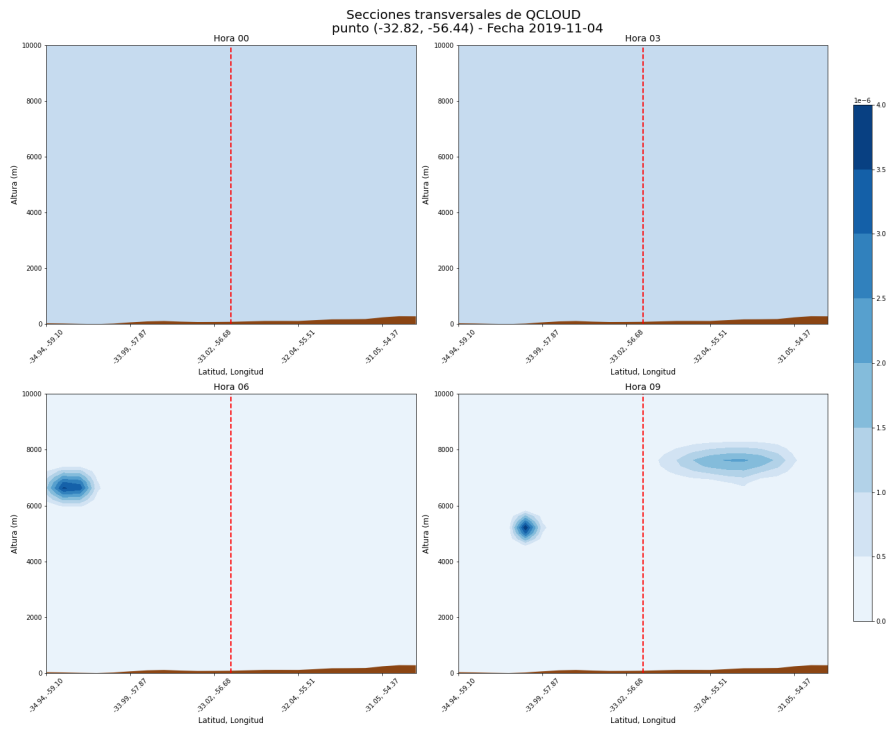
(a) Horas de la Mañana

Viento 850 hPa y Z850

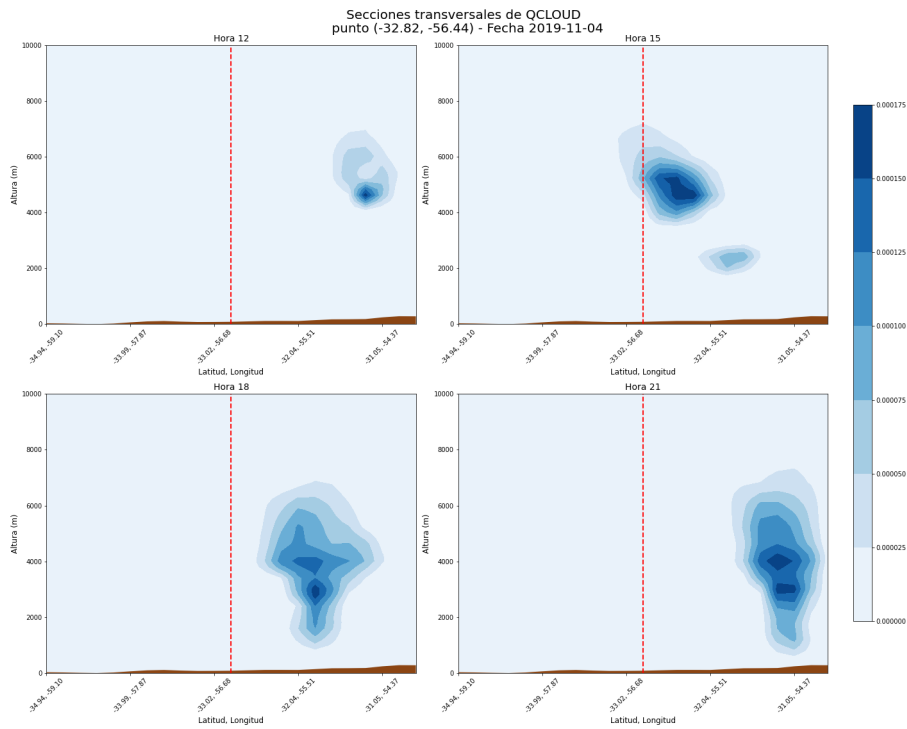


(b) Horas de la Tarde

Figura 5.23: Viento y Geopotencial en 850 hPa para el día 4 de noviembre de 2019.



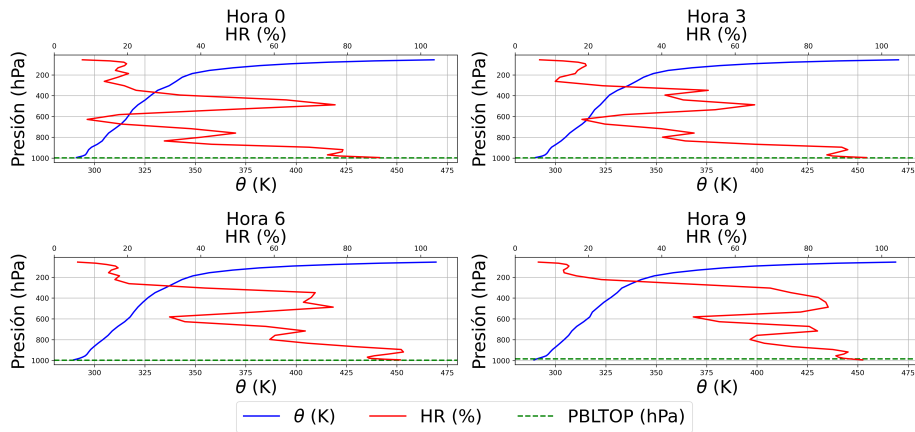
(a) Horas de la Mañana



(b) Horas de la Tarde

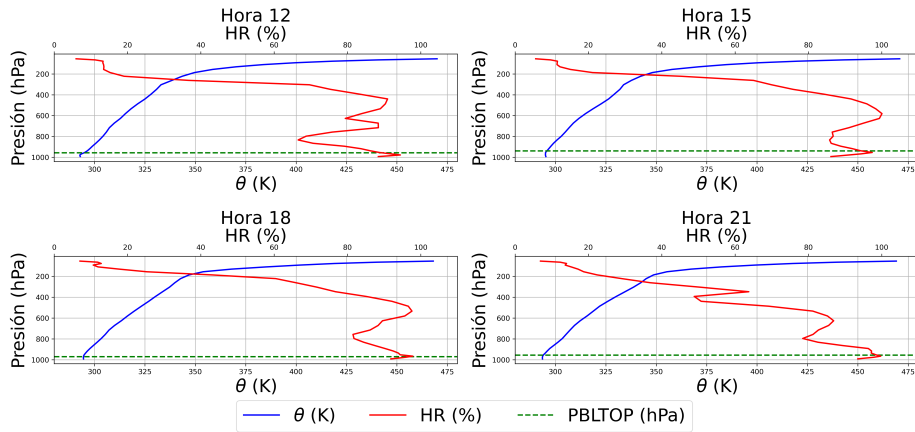
Figura 5.24: Sección transversal para el día 4 de noviembre de 2019, la línea vertical roja muestra el punto central de coordenadas -32.82°S , -56.44°W .

Perfiles Verticales para el 04-11-2019 en lat=-32.82, lon=-56.44
 Humedad Relativa (HR) y Temperatura Potencial (θ)



(a) Horas de la Mañana

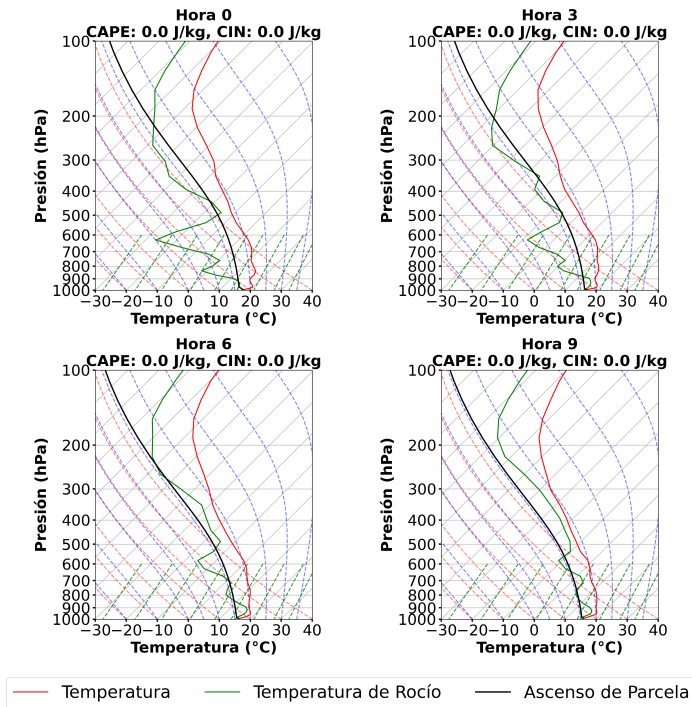
Perfiles Verticales para el 04-11-2019 en lat=-32.82, lon=-56.44
 Humedad Relativa (HR) y Temperatura Potencial (θ)



(b) Horas de la Tarde

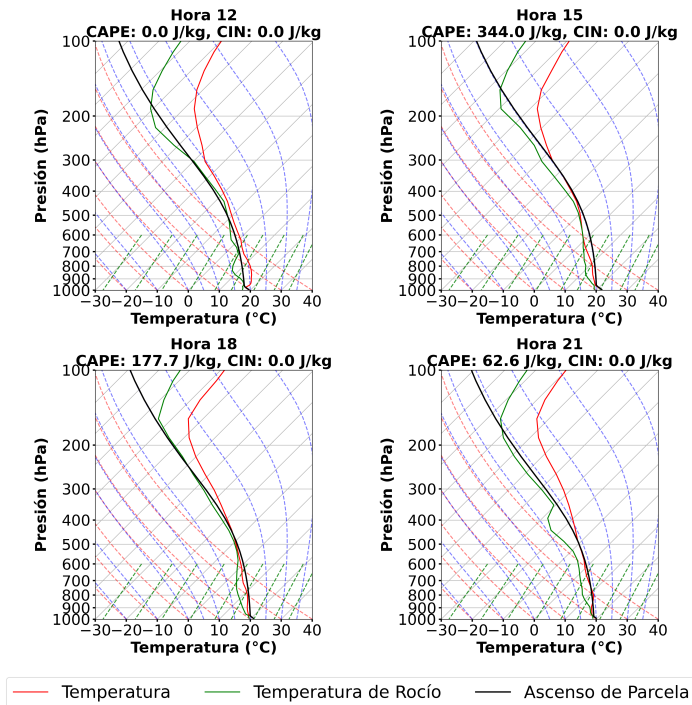
Figura 5.25: Perfiles HR, Temperatura Potencial y Tope de PBL para el día 4 de noviembre de 2019.

