

## ESTUDIO DE DIFERENTES PARAMETRIZACIONES DE ACTIVIDAD ELÉCTRICA Y SU EVALUACIÓN CON OBSERVACIONES DEL GLM

Federico Cutraro<sup>1</sup>, María Eugenia Dillon<sup>1,2</sup>, Juan Ruiz<sup>2,3,4</sup>

<sup>1</sup>Servicio Meteorológico Nacional

<sup>2</sup>CONICET

<sup>3</sup>Departamento de Ciencias de la Atmósfera y los Océanos, FCEyN, UBA.

<sup>4</sup>Centro de Investigaciones del Mar y la Atmósfera, CONICET/FCEN-UBA

**Autor correspondiente: Federico Cutraro, [fcutraro@smn.gov.ar](mailto:fcutraro@smn.gov.ar)**

Manuscrito recibido el 1° de septiembre de 2024, en su versión final el 6 de diciembre de 2024

### RESUMEN

El pronóstico de actividad eléctrica es de vital importancia para muchas actividades y la seguridad de las personas. Dicho pronóstico no es sencillo debido a que es un proceso que ocurre a escalas que no son resueltas por los modelos meteorológicos. Por este motivo, el presente trabajo evalúa el desempeño de cinco parametrizaciones aplicadas al modelo WRF con convección permitida. Para generar y evaluar las parametrizaciones se utilizaron los pronósticos determinísticos y por ensamble del Sistema de Asimilación y Pronóstico del Servicio Meteorológico Nacional (SAP.SMN) y las observaciones del sensor GLM entre noviembre de 2022 y abril de 2023. Se encontró que todas las parametrizaciones logran representar de manera correcta distintas características de las descargas observadas durante el período analizado como su distribución y ciclo diario, siendo la que se basa en el contenido de hielo integrado en la vertical, la que brinda los mejores resultados. Se observó además una sensibilidad a la parametrización de la microfísica empleada en el modelo, no así con la parametrización de la capa límite.

**Palabras clave:** parametrización, actividad eléctrica, GLM, WRF



## **STUDY OF DIFFERENT LIGHTNING PARAMETERIZATIONS AND THEIR EVALUATION WITH GLM OBSERVATIONS.**

### **ABSTRACT**

A lightning forecast is vital for many activities and people's safety. This forecast is challenging since cloud charging processes occur at scales that numerical models do not resolve. For this reason, the present work evaluates the performance of five parameterizations applied to the WRF model at convective allowed resolution. Deterministic and ensemble forecasts from the Data Assimilation and Numerical Weather Forecasting System of the Argentinian National Meteorological Service (SAP.SMN) and GLM observations from November 2022 to April 2023 were used to generate and evaluate the parametrizations. All parameterizations were found to correctly represent different characteristics of the observed lightning during the analyzed period like their distribution and daily cycle, while the one based on the vertically-integrated ice content provides the best results. The sensitivity of the parameterization to the microphysics and boundary layer schemes is also investigated. The microphysics scheme produces a relatively large impact on the estimation of the lightning densities. The boundary layer scheme produces only a marginal impact.

**Key Words:** parametrization, lightning, GLM, WRF

### **1. INTRODUCCIÓN**

La actividad eléctrica presente en las tormentas supone un elevado riesgo para los humanos representando una de las primeras causas de fatalidades asociadas a fenómenos convectivos, especialmente por la falta de alertas sobre este fenómeno (Ashley y Gilson, 2009). En particular, Argentina es una de las regiones del mundo con mayor actividad eléctrica (Zipser y otros, 2006) generando zonas de alto riesgo para las regiones más densamente pobladas (Nicora y otros, 2013). Por este motivo, es importante tener una previsión de la ocurrencia de la actividad eléctrica. Numerosos trabajos abordaron el pronóstico de la actividad eléctrica en base a observaciones y a muy corto plazo, es decir tiempos menores a unas pocas horas (e.g. Cintineo y otros, 2022). No obstante, no abundan los estudios que analizan pronósticos a

corto y mediano plazo, es decir entre horas y algunos días, para los cuales es necesario realizar un pronóstico de las descargas a partir de simulaciones numéricas. Una opción es la de resolver numéricamente de forma explícita los procesos de carga y la cantidad de descargas resultantes (Mansell y otros, 2002; Mansell y otros, 2005). La principal desventaja de esta metodología es que incluir estos procesos puede resultar en un aumento significativo del costo computacional, aún cuando los procesos de carga se representen de una manera simplificada (Fierro y otros, 2013). Otra opción es la de parametrizar la tasa de descargas eléctricas en función de las variables que son explícitamente simuladas por los modelos. En este caso, existen en la literatura una gran diversidad de variables que fueron relacionadas con la actividad eléctrica y que pueden ser utilizadas en dichas parametrizaciones (Price y Rind 1992; McCaul y otros, 2009; entre otros). Dado que la mayoría son expresiones matemáticas deducidas empíricamente que estiman una tasa de descargas, estas podrían no ser óptimas en diferentes regiones o bajo diferentes configuraciones del modelo numérico (e.j. diferente resolución espacial, diferente forma de parametrizar los procesos microfísicos, entre otros). Por esto es importante realizar validaciones de las mismas antes de utilizarlas en una determinada región y realizar las correcciones necesarias para ajustarlas a las observaciones locales.

Las observaciones de descargas eléctricas se pueden obtener mediante redes de superficie o desde sensores montados en satélites (Nag y otros, 2015). Estos últimos, suelen estimar la actividad eléctrica a partir de cambios en la luminosidad del tope de las nubes. Los primeros sensores de este tipo fueron el Optical Transient Detector (OTD) y el Lightning Imaging Sensor (LIS) y estaban montados en satélites de órbita baja, por lo que medían durante poco tiempo una región específica del planeta. Recientemente, con la inclusión de sensores de descargas eléctricas en satélites geoestacionarios, como es el caso del Geostationary Lightning Mapper (GLM) montado en el GOES-16 (Goodman y otros, 2013), la cantidad de observaciones creció enormemente, lo cual consiste en una oportunidad única para la evaluación y optimización de los pronósticos de descargas eléctricas.

El presente trabajo busca avanzar en la evaluación de distintas parametrizaciones de la actividad eléctrica aplicadas al modelo WRF (Skamarock y otros, 2019) para diferentes configuraciones del mismo, comparándolas con observaciones del sensor GLM. Para ello se emplearon simulaciones con distintas combinaciones de esquemas de microfísica y capa límite provistos por los pronósticos determinísticos y por ensamble del Sistema de

Asimilación y Pronóstico Numérico del Servicio Meteorológico Nacional argentino (SAP.SMN).

En la sección 2 se describen los datos, parametrizaciones de la actividad eléctrica utilizadas y la metodología para comparar los pronósticos con las observaciones. En la sección 3 se muestran los resultados y por último, en la sección 4, las conclusiones.

## **2. DATOS Y METODOLOGÍA**

### **2.1 Observaciones**

Las observaciones de actividad eléctrica utilizadas corresponden a las realizadas por el sensor Geostationary Lightning Mapper (GLM) montado en el satélite GOES-16. Este sensor mide la radiación en el infrarrojo cercano que es emitida por el tope de la nubosidad y asocia los cambios de intensidad a la actividad eléctrica que se produce en las tormentas, con una resolución temporal de 2 ms y espacial de 8 km en el nadir y 14 km hacia los bordes del campo de visión, y con una eficiencia de detección cercana al 90% (Goodman y otros, 2013). El dato básico medido por este sensor es denominado evento y se corresponde con la detección de un cambio significativo de luminosidad en el campo de visión, pero debido a la extensión que puede alcanzar una misma descarga, ésta puede generar varios eventos por lo que se aplica un algoritmo de clusterización (Mach y otros, 2007). Este algoritmo permite agrupar eventos que se encuentran cercanos en tiempo y espacio en lo que se denomina un flash, que sería el equivalente a las descargas eléctricas observadas. Cada uno de los flashes tiene asociada una ubicación geográfica (latitud y longitud que resulta del promedio de los eventos que lo conforman, pesados por su intensidad).

En este trabajo se utilizó la librería de Python glmtools (Bruning y otros, 2019) para procesar los datos del GLM. Dicha librería permite estimar la variable Flash Extent Density (FED) que consiste en un conteo de la cantidad de flashes detectados en puntos de malla de una dada retícula, acumulados en un período de tiempo determinado. Dicha detección considera la extensión horizontal de los flashes, es decir, que un mismo flash puede ser contabilizado simultáneamente en más de un punto de retícula (Figura 1). La FED fue obtenida en una retícula con una resolución horizontal de 10 km que es similar a la resolución nativa del sensor. El intervalo de tiempo en que se acumularon las descargas es de 5 minutos centrados en la hora en que se tienen los pronósticos.

## 2.2 Pronósticos numéricos y período de estudio

Los pronósticos utilizados provienen del SAP.SMN operativo, el cual usa el modelo WRF v.4.0 con una resolución horizontal de 4 km y 45 niveles sigma-p en la vertical con el tope en 10 hPa, y posee una realización determinística y una por ensambles (Matsudo y otros, 2022). La ejecución determinística consiste en pronósticos a 72 horas utilizando condiciones iniciales y de contorno provistas por la simulación de alta resolución (0.25° de resolución horizontal) del modelo GFS (Zhou y otros, 2017) y emplea los siguientes esquemas: WSM6 (Hong y Lim, 2006) para la microfísica; MYJ (Janjic, 1994) para la capa límite planetaria; RRTMG para la radiación de onda larga y onda corta (Iacono y otros, 2008); Noah-MP para el suelo (Niu y otros, 2011). El pronóstico por ensambles está constituido por 20 miembros cuyas condiciones iniciales y de contorno se construyen sumando al GFS de alta resolución las perturbaciones de cada miembro del Global Ensemble Forecast System (GEFS) de 0,5° de resolución horizontal hasta un plazo de pronóstico de 48 horas. Cada miembro del ensamble, además de tener condiciones iniciales y de borde diferentes, emplea distintas combinaciones de esquemas de microfísica y capa límite (Tabla I). Además de las empleadas para la realización determinística se añaden el esquema de la microfísica Lin (Chen y Sun, 2002) y los de capa límite SH (Shin y Hong, 2015) y YSU (Hong y otros, 2006). Notar que la realización determinística y los miembros 1, 2, 3 y 10 del ensamble difieren sólo en las condiciones iniciales y de borde, ya que usan la misma configuración de esquemas. Además, todas las simulaciones resuelven la convección de manera explícita aunque con el objetivo de que los pronósticos representen procesos de una escala similar a la de las observaciones, se realizó una interpolación bilineal para transformar la retícula del modelo (de mayor resolución) a la del GLM (de 10 km).

