

SOBRE EL USO DE DATOS DE RAYOS COMO PROXY PARA LA REFLECTIVIDAD RADAR EN LA REGIÓN CENTRAL DE ARGENTINAMailén Gómez Mayol^{1,4}, Luciano Vidal¹, Paola Salio^{2,3}, Maximiliano Sacco^{1,3}¹Servicio Meteorológico Nacional²Centro de Investigaciones del Mar y la Atmósfera, CONICET-UBA UMI
3351-CNRS-CONICET-UBA³Departamento de Ciencias de la Atmósfera y los Océanos, FCEyN, UBA⁴Departamento de Física, FCEyN, UBA

(Manuscrito recibido el 14 de noviembre de 2018, en su versión final el 1 de marzo de 2019)

RESUMEN

Las tormentas y sus fenómenos asociados tienen alto impacto en el ámbito social y económico de un país. Por ello es necesario contar con herramientas que permitan garantizar la seguridad de las personas y sus bienes frente a estos eventos. Los datos aportados por una red de radares meteorológicos son de suma importancia para el pronóstico y seguimiento de las tormentas. Estas redes de radares son costosas y de difícil mantenimiento. Asimismo, no todos los países cuentan con una red de radares meteorológicos que cubra toda su extensión y aun así, a menudo tienen problemas técnicos o de conectividad que dejan sin información a amplias zonas.

En este trabajo se presenta una técnica de producción de un campo artificial de reflectividad COLMAX utilizando técnicas de aprendizaje profundo supervisado sobre información de descargas eléctricas en superficie. Este campo se puede usar como una alternativa rentable de la información de radares meteorológicos en áreas donde no hay cobertura o la misma es deficiente, o como un complemento a las imágenes de radares meteorológicos que están afectadas por atenuación, interferencias u otros problemas de observación.

Palabras clave: reflectividad radar, descargas eléctricas, redes neuronales, radar sintético.

ON THE USE OF LIGHTNING DATA AS A PROXY FOR RADAR REFLECTIVITY OVER THE CENTRAL REGION OF ARGENTINA**ABSTRACT**

Storms and their associated phenomena have a high impact in a country's social and economic environment. It is necessary to have tools that guarantee the safety of people at all times. The data provided by a network of meteorological radars is of utmost importance for storm tracking and identification. These radar networks are expensive and difficult to maintain. Not all countries have a radar network that covers all its extension and even so, they often have technical or connectivity problems that leave large areas without information about rainfall.

In this work we present a technique to generate a synthetic radar reflectivity field using supervised deep learning techniques with cloud-to-ground lightning data. This

product can be used as a cost-effective radar alternative in areas where there is no or poor coverage, or as a complement to meteorological radar images that are affected by attenuation, interference or other observation problems.

Keywords: radar reflectivity, cloud-to-ground lightning, neural networks, synthetic radar.

1. INTRODUCCIÓN

Las tormentas y sus fenómenos asociados pueden tener alto impacto en el ámbito social y económico de un país. Los fuertes vientos, las lluvias intensas en cortos periodos de tiempo, el granizo de gran tamaño y la actividad eléctrica son algunas de las características más peligrosas, especialmente para las diversas actividades humanas. El granizo daña casas, autos, aviones, cultivos y ganado; los vientos rafagosos y las cortantes asociadas cerca de superficie pueden causar problemas a los aviones especialmente en las fases de despegue y aterrizaje; la actividad eléctrica puede desencadenar incendios, daños a estructuras en ciudades y hasta explosiones, así como afectar el normal desarrollo de las actividades en los aeropuertos (Sénési y otros, 2009; Taffener y otros, 2008). En este contexto, son muchos los estudios en nuestra región basados en información de sensores remotos que muestran que la frecuencia de fenómenos severos es muy alta e incluso superior a otras regiones del planeta (Rasmussen y otros, 2014; Cecil y Blankenship, 2012; Mezher y otros, 2012; Matsudo y Salio, 2011).

Las zonas donde se desarrollan los fenómenos meteorológicos más intensos dentro de estas tormentas, suelen estar muy acotadas en espacio y tiempo. Las ascendentes capaces de formar grandes piedras de granizo se caracterizan por tener unos pocos kilómetros de radio y pocas horas de tiempo de vida, por ello el pronóstico se ve dificultado mediante el uso de los modelos numéricos de predicción del tiempo (Gagne y otros, 2017; Grubisic y otros, 2005). Así, los sensores remotos como los radares meteorológicos, los satélites y las redes de detección en superficie de descargas eléctricas atmosféricas (rayos), juegan un rol fundamental

para la detección, el seguimiento y pronóstico a muy corto plazo de este tipo de tormentas (Dance, 2010; Megenhardt y otros, 2004).

Las zonas geográficas con grandes concentraciones de rayos representan regiones de convección activa muchas veces severa. Se ha observado que un incremento súbito de la tasa de descargas eléctricas totales (lightning jumps) se asocia a la presencia de tormentas severas y tornádicas. La hipótesis que sobre la cual se basa esta tendencia está relacionada con la respuesta eléctrica debido a la carga de hielo y agua sobre-enfriada asociada al comportamiento de la ascendente en la evolución de las tormentas (Schultz y otros, 2011; Gatlin, 2006; Steiger y otros, 2005, 2007; Bridenstine y otros, 2005; Goodman y otros, 2005; Wiens y otros, 2005).

Numerosas investigaciones han centrado sus esfuerzos en la determinación de una relación sólida entre la actividad eléctrica atmosférica asociada a las nubes de tormentas y los campos de reflectividad generados por los radares meteorológicos. Tapia y otros (1998) estudiaron 22 tormentas de verano en la península de la Florida (EE.UU.) y desarrollaron un modelo que relaciona la precipitación convectiva y la actividad eléctrica en la escala de la tormenta. Los autores luego aplicaron su modelo para estimar la lluvia con el radar a partir de los datos de las descargas eléctricas y concluyeron que era posible obtener estimaciones de precipitación razonables en espacio y tiempo en eventos de lluvias intensas. Weber y otros (1998) realizaron una regresión entre la variable derivada del radar denominada agua líquida integrada en la vertical (VIL, por su sigla en inglés) y datos de actividad eléctrica provenientes de la red National Lightning Detection Network para tres casos convectivos en EE.UU. y encontraron

una correlación modesta entre los dos campos, señalando que los datos de actividad eléctrica podrían proporcionar beneficios para completar los vacíos presentes en la cobertura espacial de la red de radares disponible en la región. Luego, Mueller y otros (1999) desarrollaron, también para EE.UU., un producto denominado NCWF (National Convective Weather Forecast) que indica posibles riesgos de origen convectivo para la aviación. El NCWF se deriva de datos de radar (producto VIL) del compuesto nacional de la NEXRAD y datos de actividad eléctrica de Global Inc. Por su parte, Megenhardt y otros (2000) llevaron a cabo un mapeo de los datos de actividad eléctrica a una retícula de 4 km para crear relaciones entre las descargas nube-tierra integradas en 10 minutos y en cajas de 8 km de lado, y VIL para servir como entrada a un campo de detección de peligros denominado Detección Nacional del Tiempo Convectivo (NCWD, por sus siglas en inglés). Ellos utilizan los datos de actividad eléctrica para mejorar el NCWD a lo largo de los bordes principales de las tormentas y en las regiones donde faltan datos de radar. Por su parte, Soula y Chauzy (2001) encontraron buenas correlaciones entre las descargas eléctricas y los datos de radar en la escala de la tormenta para casos de tormentas eléctricas en Francia. Si bien estos estudios de casos muestran el potencial de utilizar datos de actividad eléctrica como un estimador del campo de reflectividad radar o de precipitación cerca de superficie, en regímenes localizados, sus relaciones no son directamente aplicables a los sistemas de predicción inmediata que deben operar en grandes dominios espaciales donde los regímenes de precipitación pueden variar notablemente. Mientras que Zhou y otros (2002) obtuvieron una relación similar para las tormentas en China y concluyeron que las descargas nube-tierra podrían usarse para estimar la precipitación convectiva, especialmente en regiones con una cobertura de radar insuficiente. Esta herramientas asociadas a las mejoras y calibraciones de los campos de precipitación estimados usando información de descargas eléctricas presenta notables resultados en las

variaciones entre la precipitación convectiva y estratiforme (Xu y otros 2014; Morales y Anagnostou, 2003).

En los últimos años han surgido nuevos métodos de pronóstico y alerta temprana de variables meteorológicas basadas en técnicas de aprendizaje automático (comúnmente llamado machine learning). Las investigaciones en el ámbito del estudio de variables meteorológicas mediante redes neuronales y aprendizaje profundo son cada vez más numerosas. La literatura es amplia y en constante crecimiento. Se han implementado algunas técnicas básicas como el análisis de componentes principales para la reducción de dimensionalidad (Monahan y otros, 2009) y k-means para clustering (Steinhaeuser y otros, 2011). Por mencionar algunos, Atkinson y Tatnall (1997) estudiaron la aplicación de las redes neuronales en datos de sensores remotos y Krasnopolsky y Schiller (2003) las aplicaron a datos satelitales. Hong y otros (2004) usaron las redes neuronales para calcular estimados de lluvia de alta densidad a partir de mediciones de sensores remotos. Mientras que Marzban y Stumpf (1996) entrenaron una red neuronal usando propiedades de la circulación para identificar tornados, por su parte Marzban y Witt (2001) entrenaron una red usando propiedades de las tormentas para hacer un pronóstico de granizo. Con respecto al aprendizaje profundo, Xingjian y otros (2015) utilizaron redes convolucionales LSTM (Long Short-Term Memory) para predecir eventos de precipitación en escala local usando datos de ecos de radar. En el ámbito del clima, Prabhat y otros (2012, 2015) desarrollaron una aplicación llamada TECA (Toolkit for Extreme Climate Analysis) que detecta patrones en datos climáticos. A partir de esta herramienta, Liu y otros (2016) demostraron que las redes neuronales convolucionales sobre imágenes con múltiples canales pueden utilizarse para predecir distintos tipos de eventos extremos.

El objetivo del presente trabajo es presentar un modelo que mediante técnicas de aprendizaje profundo supervisado generativo, simula un

campo de reflectividad de radar (en adelante, radar sintético) a partir de datos de descargas eléctricas nube-tierra obtenidos por una red de detección en superficie. Como dato de entrenamiento se utiliza un par de imágenes donde están representadas la información de tasa de descargas eléctricas nube-tierra y el campo de reflectividad máximo de la columna observado. Una vez entrenado, el modelo puede usarse para generar el producto radar sintético extrapolado a cualquier región donde se disponga de información de actividad eléctrica. Estos datos pueden ser utilizados en regiones donde la cobertura de radares es pobre o bien los datos están fuertemente degradados, por ejemplo, por bloqueos topográficos, atenuación por precipitación sólida o líquida, entre otros.

El trabajo se organiza de la siguiente manera. En la sección 2 se introduce el concepto de redes neuronales. Los datos que se usaron para entrenar el algoritmo: reflectividad radar y tasa de descargas eléctricas nube-tierra, se describen en la sección 3. En la sección 4 se detallan la metodología y todos los procesos por los que deben pasar los datos de radar y descargas eléctricas para poder ser aptos para el entrenamiento y luego para el análisis de los resultados. En la sección 5 se muestran los resultados y validación. Las conclusiones y trabajos futuros se describen en la sección 6.

2. SOBRE LAS REDES NEURONALES

Las redes neuronales (ANNs del inglés artificial neural networks) son algoritmos inspirados en sistemas biológicos, más precisamente en la investigación del cerebro humano, que son capaces de aprender y generalizar a partir de la experiencia (Bishop, 1995). Han demostrado ser excelentes herramientas en tareas de clasificación y reconocimiento de patrones. Actualmente se utilizan en muchos ámbitos de la ciencia, industria y los negocios (Paliwal y Kumar, 2009; Cherkassky y otros, 2006).

La red neuronal es un modelo que busca inferir una función que mapee un conjunto de entrada

en un conjunto de salida u objetivo (target, en inglés). El entrenamiento se hace por pasos. El primer paso consiste en aplicar una función con parámetros entrenables a un subconjunto de entrada y luego una función de activación. El resultado se compara con los datos del conjunto de salida utilizando una función de error (cuadrados mínimos, norma L1). Luego un algoritmo de tipo gradiente descendente (Ruder, 2016) optimiza el valor de los parámetros entrenables de forma tal que el error sea mínimo.

El segundo paso será repetir la misma operación a otro subconjunto de entrada y así hasta que todos los datos del conjunto de entrada hayan pasado por la red. Esto se lo conoce como haber completado una época. Al término de una cierta cantidad de épocas, que varía según el tipo de datos, se hace una verificación. La verificación consiste en realizar un paso de entrenamiento sobre datos que sean compatibles con los del conjunto de entrada pero que la red no haya procesado nunca (conjunto independiente).