SkewT-logP : 04-11-2019
Lat: -32.82, Lon: -56.44



(a) Horas de la Mañana

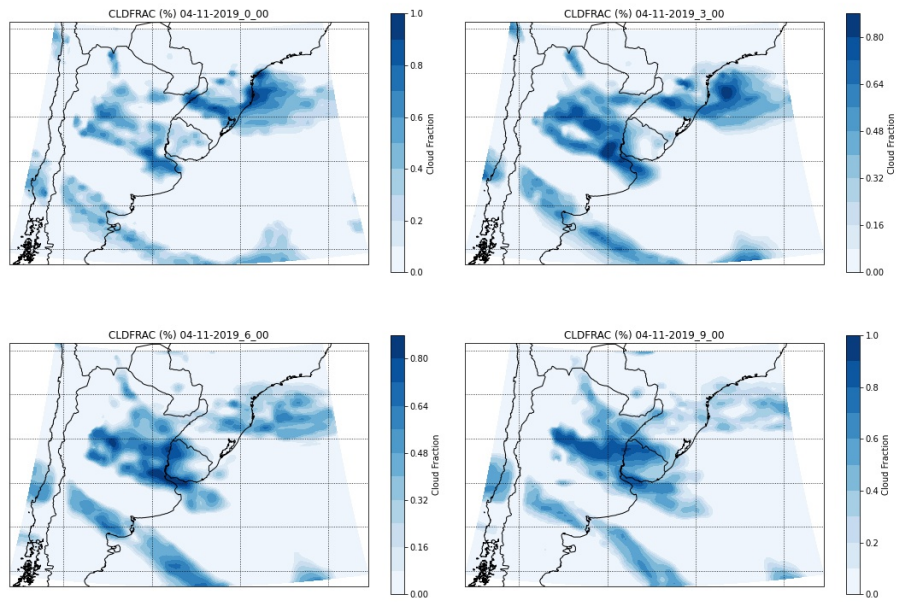
SkewT-logP : 04-11-2019
Lat: -32.82, Lon: -56.44



(b) Horas de la Tarde

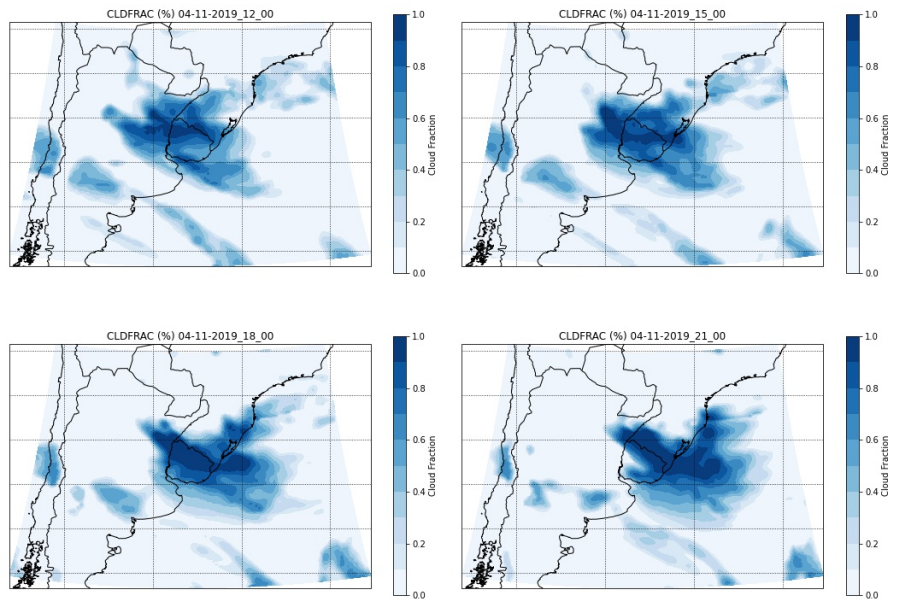
Figura 5.26: Diagramas SkewT-logP para el día 4 de noviembre de 2019.

CLDFRAC for h1 - High-Level Clouds



(a) Horas de la Mañana

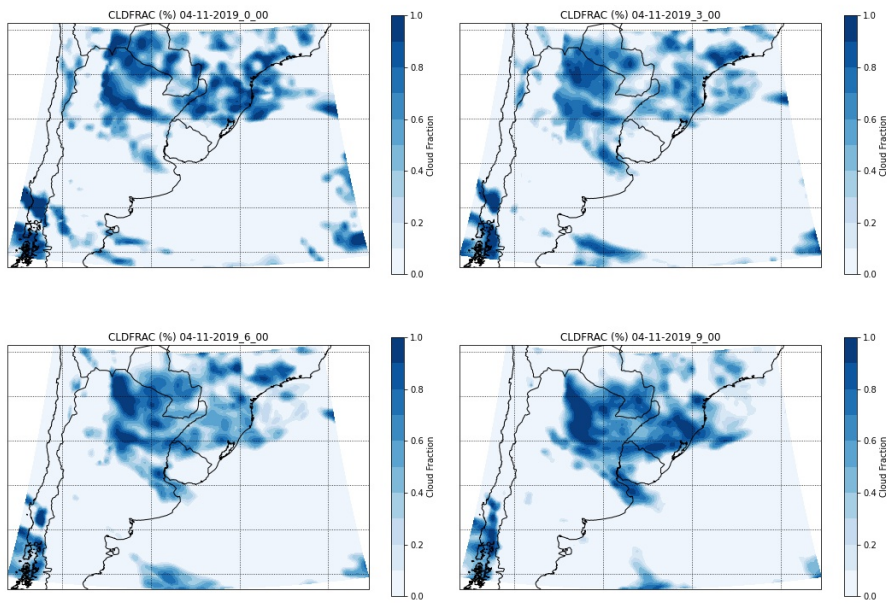
CLDFRAC for h2 - High-Level Clouds



(b) Horas de la Tarde

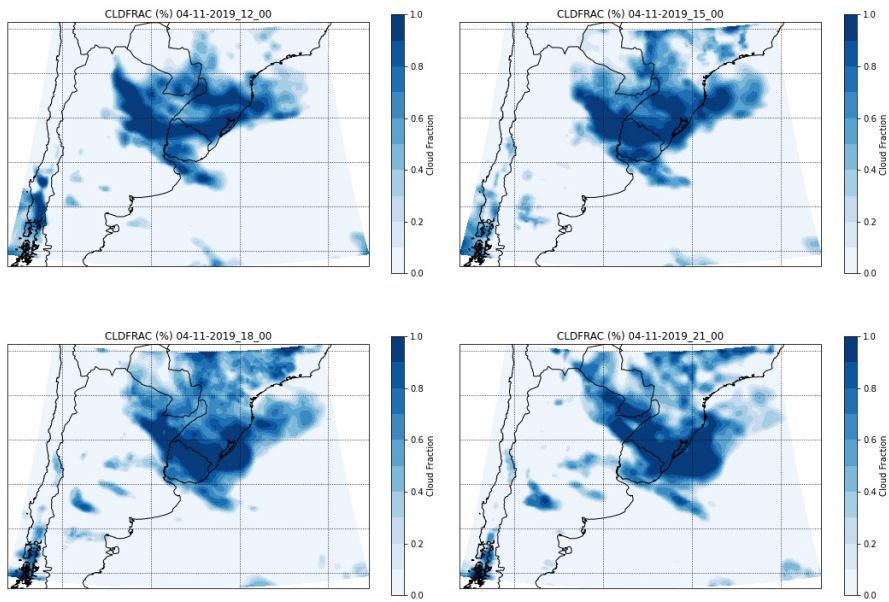
Figura 5.27: Fracción de nubes altas para el día 4 de noviembre de 2019.

CLDFRAC for h1 - Mid-Level Clouds



(a) Horas de la Mañana

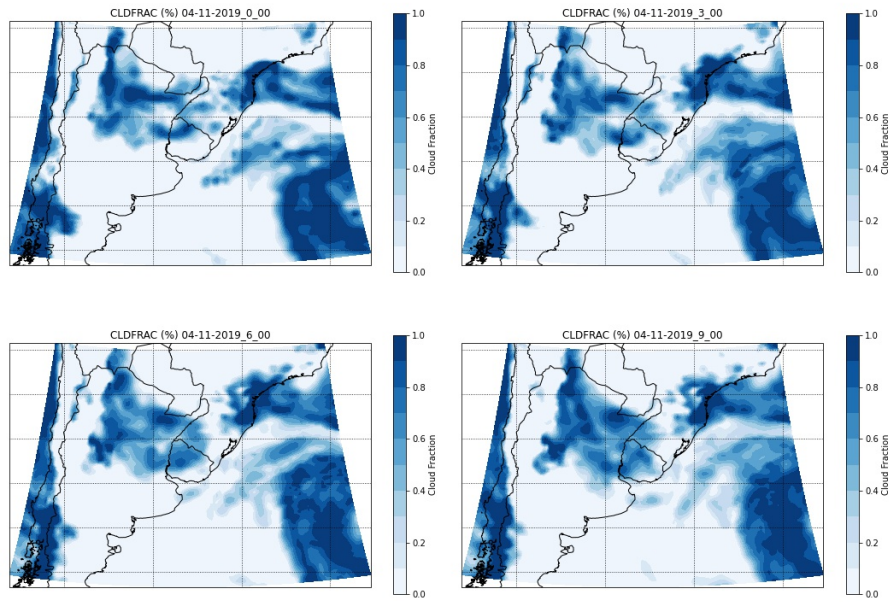
CLDFRAC for h2 - Mid-Level Clouds



(b) Horas de la Tarde

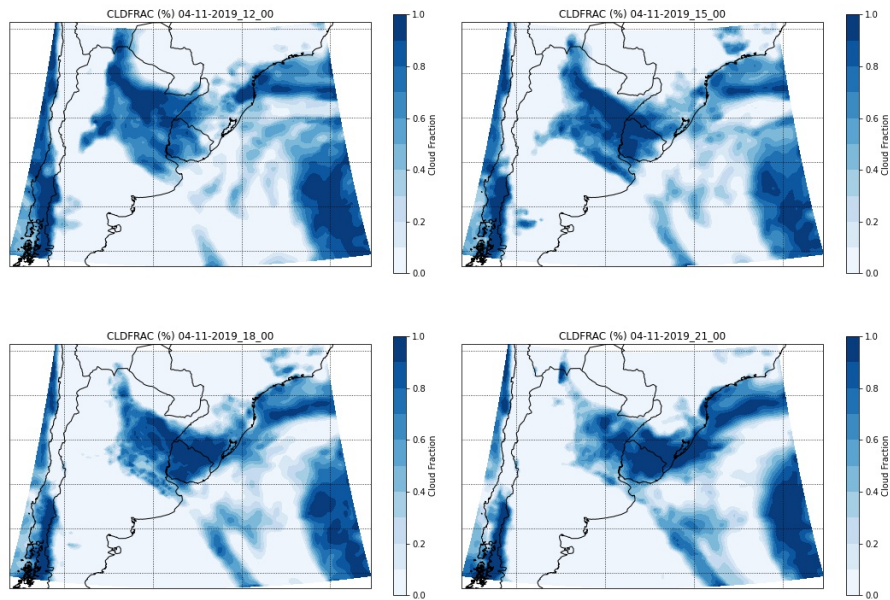
Figura 5.28: Fracción de nubes medias para el día 4 de noviembre de 2019.