Diariamente se ejecutan 4 conjuntos de pronósticos inicializados a las 00, 06, 12 y 18 UTC. Este trabajo se focaliza en el desempeño de los pronósticos a corto plazo, dado que son los más relevantes para aplicaciones tales como el pronóstico de eventos de alto impacto o la asimilación de datos, para lo cual para cada inicialización se consideran los plazos de pronóstico de 6 a 11 horas. Las primeras 6 horas de pronóstico no se evalúan debido a que se considera que están afectadas por el proceso de spin-up de la simulación, durante el cual procesos fundamentales para la predicción de la actividad eléctrica, como el desarrollo de nubes convectivas profundas, podrían no estar adecuadamente representados. Por otra parte,

si bien estudiar plazos mayores de pronóstico resultaría de interés, añade la dificultad de que la convección en el modelo puede estar mal representada, complejizando los análisis a realizar, escapándose del objetivo de este trabajo. Al considerar solamente 6 horas de cada una de las cuatro inicializaciones diarias del SAP.SMN se cuenta entonces con una serie continua de pronósticos horarios, para ser comparados con las observaciones de GLM centradas alrededor de cada hora, tal como se representa en la Figura 2.

En cuanto al dominio analizado, dado que el pronóstico del SAP.SMN ocupa regiones en las que climatológicamente la actividad eléctrica es escasa (Bertone y otros 2022), se optó por reducir la región de estudio a donde ésta ocurre con mayor regularidad, indicada en la Figura 3. Dicha región se subdividió en cuatro cuadrantes que corresponden a regiones en donde la forma en la que ocurre la convección húmeda profunda y su climatología son diferentes: NO (noroeste), NE (noreste), SO (suroeste) y SE (sudeste) (Rasmussen y otros, 2016). De esta manera se busca evaluar si dichas diferencias afectan el desempeño de las parametrizaciones de actividad eléctrica. Si bien el dominio analizado incluye parte de los bordes del dominio de ejecución del modelo WRF, se verificó que los resultados no son afectados por la inclusión de estos (no se muestra).

El período de estudio fue entre el 1 de noviembre de 2022 y el 3 de abril de 2023, el cual se corresponde con la estación cálida y es cuando se produce la mayor actividad convectiva en esta región (Vidal, 2014). Durante este período se contó en total con pronósticos para 132 días. Multiplicando este valor por las 6 horas de pronóstico almacenadas de cada una de las 4 inicializaciones del SAP.SMN, se obtuvieron 3168 horas de cada simulación (la determinística y los 20 miembros del ensamble).

## **2.3 Parametrizaciones de la actividad eléctrica**

### **2.3.1 Predictores**

A partir de distintas variables pronosticadas por los modelos meteorológicos que en la literatura fueron relacionadas con la actividad eléctrica, se evaluaron los siguientes predictores:

- $W_{\max} [\frac{m}{s}]$ : la velocidad máxima vertical en cada columna de la retícula del modelo fue utilizada por Price y Rind (1992) y en una gran cantidad de trabajos que buscan

pronosticar observaciones relacionadas con la actividad eléctrica. Se basa en que observacionalmente se encontró una relación entre la velocidad vertical y la tasa de descargas que se corresponde con que las ascendentes más intensas están vinculadas a la convección y la separación de cargas dentro de las nubes.

- **ICE\_INT**, [ $\frac{kg}{m^2}$ ]: el contenido de hielo integrado verticalmente en cada columna de la retícula del modelo, es una variable utilizada por Petersen y otros (2005) quienes encontraron una alta correlación entre esta variable y la densidad de descargas eléctricas estimadas a partir del sensor LIS a bordo del satélite TRMM durante un período de 3 años. En particular, para la región subtropical de Sudamérica, los autores notaron que para una dada densidad de descargas, la cantidad de hielo observada es mayor que en otras regiones del mundo.
- **WQ<sub>g</sub>** [ $\frac{mg}{s kg}$ ]: el flujo vertical de graupel en la isoterma de  $-15^{\circ}C$  fue utilizado por McCaul y otros (2009) basándose en que el flujo vertical de partículas grandes de hielo en temperaturas a las que éstas conviven con el agua sobreenfriada (factor clave en la electrificación de las tormentas) está relacionado con la densidad de descargas eléctricas.
- **LPI** [ $\frac{J}{kg}$ ]: el Lightning Potential Index es un índice desarrollado por Yair y otros (2010) que relaciona la velocidad vertical con el contenido de hidrometeoros entre las isotermas de  $0$  y  $-20^{\circ}C$ , que es donde es máximo el proceso de separación de cargas que conducen a la electrificación de las tormentas, y busca identificar las regiones donde es más probable encontrar las descargas eléctricas.

### 2.3.2 Regresión cuantil-cuantil

Las descargas observadas por el sensor GLM no pueden ser directamente comparadas con los pronósticos de las variables utilizadas como predictores debido a que los errores en la ubicación de las tormentas en los pronósticos dificultarían el análisis. Por este motivo, se empleó la metodología de regresión cuantil-cuantil para transformar la distribución de estas variables a la de las observaciones (Combarous y otros, 2022). Al realizar esta comparación se asume que existe una relación monótona creciente entre las variables pronosticadas y observadas, es decir que, por ejemplo, a mayor cantidad de hielo integrado en la vertical, la

tasa de rayos es mayor. Esta suposición es consistente con las hipótesis subyacentes a la definición de las parametrizaciones. Al considerar la correspondencia entre la distribución del predictor y del predictando (FED), las relaciones entre ambos se establecen independientemente de que exista una coincidencia espacial entre pronósticos y observaciones. Esto reduce el impacto de los errores de los pronósticos en términos de posicionamiento o tiempo de ocurrencia de las celdas convectivas en la relación que se obtiene entre el predictor y el predictando. Por otro lado, la regresión cuantil-cuantil permite reducir el impacto de los errores sistemáticos en la intensidad de la convección o en la frecuencia de ocurrencia de la misma.

Previo a realizar la regresión, se determinó el cuantil correspondiente a 1 descarga por punto de retícula en 5 minutos ( $f_{umb}$ ) que es el mínimo valor de FED considerado en este trabajo ( $min_o$ ). Luego se tomó el valor del predictor correspondiente a dicho cuantil en la distribución del predictor ( $min_p$ ). Este valor se considera como el valor del predictor a partir del cual se espera la ocurrencia de descargas. De esta manera, se busca reducir los sesgos entre las tasas de descargas estimadas a partir del modelo y las observadas a partir de GLM. En la Figura 4 se muestra como ejemplo dicho cálculo considerando los pronósticos de la ejecución determinística del predictor ICE\_INT. En la distribución de probabilidad de la FED, el valor de una descarga cada 5 minutos ( $min_o$ ) se corresponde con el cuantil 0,9882 ( $f_{umb}$ ). El valor de ICE\_INT asociado a este cuantil es de  $1,18 \frac{kg}{m^2}$  ( $min_p$ ), por lo que no es de esperarse actividad eléctrica por debajo de este umbral. El mismo procedimiento se utilizó para cada miembro del ensamble y para cada predictor.

Considerando las observaciones mayores a  $min_o$  y los predictores mayores a  $min_p$ , se calcularon las funciones de distribución acumulada (CDF de su sigla en inglés) dividiendo el espacio de probabilidades en  $10^4$  intervalos. El gran número de intervalos utilizados permite un buen muestreo de la CDF en los valores de FED más extremos. A partir de las distribuciones pronosticadas y observadas se generó la curva cuantil-cuantil la cual se ajustó con una función lineal a trozos que es la que se utilizará posteriormente para realizar la predicción de la tasa de descargas.

Para la evaluación de las diferentes parametrizaciones de descargas se separó el conjunto de datos disponible en una parte de entrenamiento, la cual se usó para realizar la regresión cuantil-cuantil entre los predictores y las densidades de descargas observadas, y otra de validación que se usó para evaluar el desempeño de las parametrizaciones. De esta manera,

las parametrizaciones se evalúan en un período independiente del que se usó para generarlas. La primera parte contó con los datos entre el 1 de noviembre de 2022 y el 14 de febrero de 2023 y la segunda entre el 15 de febrero y el 3 de abril de 2023. La fecha que divide ambos períodos se definió de manera tal que para el entrenamiento se consideren dos tercios del total de los datos y para la validación el tercio restante.

### 2.3.3 Parametrizaciones propuestas

Considerando los predictores mencionados en la sección 2.3.1 y aplicando la regresión cuantil-cuantil para transformar sus distribuciones a la de FED observada (sección 2.3.2), se generaron las parametrizaciones  $FED_{w}$ ,  $FED_{ICE}$ ,  $FED_{WQ}$  y  $FED_{LPI}$  que utilizan los predictores  $W_{max}$ ,  $ICE\_INT$ ,  $WQ_g$  y  $LPI$  respectivamente. Además, se generó la parametrización  $FED_{McCaul}$  que realiza una combinación de  $FED_{ICE}$  y  $FED_{WQ}$  a partir de la ecuación 1, siguiendo lo propuesto por McCaul y otros (2009).

$$FED_{McCaul} = \alpha * FED_{ICE} + (1 - \alpha) FED_{WQ} \quad (1)$$

Esta última parametrización en particular, tiene además el coeficiente  $\alpha$  que determina la mejor combinación de las dos parametrizaciones que emplea. En el trabajo de McCaul y otros (2009), este valor fue definido en 0,05; es decir, se le da mucho más peso a la estimación realizada por el flujo de graupel que a la del contenido de hielo integrado. Para evaluar cuál es el valor que ofrece los mejores resultados en este trabajo, se comparó en el período de validación la regresión obtenida (del período de entrenamiento) con las observaciones, empleando valores de  $\alpha$  variando entre 0,05 y 0,95 cada 0,05; para determinar cuál es el valor que minimiza las diferencias según la ecuación (2)

$$\sum_i \frac{(o_i - f_i)^2}{o_i} \quad (2)$$

donde  $o_i$  es el i-ésimo cuantil de las observaciones y  $f_i$  el del pronóstico. El valor de  $\alpha$  que minimizó las diferencias fue 0,4; es decir, que se le está dando un peso casi similar a  $FED_{ICE}$  y  $FED_{WQ}$ . La diferencia en el valor de  $\alpha$  respecto del hallado por McCaul y otros (2009) puede deberse a que tanto la formulación de las parametrizaciones  $FED_{ICE}$  y  $FED_{WQ}$  como la metodología para realizar la estimación difieren con lo realizado en este trabajo: por ejemplo Mc Caul y otros (2009) utilizan un ajuste lineal en base al máximo de las variables pronosticadas en casos de estudio, mientras que aquí se propone una relación cuantil-cuantil.

A su vez, la aplicación es en una región distinta, lo cual también puede influir. En los resultados que se presentan en la próxima sección, el valor de 0,4 es el utilizado para el coeficiente  $\alpha$  de  $FED_{McCaul}$ .

A modo de resumen, en la Figura 5 se presenta un diagrama de flujo de los datos y pasos involucrados en los procesos descritos en la sección 2.

### 3. RESULTADOS

En la sección 3.1 se estudiará la sensibilidad de las parametrizaciones de la actividad eléctrica al dominio considerado y a la configuración de esquemas empleada en el modelo, utilizando el período de entrenamiento. Luego se estimará el desempeño de dichas parametrizaciones durante el período de validación (sección 3.2) y por último se las evaluará para un caso de estudio (sección 3.3).