En el presente trabajo se utilizan redes neuronales que tienen como base bloques de redes convolucionales (CNN, del inglés Convolutional Neural Networks) que están diseñadas para trabajar con imágenes como dato de entrada (LeCun, 2015). Una red convolucional simple consiste en una secuencia de capas, y cada capa tiene como entrada la salida de la anterior. Las redes convolucionales profundas consisten en una primer capa formada por un conjunto de filtros que se entrenan para reconocer un patrón determinado en una imagen. Las redes profundas son una sucesión de estas capas. Cada una interpreta la salida de la capa anterior logrando así identificar patrones cada vez más complejos.

En particular se aplicó una red neuronal generativa antagónica condicional (cGAN), que es una variante de las redes GAN. Las GAN fueron desarrolladas en el trabajo de Goodfellow y otros (2014) y se las considera el estado del arte en redes neuronales ya que son capaces de producir datos sintéticos muy realistas. El algoritmo propuesto en su

trabajo se presenta como un método para estimar modelos generativos donde se entrenan dos modelos en simultáneo. Estos modelos son redes neuronales que compiten entre sí en un juego de suma cero: un modelo “generativo” que aprende la distribución de los datos, y un modelo “discriminatorio” que estima la probabilidad que una muestra haya salido del conjunto de entrenamiento o del generador. Ellos proponen ambos modelos sean perceptrones multicapa. El generador se entrena para engañar al discriminador y el discriminador se entrena para minimizar sus errores y distinguir mejor las imágenes simuladas de las reales, un esquema del algoritmo se puede ver en la figura 1. Como se mencionó, el modelo que se usa en este trabajo es un cGAN, que a diferencia del primero, aprende a inferir una función que parte de un conjunto de imágenes de entrada a un conjunto de imágenes de salida. Como punto de partida para el entrenamiento del discriminador se usa un conjunto de imágenes reales. En este caso se le presentan imágenes del producto COLMAX de reflectividad radar. Luego se entrena el generador, que toma las imágenes del campo preliminar de radar y genera una nueva imagen. Esta imagen es evaluada por el discriminador. En ambas redes se entrenan de modo que el generador produce progresivamente mejores imágenes, mientras que el discriminador distingue con mejor precisión las imágenes sintéticas de las reales.

3. DATOS

3.1 Reflectividad radar

Los datos de reflectividad utilizados provienen del radar meteorológico Selex SI Gematronik Doppler de doble polarización en banda C (5.6 GHz) instalado en la Estación Experimental Agropecuaria INTA Anguil (La Pampa), ubicado en 36° 32' 23"S y 63° 59' 24"O. Los datos corresponden a la estrategia de escaneo que genera un volumen de datos cada 10 minutos conformado por un total de 12 elevaciones de antena que varían entre 0.5 y 15.1 grados, con una resolución en rango de 0.5 km y 1° en azimut y un alcance máximo de 240 km. Los datos de

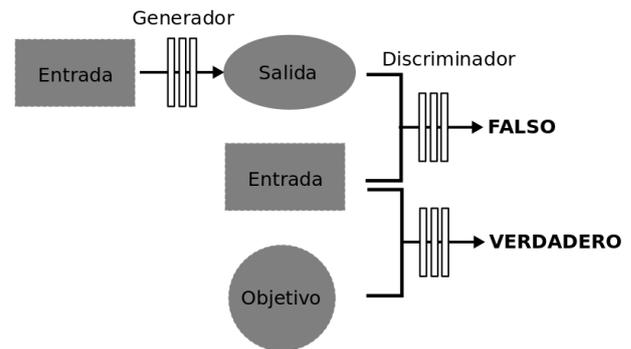


Figure 1: Esquema del algoritmo de una red neuronal generativa antagónica condicional (cGAN).

entrenamiento fueron tomados para el período de verano entre el 1 de noviembre de 2016 y el 28 de febrero de 2017, mientras que los datos de validación fueron tomados en el período enero y febrero de 2016.

3.2 Descargas eléctricas nube-tierra

Los datos de actividad eléctrica atmosférica empleados provienen de la red Vaisala Global Lightning Dataset (GLD360) puesta en servicio en septiembre de 2009, estando los datos disponibles para los usuarios a partir de mayo de 2011. La misma consta de sensores ubicados estratégicamente en todo el mundo para la detección óptima de descargas eléctricas desde las nubes a superficie (cloud-to-ground, CG). Las ubicaciones de las descargas se obtienen utilizando diferentes métodos como el tiempo de llegada, búsqueda de la dirección magnética y un algoritmo de reconocimiento de forma de la onda del rayo (Said y otros, 2010). De este modo se cuenta con datos como la posición en tiempo (hora, minuto y segundo) y espacio (latitud y longitud), y la polaridad (positiva o negativa) de cada descarga.

La eficiencia de detección (en adelante, DE) esperada y la precisión de la ubicación media (en adelante, LE) de esta red es 70% y 5-10

km, respectivamente. No obstante, varios son los trabajos realizados, mayormente en EE.UU. y Europa, tendientes a evaluar el desempeño de la red. Demetriades y otros (2010) llevaron a cabo una comparación de GLD360 con la Red Nacional de Detección de Rayos (NLDN) de EE.UU. para el periodo comprendido entre el 1 de diciembre de 2009 y el 31 de enero de 2010, y encontraron valores para DE de entre 86 y 92 %, mientras que para LE el valor medio encontrado fue de 10,8 km. Para el mismo periodo, pero usando la Red Brasileña de Detección de Rayos (BrasilDAT), Naccarato y otros (2010) encuentran para el sudeste de Brasil una DE de 16 % y un LE de 12,5 km. Más tarde, Said y otros (2013) realizan una evaluación contrastando GLD360 con la red NLDN pero para el periodo comprendido entre el 21 de julio de 2011 y el 21 de julio de 2012. Los resultados encontrados muestran que la DE fue del 57 % con un LE de 2,5 km. Estudios comparativos similares se realizaron en otras partes del mundo como Europa. Pohjola y Makela (2013) encuentran un valor para DE del 48 % y de 1,5 km para LE cuando comparan GLD360 contra EUCLID (European Cooperation for Lightning Detection) en Austria durante julio de 2011. Más recientemente, Mallick y otros (2014) estimaron las características de rendimiento de GLD360 utilizando datos de rayos activados por cohete y cable en Camp Blanding (Florida, EE.UU.) entre 2011 y 2013. Esta es la primera evaluación exhaustiva de las características de rendimiento de GLD360 en relación con la verdad absoluta, ya que todas las evaluaciones previas discutidas anteriormente son al menos en parte relativas a otros sistemas de detección. No obstante es importante remarcar que estos resultados son sólo aplicables a esta región. Los autores encuentran una DE de entre 67 y 37 %, para flashes y strokes respectivamente, y un LE de 2 km. Otro tipo de análisis comparativo tendiente a evaluar la red son los que llevaron a cabo Rudlosky y otros (2017). Los autores evaluaron el rendimiento de los datos GLD360 operacionales y reprocesados en relación con el sensor de imágenes de rayos LIS a bordo del satélite TRMM durante el periodo 2012–2014.

El análisis compara las observaciones de rayos basadas en superficie y las vistas desde el espacio con el objetivo principal de realizar una mejor caracterización de la red GLD360 antes y después de la actualización. La red GLD360 se sometió a una actualización el 18 de agosto de 2015 que incluyó cambios del algoritmo en el procesador central (Said y Murphy, 2016). Estos resultados muestran que la DE se incrementó de 32,6 % (2012) a 38 % (2013) a 43,2 % (2014) sobre Sudamérica.

Para el presente trabajo se utilizaron datos de la red GLD360 (fecha/hora/lat/lon) disponibles en el banco de datos del Departamento de Investigación y Desarrollo del Servicio Meteorológico Nacional para el período mencionado en la sección anterior de los datos de radar.

4. METODOLOGIA

Como se mencionó en la introducción, el objetivo del presente trabajo es obtener un producto de reflectividad simulado a partir de datos de actividad eléctrica (rayos nube-tierra). Para ello el flujo de trabajo incluye una primer etapa donde se sigue la metodología propuesta por Iskenderian (2008) con el objetivo de obtener un producto preliminar de reflectividad COLMAX a partir de la tasa de descargas eléctricas (en adelante, radar preliminar). Luego, en una segunda etapa se utiliza el framework "pix2pix" propuesto por Isola y otros (2016) que emplea un tipo de Red Generativa Antagónica (GAN - del inglés Generative Adversarial Network) para aprender un mapeo entre imágenes de entrada (radar preliminar) a imágenes de salida (radar observado). Este algoritmo es de tipo supervisado por lo que es necesario contar con un conjunto de datos de entrenamiento. A continuación se describen cada uno de los pasos dentro del flujo de trabajo mencionado.

4.1 Imágenes objetivo: producto COLMAX

Se generaron imágenes COLMAX con resolución de 2 km² para los datos del radar utilizando el

paquete PyART (Hellmus y Collis, 2016). Este paquete lee un archivo donde se encuentran las mediciones del radar (campos medidos, información de las coordenadas del radar, información sobre las elevaciones y estrategias de escaneo, entre otros) y genera un objeto radar. Los archivos que se usaron en este trabajo contienen la variable reflectividad radar (dBZ) para las 12 elevaciones que corresponden a un intervalo de tiempo de medición de 4 minutos aproximadamente (volumen completo).

Se calcularon 30 campos CAPPI (Constant Altitude Plan Position Indicator) entre el nivel del mar y los 15 km de altura con una resolución vertical de 0.5 km y horizontal de 2 km. Los CAPPI se generaron mediante la función `grid_from_radars` del módulo `map` de PyART. Esta función toma como entrada un objeto radar y lo mapea a una retícula regular mediante el algoritmo de reticulado `map_gates_to_grid`, que se eligió por sobre la otra opción (`map_to_grid`) por ser más rápida computacionalmente. El algoritmo `map_gates_to_grid` genera una retícula cartesiana de un campo de mediciones de radar a partir de los puntos de uno o varios radares. Para cada punto define un radio de influencia y pesa la contribución de cada punto usando una función de peso. El radio de influencia se eligió para que crezca con la distancia al radar, siendo así más importantes los puntos que están cerca. Finalmente, la imagen COLMAX se obtuvo buscando el valor máximo de reflectividad en la altura para cada punto de la retícula (figura 2).

4.2 Filtrado de volúmenes atenuados

La señal de los radares meteorológicos en banda C puede verse afectada fuertemente por problemas de atenuación cuando la señal atraviesa zonas de precipitación intensa y/o presencia de granizo, llegando a casos extremos en los cuales la misma puede extinguirse totalmente. En esta condición, la señal del radar puede verse muy debilitada lo cual puede llevar a interpretaciones erróneas de la información

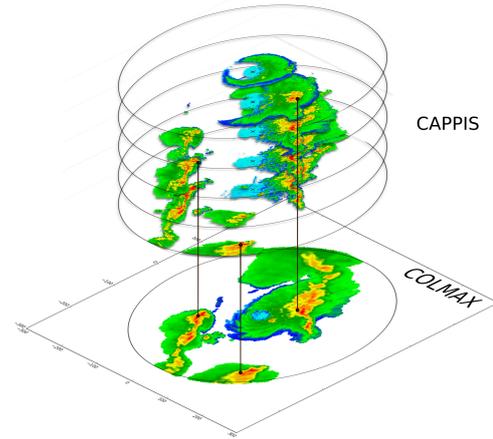


Figure 2: Esquema 3D de un volumen de reflectividad radar (dBZ) donde se muestran los cortes CAPPI (Constant Altitude Plan Position Indicator) y como se realiza la proyección para obtener el producto COLMAX.

provista por el radar (Gu y otros, 2011; Bringi y otros, 2001; Carey y otros, 2000).

