

## **Precipitación Integrada Satelital: Combinación de Productos, Temperatura de Brillo Infrarroja y Actividad Eléctrica mediante Redes Neuronales Convolucionales**

Sergio Hernán González<sup>1,2,3,4</sup> [0009-0009-1531-2182], Juan Jose Ruiz<sup>1,2,3</sup> [0000-0002-5079-641X], Pablo Negri<sup>5,6</sup> [0000-0003-0250-5208], Luciano Vidal<sup>4</sup> [0000-0001-6715-2648], Ezequiel Geslin<sup>3</sup>

<sup>1</sup> Centro de Investigaciones del Mar y la Atmósfera (CIMA/CONICET-UBA), Buenos Aires, Argentina,

<sup>2</sup> Instituto Franco-Argentino de Estudios sobre el Clima y sus Impactos (IRL IFAEI/CNRS-IRD-CONICET-UBA), Buenos Aires, Argentina,

<sup>3</sup> Departamento de Ciencias de la Atmósfera y los Océanos (DCAO-FCEN-UBA), Buenos Aires, Argentina

<sup>4</sup> Servicio Meteorológico Nacional de Argentina, Buenos Aires, Argentina

<sup>5</sup> Instituto de Investigación en Ciencias de la Computación (ICC/CONICET-UBA), (DC,FCEyN/UBA), Buenos Aires, Argentina

<sup>6</sup> Departamento de Computación (DC-FCEN-UBA), Buenos Aires, Argentina  
sergio.gonzalez@cima.fcen.uba.ar, jruiz@cima.fcen.uba.ar,  
lvidal@smn.gov.ar, pnegri@gmail.com, ezegeslin@gmail.com

**Resumen.** El monitoreo de la precipitación es sumamente crucial para la actividad agropecuaria, ya que es un componente fundamental del balance hidrológico que tiene un gran impacto en los rindes. Las observaciones in-situ a través de pluviómetros son escasas, por lo cual se complementa con estimaciones de precipitación provenientes de sensores remotos (i.e. satélites y radares meteorológicos) que incrementan la cobertura espacial y temporal. En este trabajo se propone utilizar un modelo de redes neuronales convolucionales con una arquitectura de tipo UNet a partir de datos provistos por el satélite GOES-16. En particular, se evaluará el uso combinado de temperatura de brillo en infrarrojo (que brinda información de la temperatura del tope de las nubes) y la actividad eléctrica (que brinda información sobre la intensidad de la convección). El entrenamiento del modelo se realiza utilizando datos de precipitación estimada por el radar meteorológico a bordo del satélite GPM.

**Palabras clave:** Estimación de Precipitación Satelital, Actividad Eléctrica, UNet.

## **Integrated Satellite Precipitation: Combining Products, Infrared Brightness Temperature and Electrical Activity by Convolutional Neural Networks**

**Abstract.** Precipitation monitoring is extremely crucial for agricultural activities, since it is a fundamental component of the hydrological balance that has a great impact on yields. In-situ observations through rain gauges are scarce, so it is complemented with precipitation estimates from remote sensors (i.e. satellites and meteorological radars) that increase the spatial and temporal coverage. In this work we propose to use a convolutional neural network model with a UNet type architecture, based on data provided by the GOES-16 satellite. In particular, the combined use of infrared brightness temperature (which provides information on cloud top temperature) and electrical activity (which provides information on convection intensity) will be evaluated. Model training is performed using precipitation data estimated by the GPM satellite-borne weather radar.

**Key words:** Satellite Precipitation Estimation, Electrical Activity, UNet.

## 1 Introducción

La variabilidad de la precipitación a escalas espaciales y temporales no puede resolverse completamente con pluviómetros convencionales, especialmente en regiones con escasez de datos de superficie. Con el objetivo de complementar estas observaciones se han desarrollado diferentes metodologías para aprovechar el potencial de los sensores remotos. Los radares pueden proporcionar una estimación de la precipitación en tiempo real y con alta resolución espacial, pero esta solo está disponible en su rango de medición. Los satélites geoestacionarios tienen una amplia cobertura espacial y disponibilidad en tiempo real, con una resolución espacio-temporal relativamente alta. Estas características son muy relevantes para la detección y el monitoreo de nubes precipitantes.