#### 3.1 Sensibilidad de las parametrizaciones

El primer elemento que se analizó es cuán diferente es cada relación en función de la región del dominio estudiada. Para esto se calcularon las relaciones entre las variables pronosticadas por el modelo y las observaciones en el período de entrenamiento para el dominio total (Figura 3), dividiéndolo en los cuadrantes noroeste (NO), noreste (NE), sudoeste (SO), y sudeste (SE) y también dividiéndolo en si la superficie es tierra o agua. Esta última distinción se hace debido a la gran cantidad de estudios que encontraron una clara diferencia en la actividad eléctrica que ocurre sobre el mar y sobre tierra (e.g. Price y Rind 1992; Williams y Stanfill 2002). En la Figura 6 se muestran las relaciones halladas en cada subdominio para la simulación determinística del SAP.SMN. Se puede ver que las mayores diferencias se encuentran para los valores más bajos de descargas, mientras que para valores mayores a 10 descargas en 5 minutos las diferencias son mínimas.  $W_{max}$  presenta la mayor similitud entre los subdominios, en comparación con las otras parametrizaciones.

Centrándose en los valores más bajos de actividad eléctrica, para ICE\_INT,  $WQ_g$  y LPI se puede ver que existe una clara distinción entre la parte norte y sur del dominio: los cuadrantes NO y NE presentan mayores descargas para un mismo valor de la variable pronosticada que los cuadrantes SO y SE. Lo mismo se observa separando por tipo de superficie: sobre tierra

los umbrales son más bajos que sobre el océano, coincidiendo con los trabajos previos que estudiaron el efecto de la superficie en la actividad eléctrica.

Las diferencias halladas para las diferentes regiones podrían ser relevantes dependiendo del estudio o aplicación que se quiera llevar a cabo. En particular, en lo que resta del trabajo se considerará el dominio completo para simplificar el análisis, y se deja para un trabajo futuro profundizar en los resultados por regiones.

El segundo elemento que se analizó es el impacto de los distintos esquemas utilizados en el modelo WRF en las parametrizaciones de la actividad eléctrica evaluadas. Para ello, se llevó adelante la regresión cuantil-cuantil entre la FED observada y los diferentes predictores en base al pronóstico determinístico y a cada uno de los miembros del ensamble del SAP.SMN. En la Figura 7, el primer rasgo que se nota es que las curvas se agrupan en función del esquema de la microfísica que se utilizó en el modelo, especialmente para los valores bajos de tasa de descargas eléctricas, más allá de que se haya tratado cada miembro del ensamble de manera particular. Esto indica que la variabilidad introducida por las diferentes configuraciones de esquemas es mayor que la generada por las distintas condiciones iniciales de cada simulación. La excepción a este comportamiento es la velocidad máxima vertical, lo cual está indicando que esta variable es menos sensible a las dos parametrizaciones de microfísica utilizadas. En el caso del contenido de hielo integrado en la vertical, las diferencias son notorias incluso para los valores más altos. Los miembros que utilizan el esquema de Lin requieren menor cantidad de hielo para generar una igual tasa de descargas que los que utilizan el esquema WSM6. Por ejemplo, para producir una tasa de 10 descargas en 5 minutos los miembros que emplean el esquema de Lin requieren aproximadamente  $3 \frac{kg}{m^2}$ , mientras que los que emplean WSM6 requieren aproximadamente  $6.3 \frac{kg}{m^2}$ .

A partir de la Figura 7 también se puede apreciar cuál es el valor mínimo del predictor a partir del cual se asume la presencia de actividad eléctrica en el modelo. Analizando los valores hallados, los mismos son coherentes con las teorías de fondo que respaldan su uso en las distintas parametrizaciones. La presencia de hielo y graupel son claves en el proceso no inductivo de carga de las nubes y se ve que los umbrales para las variables que las incluyen están en valores cercanos a cero. En los miembros que usan el esquema WSM6 los umbrales son de aproximadamente  $1,3 \frac{kg}{m^2}$  y  $4 \times 10^{-4} \frac{mg}{s kg}$  y en los que usan Lin de  $0,5 \frac{kg}{m^2}$  y  $2 \times 10^{-3} \frac{mg}{s kg}$  para los predictores ICE\_INT y WQ<sub>g</sub> respectivamente. Una conclusión similar se puede obtener

para el LPI: esta variable al delimitar zonas en que existe el potencial para que se produzca la electrificación tiene un umbral muy cercano a cero ( $1 \times 10^{-2} \frac{J}{kg}$  en los miembros que usan el esquema WSM6 y  $6 \times 10^{-2} \frac{J}{kg}$  en los que usan Lin). Para la velocidad máxima vertical, en cambio, el valor mínimo para la ocurrencia de descargas eléctricas tiene un valor de  $0.9 \frac{m}{s}$ , que se podría considerar bajo, dada la intensidad típica de esta variable en una nube convectiva, y como se verá más adelante, puede generar problemas en regiones con topografía.

En base a las diferencias halladas relacionadas especialmente al esquema de microfísica empleado en el modelo, se realizó el test de Kolmogorov-Smirnov para evaluar si cada una de las distribuciones obtenidas son estadísticamente diferentes entre sí (Wilks, 2011). El test resaltó lo que se evidenciaba a partir de los gráficos, las distribuciones son significativamente diferentes cuando lo que varía es el esquema de microfísica (excepto para  $W_{max}$ ), pero en general no cuando lo hace el esquema de capa límite. Dado que no se puede considerar que las distribuciones son iguales, se optó por utilizar una relación cuantil-cuantil específica para la simulación determinística y cada uno de los 20 miembros del ensamble. Es necesario tener en cuenta que esta solución tiene la problemática de que ante la utilización de un nuevo esquema o un cambio de modelo, se requiere evaluar si las relaciones halladas previamente pueden seguir siendo aplicadas o hay que hallar una nueva relación entre la tasa de descargas y las variables pronosticadas.

### **3.2 Evaluación de las parametrizaciones**

Para evaluar las relaciones halladas se utilizaron los datos del período de validación, con el fin de que sean independientes de los usados para generarlas. En particular, se mostrarán solo los resultados para la simulación determinística del SAP.SMN, ya que las conclusiones que se obtienen de analizar los diferentes miembros del ensamble son en líneas generales equivalentes. Sin embargo se mencionarán en el texto detalles relevantes que se hayan observado en el ensamble.

Como la metodología empleada busca transformar la distribución de las variables pronosticadas en la de las observadas, se comenzó por comparar las distribuciones de las descargas observadas y pronosticadas por las distintas parametrizaciones (Figura 8). Se puede

## Artículo en edición

ver, en el panel izquierdo, que para todas las parametrizaciones, excepto  $FED_{McCaul}$ , el valor correspondiente a los cuantiles de la densidad de descarga pronosticada está cerca del correspondiente a los cuantiles de la densidad de descargas observadas (curva de relación 1:1). No obstante, se observa una sobrestimación para los valores más bajos de la distribución con errores en torno al 20% (panel derecho, Figura 8). Para valores mayores, las parametrizaciones presentan diferentes comportamientos,  $FED_W$  sobrestima los valores,  $FED_{ICE}$  y  $FED_{WQ}$  los subestima y  $FED_{LPI}$  es la que mejor ajusta la distribución. En todos los casos los errores son menores al 10%. En el caso de  $FED_{McCaul}$ , se observa que la misma subestima las densidades de descarga con respecto a los valores observados, con errores que superan el 20% para la mayor parte de la distribución. En base a estos resultados se puede decir que la metodología aplicada logra transformar de manera correcta la distribución de las variables pronosticadas en la observada. El próximo paso es evaluar la manera en que las parametrizaciones representan distintos aspectos de la variabilidad espacial y temporal de la densidad de descargas.

Como fue mencionado previamente, no puede realizarse una comparación punto a punto entre las observaciones y pronósticos debido a que los errores en el pronóstico de la ubicación e intensidad de la convección podrían conducir a resultados incorrectos. Por lo tanto, se evaluó el desempeño global tomando medidas integrales que consideren a todo el dominio analizado, y así determinar la habilidad de cada una de las parametrizaciones. Para ello, se consideró la distribución conjunta de la cantidad total de descargas observadas y pronosticadas en todo el dominio (Figura 9). Se puede observar cómo para todas las parametrizaciones existe una gran dispersión que aumenta al incrementarse la cantidad de descargas acumuladas, lo cual se asocia a los eventos más extremos que son más difíciles de pronosticar. Entre todas las parametrizaciones,  $FED_{ICE}$  es la que presenta los mejores resultados, mostrando un mayor coeficiente de correlación lineal. Un aspecto que se destaca es que la parametrización  $FED_{WQ}$  tiende a subestimar la ocurrencia de eventos con una cantidad de descargas por encima de 7500 descargas, probablemente debido a su formulación que está más centrada al pronóstico de las regiones convectivas más intensas y que no logra representar de manera correcta el área total afectada por la actividad eléctrica (McCaul y otros, 2009). Este mismo sesgo se observa en la parametrización de McCaul, pero disminuído, ya que al incluir  $FED_{ICE}$  también en su formulación mejora su desempeño. Este mismo análisis se realizó en los 4 cuadrantes previamente utilizados (Figuras no mostradas). Se encontró que excepto para el cuadrante NE donde  $FED_{ICE}$  es la parametrización con un

mejor desempeño, en los demás lo es  $FED_{LPI}$ . También se analizó el desempeño según el miembro del ensamble, en el que se observó que para los que usan la parametrización WSM6 la mejor parametrización resulta  $FED_{ICE}$ , mientras que para los que usan Lin es  $FED_{LPI}$  (Figuras no mostradas).

Otro elemento que se analizó fue cómo cada parametrización logra representar el ciclo diario de descargas. Para ello se utilizó nuevamente el dato de descargas acumuladas en todo el dominio para cada hora de pronóstico y se realizó el promedio según la hora del día (Figura 10). El rasgo más notorio que se puede ver es que el máximo de descargas pronosticadas ocurre más temprano que el máximo observado, aunque esto podría estar más relacionado a una deficiencia en el modelo en representar la convección que a las parametrizaciones de la actividad eléctrica en sí, ya que se observa en mayor o menor medida en todas por igual. Por otro lado, pareciera haber 2 grupos de parametrizaciones que se comportan de manera similar. Un primer grupo se encuentra conformado por  $FED_W$  y  $FED_{LPI}$  y un segundo grupo por  $FED_{ICE}$ ,  $FED_{WQ}$  y  $FED_{McCaul}$ . El primer grupo tiende claramente a subestimar la cantidad de descargas durante las primeras horas del día pero representa de mejor manera el valor mínimo observado. En cuanto al máximo, éste se encuentra sobrestimado y adelantado 3 horas. El segundo grupo, representa mejor las descargas de las primeras horas pero no tiene un marcado mínimo como las observaciones. El máximo en cambio está mejor pronosticado respecto al primer grupo, tanto en cantidad de descargas como en tiempo ya que el adelantamiento es de solo 2 horas. Esta diferencia podría estar relacionada a que las parametrizaciones del primer grupo, en las que la velocidad vertical tiene una gran influencia, pueden llegar a presentar valores más altos durante las etapas iniciales del desarrollo de la convección con la aparición de las primeras ascendentes intensas. En cambio las del segundo grupo, basadas principalmente en el contenido de hidrometeoros, requieren que la convección esté más desarrollada para presentar una mayor concentración de partículas de hielo, retrasando así el máximo de actividad eléctrica. Esto podría explicar también por qué las parametrizaciones del segundo grupo pronostican mayor cantidad de descargas eléctricas durante la noche, ya que la presencia de sistemas convectivos de mesoescala que exhiben un alto contenido de hielo sobre una gran área alcanzan su máxima extensión en estas horas para la región de estudio (Salio y otros, 2007).