Los pioneros en el estudio de los efectos de atenuación fueron Atlas y Bank (1951) cuando encontraron que los patrones de reflectividad del radar eran distorsionados por la presencia de patrones de lluvia. Más tarde Hitschfeld y Bordan (1954) derivaron y estudiaron la ecuación de la atenuación, que relaciona los perfiles de lluvia con la reflectividad medida. Esta ecuación tenía algunos problemas de estabilidad que luego solucionaron por Marzoug y Amayenc (1991). De este modo se define el factor de reflectividad medido en un rango dado r para una longitud de onda atenuada (Z_m) como:

$$Z_m(r) = Z_e(r) * \delta C * A(r_0) * \exp \left(-0,46 * \int_{r_0}^r k(s) ds \right) \quad (1)$$

donde Z_e es el factor de reflectividad equivalente en el mismo rango., δC representa el posible error de calibración del radar, y A es el factor PIA que

se define como:

$$A(r) = A(r_0) * \exp\left(-0,46 * \int_{r_0}^r k(s) ds\right) \quad (2)$$

Aquí, k es el coeficiente de atenuación (dB km^{-1}), que depende de la longitud de onda del radar y de la distribución del tamaño y la temperatura de las gotas de lluvia. Se define el $PIA(\text{dB})$ como el doble de la integral de k entre el rango 0 y el rango r , o:

$$PIA = -10 * \log_{10}[A(r)] \quad (3)$$

donde \log hace referencia al logaritmo en base 10. La variable r_0 representa el llamado rango ciego, es decir, el rango donde se inicia el muestreo de reflectividad o donde las reflectividades pueden considerarse libres de ruido. Aquí, $A(r_0)$ es el factor de atenuación de rango ciego, por lo tanto, permite la descripción de los efectos de atenuación que ocurren en caso de lluvia cerca del sitio del radar (por ejemplo, radomo mojado).

Si se asume que las relaciones entre el factor de reflectividad equivalente Z_e , el coeficiente de atenuación k y la tasa de lluvia R pueden representarse satisfactoriamente mediante modelos de ley de potencia con las siguientes ecuaciones:

$$Z_e = aR^n, \quad k = cR^d \quad \text{y} \quad Z_e = aR^\beta \quad (4)$$

Luego el factor de atenuación puede ser estimado combinando las ecuaciones (1) y (4) con $a=543$, $b=136$, $c= 1,55 \times 10^{-3}$ y $d=1,30$. Estos valores fueron elegidos usando las curvas de la evolución de los componentes para radar de banda C que se presenta en el trabajo de Delrieu y otros (1999). Estas curvas fueron calculadas usando un modelo de distribución de gotas en una zona de latitudes medias.

En este trabajo las imágenes de radar sirven como objetivo de un entrenamiento y es de suma importancia contar con imágenes de buena calidad que representen lo más fielmente posible la reflectividad real de los ecos meteorológicos. Por esta razón, se seleccionaron aquellos datos

que no presentan atenuación significativa. Para ello se aplicó un filtro basado en el cálculo de la atenuación para cada píxel en función de la reflectividad de acuerdo a lo mostrado anteriormente. Como resultado se descartan todos aquellos volúmenes que tengan algún píxel atenuado por encima de un umbral de 1 dB. Para definir este umbral se analizó un conjunto de imágenes de radar que presentaban ecos atenuados y se determinó el valor a partir del cual los ecos empezaban a deformarse debido a la atenuación. Como umbral se eligió el menor de dichos valores. Si bien el radar utilizado es de doble polarización, el filtro propuesto no pretende hacer una corrección de la reflectividad por atenuación sino identificar regiones donde la misma es muy notoria. Por lo tanto no se utilizó la variable diferencial de fase específica para el cálculo de la atenuación como propone Bringi y otros (2001).

4.3 Preparación del dato de entrada: radar preliminar COLMAX

Las imágenes de entrada a la red neuronal son campos preliminares del producto COLMAX que surgen de encontrar una relación entre la tasa de descargas eléctricas CG (en adelante, TCG) de un punto (figura 3b) y su valor correspondiente observado en el campo de reflectividad COLMAX (figura 3c). Estas relaciones se crean usando un método de comparación de probabilidades (PMM, del inglés Probability Matching Method; Calheiros y Zawadzki, 1987). El método PMM ha sido utilizado en la bibliografía para derivar relaciones entre reflectividad radar y tasa de lluvias (Atlas y otros, 1990, 1993; Crosson y otros, 1996; Marks y otros, 1993; Rosenfeld y otros, 1993; entre muchos otros) y para encontrar la relación entre descargas eléctricas y lluvia convectiva con el fin de asimilar información de actividad eléctrica en modelos de pronóstico numérico (Chang y otros, 2001).

Siguiendo la metodología propuesta por Iskenderian (2008), la posición de cada descarga eléctrica CG (figura 3a) se alinea con la retícula de 2 km² de los datos de radar (figura 3c) y

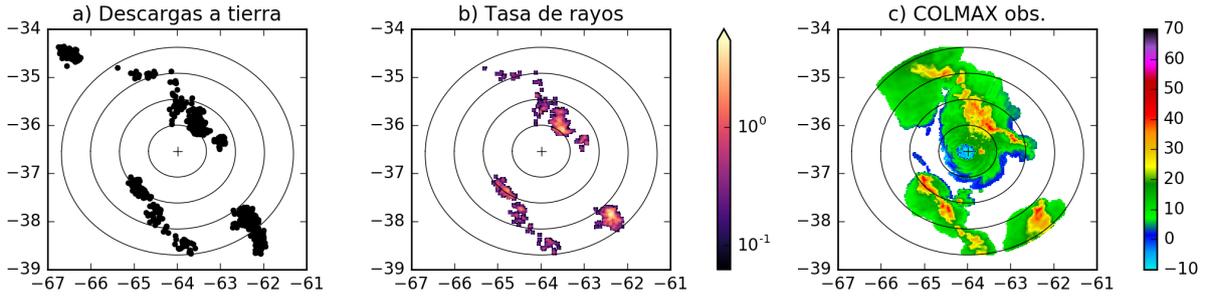


Figure 3: (a) Localización de las descargas CG detectados por la red Vaisala GLD360 el día 5/12/2016 entre las 00:50 y las 00:54 UTC. (b) tasa de descargas CG (TCG, min^{-1}). (c) Imagen COLMAX (dBZ) de reflectividad correspondiente al 5/12/2016 a las 00:50 UTC. El espaciamiento entre anillos es de 60 km. El centro del radar se indica con el símbolo +.

se calcula la TCG en cada celda (figura 3b). No se consideran aquellas descargas que estén por fuera del rango máximo del volumen radar empleado en el presente trabajo (240 km). Luego se hace un suavizado espacial que asigna a cada punto la suma de una contribución del valor de la tasa de actividad eléctrica en un radio de 8 km conforme a la bibliografía citada anteriormente, de modo tal que el resultado tenga un aspecto similar a las imágenes de radar. Para obtener el suavizado se convoluciona la imagen con un filtro que suma la TCG de los píxeles vecinos dependiendo de la distancia a la que se encuentra. Se utilizó un filtro de tipo gaussiano que suma a cada píxel el valor de sus primeros vecinos, un 0.75 de la suma de los segundos vecinos (8 píxeles) y un 0.25 de la suma de los terceros vecinos (16 píxeles).

Luego, el método PMM consiste en construir relaciones TCG [min^{-1}] - ZCOLMAX [dBZ] basados en pares de datos (TCG_i , $ZCOLMAX_i$) de forma que la función de distribución acumulada (CDF) de tasa de descargas eléctricas y reflectividad en el i -ésimo intervalo sean iguales:

$$\int_{ZCOLMAX,\tau}^{ZCOLMAX,i} P(ZCOLMAX) dZCOLMAX = \int_{TCG,\tau}^{TCG_i} P(TCG) dTCG \quad (5)$$

donde $P()$ representa una función de densidad de probabilidad y $ZCOLMAX$, τ son los valores más bajos que toma cada CDF. Para encontrar los valores mínimos se usa todo el set de datos TCG y COLMAX observado y se computa la función de distribución acumulada para cada uno de ellos. El menor valor detectable de tasa de descargas es 1 por lo tanto se puede especificar $TCG_\tau = 0,25$ de antemano. Luego los demás valores se usan para emparejar con la probabilidad condicional de la reflectividad.

Una vez que se determina el umbral de reflectividad se aplica la ecuación (5) al conjunto de datos para encontrar las relación tasa de descargas-reflectividad (figura 4c). Luego se ajusta la curva de la figura 4c obteniéndose la relación entre la tasa de descargas eléctricas y la reflectividad. La relación obtenida para los datos de este trabajo es:

$$COLMAX_{prel} = 5,90 * \ln(TCG) + 33,41 \quad (6)$$

4.4 Imágenes de entrada

El algoritmo de la red neuronal utiliza como dato de entrada una imagen compuesta como el ejemplo que se muestra en la figura 5. Para poder utilizar tanto el campo preliminar COLMAX como el observado como imágenes de entrada es necesario aplicarles una normalización. Luego se grafican ambas imágenes con una paleta de colores lineal y se unifican los límites a -10 dBZ (mínimo) y 80 dBZ (máximo). Si bien en general

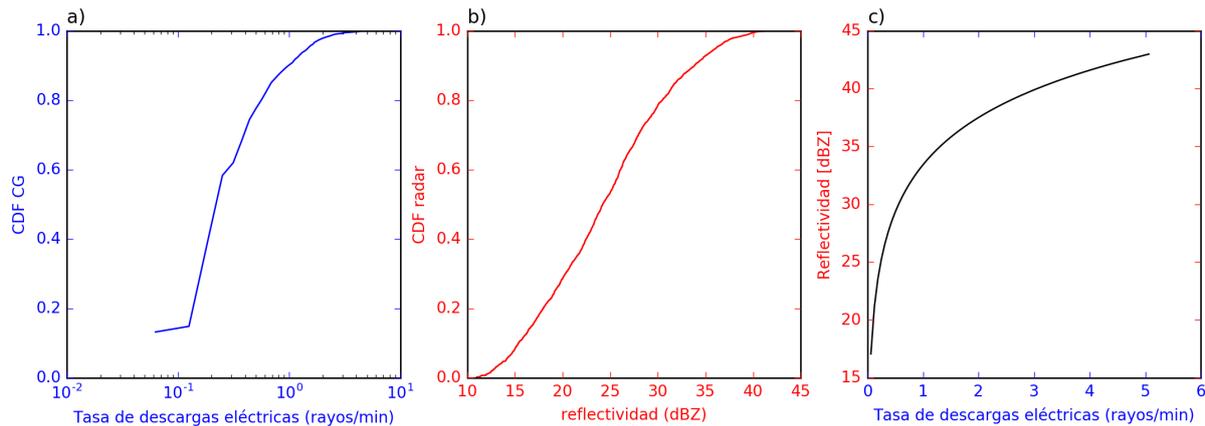


Figure 4: (a) Función de distribución acumulada de tasa de descargas CGs (TCG, min⁻¹). (b) Función de distribución acumulada de reflectividad COLMAX (dBZ). (c) Relación TCG-ZCOLMAX obtenida mediante la técnica PMM. Los puntos con igual probabilidad acumulada se combinan para crear esta última curva.

en el uso operativo se utilizan diferentes paletas de colores con rangos variables dependiendo del uso que se le dé a los datos de radar, en el caso particular del presente trabajo es conveniente aplicar una paleta de colores lineal (Borland y otros, 2007) para poder luego recuperar los valores de reflectividad a partir de las imágenes. Al usar una paleta lineal luego es sencillo hacer una traducción de la imagen a datos crudos (0 a 255), ya que si se emplea una paleta cualitativa luego se vuelve complejo recuperar el valor del píxel en dBZ a partir de su valor RGB.

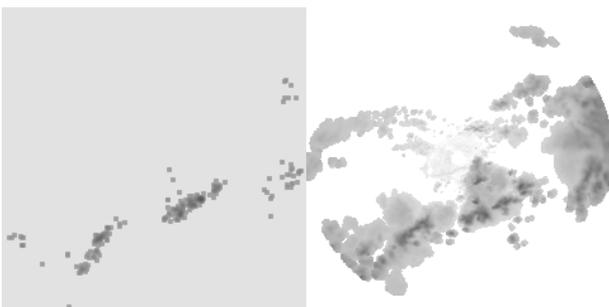


Figure 5: Ejemplo de imagen de entrenamiento correspondiente al día 26 de febrero de 2016 de 20:00 a 20:04 UTC. La misma está conformada por la imagen COLMAX preliminar (izq.) y la imagen COLMAX observado (der.).

4.5 Conjunto de entrenamiento y validación

El conjunto de entrada para la red está compuesto por 3422 pares de imágenes (figura 5) de las cuales el 75% de las mismas dentro de la red se utilizan para entrenamiento y el resto para verificación. La red se entrenó durante 100 épocas. El criterio de corte de entrenamiento consiste en evaluar el error de verificación. Se calcula el error del conjunto de verificación cada 25 épocas. Cuando el error de verificación deja de decrecer, se termina el entrenamiento. Esta técnica evita el sobre entrenamiento de las redes (Piotrowski, 2013). Una vez entrenada la red se la aplicó a un conjunto de validación que consta de 100 pares de imágenes que corresponden a eventos con gran actividad eléctrica en el período correspondiente a enero y febrero de 2016. Se eligieron los volúmenes, se generaron las retículas y se aplicaron los mismos métodos que al conjunto de entrenamiento. Una vez obtenidas las imágenes compuestas se le aplicó el modelo y se las comparó con las imágenes reales. El costo de entrenar la red es relativamente bajo, para completar las 100 épocas bastó entrenar durante 24 horas en un clúster de 24 procesadores. Por su parte el costo de aplicar el modelo al conjunto de validación es aún menor; se corrió el modelo sobre el conjunto completo de validación en menos de

un minuto en una computadora de escritorio con un procesador i5 y 4GB de memoria RAM.