Los datos provenientes de satélites de órbita geoestacionaria que utilizan sensores en el rango infrarrojo (IR) ofrecen ventajas como una alta frecuencia temporal y cobertura espacial sobre continentes y océanos. Esto permite capturar el ciclo de vida de las nubes precipitantes, a diferencia de los satélites de órbita terrestre baja, cuya frecuencia y cobertura de observación son limitadas debido a su tiempo de revisita y ancho de escaneo. Asimismo, satélites geoestacionarios como el Geostationary Operational Environmental Satellites (GOES) - 16 y el actual GOES-19 poseen sensores ópticos que detectan los destellos provenientes de la actividad eléctrica asociada a tormentas.

Estudios previos han demostrado la relación entre la temperatura de brillo (TB) infrarroja, la actividad eléctrica y la precipitación [1,2,3]. Los datos de IR solo proporcionan información sobre las características del tope de las nubes, mientras que los destellos ubican las regiones donde la convección húmeda profunda es más activa.

En este trabajo se propone utilizar un modelo de aprendizaje automático basado en una arquitectura UNet [4] que estime la precipitación a partir del algoritmo de ECP Rainfall Rate and Quantitative Precipitation Estimation (RRQPE) proveniente del satélite GOES-16 [5], y evaluar su desempeño incluyendo datos de TB en infrarrojo y actividad eléctrica. La red se entrena con estimaciones de precipitación del sensor Dual-frequency Precipitation Radar (DPR) [6] a bordo del satélite de órbita baja Global Precipitation Measurement (GPM).

## 2 Datos

El objetivo del trabajo es el desarrollo de modelos de estimaciones de precipitación satelitales mediante redes neuronales convolucionales. Como datos de entrada se utilizan el algoritmo de ECP del satélite GOES-16 (RRQPE), derivado de datos en infrarrojo; la temperatura de brillo (TB) del canal 13 (10,3  $\mu\text{m}$ ) del sensor Advance Baseline Imager (ABI); y del sensor óptico Geostationary Lightning Mapper (GLM) [7] la variable Flash Extent Density (FED), que es el conjunto de diversos grupos de pulsos ópticos, que cada uno representa a un flash individual. Los datos del GLM son provistos en pares de datos latitud/longitud con una resolución temporal de 20 segundos, a lo cual se utilizó la herramienta de GLMTools [8] para proyectar los datos al reticulado del sensor ABI y en acumulados de un minuto. Finalmente estos datos poseen una resolución temporal de 10 minutos (1 minuto para el GLM) y espacial de 2 km, con una cobertura espacial de América y los océanos adyacentes.

El algoritmo de estimación de precipitación en superficie del sensor DPR del satélite GPM constituye el target del modelo. Esta estimación está basada en datos de radar (en microondas activas), pero con un periodo de revisita de 1 o 2 días, con un ancho de sensado de 245 km<sup>2</sup> y una resolución espacial de 5 km. Las estadísticas de validación de este sensor son un error cuadrático medio de 1.83 mm h<sup>-1</sup>, una correlación lineal de Pearson de 0.61 y un sesgo positivo de 0.2 mm h<sup>-1</sup> [9].

### 3 Metodología

#### 3.A Construcción del Conjunto de Datos

Los datos comprenden el periodo de enero 2020 a junio 2021 inclusive, sobre la región de Sudamérica (90°O-30°O | 50°S-10°N). El conjunto de datos para el entrenamiento del modelo consiste en sectores cuadrados de 240 km<sup>2</sup> tomados a lo largo de la pasada del DPR que se superponen en un 40% (para incrementar el tamaño del conjunto de entrenamiento). Por otra parte, se cuenta con estimaciones del RRQPE, la TB y el FED interpoladas a la retícula de los sectores del DPR utilizando interpolación bi-lineal y el dato más cercano en tiempo. Para el FED se realizó un acumulado de 10 minutos anteriores al minuto del DPR.

Las imágenes utilizadas para el entrenamiento son seleccionadas mediante el siguiente criterio: que la imagen cuente con al menos un 20% de los píxeles con PP mayor a 0.1 mm.h<sup>-1</sup> del DPR. Mediante este criterio se obtuvieron un total de 6160 muestras de entrenamiento. Este criterio solo se refiere al RRQPE, debido a que es el objetivo evaluar la inclusión de más variables de entrada.

#### 3.B UNet

El modelo implementado es un modelo de redes neuronales convolucionales con una arquitectura del tipo UNet propuesta en [4]. Para entrenar y evaluar el desempeño de la red se utilizó un conjunto de datos de entrenamiento (90%) y validación (10%) de los datos de enero 2020 a junio 2021; se excluyen los datos de enero y junio de 2021 para el periodo de testeo.