Este mismo desfase en la hora del máximo se observa analizando los cuatro cuadrantes en que se dividió el dominio completo y para los distintos miembros del ensamble, con la excepción de que para la parametrización  $FED_{LPI}$  en los miembros que usan el esquema de

microfísica Lin el horario el máximo se encuentra adelantado 2 horas en vez de 3 (figuras no mostradas).

### 3.3 Aplicación a un caso de estudio

Por último se analizó un caso de estudio fuera del período de validación y entrenamiento, en el que el pronóstico tuvo un buen correlato con las observaciones, para analizar el desempeño de las distintas parametrizaciones. En la Figura 11 se puede ver el pronóstico obtenido con cada parametrización y la observación para el 27 de octubre de 2023 a las 19 hora local (21 UTC) empleando la ejecución determinística del SAP.SMN inicializada a las 00 UTC y utilizando el dominio completo del mismo, que cubre toda la Argentina, para así también evaluar el desempeño de las parametrizaciones por fuera de donde fueron validadas. El plazo de pronóstico mostrado se seleccionó luego de una comparación cualitativa entre la simulación y las observaciones, con el fin de asegurarse una representación razonable de la convección por parte del modelo (no mostrado). Si bien las 5 parametrizaciones pronostican descargas en las regiones donde la convección húmeda profunda está presente en la simulación, algunas predicen la presencia de descargas en regiones donde no se pronostica convección profunda. Por ejemplo, todas las parametrizaciones pronostican descargas en el sur del océano Pacífico asociadas a convección de poco desarrollo vertical para las cuales no se observó ninguna descarga. Mismo es el caso para el pronóstico que emplea  $W_{\max}$  como predictor, en la región de la cordillera de los Andes: el bajo umbral encontrado para considerar la presencia de descargas hace que éstas se observen en toda su extensión, por los movimientos verticales asociados a la orografía de la región.

Para profundizar el análisis en la región con convección, se focalizó en el área central del país, donde tanto en el pronóstico como en la observación se encontraron valores altos de descargas (Figura 12). En este caso se puede hacer nuevamente la diferenciación entre las parametrizaciones  $FED_W$  y  $FED_{LPI}$  respecto de  $FED_{ICE}$ ,  $FED_{WQ}$  y  $FED_{McCaul}$ . Las primeras, al estar ligadas a la velocidad vertical, pronostican más descargas en las regiones donde los movimientos verticales presentan valores más altos, especialmente para  $FED_W$  en la que se notan claramente los frentes de ráfagas, y una distribución poco homogénea. Las segundas, como dependen del contenido de hidrometeoros presentan una estructura más homogénea y con mayor extensión areal, que se asemeja más a las observaciones. Esto se nota principalmente en  $FED_{ICE}$  ya que  $FED_{WQ}$  al incluir la velocidad vertical en el cálculo presenta

un aspecto más similar al de  $FED_W$  y  $FED_{LPI}$ , pero sí colabora en el pronóstico de máximos que  $FED_{ICE}$  no presenta, al combinarse en la parametrización  $FED_{McCaul}$ .

Por último, y para dar una idea de los valores pronosticados por cada parametrización, se cuantificó el máximo observado y pronosticado de FED. Para este caso, el máximo observado se encuentra en las 85 descargas en 5 minutos y  $FED_{ICE}$  fue la que mejor lo estimó con un máximo de 90 descargas. Las demás parametrizaciones presentaron una clara sobrestimación con 116, 138, 153 y 157 descargas para  $FED_{McCaul}$ ,  $FED_{Wmax}$ ,  $FED_{WQ}$  y  $FED_{LPI}$  respectivamente.

#### **4. CONCLUSIONES**

Con el objetivo de contar con un pronóstico de la actividad eléctrica empleando el modelo WRF, se evaluó el desempeño y la sensibilidad a la configuración de éste utilizando cinco parametrizaciones. Las mismas fueron calibradas para transformar la distribución de variables pronosticadas por el modelo WRF, y utilizadas como predictores de la actividad eléctrica, en la de la densidad de descargas eléctricas observada por el sensor GLM montado en el satélite GOES-16. Se pudo ver que las parametrizaciones generadas logran ajustar de manera correcta las distribuciones aún utilizando un set de datos independiente, mostrando su capacidad de generalizar y poder ser utilizadas para el pronóstico. Se encontró cierta diferencia entre las parametrizaciones que dependen de la velocidad vertical y de las que están influenciadas principalmente por el contenido de hidrometeoros. Por un lado,  $FED_W$  y  $FED_{LPI}$ , que dependen de la velocidad vertical, tienden a generar campos de descargas espacialmente poco homogéneos. Estas parametrizaciones presentan además problemas en las regiones con topografía compleja (esto podría asociarse a que no se incluyeron regiones de montaña en la generación de las transformaciones ni criterios adicionales como podría ser la extensión vertical de los ascensos para el caso de  $FED_W$ ) y tener la hora del máximo de actividad eléctrica más adelantado respecto de las observaciones. El otro conjunto de parametrizaciones, es decir  $FED_{ICE}$ ,  $FED_{WQ}$  y  $FED_{McCaul}$ , más influenciados por el contenido de hidrometeoros lograron un mejor desempeño al presentar distribuciones de densidad de descargas espacialmente más homogéneas y lograr representar mejor el ciclo diario de la actividad eléctrica.

En particular,  $FED_{ICE}$  es la que tuvo el mejor desempeño de las 5 parametrizaciones y la que se plantea como mejor opción para uso en futuras aplicaciones, entre las que se encuentran el

pronóstico de las descargas para su uso en las oficinas de pronóstico y la asimilación de datos de descargas con el objetivo de mejorar los pronósticos a muy corto plazo generados por el modelo. Cabe destacar que el costo computacional de calcular las parametrizaciones es bajo y accesible.

No obstante, la metodología planteada para generar las parametrizaciones presenta algunas limitaciones. En primer lugar se encontró que las parametrizaciones generadas son sensibles tanto a la región en que se emplea como a la configuración de esquemas utilizados en el modelo. Por lo tanto, se reduce tanto la capacidad de emplearlas en otro modelo o en el mismo pero utilizando una configuración distinta como de emplearlas en otras regiones. Además, en el presente trabajo no se abordó la cuantificación de la incertidumbre de la estimación que ofrecen las parametrizaciones y que es un desafío todavía presente. Se propone abordar estas limitaciones en trabajos futuros con el objetivo de reducir su impacto, ya sea al mejorar las parametrizaciones presentadas o evaluando el desempeño de otras parametrizaciones. Por otra parte en futuros estudios es necesario abordar aspectos centrales para la predicción, tales como el desempeño de los pronósticos en función del plazo de pronóstico y en función de los diferentes modos en los que se organiza la convección húmeda profunda.

## **5. AGRADECIMIENTOS**

Los autores quieren agradecer los proyectos PICT-2020-SERIEA-I-A de la Agencia Nacional de Promoción Científica y Tecnológica y PIP Nro 11220210100747 del CONICET, al Servicio Meteorológico Nacional por proveer los pronósticos y recursos computacionales utilizados para el desarrollo del trabajo, al proyecto PREVENIR SATREPS financiado por JICA y JST, a Martín Rugna por su ayuda en el uso de los datos del sensor GLM y a los/as revisores/as que con sus comentarios ayudaron a mejorar el trabajo.

## **6. REFERENCIAS**

Allen, B. J., E. R. Mansell, D. C. Dowell, y W. Deierling, 2016: Assimilation of Pseudo-GLM Data Using the Ensemble Kalman Filter. *Mon. Wea. Rev.*, 144, 3465–3486, <https://doi.org/10.1175/MWR-D-16-0117.1>.

Ashley, W. S., y C. W. Gilson, 2009: A Reassessment of U.S. Lightning Mortality. *Bull. Amer. Meteor. Soc.*, 90, 1501–1518, <https://doi.org/10.1175/2009BAMS2765.1>.

Bertone, F., Nicora, G., y Vidal, L., 2022: Thunderstorm days over Argentina: Integration between human observations of thunder and the world wide lightning location network lightning data, *International Journal of Climatology*, 42(16), 9072-9087, <https://doi.org/10.1002/joc.7800>

Bruning, E. C., Tillier, C. E., Edgington, S. F., Rudlosky, S. D., Zajic, J., Gravelle, C., Foster, M., Calhoun, K., Campbell, A., Stano, G., Schultz, C., y Meyer, T. C., 2019: Meteorological imagery for the geostationary lightning mapper, *Journal of Geophysical Research: Atmospheres*, 124(24), 14285-14309, <https://doi.org/10.1029/2019JD030874>

Chen, S.-H. y W.-Y. Sun, 2002: A one-dimensional time dependent cloud model. *J. Meteor. Soc. Japan.*, 80(1), 99–118, <https://doi.org/10.2151/jmsj.80.99>

Cintineo, J. L., M. J. Pavolonis, y J. M. Sieglaff, 2022: ProbSevere LightningCast: A Deep-Learning Model for Satellite-Based Lightning Nowcasting. *Wea. Forecasting*, 37, 1239–1257, <https://doi.org/10.1175/WAF-D-22-0019.1>.

Combarous, P., Erdmann, F., Caumont, O., Defer, É., y Martet, M. 2022: A satellite lightning observation operator for storm-scale numerical weather prediction. *Natural Hazards and Earth System Sciences*, 22(9), 2943-2962, <https://doi.org/10.5194/nhess-22-2943-2022>

Federico, S., Torcasio, R.C., Puca, S., Vulpiani, G., Comellas Prat, A., Dietrich, S., Avolio, E., 2021: Impact of Radar Reflectivity and Lightning Data Assimilation on the Rainfall Forecast and Predictability of a Summer Convective Thunderstorm in Southern Italy. *Atmosphere*, 12, 958. <https://doi.org/10.3390/atmos12080958>

Fierro, A. O., Mansell, E. R., MacGorman, D. R., y Ziegler, C. L., 2013: The implementation of an explicit charging and discharge lightning scheme within the WRF-ARW model:

Benchmark simulations of a continental squall line, a tropical cyclone, and a winter storm. *Monthly Weather Review*, 141(7), 2390-2415, <https://doi.org/10.1175/MWR-D-12-00278.1>

Goodman, S, Blakeslee, R., Koshak, W., Mach, D., Bailey, J., Buechler, D., Carey, L., Schultz, C., Bateman, M., McCaul, E. y Stano, G., 2013: The GOES-R Geostationary Lightning Mapper (GLM). *Atmos. Res.*, 125–126, 34–49, <https://doi.org/10.1016/j.atmosres.2013.01.006>.