4.6 Verificación de los resultados

A fin de determinar el desempeño del modelo a lo largo de todo el periodo considerado para la validación, se calcularon los índices de verificación estadística que miden la correspondencia entre la ocurrencia de eventos observados y estimados. Este método de verificación se basa en la confección de una tabla de contingencia de 2x2 donde se computan eventos que ocurrieron o no ocurrieron, es decir hubo eco/no hubo eco de radar. Se calcularon estas tablas para distintos umbrales de reflectividad usando los datos COLMAX preliminar, COLMAX sintético y COLMAX observado. Luego, se calculó la probabilidad de detección (POD), que mide la fracción de eventos ocurridos que fueron estimados correctamente; la tasa de falsa alarma (FAR), que mide la fracción de eventos estimados que no ocurrieron; el BIAS, que da la relación entre la reflectividad estimada (preliminar y sintética) y la observada; y el índice de éxito crítico (ETS) que estima la fracción de todos los eventos estimados y/o observados que fueron detectados correctamente. Estos índices se definen como,

$$\begin{aligned}
 POD &= \frac{n_{\text{aciertos}}}{n_{\text{aciertos}} + n_{\text{fallas}}} \\
 FAR &= \frac{n_{\text{falsa alarma}}}{n_{\text{aciertos}} + n_{\text{falsa alarma}}} \\
 BIAS &= \frac{n_{\text{aciertos}} + n_{\text{falsa alarma}}}{n_{\text{aciertos}} + n_{\text{fallas}}} \\
 ETS &= \frac{n_{\text{aciertos}}}{n_{\text{aciertos}} + n_{\text{fallas}} + n_{\text{falsa alarma}}}
 \end{aligned}$$

Donde n_{aciertos} es el número de puntos donde se observa reflectividad por encima del umbral considerado tanto en el COLMAX sintético como en el COLMAX observado; n_{fallas} es el número de puntos donde hay datos válidos en el COLMAX observado pero no en el sintético; y $n_{\text{falsa alarma}}$ es el número de puntos donde se observan datos válidos en el COLMAX sintético pero no en el observado. El ETS resulta proporcional a la cantidad de aciertos mientras que el BIAS compara el tamaño de las áreas donde hay

datos válidos de reflectividad por encima de un umbral, sin importar la ubicación de la misma. Los valores de BIAS mayores (menores) a uno indican que el área donde se observan datos de reflectividad en el COLMAX sintético es mayor (menor) al área correspondiente en el COLMAX observado. En una estimación perfecta del COLMAX, el ETS y el BIAS serían iguales a uno, al igual que el POD, mientras que el FAR debería ser cero. Todos estos estadísticos se calcularon para nueve umbrales de reflectividad COLMAX (0, 10, 15, 20, 25, 30, 35, 40 y 50 dBZ).

5. RESULTADOS

5.1 Análisis estadístico

En la figura 6 se muestran los histogramas de los valores de reflectividad COLMAX obtenidos al aplicar el modelo de radar sintético sobre todo el conjunto de situaciones elegidas para la de validación. Se observa que las distribuciones de ambas variables son similares. La distribución de la diferencia presenta un máximo en cero resultante de una compensación entre casos donde el radar sintético sobrestima o subestima los valores de reflectividad observada. En el histograma 2D que se muestra en la figura 6b se puede ver que hay una buena correlación entre los valores observados y sintéticos, observándose una ligera subestimación de los valores de reflectividad COLMAX por parte del radar sintético en la región a partir de los 20 dBZ. Las funciones de distribución muestran una sobreestimación del máximo de frecuencia en 18 dBZ, asimismo son subestimadas las frecuencias inferiores a 18 dBZ y por encima de 35 dBZ. Es importante destacar, que el radar meteorológico sufre de diferentes procesos de atenuación, que si bien durante el entrenamiento fue tenida en cuenta la atenuación por la presencia de la precipitación en el camino del radar, esto puede generar errores dado la falta de conocimiento de la reflectividad real. La función distribución de las diferencias muestra un histograma muy balanceado que denota un adecuado balance entre subestimaciones y sobreestimaciones. Eso es claro en un ligero bias de 0.69 muy cercano a

cero.

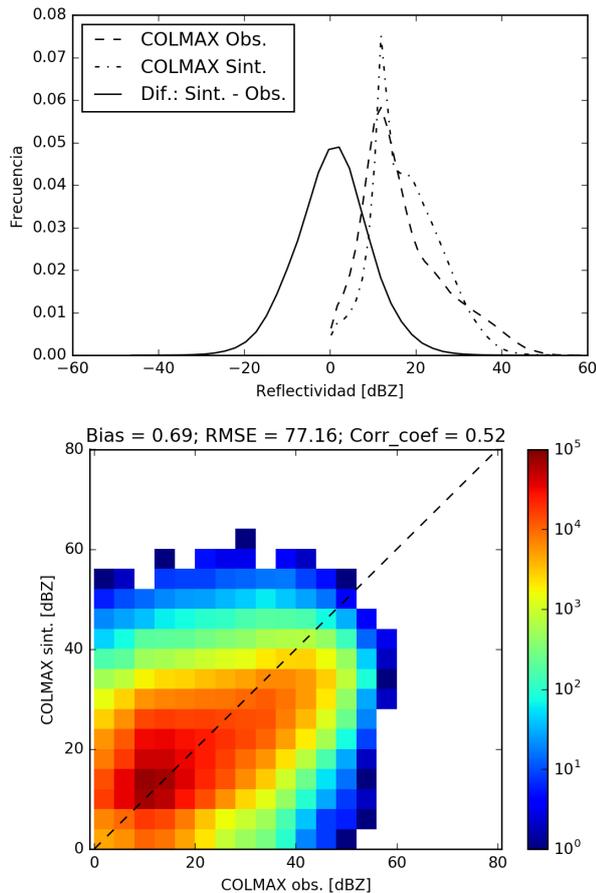


Figure 6: Gráfico de la distribución e histograma de valores de reflectividad observada y sintética para todo el conjunto de validación (sup.) e histograma 2D (inf.).

En la figura 7 se muestran los índices estadísticos para distintos umbrales de reflectividad para el COLMAX sintético. Todos los índices se calcularon con un total de 1.267.401 pares de valores de reflectividad observada-sintética. El índice POD indica que el COLMAX sintético tiene un buen desempeño en reflectividades medias y bajas, hasta 30 dBZ, donde también el índice de falsas alarmas es menor. El ETS un rendimiento adecuado en reflectividades hasta 30 dBZ. Finalmente, el BIAS es siempre ligeramente menor a 1 lo que indica que el COLMAX sintético subestima los valores de reflectividad. Esta subestimación es máxima en valores altos y mínima en valores medios (20 dBZ). En conclusión, el COLMAX sintético es capaz de

reconstruir los valores de reflectividad media. Sin embargo a medida que los valores de reflectividad por encima de 40 dBZ se observan valores inferiores en los estadísticos lo cual puede estar debido a un número inferior de pares de puntos que vuelve menos confiable los valores de los estadísticos, dado que las tasas de actividad eléctrica en presencia de convección es grande y eso mejora el ajuste entre los parámetros.

5.2 Análisis de casos

Como se mencionó en la introducción, los datos del radar muchas veces están degradados por problemas como atenuación de la señal del radar por presencia de núcleos intensos de precipitación o bien por la presencia de ecos no meteorológicos que introducen ruido al momento de una correcta interpretación de la información por parte del usuario. Con el objetivo de mostrar los beneficios asociados al producto radar sintético en situaciones particulares donde este tipo de artilugios está presente en los datos, se muestran los siguientes ejemplos seleccionados.

a) Línea convectiva En la figura 8 se muestra el caso del día 16 de enero de 2016 a las 19:30 UTC. Este caso corresponde a una línea convectiva ubicada al noreste del radar y dos celdas aisladas al noroeste. El campo COLMAX preliminar (figura 8b) muestra una buena correspondencia con la región más convectiva, tanto de la línea como de las celdas aisladas ($ZCOLMAX > 30$ dBZ) cuando se la compara con el campo de COLMAX observado (figura 8d). Sin embargo los valores están sobreestimados, con diferencias del orden de 10 dBZ en algunos sectores, y no es capaz de capturar las regiones con reflectividades menores a 30 dBZ (región estratiforme o asociada al yunque del sistema) ya que se centra en las zonas con baja actividad eléctrica. El campo de COLMAX sintético (figura 8c) en cambio es capaz de reproducir los máximos de reflectividad en la zona convectiva con diferencias máximas de 5 dBZ y logra capturar muy bien la región estratiforme que rodea a los núcleos convectivos. Esto se refleja en los valores estadísticos calculados para el COLMAX

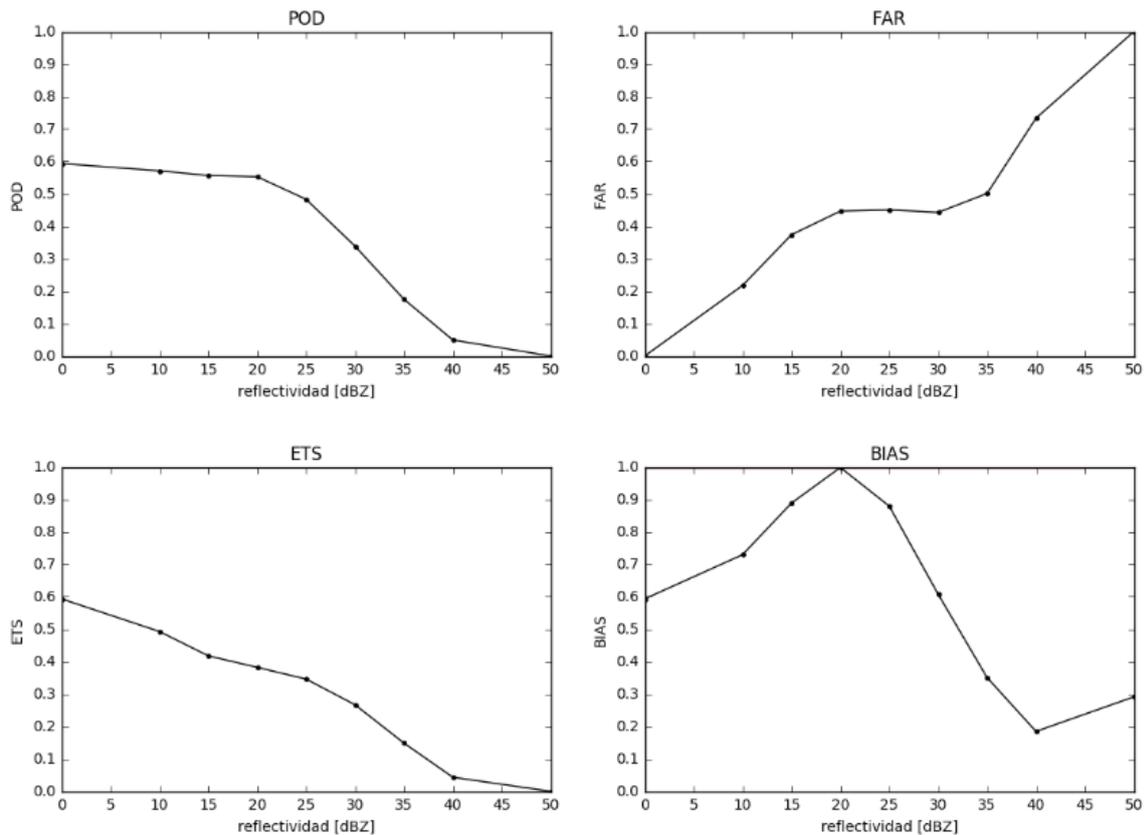


Figure 7: Resultado de los índices de validación estadísticos de COLMAX sintético para distintos umbrales de reflectividad.

observado y COLMAX sintético. El valor del coeficiente de correlación muestra que hay una buena correlación entre los valores sintéticos y observados; el bias positivo indica que hay una ligera sobrestimación por parte de los valores sintéticos. En la figura 8 además se grafican los histogramas y distribuciones de los valores para el COLMAX sintético, observado y la diferencia. Se observa que hay más píxeles con valores por encima de los 30 dBZ en el COLMAX observado que en el sintético, sin embargo, la distribución de la diferencia está centrada en cero y se ven algunas diferencias por encima de los 10 dBZ justificando el bias positivo. En el gráfico de la dispersión se ve que el COLMAX sintético sobrestima la reflectividad observada por debajo de los 20 dBZ y subestima los valores observados por encima de este umbral. Algo a destacar es la ausencia del clutter (eco de terreno) cerca del radar en la imagen sintética que se observa en

la figura 8d hacia el oeste.