La arquitectura UNet tiene por propósito codificar la información de entrada a un espacio de mínima dimensión, la cual posteriormente es decodificada en conjunto con la información del mismo nivel de codificación. Para el entrenamiento de la red, se utiliza una función de costo de error cuadrático medio, un optimizador ADAM con un batch size de 128 imágenes, con una tasa de aprendizaje de 0.001, y se define un máximo de 1000 épocas de entrenamiento, implementando un método de reducción de la tasa de aprendizaje en función de la optimización de la función de costo para el conjunto de validación.

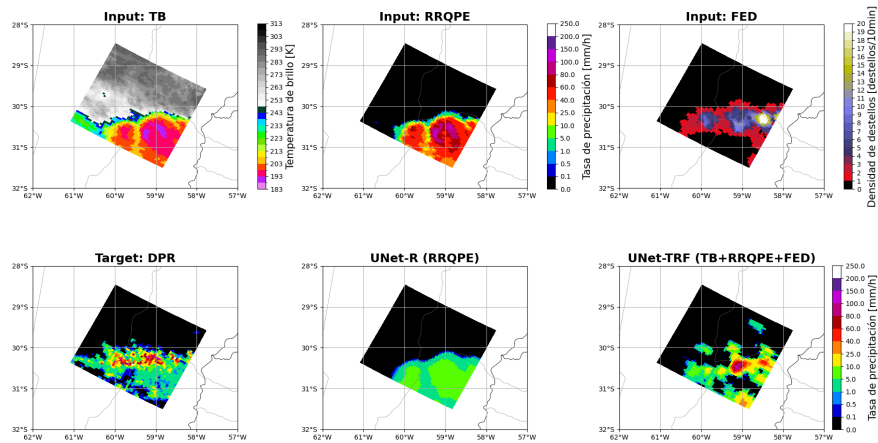
### 4 Resultados y Conclusiones

Se desarrolla un modelo de UNet incluyendo sucesivamente diferentes conjuntos de datos de entrada. En primera instancia se utiliza el RRQPE; posteriormente se incluye la TB y el FED. En la Tabla 1 se presentan las métricas del error cuadrático medio (MSE), el sesgo (BIAS) y la correlación de Pearson (CORR) obtenidas. También se

incluye el Equitable Skill Score (ETS) como métrica derivada de la matriz de confusión para los umbrales de 1 y 5 mm.h<sup>-1</sup>, que mide la habilidad de detección con respecto a un modelo aleatorio.

Estimación	MSE [mm.h <sup>-1</sup> ]	BIAS [mm.h <sup>-1</sup> ]	CORR	ETS - 1 mm.h <sup>-1</sup>	ETS - 5 mm.h <sup>-1</sup>
RRQPE	5.48	0.07	0.20	0.16	0.10
UNet-R	2.29	-0.01	0.24	0.17	0.0
UNet-RTF	2.24	-0.05	0.26	0.17	0.03

**Tabla 1.** Métricas de error cuadrático medio (MSE), el sesgo (BIAS) la correlación de Pearson (CORR), y el Equitable Skill Score (ETS) para 1 y 5 mm para el conjunto de testeo para el modelo implementado.



**Figura 1:** Ejemplo de aplicación del modelo propuesto en la fecha 12/01/2021 06:14 UTC. Los datos de entrada RRQPE, TB y FED, el target el DPR y las salidas de la UNet utilizando todos los datos de entrada y solo utilizando el RRQPE.

Las estimaciones obtenidas por las UNets propuestas mejoran todas las métricas del RRQPE, a excepción del ETS en 5 mmh<sup>-1</sup> que es el producto disponible de manera instantánea en alta resolución temporal. Se observa que la inclusión de la TB y el FED reduce el MSE, aumenta el BIAS negativo y aumenta la correlación. Asimismo, para el ETS de 1 mm.h<sup>-1</sup> no se observan diferencias, pero en 5 mm.h<sup>-1</sup>, se observa una leve mejora a partir de la inclusión de la TB y la FED. Para un trabajo similar aplicado sobre la región de Brasil [10], obtienen métricas superadoras con un MSE, cercano a 1.5 mm.h<sup>-1</sup>, un BIAS positivo de 0.01 y una correlación levemente menor a 0.5. Sin embargo, el target que utilizan combina el DPR (basado en microondas activas) con estimaciones provenientes de microondas pasivas, lo cual podría dar resultados que, aunque numéricamente sean superiores, se basan en una referencia diferente.