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.

Hong, Song–You, Yign Noh y Jimy Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, 134, 2318–2341. <https://doi.org/10.1175/MWR3199.1>

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, y W. D. Collins, 2008: Radiative forcing by long–lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res.*, 113, D13103. <https://doi.org/10.1029/2008JD009944>

Janjic, Zavisla I., 1994: The Step–Mountain Eta Coordinate Model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.*, 122, 927–945. [https://doi.org/10.1175/1520-0493\(1994\)122%3c0927:TSMECM%3e2.0.CO;2](https://doi.org/10.1175/1520-0493(1994)122%3c0927:TSMECM%3e2.0.CO;2)

Mach, D. M., Christian, H. J., Blakeslee, R. J., Boccippio, D. J., Goodman, S. J., y Boeck, W. L., 2007: Performance assessment of the optical transient detector and lightning imaging sensor. *Journal of Geophysical Research: Atmospheres*, 112(D9), <https://doi.org/10.1029/2006JD007787>

Mansell, E. R., MacGorman, D. R., Ziegler, C. L., y Straka, J. M., 2002: Simulated three-dimensional branched lightning in a numerical thunderstorm model, *J. Geophys. Res.*, 107( D9), <https://doi.org/10.1029/2000JD000244>.

Mansell, E. R., MacGorman, D. R., Ziegler, C. L., y Straka, J. M., 2005: Charge structure and lightning sensitivity in a simulated multicell thunderstorm. *Journal of Geophysical Research: Atmospheres*, 110(D12), <https://doi.org/10.1029/2004JD005287>

Matsudo, C., García Skabar, Y., Righetti, S., Cutraro, F., Sacco, M., Dillon, M. E., Alvarez Imaz, M., Maldonado, P. y Salles, A., 2022: Sistema de pronóstico numérico del Servicio Meteorológico Nacional: componente operativa. XIV Congreso Argentino de Meteorología, <http://hdl.handle.net/20.500.12160/2372>

McCaul, E. W., Goodman, S. J., LaCasse, K. M., y Cecil, D. J., 2009: Forecasting lightning threat using cloud-resolving model simulations. *Weather and Forecasting*, 24(3), 709-729, <https://doi.org/10.1175/2008WAF2222152.1>

Nag, A, Murphy, MJ, Schulz, W, y Cummins, KL, 2015: Lightning locating systems: Insights on characteristics and validation techniques. *Earth and Space Science*, 2, 65–93. <https://doi.org/10.1002/2014EA000051>.

M. G. Nicora, E. J. Quel, R. E. Bürgesser y E. E. Ávila, 2013: "Electric atmospheric activity in Argentina, a study for estimating the annual death rate by lightning," 2013 International Symposium on Lightning Protection (XII SIPDA), Belo Horizonte, Brazil, pp. 340-344, <https://doi.org/10.1109/SIPDA.2013.6729213>.

Niu, Guo–Yue, Zong–Liang Yang, Kenneth E. Mitchell, Fei Chen, Michael B. Ek, Michael Barlage, Anil Kumar, Kevin Manning, Dev Niyogi, Enrique Rosero, Mukul Tewari y Youlong Xia, 2011: The community Noah land surface model with multiparameterization options (Noah–MP): 1. Model description and evaluation with local–scale measurements. *J. Geophys. Res.*, 116, D12109. <https://doi.org/10.1029/2010JD015139>

Petersen, W. A., Christian, H. J., y Rutledge, S. A., 2005: TRMM observations of the global relationship between ice water content and lightning. *Geophysical research letters*, 32(14), <https://doi.org/10.1029/2005GL023236>

Price, C., y Rind, D., 1992: A simple lightning parameterization for calculating global lightning distributions. *Journal of Geophysical Research: Atmospheres*, 97(D9), 9919-9933, <https://doi.org/10.1029/92JD00719>

Rasmussen, K. L., M. M. Chaplin, M. D. Zuluaga, y R. A. Houze , 2016: Contribution of Extreme Convective Storms to Rainfall in South America. *J. Hydrometeor.*, 17, 353–367, <https://doi.org/10.1175/JHM-D-15-0067.1>.

Salio, P., Nicolini, M. y Zipser, E.J., 2007: Mesoscale Convective Systems over Southeastern South America and Their Relationship with the South American Low-Level Jet. *Mon. Wea. Rev.*, 135, 1290–1309, <https://doi.org/10.1175/MWR3305.1>.

Shin, H. H., y S.-Y. Hong, 2015: Representation of the subgrid-scale turbulent transport in convective boundary layers at gray-zone resolutions. *Mon. Wea. Rev.*, 143, 250-271. <https://doi.org/10.1175/MWR-D-14-00116.1>

Skamarock, C., Klemp, B., Dudhia, J., Gill, O., Liu, Z., Berner, J., Wang, W., Powers, G., Duda, G., Barker, D.M., y Huang, X., 2019: A Description of the Advanced Research WRF Model Version 4, <http://dx.doi.org/10.5065/1dfh-6p97>

Sun, J., Xue, M., Wilson, J., Ballard, S., Onvlee, J., Joe, P., Barker, M., Li, P., Golding, B., Xu, M., Pinto, J., Sun, Wilson y Xu, D., 2013: Use of NWP for Nowcasting Convective Precipitation: Recent Progress and Challenges. *Bulletin of the American Meteorological Society*. 95. 409-426. <https://doi.org/10.1175/BAMS-D-11-00263.1>.

Vidal, Luciano. 2014: Convección extrema sobre Sudamérica : estructura interna, ciclos de vida e influencia de la topografía en la iniciación. Tesis Doctoral, Universidad de Buenos Aires. Facultad de Ciencias Exactas y Naturales, [https://hdl.handle.net/20.500.12110/tesis\\_n5573\\_Vidal](https://hdl.handle.net/20.500.12110/tesis_n5573_Vidal)

Wilks, D. S., 2011: *Statistical methods in the atmospheric sciences*. Elsevier Academic Press.

Williams E. y Stanfill S., 2002: The physical origin of the land–ocean contrast in lightning activity, *Comptes Rendus Physique*, Volume 3, Issue 10, Pages 1277-1292, ISSN 1631-0705, [https://doi.org/10.1016/S1631-0705\(02\)01407-X](https://doi.org/10.1016/S1631-0705(02)01407-X).

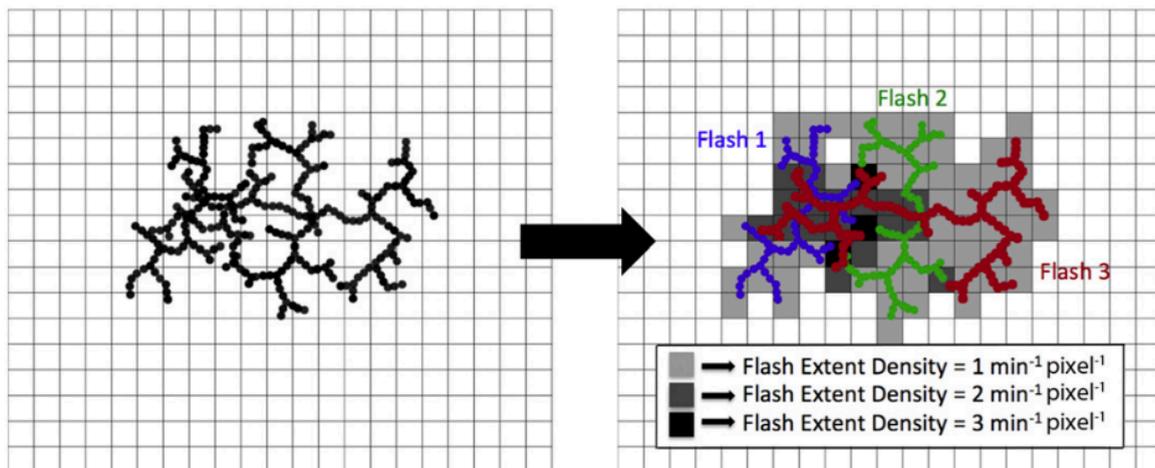
Yair, Y., B. Lynn, C. Price, V. Kotroni, K. Lagouvardos, E. Morin, A. Mugnai, y M. d. C. Llasat, 2010: Predicting the potential for lightning activity in Mediterranean storms based on the Weather Research and Forecasting (WRF) model dynamic and microphysical fields, *J. Geophys. Res.*, 115, D04205, <https://doi.org/10.1029/2008JD010868>.

Zhou, X., Y. Zhu, D. Hou, Y. Luo, J. Peng, y R. Wobus, 2017: Performance of the New NCEP Global Ensemble Forecast System in a Parallel Experiment. *Wea. Forecasting*, 32, 1989–2004, <https://doi.org/10.1175/WAF-D-17-0023.1>.

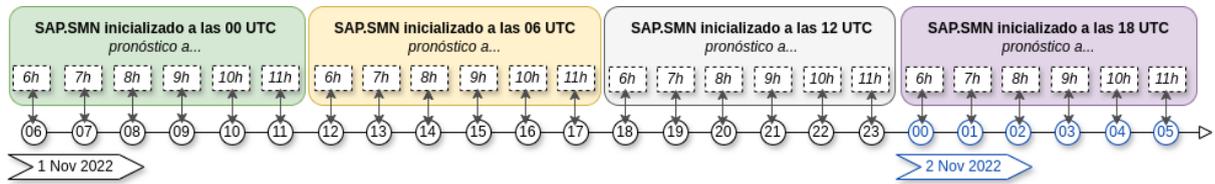
Zipser, E. J., D. J. Cecil, C. Liu, S. W. Nesbitt, and D. P. Yorty, 2006: Where are the most intense thunderstorms on earth?. *Bull. Amer. Meteor. Soc.*, 87, 1057–1072, <https://doi.org/10.1175/BAMS-87-8-1057>.