b) Atenuación de la señal por radomo mojado
 Este caso se corresponde con una situación en la que se observa precipitación intensa sobre el radar que genera atenuación en la señal dado que el radomo se encuentra mojado. La figura 9 muestra una secuencia temporal correspondiente al día 18 de febrero de 2016 entre las 13:10 y las 13:30 UTC. En particular, a las 13:20 UTC se da el pasaje de una celda convectiva con reflectividades altas (>45 dBZ) por encima del radar haciendo que la señal se atenúe como consecuencia de la intensa precipitación que genera una película de agua sobre el radomo (figura 9 der.). En la secuencia correspondiente al COLMAX sintético (figura 9 izq.) se puede ver que no presentan este tipo de anomalías mostrando el beneficio de contar con un producto de estas características que permiten minimizar

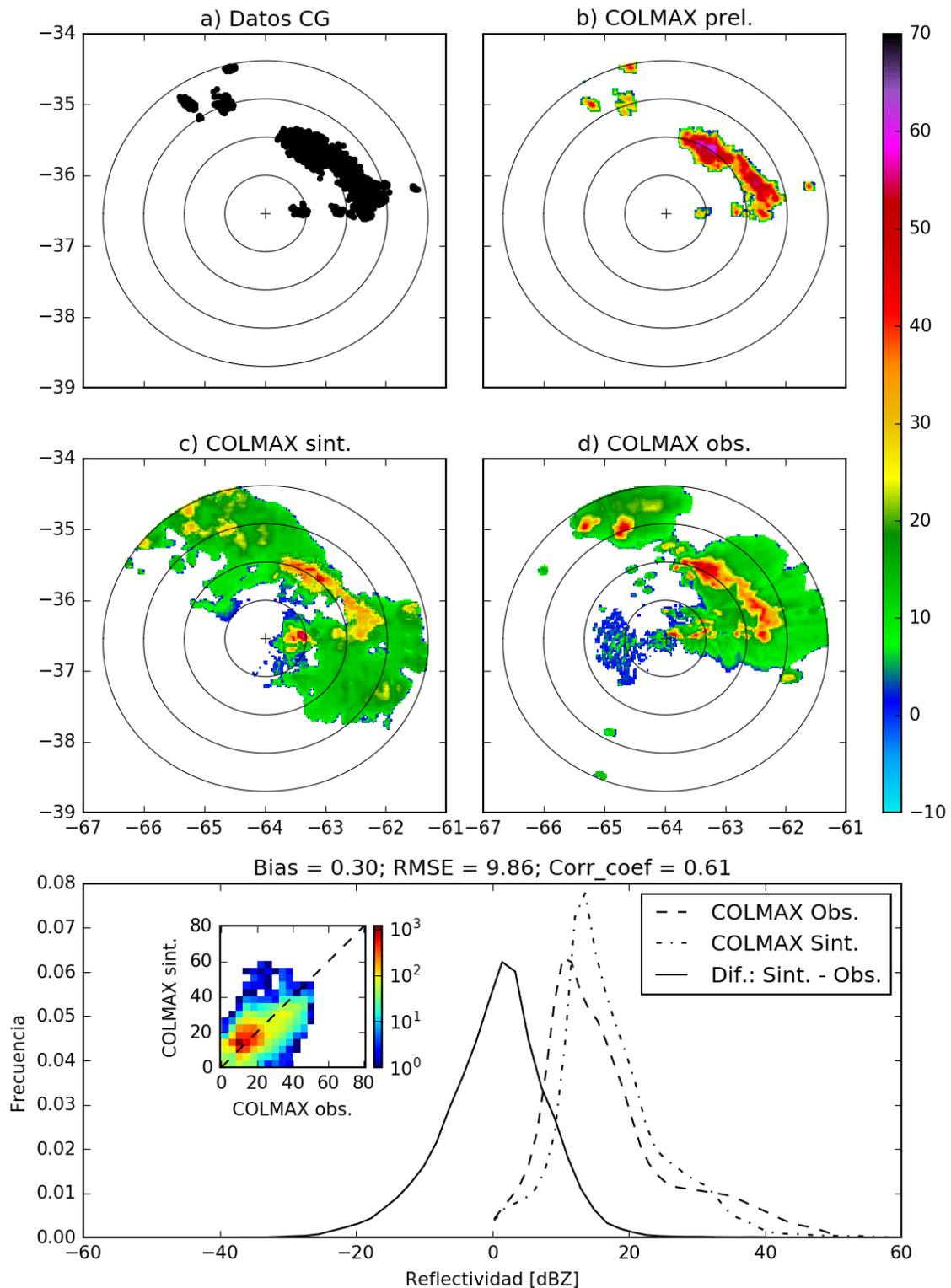


Figure 8: Ejemplo de los resultados obtenidos para el día 16 de enero de 2016 a las 19:30 UTC. (a) Localización de los rayos detectados por GLD360 entre las 19:30 y las 19:34 UTC. (b) COLMAX preliminar obtenido a partir de aplicar la ecuación (9) a la tasa de rayos derivada de a). (c) COLMAX sintético obtenido a partir del entrenamiento de la red pix2pix. (d) COLMAX observado. La separación entre anillos es de 60 kilómetros. Se muestra el histograma correspondiente a reflectividad observada COLMAX (línea punteada), reflectividad sintética (guión punto) y su diferencia (línea continua). Dentro se muestra un histograma 2D con la distribución de valores observados contra los valores sintéticos.

problemas típicos de la información radar. En la figura 10 se puede ver en detalle todos los campos que se calculan en el algoritmo. El gráfico de las distribuciones (histograma) de valores para este caso particular muestra que el COLMAX sintético parece sobrestimar la reflectividad observada, resultado esperable ya que este se ve afectado por la atenuación producto del efecto de radomo mojado. La atenuación en los datos observados también hace que el coeficiente de correlación sea pequeño. La distribución de datos sintéticos está siempre por encima de la distribución de datos observados (figura 10, ver histograma) y la distribución de la diferencia tiene un pico alrededor de los 7 dBZ. Esto indica que los valores del COLMAX observado son menores a los del COLMAX sintético, resultado que es evidente en el diagrama de dispersión. Esto también se puede ver reflejado en el valor del bias que es positivo y de valor absoluto mayor al del conjunto completo.

c) Eco meteorológico sin actividad eléctrica Este caso muestra una secuencia de dos tiempos sucesivos de imágenes COLMAX del día 19 de enero de 2016 a las 22:40 (figura 11) y 22:50 UTC (figura 12). En la figura correspondiente a las 22:40 UTC de COLMAX observado (figura 11d) se presenta en el sector oeste-noroeste una zona de ecos meteorológicos probablemente asociado al yunque de un núcleo convectivo que se encuentra por fuera del dominio de radar y que no presenta actividad eléctrica. Allí el producto COLMAX sintético (fig. 11c) no logra capturar nada del patrón de reflectividad presente, esto se refleja en el valor del bias que es negativo y de valor relativamente alto. El coeficiente de correlación es alto probablemente porque los datos sintéticos logran reproducir los ecos del sudoeste del área estudiada. En el segundo campo se observa que empieza a registrarse actividad eléctrica en dicho sector y ahí si el COLMAX sintético (fig. 12c) presenta un mejor acuerdo el COLMAX observado (fig. 12d); el valor del bias, si bien permanece negativo, aumenta. Las distribuciones (histogramas) para ambos tiempos son similares, ambas muestran que los datos observados tienen un pico muy alto en regiones de reflectividad baja

(10-15 dBZ) asociados a zonas estratiformes. En esa zona la distribución del COLMAX sintético es menor en los dos tiempos. A las 22:50 UTC (fig. 12) la distribución de los datos sintéticos aumenta levemente en la zona de los 20-30 dBZ, esto se corresponde con la aparición de actividad eléctrica en el borde noroeste del radar.

d) Atenuación asociada a celdas intensas El caso del día 7 de febrero de 2016 a las 03:40 UTC (figura 13) muestra una situación típica de atenuación por detrás de núcleos convectivos intensos. En la figura se puede ver que la reflectividad observada (fig. 13d) presenta valores atenuados al sudeste del radar (recuadro), sin embargo tanto el radar sintético (fig. 13c) como el preliminar (fig. 13b) logran representar adecuadamente esta región (fig. 13d). Este resultado es muy alentador pues demuestra la capacidad de este tipo de técnicas para minimizar problemas de atenuación tan frecuentes en los radares de banda C en presencia de precipitaciones intensas tan frecuentes en nuestro país.

6. CONCLUSIONES

Es de suma importancia para la seguridad de la población y de las actividades socioeconómicas que la misma desarrolla, contar con pronósticos a corto plazo de tormentas y demás fenómenos de clima severo cada vez más precisos. Para ello los radares meteorológicos juegan un rol fundamental pero como es sabido los mismos presentan ciertas limitaciones relacionadas con la calidad de la información que proveen así como de la cobertura espacial y temporal que ofrecen. En caso que la red de radares disponible tenga problemas de conectividad, salidas de servicios, pérdida de información, interferencias, bloqueo de terreno o poca cobertura, la calidad de los pronósticos a corto plazo empeora debido a la carencia de información. Pero en los últimos años numerosos estudios han mostrado que la intensidad de la actividad eléctrica asociada a las nubes de tormenta está altamente correlacionada con los campos de reflectividad medido por los radares meteorológicos y por ende

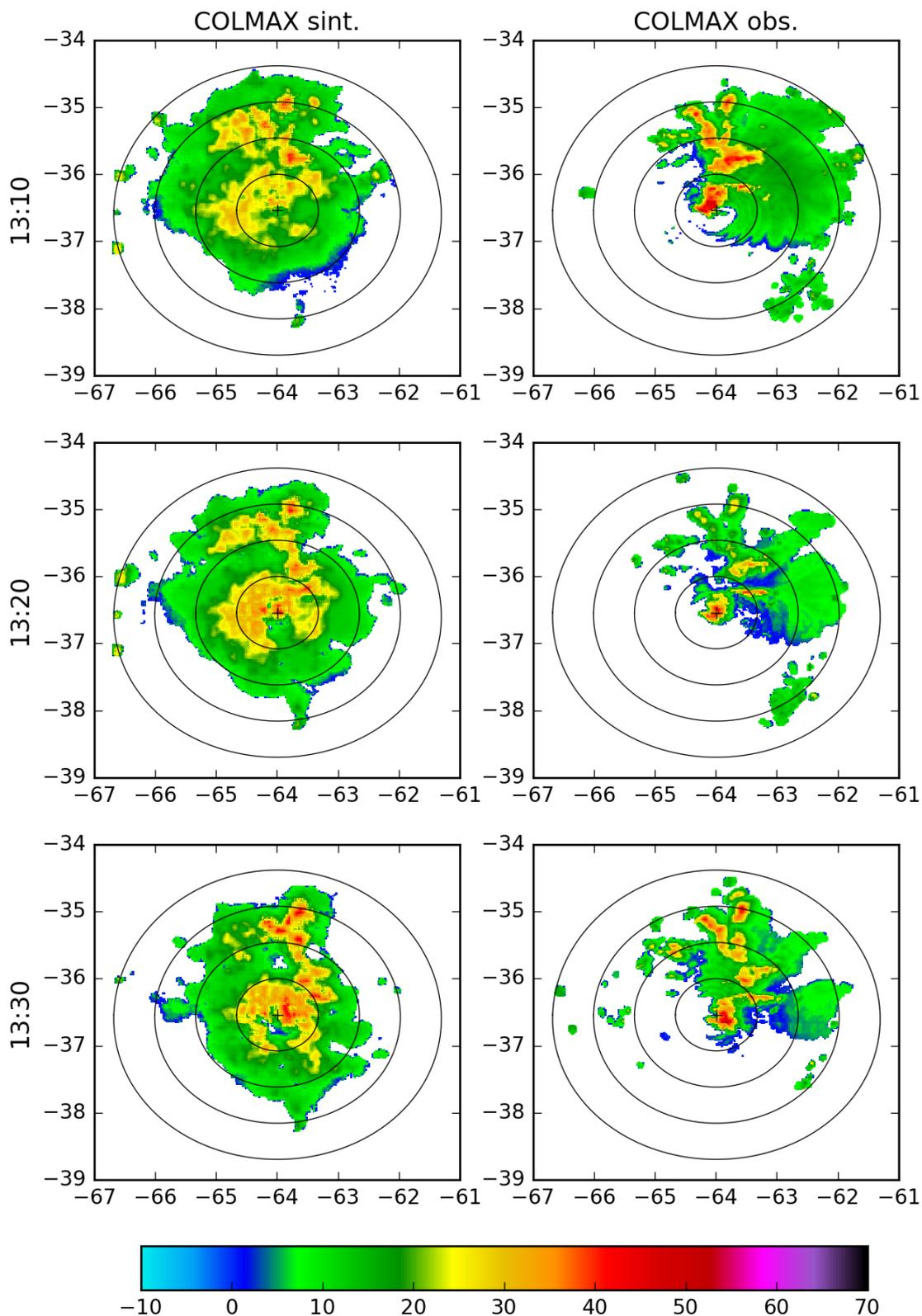


Figure 9: Secuencia de tres COLMAX consecutivos para el día 18 de febrero de 2016. Se muestra un sistema precipitante que al pasar sobre el radar atenúan la señal observada (der., COLMAX observado) pero no la señal simulada (izq., COLMAX sintético). La hora indicada es UTC. La separación entre anillos es de 60 kilómetros.