En la Figura 1 se muestra un ejemplo de aplicación, en la que se visualizan los datos de entrada, el DPR y las salidas según el conjunto de datos utilizado. Utilizando todas las variables de entrada, la UNet-TRF mejora la posición de

precipitación convectiva intensa, respecto de solo el UNet-R y el RRQPE.

Como trabajo futuro se explorarán funciones de costo que, en conjunto con los diversos datos de entrada, mejoren la detección y representación de precipitaciones intensas que son subrepresentadas por el modelo, en parte debido al uso del MSE como función de costo. Además, se incluirán verificaciones con datos observados de precipitaciones in situ, como estaciones meteorológicas automáticas, y se evaluará el desempeño con respecto a otros productos de ECP del estado del arte.

Esta investigación ha contado con el apoyo parcial del proyecto PREVENIR, ejecutado por la Agencia Japonesa de Cooperación Internacional y la Agencia Japonesa de Ciencia y Tecnología en el marco del Programa de Asociación de Investigación Científica y Tecnológica para el Desarrollo Sostenible.

## 5 Bibliografía

- [1] Arkin, P. A., & Meisner, B. N. (1987). The Relationship between Large-Scale Convective Rainfall and Cold Cloud over the Western Hemisphere during 1982-84. *Monthly Weather Review*, 115(1), 51-74. [https://doi.org/10.1175/1520-0493\(1987\)115<0051:TRBLSC>2.0.CO;2](https://doi.org/10.1175/1520-0493(1987)115<0051:TRBLSC>2.0.CO;2)
- [2] Heuscher, L., et al. (2022). Relationship between lightning, precipitation, and environmental characteristics at mid-/high latitudes from a glm and gpm perspective. *Journal of Geophysical Research: Atmospheres*, 127(13).
- [3] Thiel, K. C., Calhoun, K. M., & Reinhart, A. E. (2023). Forecast Applications of GLM Gridded Products: A Data Fusion Perspective. *Weather and Forecasting*, 38(11), 2253-2270. <https://doi.org/10.1175/WAF-D-23-0078.1>
- [4] Sadeghi, M., Nguyen, P., Hsu, K., & Sorooshian, S. (2020). Improving near real-time precipitation estimation using a U-Net convolutional neural network and geographical information. *Environmental Modelling & Software*, 134, 104856. <https://doi.org/10.1016/j.envsoft.2020.104856>
- [5] GOES-R Advanced Baseline Imager (ABI) Algorithm Theoretical Basis Document For Rainfall Rate (QPE). (2020). Retrieved June 7, 2024, from [https://www.star.nesdis.noaa.gov/goesr/documents/ATBDs/Enterprise/ATBD\\_Enterprise\\_Rainfall\\_Rate\\_v3\\_2020-07-10.pdf](https://www.star.nesdis.noaa.gov/goesr/documents/ATBDs/Enterprise/ATBD_Enterprise_Rainfall_Rate_v3_2020-07-10.pdf)
- [6] Iguchi, T., et al. (n.d.). *GPM/DPR Level-2 Algorithm Theoretical Basis Document*. Retrieved June 7, 2024, from [https://gpm.nasa.gov/sites/default/files/2022-06/ATBD\\_DPR\\_V07A.pdf](https://gpm.nasa.gov/sites/default/files/2022-06/ATBD_DPR_V07A.pdf)
- [7] Goodman, S. J., et al. (2013). The goes-r geostationary lightning mapper (glm). *Atmospheric Research*, 125-126, 34-49.
- [8] CSPP Geo. (2022). *Gridded GLM Software version 1.0*. Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin-Madison. Retrieved from <https://cimss.ssec.wisc.edu/csppgeo/history.html>
- [9] Liao, L., & Meneghini, R. (2022). GPM DPR Retrievals: Algorithm, Evaluation, and Validation. *Remote Sensing*, 14(4), 843. <https://doi.org/10.3390/rs14040843>
- [10] Pfreundschuh, S., Ingemarsson, I., Eriksson, P., Vila, D. A., & Calheiros, A. J. P. (2022). An improved near-real-time precipitation retrieval for Brazil. *Atmospheric Measurement Techniques*, 15(23), 6907-6933. <https://doi.org/10.5194/amt-15-6907-2022>