**Tabla I:** Combinación de las configuraciones físicas en el ensamble para cada miembro del 1 al 20.

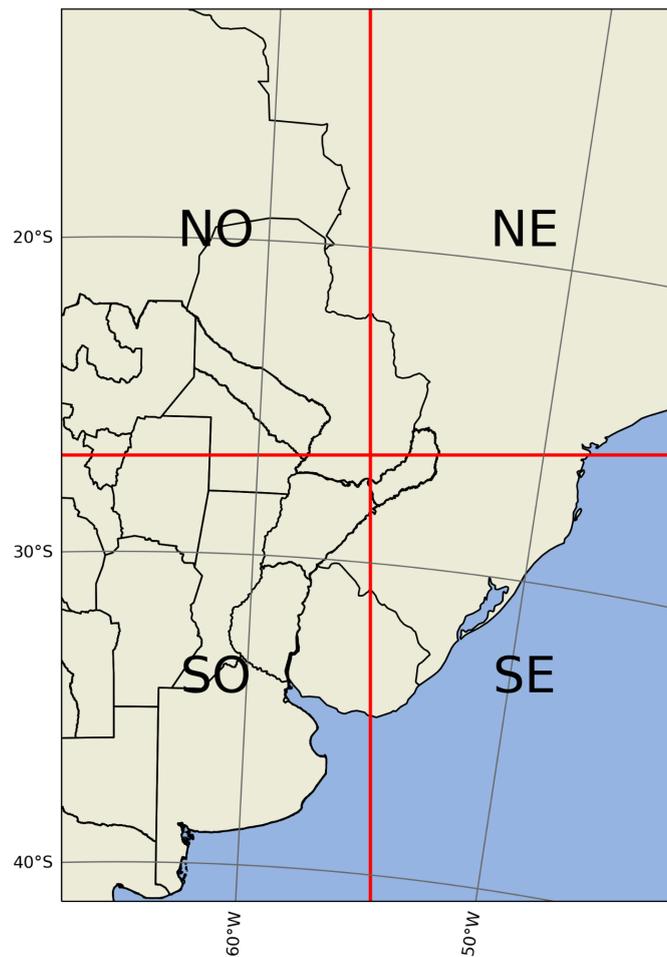
		Esquema de la microfísica	
		WSM6	Lin
Esquema de la capa límite	MYJ	1, 2, 3, 10	11, 12, 13
	SH	4, 5, 6	14, 15, 16
	YSU	7, 8, 9	17, 18, 19, 20



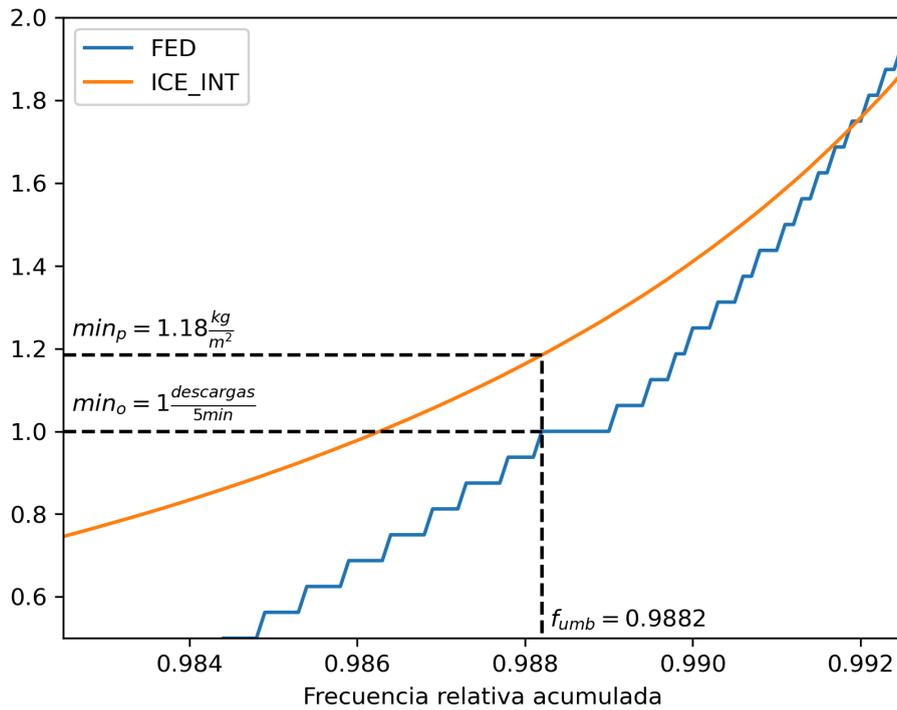
**Figura 1:** Esquema para ejemplificar la relación entre los eventos detectados, los flashes y la variable Flash Extent Density. En la figura de la izquierda se muestran todos los eventos detectados por el GLM en un período de un minuto. En la figura de la derecha, el algoritmo de clusterización asoció cada uno de los eventos a un flash distinto y a partir de esto se determinó la variable FED contando la cantidad de flashes que tienen un evento en cada punto de retícula. Figura extraída de Allen y otros (2016) © American Meteorological Society. Utilizada con permiso.



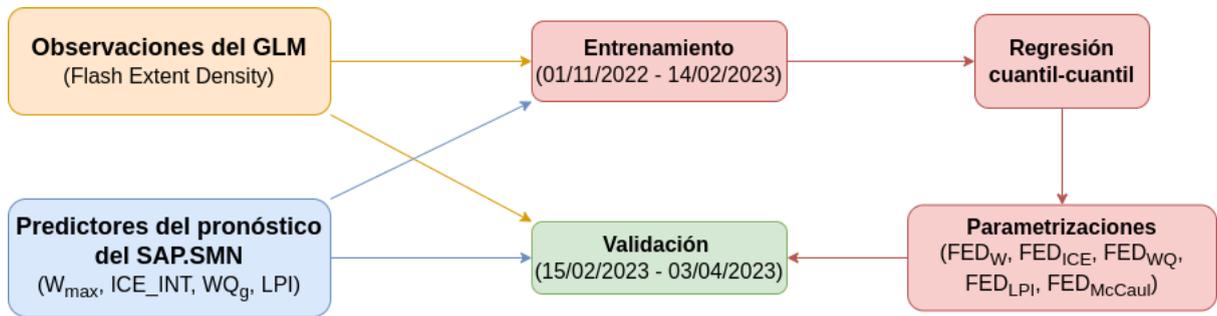
**Figura 2:** Esquema que representa la serie continua de pronósticos obtenida a partir de las distintas inicializaciones del SAP.SMN, mediante la cual se obtienen los valores horarios que son usados para las parametrizaciones de FED. Se usa de ejemplo el 1 de noviembre de 2022, día en que inicia el período de entrenamiento considerado. Las horas representadas en los círculos son en UTC.



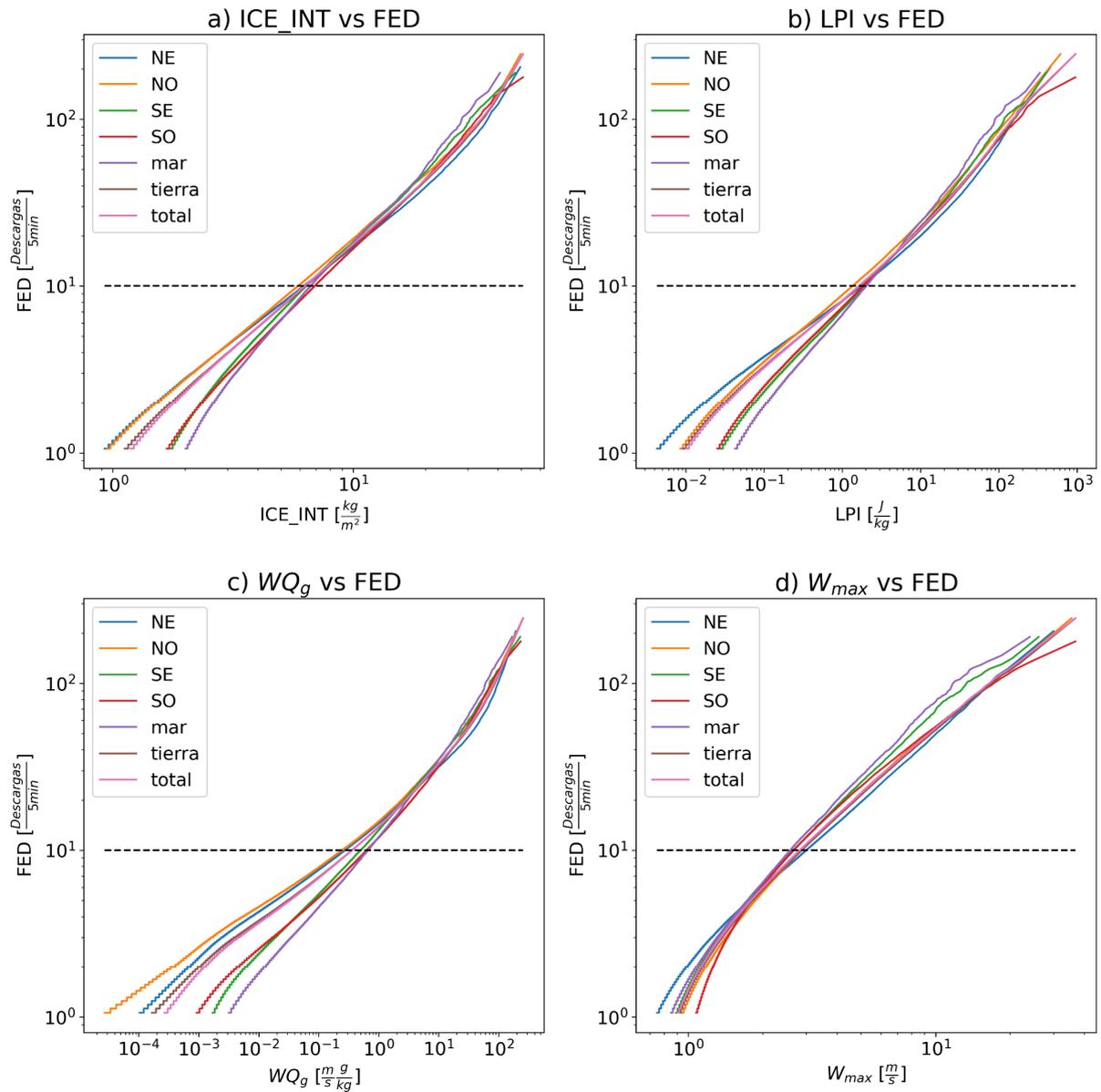
**Figura 3:** Dominio de los pronósticos utilizados en este trabajo. Cada uno de los recuadros rojos delimita las diferentes regiones consideradas, donde en el centro de cada uno se indica el nombre que se utilizará para referenciarlos