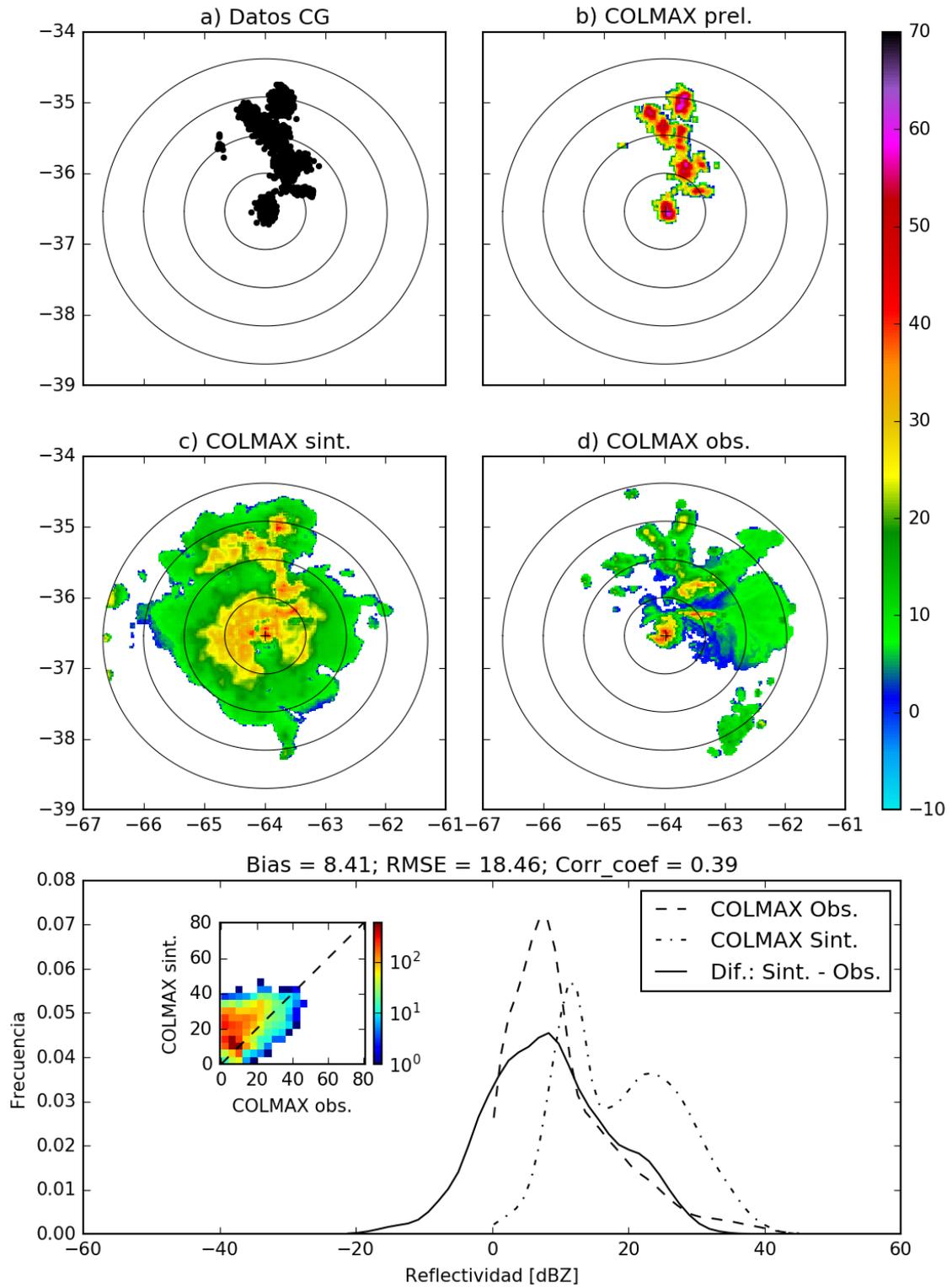


Figure 10: Ídem figura 8 pero para el día 18 de febrero de 2016 a las 13:20 UTC. Los datos mostrados en (a) corresponden al período 13:20-13:24 UTC.

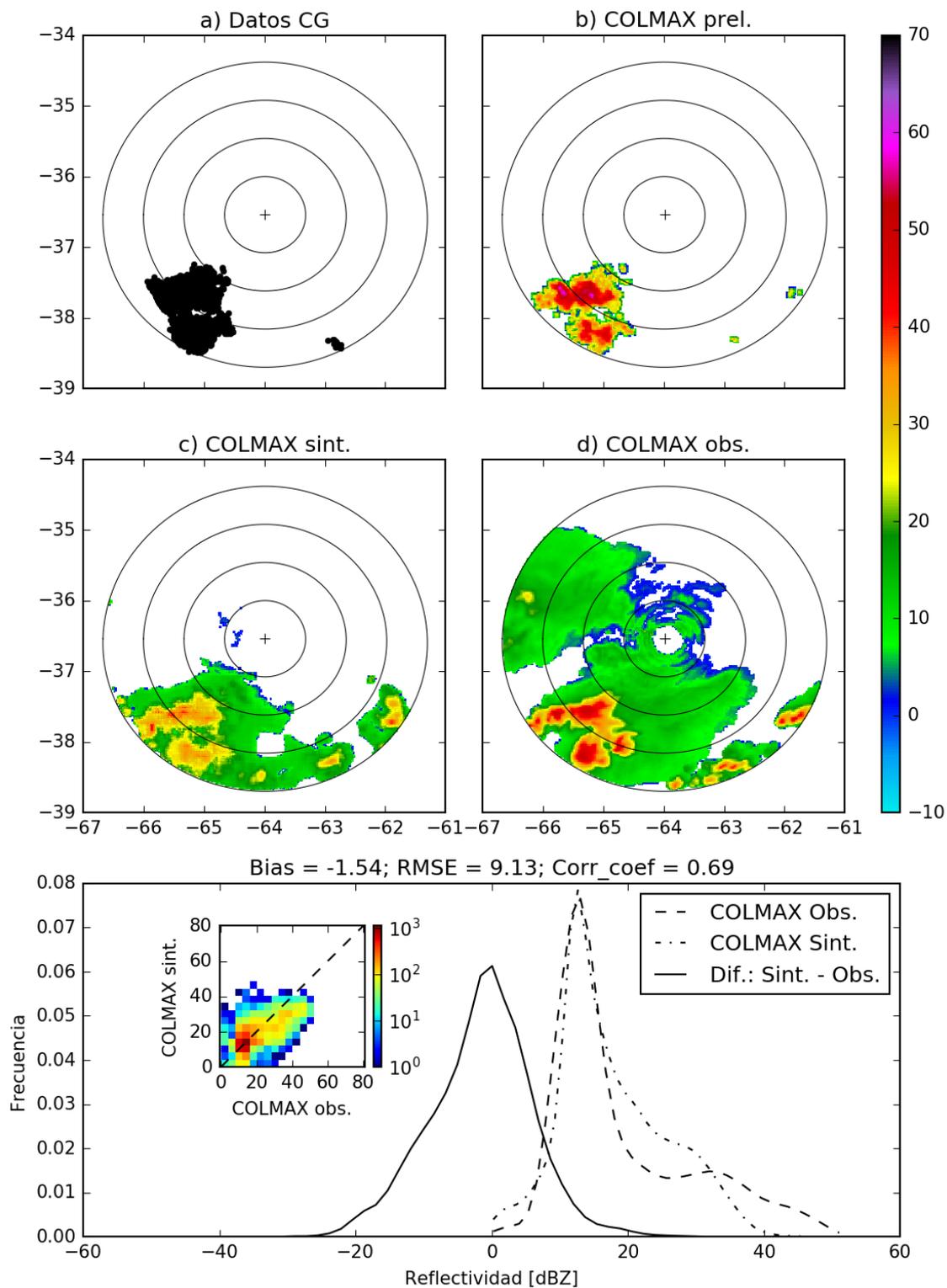


Figure 11: Ídem figura 8 pero para del día 19 de enero de 2016 a las 22:40 UTC. Los datos mostrados en (a) corresponden al período 22:40-22:42 UTC.

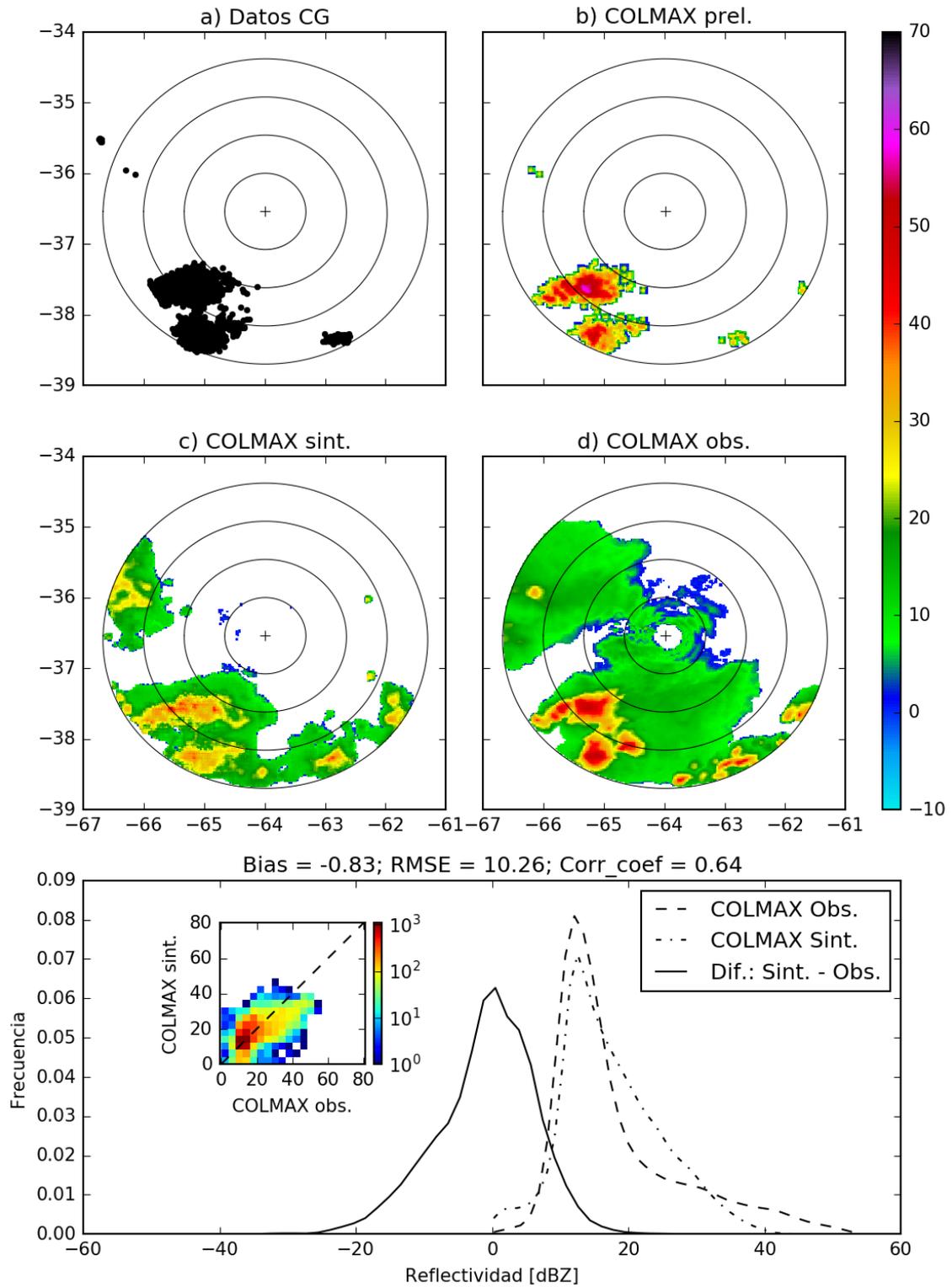


Figure 12: Ídem figura 8 pero para del día 19 de enero de 2016 a las 22:50 UTC. Los datos mostrados en (a) corresponden al período 22:50-22:54 UTC.