CLDFRAC for h1 - Low-Level Clouds



(a) Horas de la Mañana

CLDFRAC for h2 - Low-Level Clouds



(b) Horas de la Tarde

Figura 5.29: Fracción de nubes bajas para el día 4 de noviembre de 2019.

5.3.4. Evento del 17 de febrero de 2020

El día 17 de febrero de 2020 presenta un valor de Fracción Nubosa máximo de 1.00 en la categoría 3. La Figura 5.30 muestra la evolución de las fracciones nubosas a lo largo del día. Se puede observar como la nubosidad categoría 3 domina por la mañana y las primeras horas de la tarde, comenzando disminuir su fracción en horas de la tarde a la vez que aumenta la fracción de las categoría 0 principalmente.

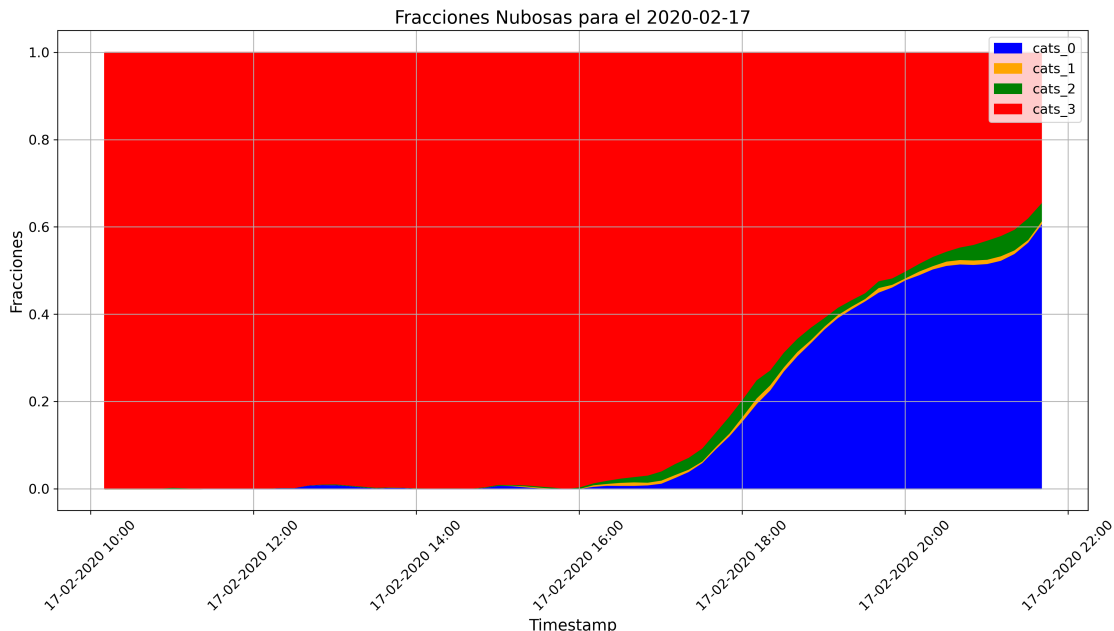


Figura 5.30: Evolución diaria de las fracciones nubosas para el día 17 de febrero de 2020.

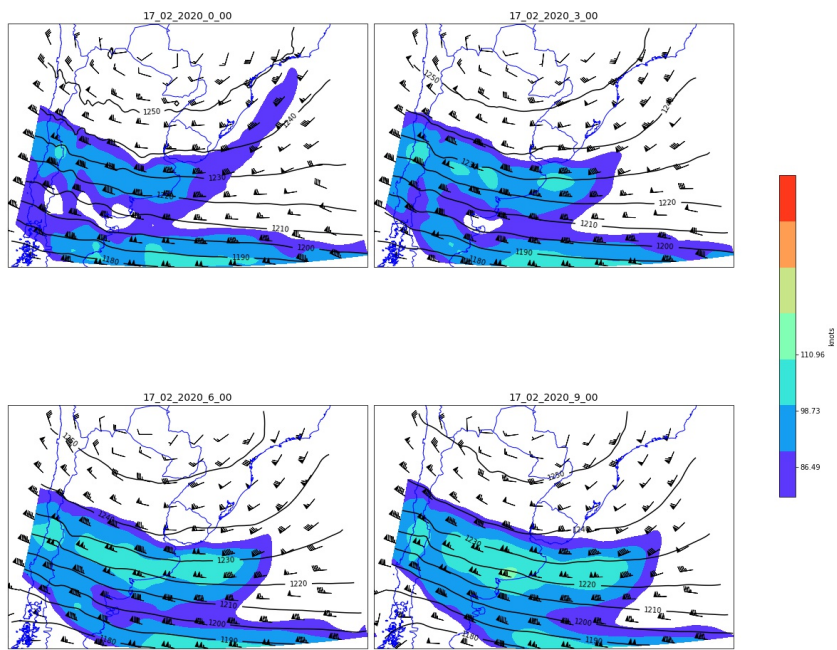
La Figura 5.31 muestra que la circulación dominante en el nivel de 200 hPa es predominantemente zonal con la Corriente en Chorro ubicada sobre el sur de Uruguay. En horas de la mañana, se observa que los máximos locales de viento de la Corriente en Chorro parecen favorecer la inestabilidad sobre el sur de Uruguay, pero a medida que pasan las horas esta contribución no parece ser significativa. En el nivel de 850 hPa (Figura 5.32) se observa que la Corriente en Chorro de Bajos Niveles afecta el territorio de Uruguay, donde converge su transporte de humedad y temperatura.

El corte transversal de nubosidad en la Figura 5.33 muestra nubes altas en todas las horas de la simulación excepto la hora 0. Del análisis de las figuras 5.36, 5.37 y 5.38 se tiene presencia de nubosidad en las tres categorías en la región de Uruguay.

De los perfiles verticales de Humedad Relativa y Temperatura Potencial, Figura 5.34, se destacan valores de humedad relativa relativamente altos en la franja de 850 a 500 hPa, coherente con la presencia de la Corriente en Chorro de Niveles Bajos. En los diagramas SkewT-Lp de la Figura 5.35 se destaca la alta inestabilidad potencial que se tiene en casi todas las horas del día.

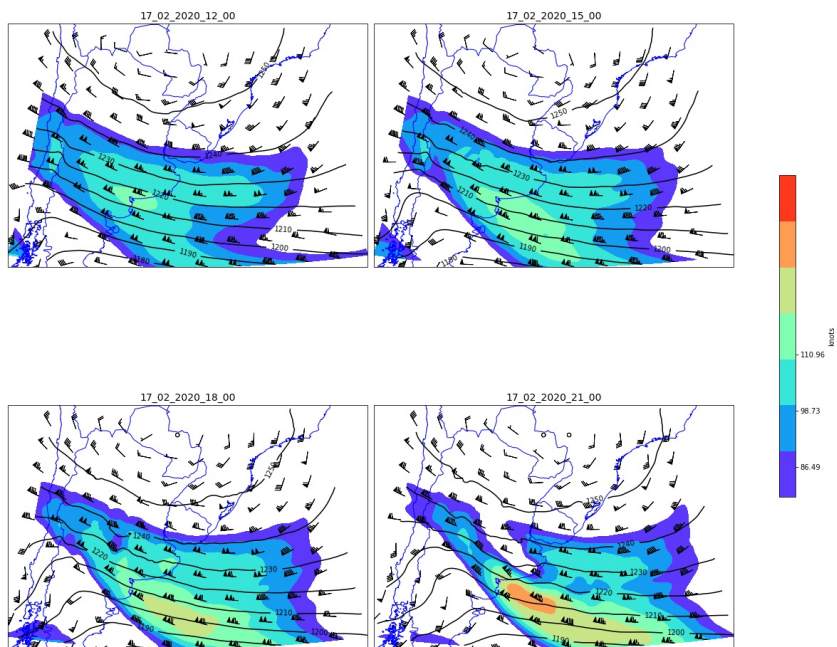
En conclusión, en este caso el modelo logra representar la nubosidad categoría 3.

Viento 200 hPa y Z200



(a) Horas de la Mañana

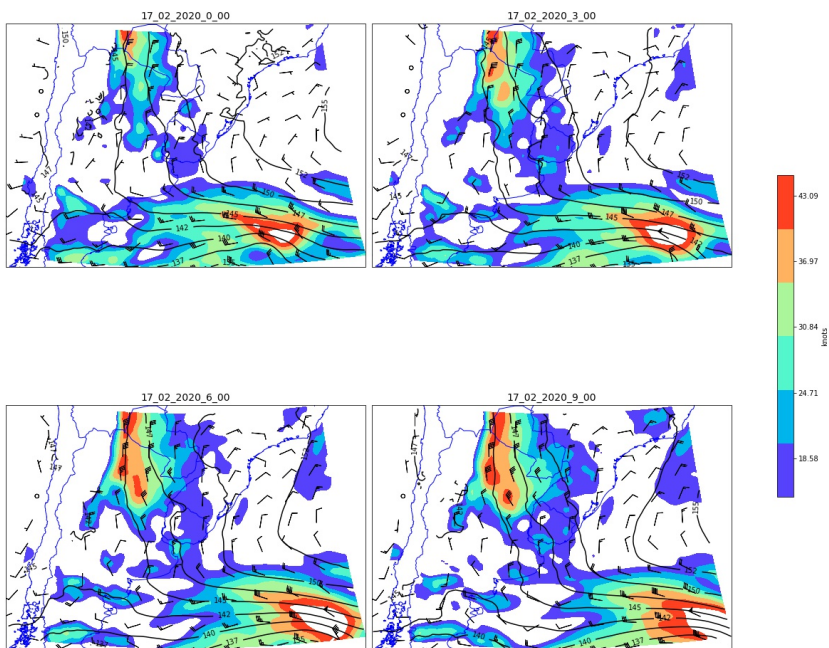
Viento 200 hPa y Z200



(b) Horas de la Tarde

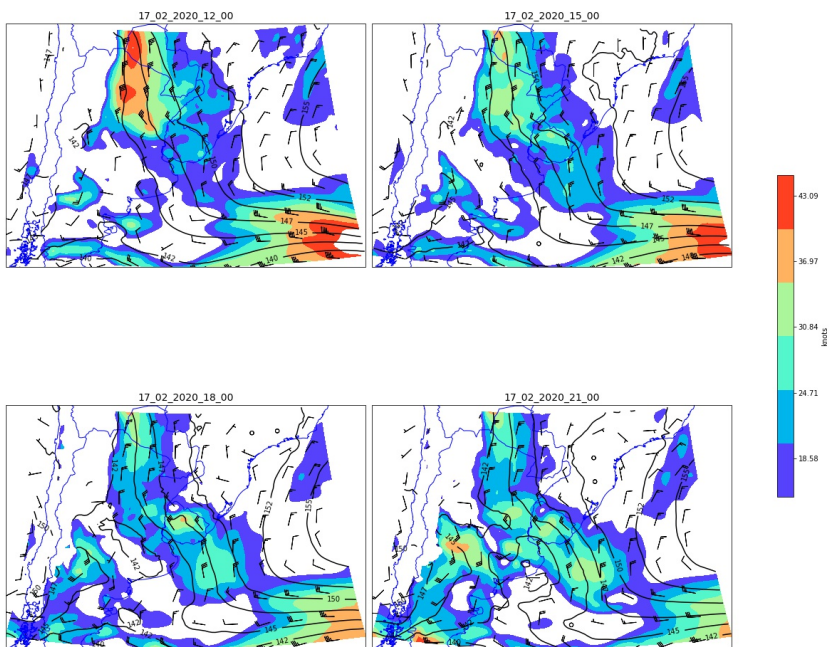
Figura 5.31: Viento y Geopotencial en 200 hPa para el día 17 de febrero de 2020.