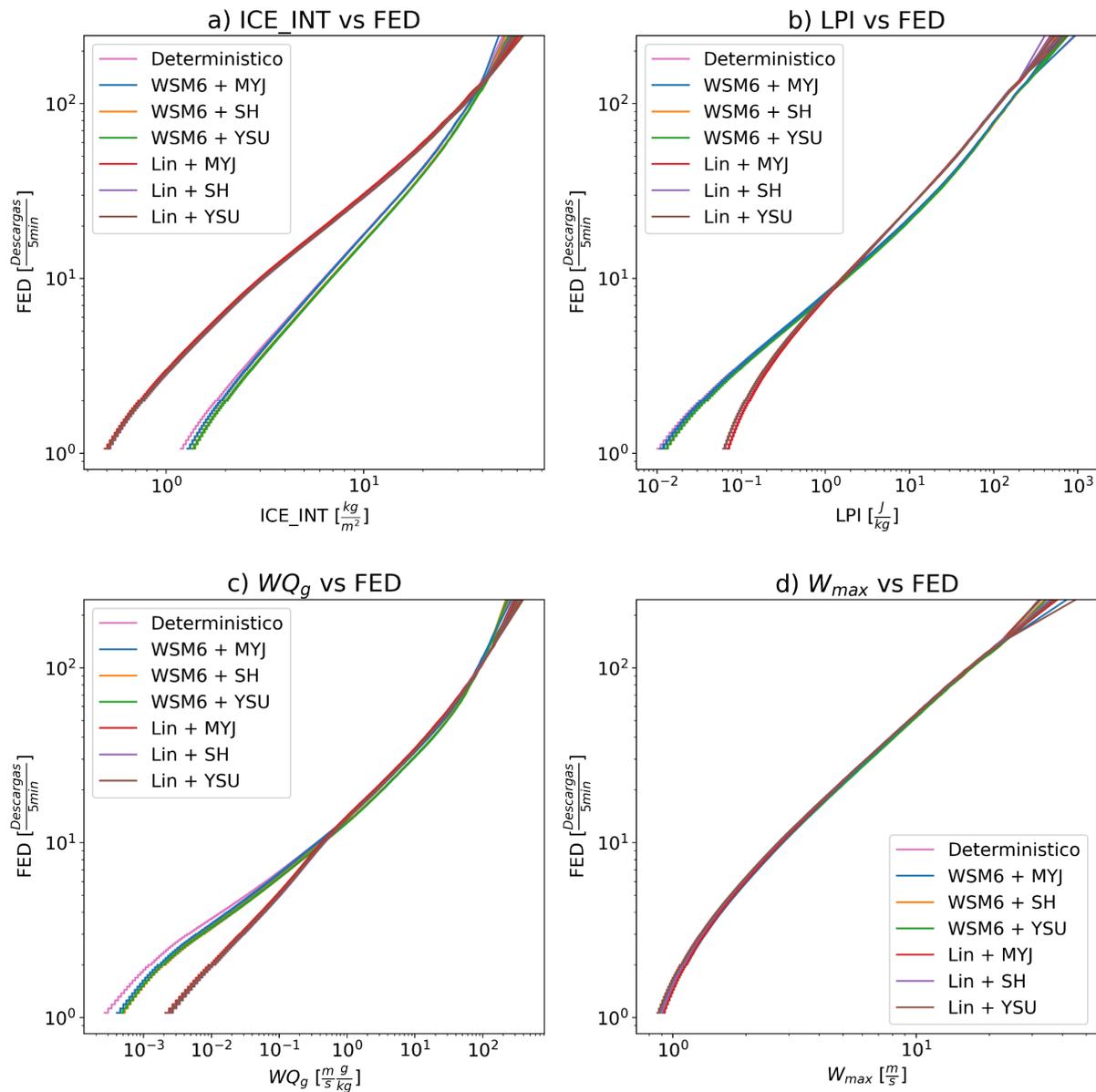
**Figura 4:** Cálculo del valor mínimo de los predictores ( $min_p$ ) a partir del cuantil  $f_{umb}$  considerado para la aparición de descargas ( $min_o$ , una descarga cada 5 minutos) tomando como ejemplo los pronósticos de la ejecución determinística de la variable ICE\_INT.



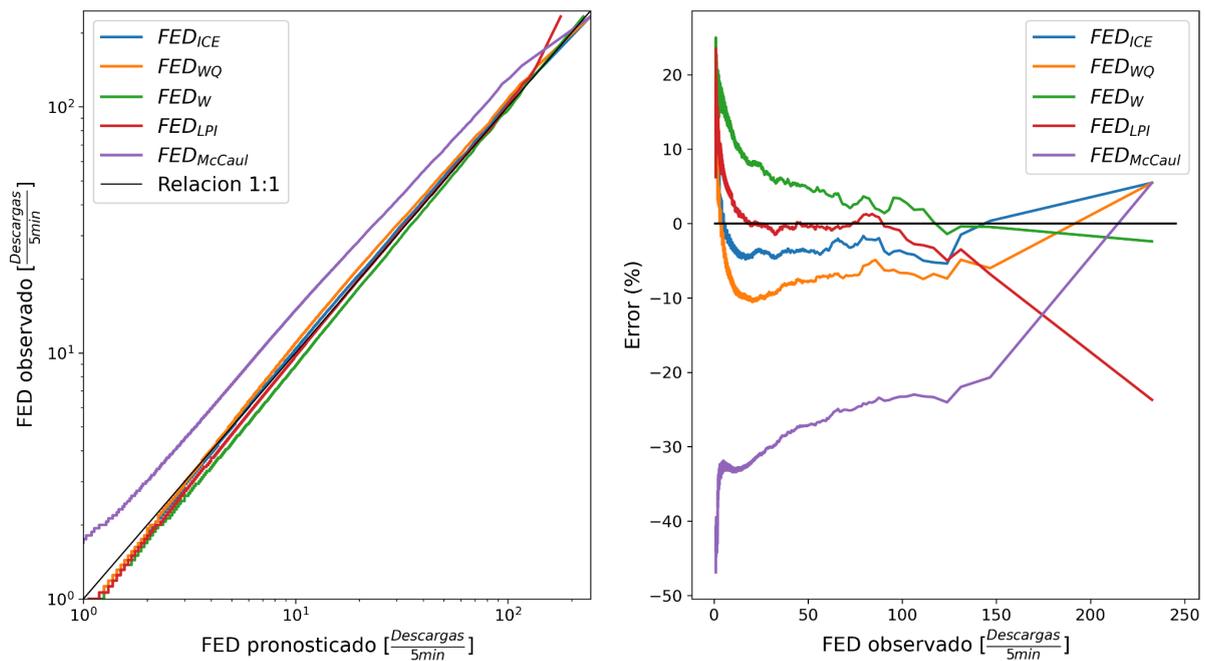
**Figura 5:** Diagrama de flujo de los datos y pasos realizados en el trabajo, explicados en la sección 2.



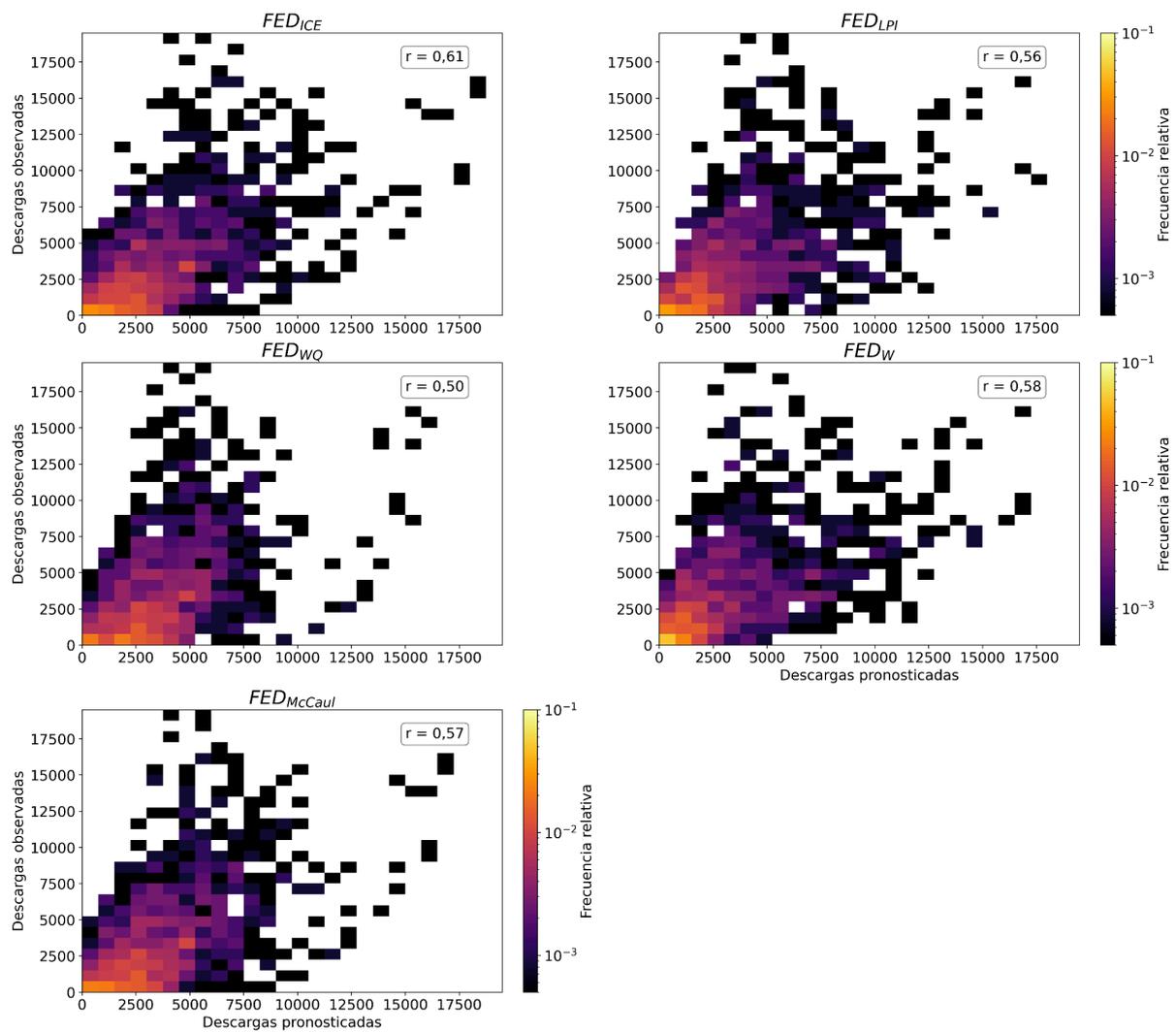
**Figura 6:** Relación entre la distribución de la variable observada Flash Extent Density y las distribución de las variables pronosticadas a) ICE\_INT, b)  $WQ_g$ , c) LPI y d)  $W_{\text{max}}$  para los distintos subdominios analizados en el período de entrenamiento. Se añade la línea de 10 descargas en 5 minutos como referencia.



**Figura 7:** Relación entre la distribución de la variable observada Flash Extent Density y las distribuciones de las variables pronosticadas a) ICE\_INT, b) WQ<sub>g</sub> c) LPI y d) W<sub>max</sub>, considerando el dominio completo. Se encuentran graficadas las relaciones para la ejecución determinística y los 20 miembros de la ejecución por ensamble del SAP.SMN, donde el color de cada línea representa la configuración de esquemas utilizados..