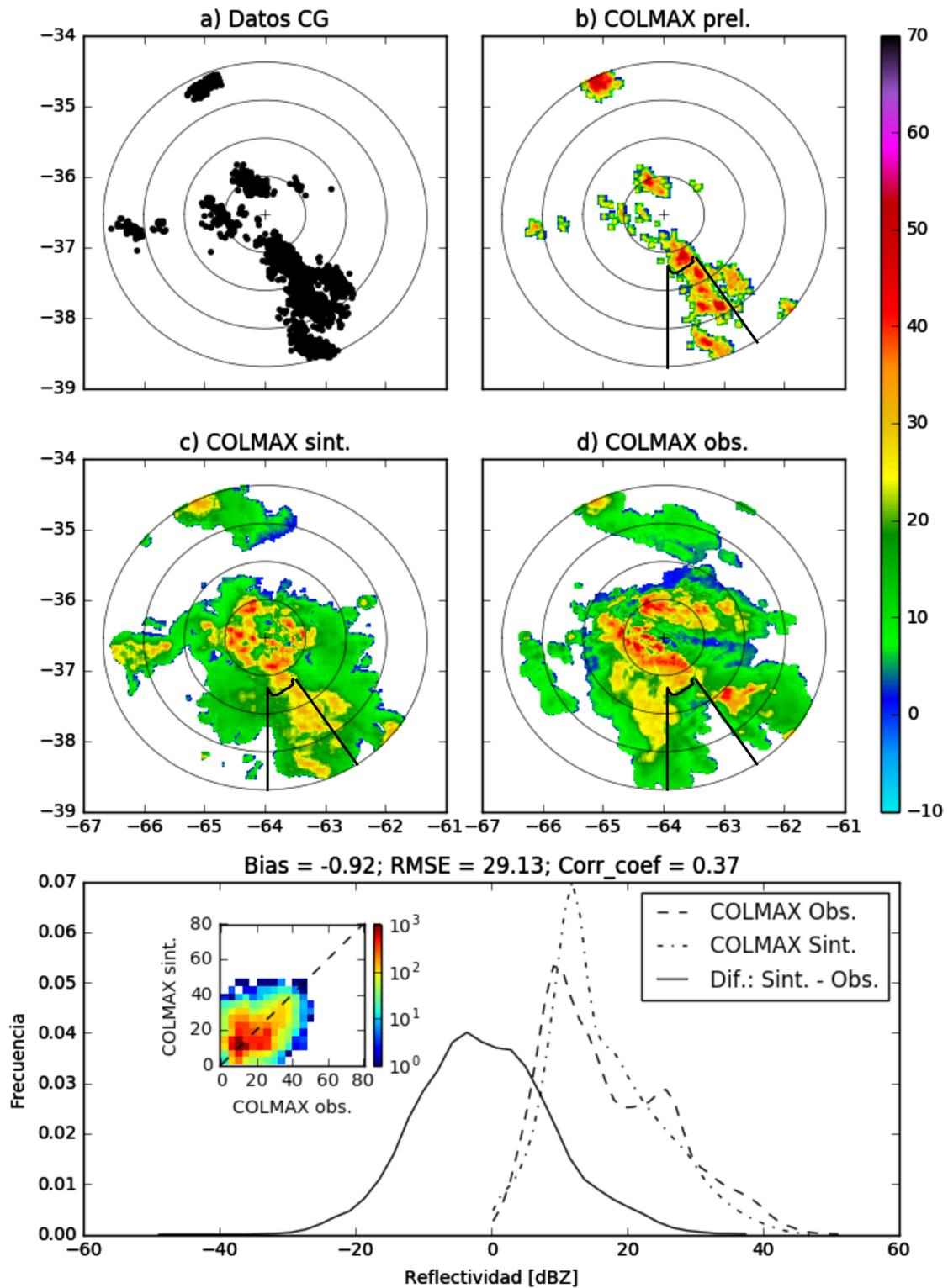


Figure 13: Ídem figura 8 pero para el 7 de febrero del 2016 a las 03:40 UTC, donde se indica con un recuadro la región donde los datos de reflectividad observada se encuentran atenuados. Los datos mostrados en (a) corresponden al período 03:40-03:44 UTC.

puede aprovecharse para desarrollar productos alternativos que complementen a los ofrecidos por los radares.

En el presente trabajo se propuso utilizar técnicas de machine learning como las redes neuronales generativas sobre datos de descargas eléctricas para lograr obtener un producto simulado de radar, denominado radar sintético. Se trabajó con datos provistos por la red global Vaisala GLD360 que detecta con una eficiencia aceptable las descargas nube-tierra (cloud-to-ground, CG) y de reflectividad provista por el radar meteorológico en banda C ubicado en la EEA INTA Anguil (La Pampa).

Los resultados encontrados son muy prometedores y son una muestra clara de la relevancia y el potencial de este tipo de técnicas de aprendizaje inteligente aplicadas en el ámbito de las ciencias de la atmósfera y en particular en la meteorología por radar. A su vez, se trata de la primera vez que se utiliza este tipo de metodología en la región para combinar dos fuentes de información con el objetivo de simular un producto crítico en las actividades operativas diarias de monitoreo y pronóstico de fenómenos meteorológicos de alto impacto tanto social como económico en el ámbito del Servicio Meteorológico Nacional de Argentina. La importancia de este radar sintético es que ofrece información del campo de reflectividad radar en cualquier lugar geográfico, tanto sobre continente como sobre el océano, donde se pueda contar datos de actividad eléctrica. Por ende, puede ser usado como un complemento a la red de radares que en su gran mayoría están desplegados sobre tierra y además cuya información bajo determinadas circunstancias discutidas con anterioridad puede encontrarse degradada. Así el área de emisión de las alertas meteorológicas a corto plazo asociadas con la ocurrencia de fenómenos severos tales como granizo y lluvias torrenciales, entre otros, se puede extender y así satisfacer las necesidades de un amplio abanico de usuarios. Por otro lado, esta técnica ofrece el potencial de utilizarse en otro tipo de aplicaciones como la estimación

cuantitativa de precipitación, donde siguiendo la misma metodología se podría encontrar una relación entre la tasa de descargas eléctricas y la intensidad de la precipitación. Ésta última podría derivarse del radar mismo como provenir de otras fuentes como las estimaciones derivadas de satélites meteorológicos.

Es importante destacar que en noviembre de 2017 se puso en órbita el sensor Geostationary Lightning Mapper (GLM) a bordo del satélite meteorológico de órbita geostacionaria GOES-16. El sensor GLM es el primero en su tipo y ofrece un monitoreo durante las 24 horas los 365 días del año de la actividad eléctrica total sobre el continente americano (Goodman y otros, 2013). De este modo, los datos provistos por dicho sensor podrán ser utilizados como entrada de la herramienta desarrollada y así generar el producto COLMAX sintético para una región mucho más amplia.

El trabajo futuro estará centrado en probar este modelo sobre otros radares de la red argentina estudiando las relaciones entre la tasa de descargas y la reflectividad radar (primer paso para obtener el radar preliminar) con el objetivo de analizar si hay cambios a nivel regional dados por los diferentes tipos de sistemas precipitantes que se dan en el centro y norte de Argentina. Este es un paso importante con el fin último de extender este producto a todo el país. Por otro lado se trabajará en evaluar los datos del sensor GLM como fuente primaria de entrada de datos de actividad eléctrica en nuestra región. La información provista por este sensor es diferente a los datos de GLD360. Mientras este último detecta mayormente descargas CG a partir de antenas posicionadas en superficie, el sensor GLM detecta la actividad eléctrica total, es decir, no tiene la capacidad de separar entre descargas CG, intra-nube o nube-nube. Entonces la pregunta a responder será si la tasa de descargas eléctricas totales está relacionada de igual manera con la reflectividad radar integrada en la vertical como si lo está la tasa de descargas CG a partir de los resultados mostrados en el presente trabajo.

Por último, entre el 1 de noviembre y el 18 de diciembre de 2018 se realizó el experimento de campo RELÁMPAGO en las Sierras de Córdoba, donde se desplegó un sinfín de instrumental entre ellos la red LMA (Lightning Mapping Array) con el objetivo de validar los datos del sensor GLM. Esta red ofrece datos tridimensionales de la actividad eléctrica con una gran resolución espacial y temporal. De este modo, en un futuro cercano se podría contar con esta información para evaluar en algún caso de estudio la metodología propuesta. Por su parte, recientemente en el Servicio Meteorológico Nacional se cuenta con datos provenientes de la red Earth Network cuya eficiencia de detección en la región centro-este de Argentina será del 70 % mientras que el error medio de localización rondará los 0,5 km.

Agradecimientos: Este trabajo se realizó en el marco del Proyecto PIDDEF N°5/2014 (Ministerio de Defensa) mediante una Beca de Capacitación Profesional. A los proyectos PDTs ALERT.AR, PICT 2013-1299, UBACyT 20020130100618BA y UBACyT 20020170100164BA.

REFERENCIAS

- Atkinson, P.M. and Tatnall, A.R.L., 1997: Introduction neural networks in remote sensing. *International Journal of Remote Sensing*, 18 (4), pp. 699-709.
- Atlas, D. y Banks, H.C., 1951: The interpretation of microwave reflections from rainfall. *Journal of Meteorology*, 8 (5), pp. 271-282.
- Atlas, D., D. Rosenfeld, and D. B. Wolff, 1990: Climatologically Tuned Reflectivity-Rain Rate Relations and Links to Area-Time Integrals. *Journal of Applied Meteorology and Climatology*, 29, 1120-1135.
- Atlas, D., D. Rosenfeld, and D. B. Wolff, 1993: C-Band Attenuation by Tropical Rainfall in Darwin, Australia, Using Climatologically Tuned Ze-R Relations. *Journal of Applied Meteorology and Climatology*, 32, 426-430.
- Bishop, C. and Bishop, C.M., 1995: Neural networks for pattern recognition. Oxford university press.
- Borland, D. y Li, R.M.T., 2007: Rainbow color map (still) considered harmful. *IEEE computer graphics and applications*, 27(2).
- Bridenstine, P.V., Darden, C.B., Burks, J. y Goodman, S.J., 2005: The application of total lightning data in the warning decision making process. Preprints. Conf. on Meteorological Applications of Lightning Data, San Diego, CA, Amer. Meteor. Soc., P1.2. [Disponible online en <http://ams.confex.com/ams/pdfpapers/83037.pdf>].
- Bringi, V.N., Keenan, T.D. and Chandrasekar, V., 2001: Correcting C-band radar reflectivity and differential reflectivity data for rain attenuation: A self-consistent method with constraints. *IEEE Transactions on Geoscience and Remote Sensing*, 39(9), pp.1906-1915.
- Calheiros, R. V., and I. Zawadzki, 1987: Reflectivity-Rain Rate Relationships for Radar Hydrology in Brazil. *Journal of Applied Meteorology and Climatology*, 26, 118-132.
- Carey, L.D., Rutledge, S.A., Ahijevych, D.A. and Keenan, T.D., 2000: Correcting propagation effects in C-band polarimetric radar observations of tropical convection using differential propagation phase. *Journal of Applied Meteorology*, 39 (9), pp. 1405-1433.
- Cecil, D.J. y Blankenship, C.B. , 2012: Toward a global climatology of severe hailstorms as estimated by satellite passive microwave imagers. *Journal of Climate*, 25 (2), pp. 687-703.
- Chang, D.-E., J. A. Weinman, C. A. Morales, and W. S. Olson, 2001: The Effect of Spaceborne Microwave and Ground-Based Continuous Lightning Measurements on Forecasts of the 1998 Groundhog Day Storm. *Monthly Weather Review*, 129, 1809-1833.
- Cherkassky, V., Krasnopolsky, V., Solomatine, D.P. and Valdes, J., 2006: Computational intelligence in earth sciences and environmental applications: Issues and challenges. *Neural Networks*, 19, 113-121.
- Crosson, W. L., C. E. Duchon, R. Raghavan, and S. J. Goodman, 1996: Assessment of Rainfall Estimates Using a Standard Z-R