Viento 850 hPa y Z850



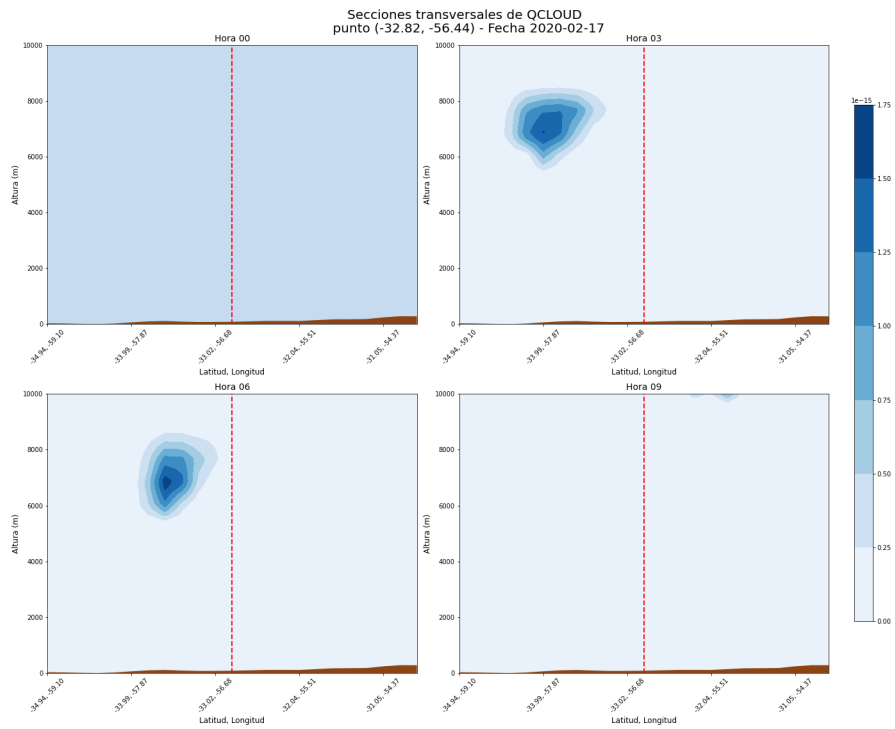
(a) Horas de la Mañana

Viento 850 hPa y Z850

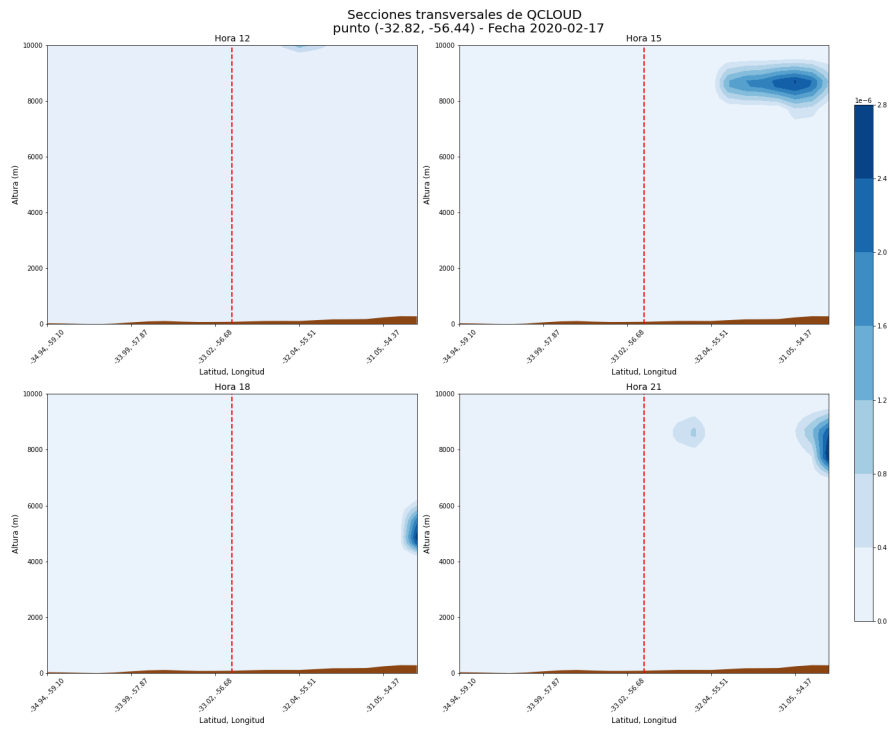


(b) Horas de la Tarde

Figura 5.32: Viento y Geopotencial en 850 hPa para el día 17 de febrero de 2020.



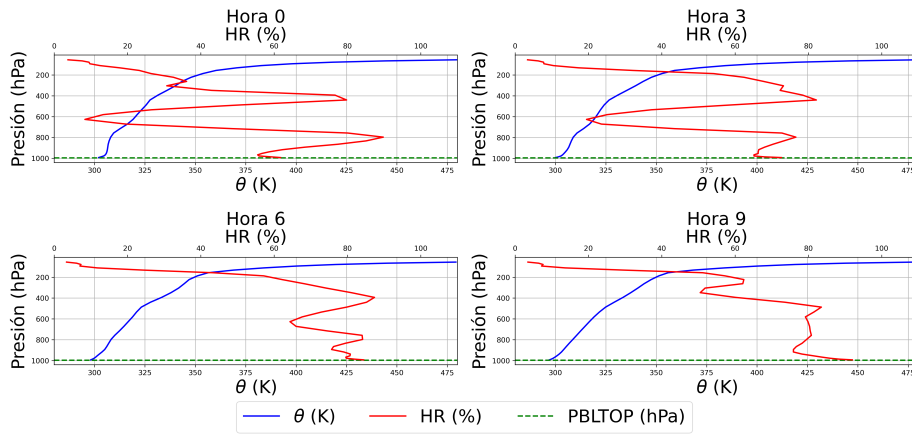
(a) Horas de la Mañana



(b) Horas de la Tarde

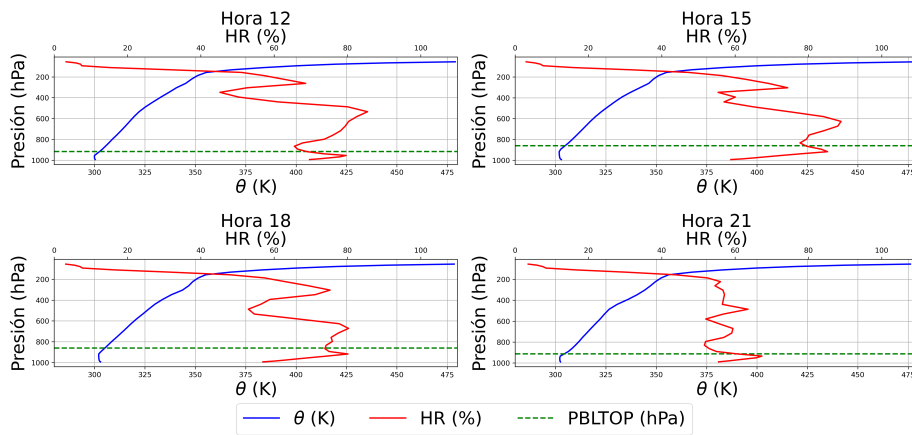
Figura 5.33: Sección transversal para el día 17 de febrero de 2020, la línea vertical roja muestra el punto central de coordenadas -32.82°S , -56.44°W .

Perfiles Verticales para el 17-02-2020 en lat=-32.82, lon=-56.44
 Humedad Relativa (HR) y Temperatura Potencial (θ)



(a) Horas de la Mañana

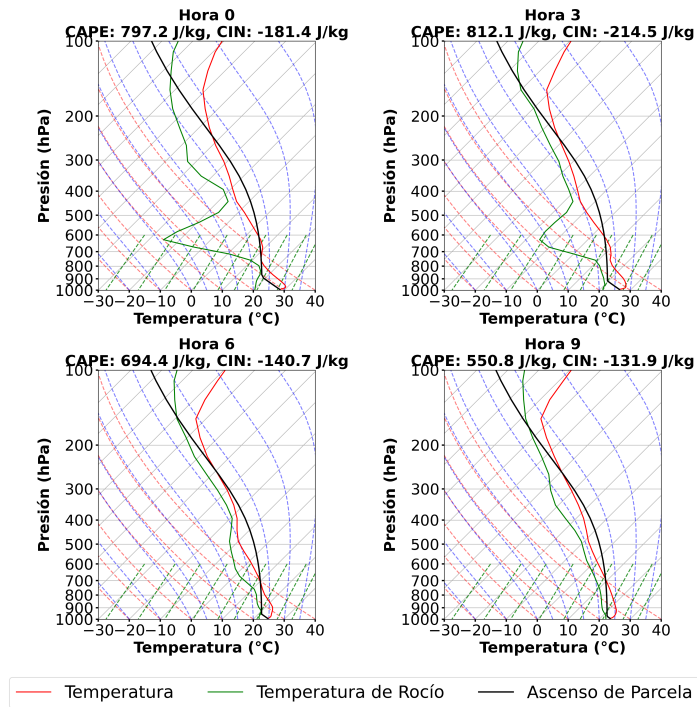
Perfiles Verticales para el 17-02-2020 en lat=-32.82, lon=-56.44
 Humedad Relativa (HR) y Temperatura Potencial (θ)



(b) Horas de la Tarde

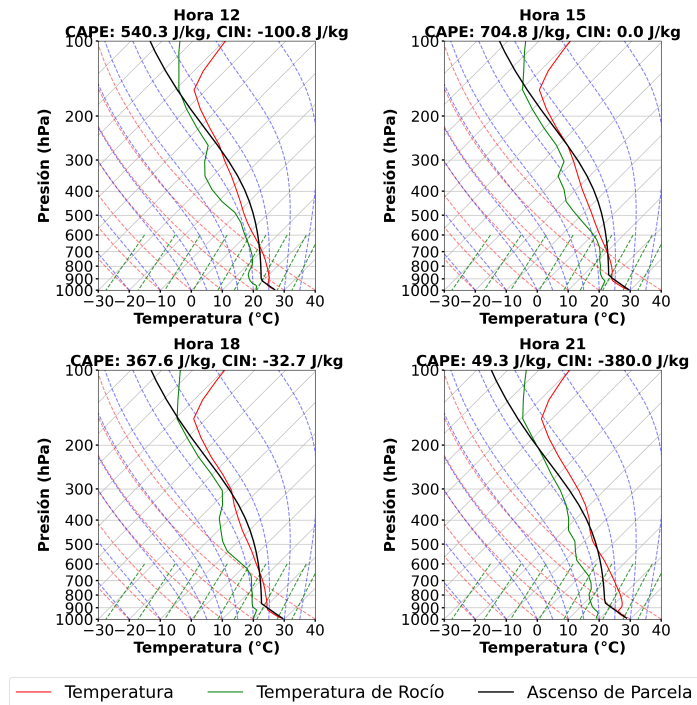
Figura 5.34: Perfiles HR, Temperatura Potencial y Tope de PBL para el día 17 de febrero de 2020.