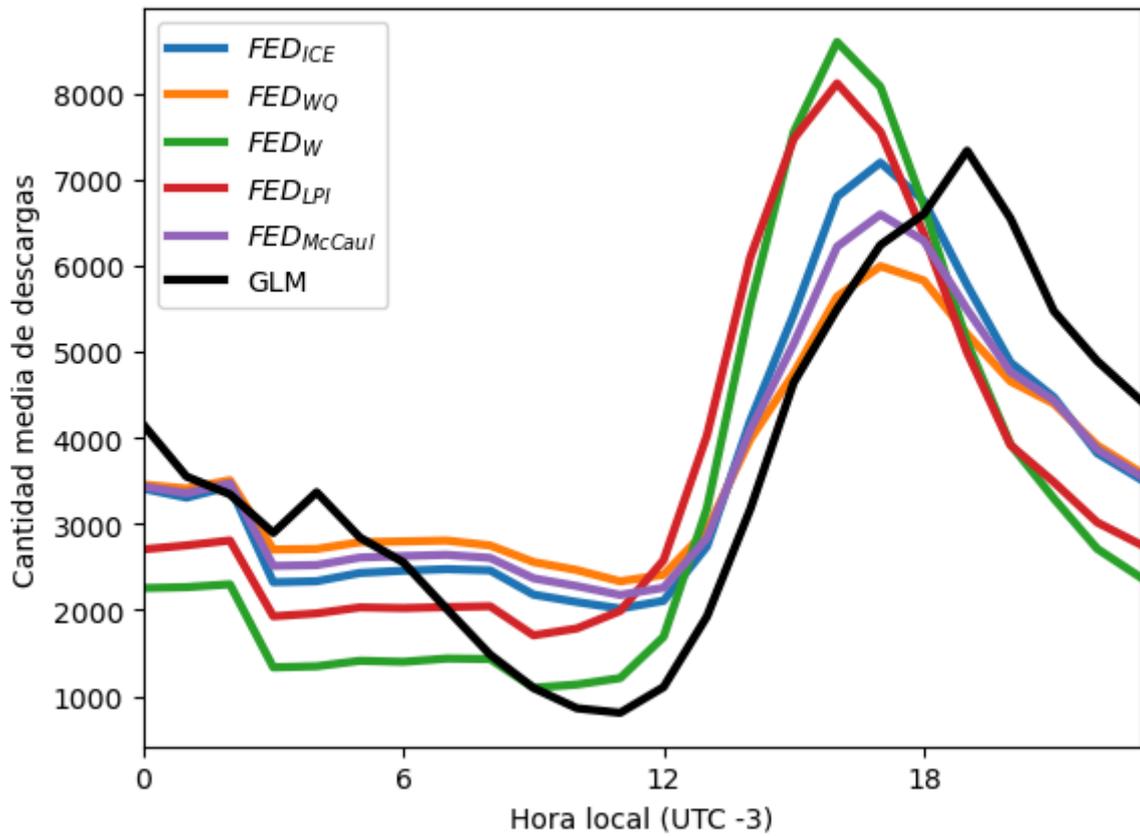


**Figura 8:** Relación entre la distribución de la Flash Extent Density observada y la distribución de descargas pronosticadas por cada parametrización en el período de validación, considerando todo el dominio, utilizando la ejecución determinística del SAP.SMN. En el panel izquierdo se muestra el diagrama cuantil-cuantil entre la distribución observada y pronosticada y en el de la derecha el porcentaje de error cometido por la distribución pronosticada.

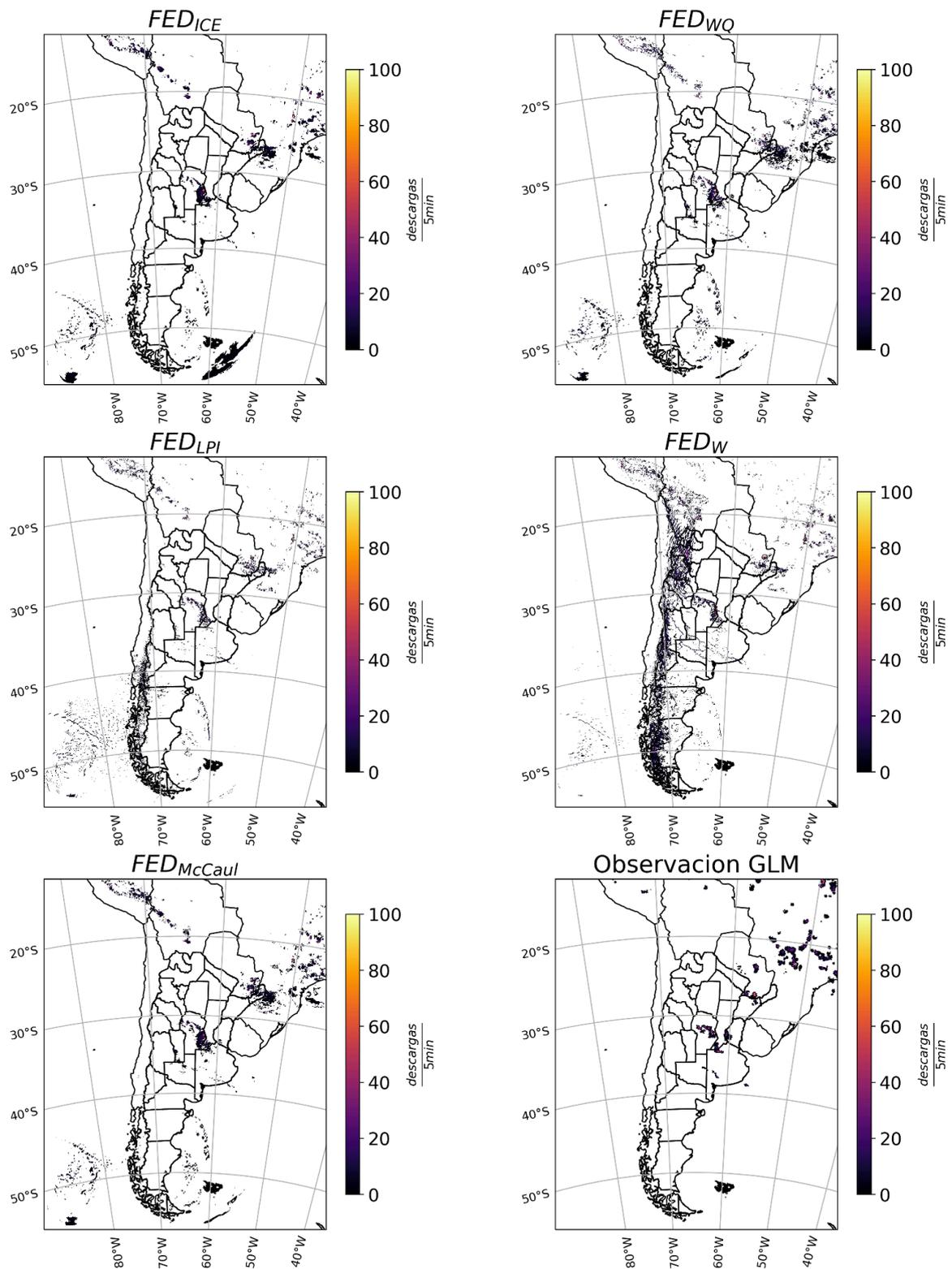


**Figura 9:** Distribución conjunta entre el acumulado horario de descargas pronosticadas por cada parametrización y observadas, considerando todo el dominio para la ejecución determinística del SAP.SMN. En el borde superior derecho de la figura de cada parametrización se muestra el coeficiente de correlación lineal.

Ciclo diario de descargas



**Figura 10:** Ciclo diario medio de descargas acumuladas en todo el dominio observado (GLM) y pronosticado por cada parametrización, utilizando la ejecución determinística del SAP.SMN.



**Figura 11:** Pronóstico de actividad eléctrica de las parametrizaciones y observación del GLM para el día 27 de octubre de 2023 a las 19 hora local en todo el dominio del SAP.SMN, utilizando su ejecución determinística.

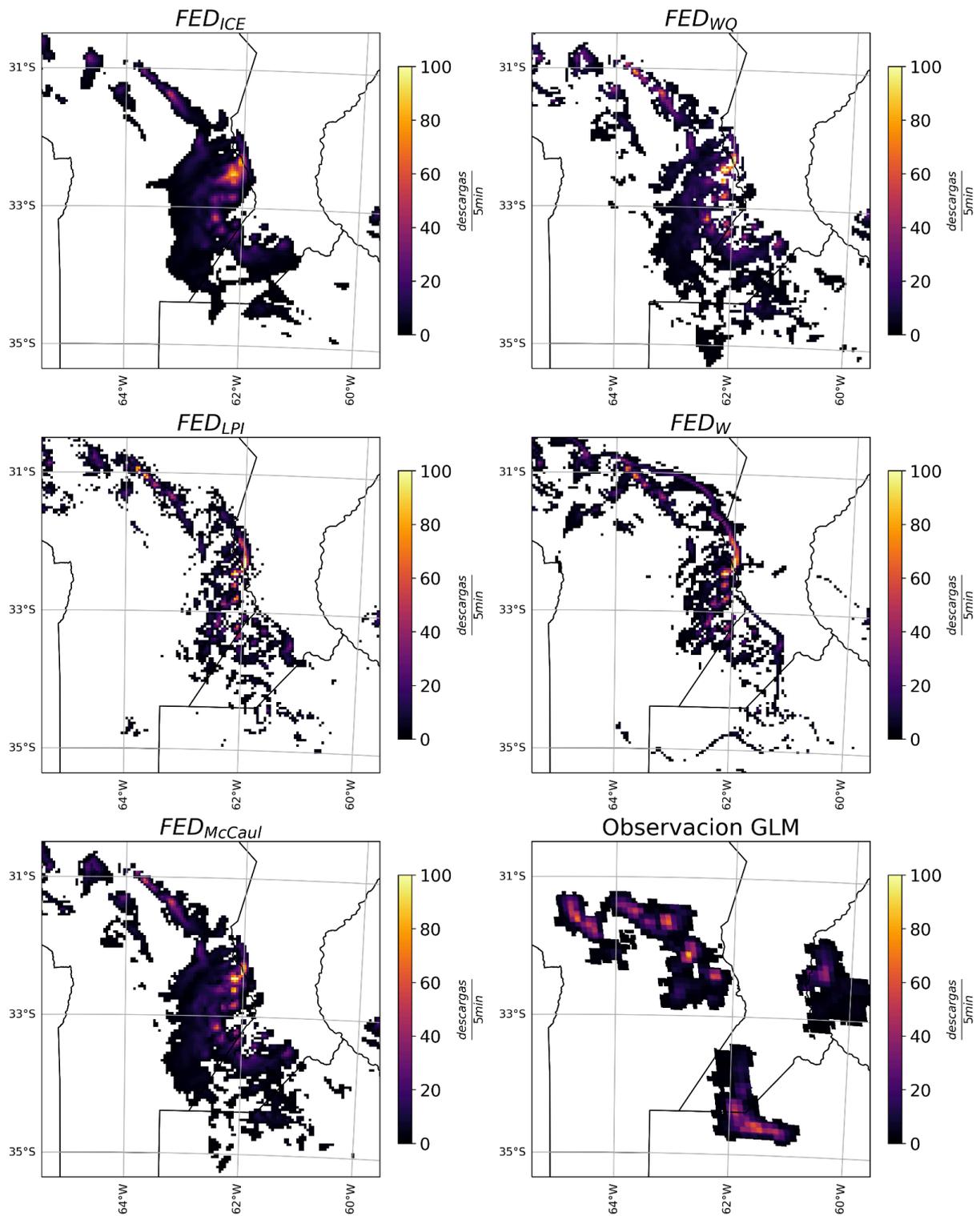


Figura 12: Idem Figura 11 pero haciendo foco en la convección del centro del país.