- Relationship and the Probability Matching Method Applied to Composite Radar Data in Central Florida. *Journal of Applied Meteorology and Climatology*, 35, 1203-1219.
- Dance, S., Ebert, E. and Scurrah, D., 2010: Thunderstorm strike probability nowcasting. *Journal of Atmospheric and Oceanic Technology*, 27, 79-93.
- Delrieu, G., Huc, L. and Creutin, J.D., 1999: Attenuation in rain for X- and C-band weather radar systems: Sensitivity with respect to the drop size distribution. *Journal of Applied Meteorology and Climatology*, 38, 57-68.
- Demetriades, N.W.S., M.J. Murphy y J.A. Cramer, 2010: Validation of Vaisala's Global Lightning Dataset (GLD360) over the continental United States, paper presented at 21st Int. Light. Detect. Conf. (ILDC), Orlando, Fla.
- Gagne, D.J., A. McGovern, S.E. Haupt, R.A. Sobash, J.K. Williams, and M. Xue, 2017: Storm-Based Probabilistic Hail Forecasting with Machine Learning Applied to Convection-Allowing Ensembles. *Weather and Forecasting*, 32, 1819-1840.
- Gatlin, P., 2006: Severe weather precursors in the lightning activity of Tennessee Valley thunderstorms. M.S. thesis, The University of Alabama in Huntsville, 87 pp.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. y Bengio, Y., 2014: Generative adversarial nets. In *Advances in neural information processing systems* (pp. 2672-2680).
- Goodman, S.J., Blakeslee, R., Christian, H., Koshak, W., Bailey, J., Hall, J., McCaul, E., Buechler, D., Darden, C., Burks, J. y Bradshaw, T., 2005: The North Alabama Lightning Mapping Array: Recent severe storm observations and future prospects. *Geophysical Research Letters*, 15, 1185-1188.
- Goodman, S.J., R. Blakeslee, W. Koshak, D. Mach, J. Bailey, D. Buechler, L. Carey, C. Schultz, M. Bateman, E. McCaul, and G. Stano, 2013: The GOES-R Geostationary Lightning Mapper (GLM). *Atmospheric Research* 125-126, 34-49.
- Gu, J.Y., Ryzhkov, A., Zhang, P., Neilley, P., Knight, M., Wolf, B. and Lee, D.I., 2011: Polarimetric attenuation correction in heavy rain at C band. *Journal of Applied Meteorology and Climatology*, 50, 39-58.
- Grubišić, V., Vellore, R.K. and Huggins, A.W., 2005: Quantitative precipitation forecasting of wintertime storms in the Sierra Nevada: Sensitivity to the microphysical parameterization and horizontal resolution. *Monthly Weather Review*, 133, 2834-2859.
- Helmus, J. and Collis, S., 2016: The Python ARM Radar Toolkit (Py-ART), a library for working with weather radar data in the Python programming language. *Journal of Open Research Software*, 4 (1).
- Hitschfeld, W. y Bordan, J., 1954: Errors inherent in the radar measurement of rainfall at attenuating wavelengths. *Journal of Meteorology*, 11, 58-67.
- Hong, Y., K. Hsu, S. Sorooshian, y X. Gao, 2004: Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *Journal of Applied Meteorology and Climatology*, 43, 1834-1853.
- Iskenderian, H., 2008: Cloud-to-ground lightning as a proxy for nowcasts of VIL and echo tops. In *13th Conference on Aviation, Range and Aerospace Meteorology* (pp. 21-24).
- Isola, P., Zhu, J.Y., Zhou, T. y Efros, A.A., 2016: Image-to-image translation with conditional adversarial networks. *arXiv preprint arXiv:1611.07004*.
- Krasnopolsky, V.M. and Schiller, H., 2003: Some neural network applications in environmental sciences. Part I: forward and inverse problems in geophysical remote measurements. *Neural Networks*, 16 (3-4), 321-334.
- LeCun, Y., Bengio, Y. and Hinton, G., 2015: Deep learning. *Nature*, 521 (7553), 436.
- Liu, Y., Racah, E., Correa, J., Khosrowshahi, A., Lavers, D., Kunkel, K., Wehner, M. and Collins, W., 2016: Application of deep convolutional neural networks for detecting extreme weather in climate datasets. *arXiv preprint arXiv:1605.01156*.
- Mallick, S., y coautores, 2014: Evaluation of the GLD360 performance characteristics

- using rocket-and-wire triggered lightning data. *Geophysical Research Letters*, 41, 3636–3642.
- Marks, F. D., D. Atlas, and P. T. Willis, 1993: Probability-matched Reflectivity-Rainfall Relations for a Hurricane from Aircraft Observations. *Journal of Applied Meteorology and Climatology*, 32, 1134-1141.
- Marzban, C., y G. Stumpf, 1996: A neural network for tornado prediction based on Doppler radar-derived attributes. *Journal of Applied Meteorology and Climatology*, 35, 617–626.
- Marzban, C., y A. Witt, 2001: A Bayesian neural network for severe-hail size prediction. *Weather and Forecasting*, 16,600–610.
- Marzoug, M. y Amayenc, P., 1991: Improved range-profiling algorithm of rainfall rate from a spaceborne radar with path-integrated attenuation constraint. *IEEE Transactions on Geoscience and Remote Sensing*, 29, 584-592.
- Matsudo, C.M. and Salio, P.V., 2011: Severe weather reports and proximity to deep convection over Northern Argentina. *Atmospheric research*, 100, 523-537.
- Megenhardt, D., Mueller, C.K., Rehak, N. and Cunning, G., 2000: Evaluation of the national convective weather forecast product. In *Conference on Aviation, Range, and Aerospace Meteorology*, 9 th, Orlando, FL (pp. 171-176).
- Megenhardt, D., Mueller, C., Trier, S., Ahijevych, D. y Rehak, N., 2004: NCWF-2 probabilistic nowcasts. In *11th Conference on Aviation, Range, and Aerospace Meteorology*, Hyannis, Massachusetts, 4–8 Oct 2004.
- Mezher, R.N., Doyle, M. and Barros, V., 2012: Climatology of hail in Argentina. *Atmospheric Research*, 114, 70-82.
- Monahan AH, Fyfe JC, Ambaum MHP, Stephenson DB, North GR., 2009: Empirical orthogonal functions: the medium is the message. *Journal of Climate*, 22, 6501–6514.
- Morales, C.A. and E.N. Anagnostou, 2003: Extending the Capabilities of High-Frequency Rainfall Estimation from Geostationary-Based Satellite Infrared via a Network of Long-Range Lightning Observations. *Journal of Hydrometeorology*, 4, 141–159.
- Mueller, C. K., C. B. Fidalgo, D. W. McCann, D. Meganhardt, N. Rehak, T. Carty, 1999: National Convective Weather Forecast Product, Preprints, Eighth Conference on Aviation, Range, and Aerospace Meteorology, Dallas, TX, 10-15 January 1999, American Meteorological Society.
- Naccarato, K. P., O. Pinto Jr., S. A. M. Garcia, M. Murphy, N. Demetriades, and J. Cramer, 2010: Validation of the new GLD360 dataset in Brazil: First results, paper presented at 21st Int. Light. Detect. Conf. (ILDC), Orlando, Fla.
- Paliwal, M. and Kumar, U.A., 2009: Neural networks and statistical techniques: A review of applications. *Expert Systems With Applications*, 36, 2-17.
- Piotrowski, A.P. and Napiorkowski, J.J., 2013: A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modelling. *Journal of Hydrology*, 476, 97-111.
- Pohjola, H., and A. Makela, 2013: The comparison of GLD360 and EUCLID lightning location systems in Europe, *Atmospheric Research*, 123, 117–128.
- Prabhat, Oliver Rubel, Surendra Byna, Kesheng Wu, Fuyu Li, Michael Wehner, and Wes Bethel, 2012: Teca: A parallel toolkit for extreme climate analysis. ICCS, 2012.
- Prabhat, Surendra Byna, Venkatram Vishwanath, Eli Dart, Michael Wehner, and William D. Collins, 2015: Teca: Petascale pattern recognition for climate science. CAIP, 2015.
- Rasmussen, K.L., Zuluaga, M.D. y Houze, R.A., 2014: Severe convection and lightning in subtropical South America. *Geophysical Research Letters*, 41, 7359-7366.
- Rosenfeld, D., D. B. Wolff, and D. Atlas, 1993: General Probability-matched Relations between Radar Reflectivity and Rain Rate. *Journal of Applied Meteorology and Climatology*, 32, 50-72.
- Ruder, S., 2016: An overview of gradient descent optimization algorithms. arXiv preprint arXiv:1609.04747.

- Rudlosky, S.D., M.J. Peterson, and D.T. Kahn, 2017: GLD360 Performance Relative to TRMM LIS. *J. Atmos. Oceanic Technol.*, 34, 1307–1322.
- Said, R. K., U. S. Inan, and K. L. Cummins, 2010: Long-range lightning geolocation using a VLF radio atmospheric waveform bank, *Journal of Geophysical Research*, 115, D23108.
- Said, R. K., and M. Murphy, 2016: GLD360 upgrade: Performance analysis and applications. *Extended Abstracts*, 24th Int. Lightning Detection Conf. and Sixth Int. Lightning Meteorology Conf., San Diego, CA, Vaisala, 8 pp.
- Sénési, S., Guillou, Y., Forster, C. and Tafferner, A., 2009: Cb nowcasting in FLYSAFE: improving flight safety regarding thunderstorm hazards, in: WMO Symposium on Nowcasting, Whistler, B.C., Canada, 30 August–4 September 2009, No. 4.12, 14 pp., 2009.
- Schultz, C.J., Petersen, W.A. and Carey, L.D., 2011: Lightning and severe weather: A comparison between total and cloud-to-ground lightning trends. *Weather and Forecasting*, 26, 744-755.
- Soula, S. y Chauzy, S., 2001: Some aspects of the correlation between lightning and rain activities in thunderstorms. *Atmospheric Research*, 56, 355-373.
- Steiger, S. M., Orville R. E., Murphy M. J., Demetriades N. W. S., 2005: Total lightning and radar characteristics of supercells: Insights on electrification and severe weather forecasting. *Preprints, Conf. on Meteorological Applications of Lightning Data*, San Diego, CA, Amer. Meteor. Soc., P1.7. [Disponible online en <http://ams.confex.com/ams/pdfpapers/84908.pdf>.]
- Steiger, S.M., Orville, R.E. y Carey, L.D., 2007: Total lightning signatures of thunderstorm intensity over north Texas. Part I: Supercells. *Monthly Weather Review*, 135, 3281-3302.
- Steinhaeuser, K., Chawla, N.V. and Ganguly, A.R., 2011: Comparing predictive power in climate data: Clustering matters. In *International Symposium on Spatial and Temporal Databases* (pp. 39-55). Springer, Berlin, Heidelberg.
- Tafferner, A., C. Forster, S. Senesi, Y. Guillou, P. Tabary, P. Laroche, A. Delannoy, B. Lunnun, D. Turp, T. Hauf, D. Makovic, 2008: Nowcasting thunderstorm hazards for flight operations: the CB WIMS approach in FLYSAFE. *ICAS 2008 Conference, International Council of the Aeronautical Sciences Conf. Proc.*(8.6. 2), Optimage Ltd., Edinburgh, UK.
- Tapia, A., Smith, J.A. y Dixon, M., 1998: Estimation of convective rainfall from lightning observations. *Journal of Applied Meteorology*, 37, 1497-1509.
- Weber, M.E., Williams, E.R., Wolfson, M.M. y Goodman, S.J., 1998: An Assessment of the Operational Utility of a GOES Lightning Map Sensor. *Project Report NOAA-18*, Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, MA.
- Wiens, K.C., Rutledge, S.A. y Tessendorf, S.A., 2005: The 29 June 2000 supercell observed during STEPS. Part II: Lightning and charge structure. *Journal of the atmospheric sciences*, 62, 4151-4177.
- Xingjian, SHI, Chen, Zhouong, Wang, Hao, Yeung, Dit-Yan, Wong, Wai-kin, and Woo, Wangchun, 2015: Convolutional lstm network: A machine learning approach for precipitation nowcasting. In *Advances in Neural Information Processing Systems*.
- Xu, W., R.F. Adler, and N. Wang, 2014: Combining Satellite Infrared and Lightning Information to Estimate Warm-Season Convective and Stratiform Rainfall. *Journal of Applied Meteorology and Climatology*, 53, 180–199.
- Zhou, Y., Qie, X. y Soula, S., 2002: A study of the relationship between cloud-to-ground lightning and precipitation in the convective weather system in China. *Annales Geophysicae* 20, 107-113.

Este es un artículo de acceso abierto distribuido bajo la licencia Creative Commons, que permite el uso ilimitado, distribución y reproducción en cualquier medio, siempre que la obra original sea debidamente citada.