SkewT-logP : 17-02-2020
Lat: -32.82, Lon: -56.44



(a) Horas de la Mañana

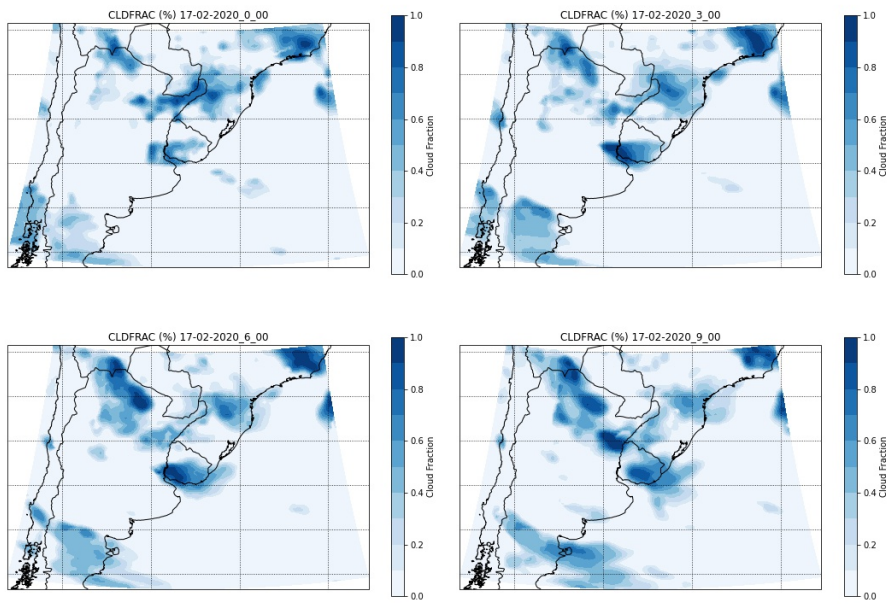
SkewT-logP : 17-02-2020
Lat: -32.82, Lon: -56.44



(b) Horas de la Tarde

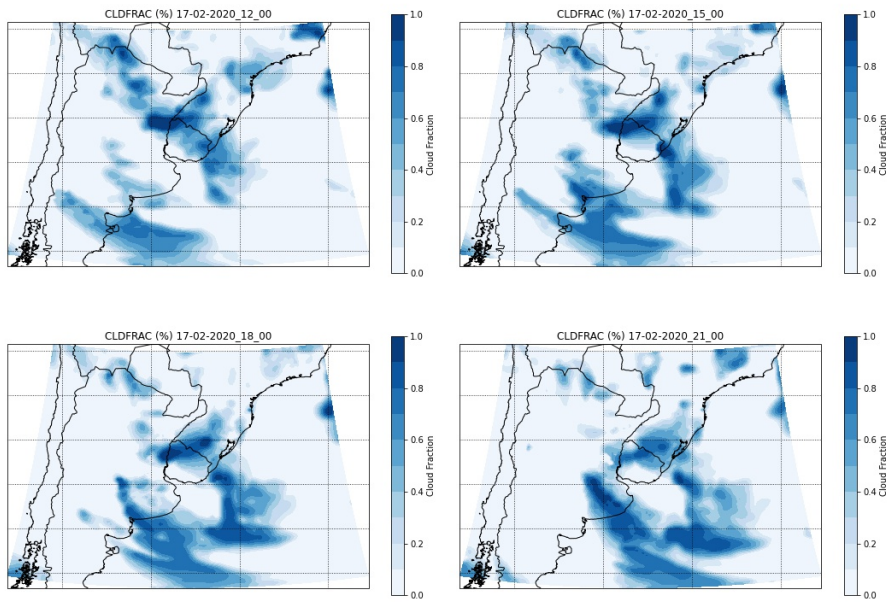
Figura 5.35: Diagramas SkewT-logP para el día 17 de febrero de 2020.

CLDFRAC for h1 - High-Level Clouds



(a) Horas de la Mañana

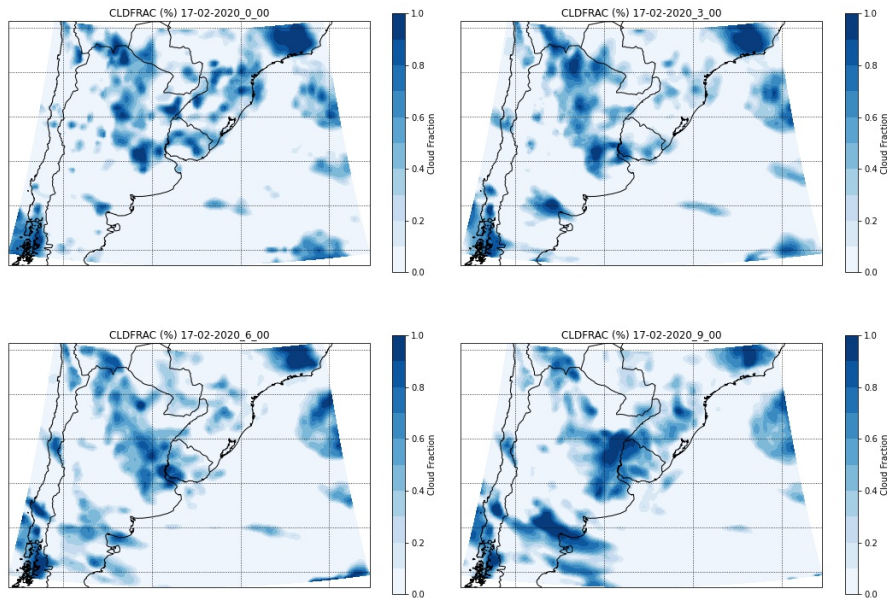
CLDFRAC for h2 - High-Level Clouds



(b) Horas de la Tarde

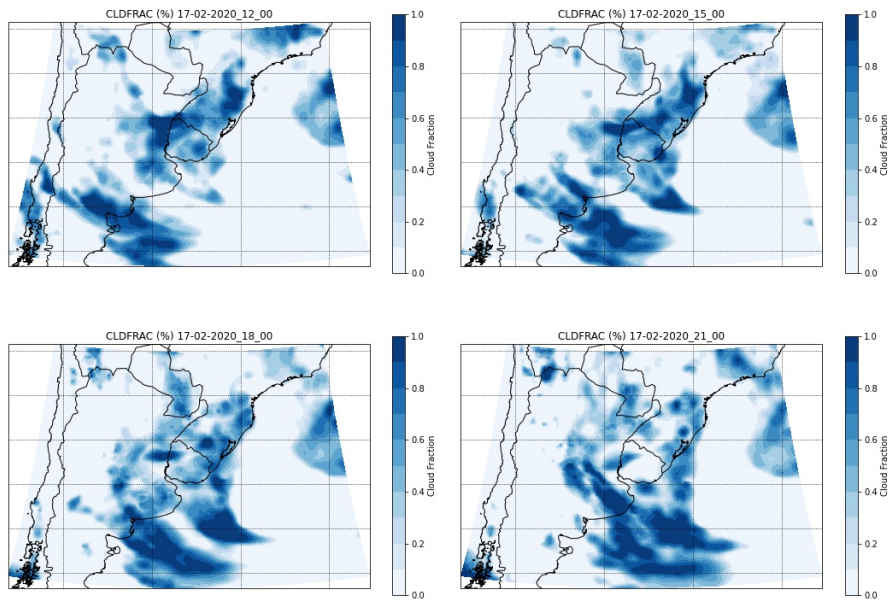
Figura 5.36: Fracción de nubes altas para el día 17 de febrero de 2020.

CLDFRAC for h1 - Mid-Level Clouds



(a) Horas de la Mañana

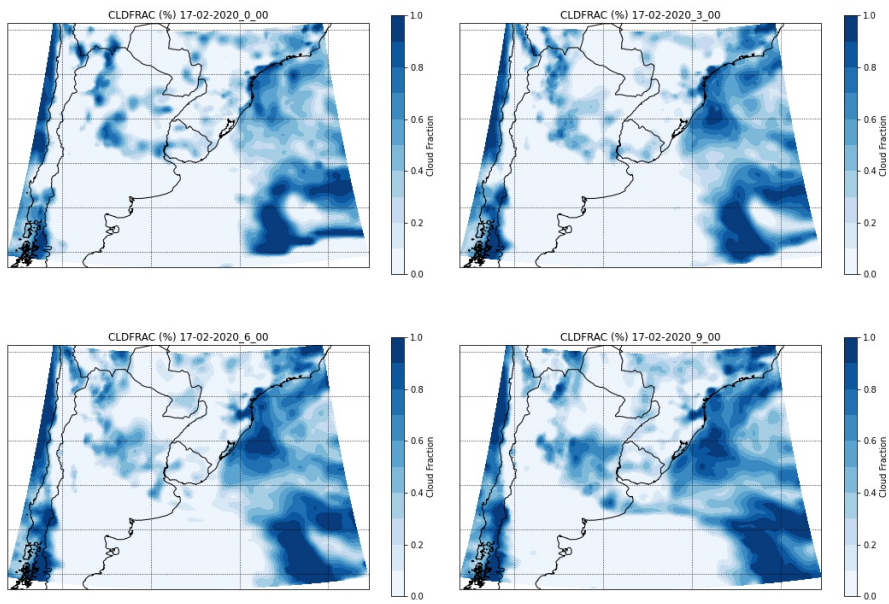
CLDFRAC for h2 - Mid-Level Clouds



(b) Horas de la Tarde

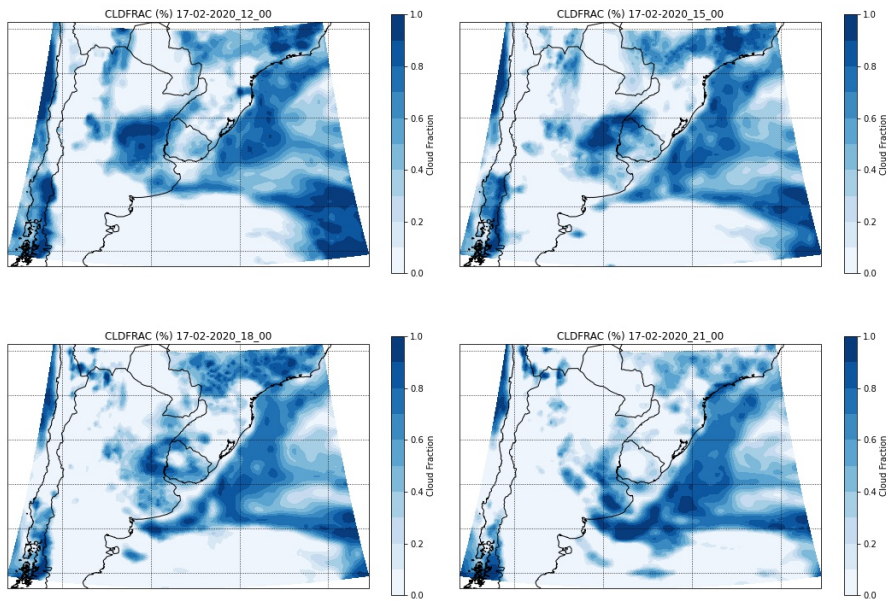
Figura 5.37: Fracción de nubes medias para el día 17 de febrero de 2020.

CLDFRAC for h1 - Low-Level Clouds



(a) Horas de la Mañana

CLDFRAC for h2 - Low-Level Clouds



(b) Horas de la Tarde

Figura 5.38: Fracción de nubes bajas para el día 17 de febrero de 2020.

Página intencionalmente dejada en blanco.

Capítulo 6

Conclusiones y Perspectivas

Este capítulo presenta las conclusiones principales del trabajo. Se consideran las ventajas y desventajas del clasificador desarrollado así como sus posibilidades de mejora. También se desarrollan las conclusiones sobre el desempeño del pronóstico WRF en los eventos seleccionados, así como las perspectivas de análisis futuros considerando la herramienta de clasificación desarrollada y otras herramientas que podrían ser desarrolladas.

Este trabajo desarrolló y evaluó tres clasificadores de altura de tope de nube. Estos están basados en las técnicas de aprendizaje automático presentadas y aprovechan la información satelital operativa GOES16. Éste es el primer trabajo de esta naturaleza en Uruguay, siendo además una temática con escasos antecedentes regionales e internacionales, y con buen grado de espacio de mejora. La realización de este trabajo requirió adquirir nuevas habilidades, como el uso, manejo y procesamiento de imágenes de satélite de distinto origen, el entrenamiento y validación de clasificadores multi-clase basados en aprendizaje automático, y el uso del modelo de predicción numérica de atmósfera WRF. Se generó además evidencia novedosa sobre el diagnóstico de predicciones numéricas de nubosidad en el país, que es central en la predicción de la radiación solar y la generación solar fotovoltaica.

Se desarrollaron clasificadores basados en tres técnicas clásicas de aprendizaje por computadora: DTC, MLP y RNDF. Al ser utilizados con las imágenes GOES16 como entrada estos logran un desempeño similar y en algunos casos superior al que se tiene en la clasificación derivada del producto CTH ABI. La información de entrada para los clasificadores incluye las imágenes propiamente dichas y algunas cantidades derivadas que surgen de su procesamiento, que se entendieron que podían ser relevantes para la tarea. El producto CTH ABI se utiliza como base de desempeño, para establecer desempeños típicos para esta clasificación en Uruguay y comparar con los obtenidos. En particular, se logró una Exactitud Balanceada superadora de la base de desempeño con el algoritmo RNDF, que logra obtener un valor de 81 %, unos dos puntos porcentuales por encima del CTH ABI. Este algoritmo logra además un mejor nivel de acierto en la clasificación de nubes bajas, foco de interés de este trabajo, alcanzando un $\simeq 80\%$ y superando en $\simeq 3\%$ al CTH ABI en esta tarea.

De la evaluación de desempeño por categorías, surge que la clasificación derivada de CTH ABI presenta tasas de acierto dispares, alcanzando para la categoría 3 valores, $AP_{3,3} = 98.3\%$, mientras que para las categorías 1 y 2 se tienen $AP_{1,1} = 77.6\%$ y $AP_{2,2} = 64.3\%$. Esto indica una gran capacidad de CTH ABI para identificar la categoría 3, penalizando la habilidad de reconocer las categorías 1 y 2. En cambio, todos los clasificadores desarrollados por aprendizaje automático en este trabajo muestran tasas $AP_{1,1}$ y $AP_{2,2}$ superiores a las que ofrece CTH ABI, volviendo más simétricas las matrices de confusión, es decir, distribuyendo mejor las tasas de EP. De esta forma, logran una mejora de desempeño ya que estos clasificadores ofrecen una habilidad más uniforme en todas las categorías de la clasificación.

No obstante, se lograron clasificadores superadores de lo anterior. Se exploró el uso de los algoritmos DTC, MLP y RNDF con el producto CTH ABI como característica de entrada adicional, logrando mejorar significativamente el desempeño de CTH ABI y de los algoritmos anteriores en todas las métricas evaluadas. En el análisis de las matrices de confusión se destaca el fortalecimiento de la tasa AP en todas las categorías y, nuevamente, se observa que los clasificadores basados en algoritmos de aprendizaje automático vuelven más simétricas las matrices de confusión, mejorando

el desempeño de la clasificación para las categorías 1 y 2, mientras baja levemente el desempeño en la categoría 3 respecto a la clasificación CTH ABI. Por ejemplo, para el caso de CTH RNDF, se tiene que $AP_{3,3,CTHABI} = 98.3\%$, mientras que $AP_{3,3,CTHRNDF} = 95.1\%$. Este algoritmo (CTH RNDF) es el algoritmo de mejor desempeño desarrollado, con una Exactitud Balanceada de 86 %, y tasas de acierto en las clasificaciones de nubes bajas, medias y altas de 86 %, 78 % y 96 %, respectivamente. Las clasificaciones obtenidas son, en todos los casos, espacialmente consistentes, permitiendo clasificar la altura de topes nubosos en variadas escenas.

Se destaca que este trabajo es una de las pocas experiencias regionales en el campo de la clasificación de nubosidad y la primera en Uruguay, donde se logró el desarrollo y validación de un producto de clasificación de nubosidad con capacidad operativa en horas diurnas y alta capacidad de acierto. La consolidación de la base de datos implicó el manejo del producto CTH ABI también por primera vez en Uruguay. En el futuro, sería necesario estudiar el comportamiento del producto CTH ABI y los clasificadores desarrollados, con medidas en tierra del ceilómetro instalado en el Aeropuerto Internacional de Carrasco. También, se podrían explorar oportunidades de mejora del producto, tales como la optimización del proceso de entrenamiento de los clasificadores, el reentrenamiento de estos clasificadores con una BDT ampliada y mejorada, y la incorporación de otras características o de la física del problema.

En particular se utilizó el mejor clasificador desarrollado como herramienta para identificar días con nubosidad dominante de las categorías 1 y 3. Estos días fueron simulados con el modelo WRF utilizando la configuración descrita en la sección 5.2. En general, se tiene que los dos eventos de nubosidad categoría 3 (4 de noviembre de 2019 y 17 de febrero de 2020), como el evento de nubosidad categoría 1 (21 de diciembre de 2019), son bien representados por el modelo WRF. En estos tres casos se observa nubosidad pronosticada por el modelo acorde con las categorías identificadas. En cambio, el evento del 7 de enero de 2020 no logra ser bien representado por el WRF. Se plantea la hipótesis de que en la simulación haya un exceso de difusión de humedad en el tope de la PBL, lo que no permite representar el LCL dentro de esta, frenando así la generación de Stratocúmulus en la simulación. También se plantean otras hipótesis, como que las condiciones de borde proporcionadas por el modelo GFS no representen correctamente la humedad y esto afecte la simulación del WRF en estos casos, o que la simulación WRF esté suponiendo un entrainment mayor al que realmente ocurre y por este motivo no se llegue a la condensación en niveles cercanos al tope de la PBL. En cualquier caso, se constata una falencia en la predicción de nubes bajas del modelo, asistido por el uso de la herramienta de clasificación desarrollada en este trabajo. El trabajo de [Porrini \(2017\)](#) cuantificó una subestimación de HR por parte del modelo WRF, correlacionada con sobreestimaciones de radiación. Este caso se encuadra en ese análisis. Al respecto de las hipótesis propuestas para el caso del 7 de enero de 2020, sería deseable realizar campañas de mediciones intensivas de perfiles verticales en un cierto periodo de tiempo, que permitan estudiar en detalle los perfiles termodinámicos observados en la PBL y sus respectivas simulaciones con el modelo WRF. Nuevamente, en un enfoque experimental de este tipo, sería deseable considerar las medidas del ceilómetro instalado en el Aeropuerto Internacional de Carrasco.

Finalmente, existen varios elementos que podrían mejorar el trabajo actual, entre los que se encuentra la mejora de consistencia espacio-temporal de la información satelital, la disponibilidad de observaciones terrestres de base y tope de nube, incluir la variabilidad temporal (no sólo la espacial) como característica de entrada, optimizar los clasificadores, ampliar la base de datos a otros productos meteorológicos, modelos atmosféricos o canales espectrales, entre muchos otros. Iniciar una línea de investigación que se enfoque en la física de este problema seguramente resultará en insumos claves para mejorar las herramientas de clasificación. Este trabajo abordó una problemática inicial y acotada de dos problemas de significativo mayor volumen, que son la clasificación de tipo de nubes por satélite y la estimación de propiedades físicas de nubes por satélite. No se ha abordado, por ejemplo, el desarrollo de una herramienta para estimar la altura del tope nuboso propiamente dicho, solo su clasificación en categorías, o poder distinguir una nube tipo Cirrus de un Strato, o cualquier otra distinción. Se espera que este trabajo pueda ser una primera contribución en esa dirección.

Referencias

- Abal, G., Alonso-Suárez, R., y Laguarda, A. (2020). *Radiación solar. notas del curso fundamentos del recurso solar*.
- Ackerman, S. A., Holz, R. E., Frey, R., Eloranta, E. W., Maddux, B., y McGill, M. (2008). Cloud detection with modis. part ii: validation. *Journal of Atmospheric and Oceanic Technology*, 25(7), 1073–1086.
- Ackerman, S. A., Strabala, K. I., Menzel, W. P., Frey, R. A., Moeller, C. C., y Gumley, L. E. (1998). Discriminating clear sky from clouds with modis. *Journal of Geophysical Research: Atmospheres*, 103(D24), 32141–32157.
- 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 Musé, P. (2012). Brightness-dependent tarpley model for global solar radiation estimation using goes satellite images: Application to uruguay. *Solar Energy*, 86(11), 3205–3215. doi: <https://doi.org/10.1016/j.solener.2012.08.012>
- Alonso-Suárez, R., Abal, G., Musé, P., y Siri, R. (2014). Satellite-derived solar irradiation map for uruguay. *Energy Procedia*, 57, 1237–1246.
- Alonso-Suárez, R. (2017). *Estimación del recurso solar en uruguay mediante imágenes satelitales* (Tesis Doctoral no publicada). Facultad de Ingeniería, Universidad de la República.
- Arce, A. M. G., de Almeida, E., Porrini, C., Romero, M., Ortelli, S., Cazes, G., ... others (2018). Pronósticos operativos de energía eólica, para la gestión de los sistemas eléctricos en la región. *ENERLAC. Revista de energía de Latinoamérica y el Caribe*, 2(1), 28–41.
- Baum, B. A., Menzel, W. P., Frey, R. A., Tobin, D. C., Holz, R. E., Ackerman, S. A., ... Yang, P. (2012). Modis cloud-top property refinements for collection 6. *Journal of applied meteorology and climatology*, 51(6), 1145–1163.
- Bottino, M. J., y Ceballos, J. C. (2006). Classificação de imagens goes da américa do sul utilizando canais termais. En *Anais del xiv congresso brasileiro de meteorologia*.
- Breiman, L., y Cutler, R. (2001). Random forests machine learning [j]. *journal of clinical microbiology*, 2, 199–228.
- Cazes Boezio, G., y Ortelli, S. (2019). Use of the wrf-da 3d-var data assimilation system to obtain wind speed estimates in regular grids from measurements at wind farms in uruguay. *Data*, 4(4), 142.
- Ceballos, J. C., y Negri, R. G. (2022). *Classificação de imagens goes (imager e abi) na dissm/cgct/inpe: Centróides associados*. (Inf. Téc.). Brasil: División de Satélites y Sistemas Ambientales (DSA), Instituto Nacional de Pesquisas Espaciais (INPE).
- Chahine, M. T. (1974). Remote sounding of cloudy atmospheres. i. the single cloud layer. *Journal of the Atmospheric Sciences*, 31(1), 233–243.
- Chahine, M. T. (1982). Remote sensing of cloud parameters. *Journal of Atmospheric Sciences*, 39(1), 159–170.

- Council, N. R., on Earth, D., Studies, L., on Atmospheric Sciences, B., y on Weather Radar Technology Beyond NEXRAD, C. (2002). *Weather radar technology beyond nexrad*. National Academies Press.
- De Almeida, E. (2019). *Pronóstico de energía eólica para horizontes temporales de corto plazo en base a modelo numérico de mesoescala y redes neuronales artificiales* (Tesis de Master no publicada). Facultad de Ingeniería Udelar.
- de Mello, S., Cazes, G., y Gutierrez, A. (2015). Operational wind energy forecast with power assimilation. En *14th international conference on wind engineering, porto alegre, brazil*.
- Doviak, R. J., Zrnic, D. S., y Sirmans, D. S. (1979). Doppler weather radar. *Proceedings of the IEEE*, 67(11), 1522–1553.
- Dudhia, J. (2015). Overview of wrf physics. 2015 basic wrf tutorial, boulder, co. *National Center for Atmospheric Research*.
- Efron, B., y Tibshirani, R. J. (1994). *An introduction to the bootstrap*. Chapman and Hall/CRC.
- Farnebäck, G. (2003). Two-frame motion estimation based on polynomial expansion. En *Image analysis: 13th scandinavian conference, scia 2003 halmstad, sweden, june 29–july 2, 2003 proceedings 13* (pp. 363–370).
- Forsythe, M. (2007). Atmospheric motion vectors: past, present and future. En *Ecmwf annual seminar* (pp. 1–79).
- Fu, Q., y Liou, K. N. (1993). Parameterization of the radiative properties of cirrus clouds. *Journal of Atmospheric Sciences*, 50(13), 2008–2025.
- Grandini, M., Bagli, E., y Visani, G. (2020). Metrics for multi-class classification: an overview. *arXiv preprint arXiv:2008.05756*.
- Hansen, J., Sato, M., Ruedy, R., Lacis, A., Asamoah, K., Beckford, K., . . . others (1997). Forcings and chaos in interannual to decadal climate change. *Journal of Geophysical Research: Atmospheres*, 102(D22), 25679–25720.
- Haykin, S. (1994). *Neural networks: a comprehensive foundation*. Prentice Hall PTR.
- Heidinger, A. (2012, July). *Abi cloud height* (Inf. Téc.). NOAA NESDIS, Center for Satellite Applications and Research. (Algorithm Theoretical Basis Document)
- Heidinger, A. K., Pavolonis, M., Holz, R., Baum, B. A., y Berthier, S. (2010). Using CALIPSO to explore the sensitivity to cirrus height in the infrared observations from npoess/viirs and goes-r/abi. *Journal of Geophysical Research: Atmospheres*, 115(D4).
- Heidinger, A. K., Pavolonis, M. J., Calvert, C., Hoffman, J., Nebuda, S., Straka III, W., . . . Wanzong, S. (2020). Abi cloud products from the goes-r series. En *The goes-r series* (pp. 43–62). Elsevier.
- Holz, R., Ackerman, S., Nagle, F., Frey, R., Dutcher, S., Kuehn, R., . . . Baum, B. (2008). Global moderate resolution imaging spectroradiometer (modis) cloud detection and height evaluation using caliop. *Journal of Geophysical Research: Atmospheres*, 113(D8).
- Hong, S.-Y., y Lim, J.-O. J. (2006). The wrf single-moment 6-class microphysics scheme (wsm6). *Asia-Pacific Journal of Atmospheric Sciences*, 42(2), 129–151.
- Horn, B. K. P., y Schunck, B. G. (1981, Aug). Determining optical flow. *Artif. Intell.*, 17(1-3), 185–203. Descargado de [http://dx.doi.org/10.1016/0004-3702\(81\)90024-2](http://dx.doi.org/10.1016/0004-3702(81)90024-2) doi: 10.1016/0004-3702(81)90024-2
- Howard, L. (1803). Lxiv. on the modifications of clouds, and on the principles of their production, suspension, and destruction; being the substance of an essay read before the askesian society in the session 1802–3. *The Philosophical Magazine*, 16(64), 344–357.
- Huo, J., Lu, D., Duan, S., Bi, Y., y Liu, B. (2020). Comparison of the cloud top heights retrieved from modis and ahi satellite data with ground-based ka-band radar. *Atmospheric Measurement Techniques*, 13(1), 1–11.

- Inoue, T. (1985). On the temperature and effective emissivity determination of semi-transparent cirrus clouds by bi-spectral measurements in the $10\mu\text{m}$ window region. *Journal of the Meteorological Society of Japan. Ser. II*, 63(1), 88–99.
- Isaacs, R., Hoffman, R., y Kaplan, L. (1986). Satellite remote sensing of meteorological parameters for global numerical weather prediction. *Reviews of Geophysics*, 24(4), 701–743.
- Justus, C., Paris, M., y Tarpley, J. (1986). Satellite-measured insolation in the United States, Mexico, and South America. *Remote Sensing of Environment*, 20(1), 57–83. doi: 10.1016/0034-4257(86)90014-3
- Li, Z., Li, J., Menzel, W. P., Schmit, T. J., y Ackerman, S. A. (2007). Comparison between current and future environmental satellite imagers on cloud classification using modis. *Remote Sensing of Environment*, 108(3), 311–326.
- Lin, Y.-L., Farley, R. D., y Orville, H. D. (1983). Bulk parameterization of the snow field in a cloud model. *Journal of Applied Meteorology and climatology*, 22(6), 1065–1092.
- Liou, K.-N. (1974). On the radiative properties of cirrus in the window region and their influence on remote sensing of the atmosphere. *Journal of Atmospheric Sciences*, 31(2), 522–532.
- Liu, Y., Liu, Y., Muñoz-Esparza, D., Hu, F., Yan, C., y Miao, S. (2020). Simulation of flow fields in complex terrain with wrf-les: Sensitivity assessment of different pbl treatments. *Journal of Applied Meteorology and Climatology*, 59(9), 1481–1501.
- Loh, W.-Y. (2011). Classification and regression trees. *Wiley interdisciplinary reviews: data mining and knowledge discovery*, 1(1), 14–23.
- Lu, X., Mao, F., Rosenfeld, D., Zhu, Y., Pan, Z., y Gong, W. (2021). Satellite retrieval of cloud base height and geometric thickness of low-level cloud based on calipso. *Atmospheric Chemistry and Physics*, 21(15), 11979–12003.
- Lucas, B. D., y Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. En *Proceedings of the 7th international joint conference on artificial intelligence - volume 2* (pp. 674–679). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. Descargado de <http://dl.acm.org/citation.cfm?id=1623264.1623280>
- Maahn, M., Turner, D. D., Löhnert, U., Posselt, D. J., Ebell, K., Mace, G. G., y Comstock, J. M. (2020). Optimal estimation retrievals and their uncertainties: What every atmospheric scientist should know. *Bulletin of the American Meteorological Society*, 101(9), E1512–E1523.
- Mantovani Júnior, J. A., Aravéquia, J. A., Carneiro, R. G., y Fisch, G. (2023). Evaluation of pbl parameterization schemes in wrf model predictions during the dry season of the central amazon basin. *Atmosphere*, 14(5), 850.
- Marchand, R., Mace, G. G., Ackerman, T., y Stephens, G. (2008). Hydrometeor detection using cloudsat—an earth-orbiting 94-ghz cloud radar. *Journal of Atmospheric and Oceanic Technology*, 25(4), 519–533.
- Markovic, M., Jones, C. G., Vaillancourt, P. A., Paquin, D., Winger, K., y Paquin-Ricard, D. (2008). An evaluation of the surface radiation budget over north america for a suite of regional climate models against surface station observations. *Climate dynamics*, 31, 779–794.
- Matuszko, D. (2012). Influence of the extent and genera of cloud cover on solar radiation intensity. *International Journal of climatology*, 32(15), 2403–2414.
- McGill, M., Hlavka, D., Hart, W., Scott, V. S., Spinhirne, J., y Schmid, B. (2002). Cloud physics lidar: Instrument description and initial measurement results. *Applied Optics*, 41(18), 3725–3734.
- Menzel, W. P., Frey, R. A., y Baum, B. A. (2015, May). *Cloud top properties and cloud phase algorithm theoretical basis document* (Inf. Téc.). University of Wisconsin – Madison. (Version 11)
- Menzel, W. P., Frey, R. A., Zhang, H., Wylie, D. P., Moeller, C. C., Holz, R. E., ... Gumley, L. E. (2008). Modis global cloud-top pressure and amount estimation: Algorithm description and results. *Journal of Applied Meteorology and Climatology*, 47(4), 1175–1198.

- Menzel, W. P., y Strabala, K. (1997). *Cloud top properties and cloud phase algorithm theoretical basis document*. University of Wisconsin–Madison.
- Min, M., Li, J., Wang, F., Liu, Z., y Menzel, W. P. (2020). Retrieval of cloud top properties from advanced geostationary satellite imager measurements based on machine learning algorithms. *Remote Sensing of Environment*, 239, 111616.
- Minnis, P., Yost, C. R., Sun-Mack, S., y Chen, Y. (2008). Estimating the top altitude of optically thick ice clouds from thermal infrared satellite observations using calipso data. *Geophysical Research Letters*, 35(12).
- Mitra, A., Di Girolamo, L., Hong, Y., Zhan, Y., y Mueller, K. J. (2021). Assessment and error analysis of terra-modis and misr cloud-top heights through comparison with iss-cats lidar. *Journal of Geophysical Research: Atmospheres*, 126(9), e2020JD034281.
- NOAA. (2018). *Noaa goes-r series advanced baseline imager (abi) level 2 cloud top height (acha)* (Inf. Téc.). NOAA National Centers for Environmental Information. Descargado access date, de <https://doi.org/10.7289/V5HX19ZQ> (GOES-R Algorithm Working Group and GOES-R Series Program Office) doi: 10.7289/V5HX19ZQ
- Oliveira, R., Ceballos, J., y de Angelis, C. (2008). Classificação de nuvens diurna e noturna usando canais termais do satélite goes 12. En *Anais del xv congresso brasileiro de meteorologia*.
- Orteli, S. (2023). *Avances en predicciones meteorológicas para la gestión de sistemas hidroeléctricos: Un análisis comparativo entre ecmuf y gfs en el río negro, uruguay* (Tesis Doctoral no publicada). Facultad de Ingeniería Udelar.
- Osborne, S. R., Abel, S. J., Boutle, I. A., y Marengo, F. (2014). Evolution of stratocumulus over land: comparison of ground and aircraft observations with numerical weather prediction simulations. *Boundary-layer meteorology*, 153, 165–193.
- Porrini, C. (2017). *Evaluación del modelo regional wrf para pronósticos de radiación solar en superficie dentro del territorio uruguayo* (Tesis de Master no publicada). Universidad de la República.
- Porrini, C., Gutiérrez, A., Boezio, G. C., Hermida, G., Oroño, D., y Puppo, M. (2015). Development of a model output statistic and implementation of an operational solar photovoltaic energy forecast model based in wrf. En *2015 ieee pes innovative smart grid technologies latin america (isgt latam)* (pp. 248–253).
- Powers, J. G., Klemp, J. B., Skamarock, W. C., Davis, C. A., Dudhia, J., Gill, D. O., ... others (2017). The weather research and forecasting model: Overview, system efforts, and future directions. *Bulletin of the American Meteorological Society*, 98(8), 1717–1737.
- Purbantoro, B., Aminuddin, J., Manago, N., Toyoshima, K., Lagrosas, N., Sumantyo, J. T. S., y Kuze, H. (2018). Comparison of cloud type classification with split window algorithm based on different infrared band combinations of himawari-8 satellite. *Advances in Remote Sensing*, 7(3), 218–234.
- Sassen, K., y Wang, Z. (2008). Classifying clouds around the globe with the cloudsat radar: 1-year of results. *Geophysical research letters*, 35(4).
- Saunders, R. W., y Kriebel, K. T. (1988). An improved method for detecting clear sky and cloudy radiances from avhrr data. *International Journal of Remote Sensing*, 9(1), 123–150.
- Schmit, T. J., Griffith, P., Gunshor, M. M., Daniels, J. M., Goodman, S. J., y Lehair, W. J. (2017). A closer look at the abi on the goes-r series. *Bulletin of the American Meteorological Society*, 98(4), 681–698.
- Silveira, L., Chreties, C., Alonso, J., Cazes Boezio, G., De Vera, A., y Crisci, M. (2018). Sensibilidad de un sistema de alerta temprana de inundaciones. En *Anales del xxviii congreso latinoamerica de hidráulica*.
- Smith, W., y Platt, C. (1978). Comparison of satellite-deduced cloud heights with indications from radiosonde and ground-based laser measurements. *Journal of Applied Meteorology (1962-1982)*, 1796–1802.

- Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., . . . others (2002). The cloudsat mission and the a-train: A new dimension of space-based observations of clouds and precipitation. *Bulletin of the American Meteorological Society*, 83(12), 1771–1790.
- Stephens, G. L., Vane, D. G., Tanelli, S., Im, E., Durden, S., Rokey, M., . . . others (2008). Cloudsat mission: Performance and early science after the first year of operation. *Journal of Geophysical Research: Atmospheres*, 113(D8).
- Stephens, G. L., y Webster, P. J. (1981). Clouds and climate: Sensitivity of simple systems. *Journal of Atmospheric Sciences*, 38(2), 235–247.
- Teixeira-Branco, V., Alonso-Suárez, R., de Almeida, E., Porrini, C., Gutiérrez, A., y Cazes, G. (2018). Evaluación del pronóstico de irradiación solar diaria en uruguay utilizando el modelo wrf. En *Anais do vii congresso brasileiro de energia solar (cbens)*.
- Teixeira-Branco, V., Alonso-Suárez, R., de Almeida, E., Porrini, C., Gutiérrez, A., y Cazes, G. (2023). *Predicción de la irradiación solar en uruguay utilizando modelos numéricos de atmósfera* (Tesis de Master no publicada). Facultad de Ingeniería Udelar.
- Vaughan, M. A., Young, S. A., Winker, D. M., Powell, K. A., Omar, A. H., Liu, Z., . . . Hostetler, C. A. (2004). Fully automated analysis of space-based lidar data: An overview of the calipso retrieval algorithms and data products. *Laser radar techniques for atmospheric sensing*, 5575, 16–30.
- Wang, C., Platnick, S., Meyer, K., Zhang, Z., y Zhou, Y. (2020). A machine-learning-based cloud detection and thermodynamic-phase classification algorithm using passive spectral observations. *Atmospheric Measurement Techniques*, 13(5), 2257–2277.
- Winker, D. M., Hunt, W. H., y McGill, M. J. (2007). Initial performance assessment of caliop. *Geophysical Research Letters*, 34(19).
- Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., . . . Young, S. A. (2009). Overview of the calipso mission and caliop data processing algorithms. *Journal of Atmospheric and Oceanic Technology*, 26(11), 2310–2323.
- WMO. (1956). *International cloud atlas: Abridged atlas*. World Meteorological Organization.
- Yang, H., y Kleissl, J. (2016). Preprocessing wrf initial conditions for coastal stratocumulus forecasting. *Solar Energy*, 133, 180–193.
- Yucheche, A. E., Lakkis, S. G., Caferra, A., Canziani, P. O., y Muszkats, J. P. (2020). A cluster approach to cloud cover classification over south america and adjacent oceans using a k-means/k-means++ unsupervised algorithm on goes ir imagery. *Remote Sensing*, 12(18), 2991.
- Zhang, R., Xiao, H., Gao, Y., Su, H., Li, D., Wei, L., . . . Li, H. (2023). Shape classification of cloud particles recorded by the 2d-s imaging probe using a convolutional neural network. *Journal of Meteorological Research*, 37(4), 521–535.
- Zhang, Y., Hemperly, J., Meskhidze, N., y Skamarock, W. C. (2012). The global weather research and forecasting (gwrp) model: Model evaluation, sensitivity study, and future year simulation. *Scientific Research*.

Página intencionalmente dejada en blanco.

Anexo A

Abreviaciones

ABI Advanced Baseline Imager	1
ACHA ABI Cloud Height Algorithm	14
AHI Advanced Himawari Imager	20
AIN Atlas Internacional de Nubes	5
AQUA AQUA del programa Earth Observating System	10
ARW Advance Research WRF	18
BDT Base de Datos de Trabajo	1
CALIOP Cloud-Aerosol Lidar with Orthogonal Polarization	16
CALIPSO Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation	11
CATS International Space Station Cloud-Aerosol Transport System	12
CBH Altura de Base de Nube	16
CM Cloud Mask	2
CMV Cloud Motion Vectors	26
CLDFRA Fracción de Nubosidad	6
CO2 Dióxido de Carbono	9
CPR Cloud Profiling Radar	16
CPTEC Centro de Previsão de Tempo e Estudos Climáticos	20

CTH Altura de Tope de Nube	1
CTP Presión de Tope de Nube	6
CTPCPA Cloud Top Properties and Cloud Phase Algorithm de MODIS	10
CTT Temperatura de Tope de Nube	6
CZ Coseno del Ángulo Cenital Solar	29
DSA División de Sistemas Ambientales	20
DCSFCMA Discriminating Clear-Sky from Cloud With Modis Algorithm	10
DTC Arbol de Decisión simple	41
ECMWF European Centre for Medium-Range Weather Forecasts	17
FOV Campo de Visión	9
FWHM Full Width Half Maximum	12
GFS Global Forecast System	20
GOES16 Geostationary Operational Environmental Satellites 16	1
HR Humedad Relativa	1
IMFIA Instituto de Mecánica de Fluidos e Ingeniería Ambiental	19
INPE Instituto Nacional de Pesquisas Espaciais	20
INUMET Instituto Uruguayo de Meteorología	8
ISCCP International Satellite Cloud Climatology Project	17
MLP Red Neuronal Multicapa	41
MOD06L2 MODIS-TERRA Cloud Product Level 2	29
MODIS Moderate Resolution Imaging Spectroradiometer	1
MSG Meteosat Second Generation	14
NASA National Aeronautics and Space Administration	1

NCAR National Center for Atmospheric Research	18
NOAA National Oceanic and Atmosphere Administration	1
NWP Predicción Numérica del Tiempo (del inglés: Numerical Weather Prediction)	1
LCL Nivel de Condensación por Ascenso	69
LES Laboratorio de Energía Solar de la UDELAR	3
LIDAR Laser Detection And Ranging	7
LWR Radiación de Onda Larga	1
OMM Organización Meteorológica Mundial	5
PBL Capa Límite Atmosférica	1
RADAR Radio Detection And Ranging	7
RNDF Bósque aleatorio	41
RP Reflectividad Planetaria	2
RP0 Brillo de Fondo	31
RTM Modelos de Transferencias Radiativas en la Atmósfera	2
SEVIRI Spinning Enhanced Visible and Infrared Imager	14
SWR Radiación de Onda Corta	1
TB-IR Temperatura de Brillo Infrarroja	2
TERRA TERRA del programa Earth Observating System	1
UDELAR Universidad de la República	3
VC Validación Cruzada	44
VIIRS Visible Infrared Imaging Radiometer Suite	20
WCRP World Climate Research Program	17
WNORM Ángulo Horario Solar Normalizado	29

WPS WRF Preprocessing System	18
WRF Weather Research and Forecasting	1