

VERIFICACIÓN ESPACIAL DE PRONÓSTICOS DE PRECIPITACIÓN EN ALTA RESOLUCIÓN PARA LA REGIÓN SUR DE SUDAMÉRICA

Natalí Giselle Aranda¹, Yanina García Skabar^{1,2}, Cynthia Mariana Matsudo¹

¹Servicio Meteorológico Nacional, Buenos Aires, Argentina. ²CONICET, Buenos Aires, Argentina.

(Manuscrito recibido el 29 de marzo de 2021, en su versión final el 21 de septiembre de 2021)

RESUMEN

Dada la creciente demanda y utilización de modelos numéricos para pronosticar la ocurrencia de precipitación, es necesario contar con métodos de verificación que permitan cuantificar la calidad de los pronósticos, a partir de la evaluación de sus errores y sesgos. El método de evaluación diagnóstica basado en objetos (MODE, por sus siglas en inglés), es una herramienta de verificación espacial que identifica regiones de interés, como áreas con precipitación, y define objetos en los campos del pronóstico y de la observación a partir de parámetros definidos por el usuario. MODE fue utilizado para evaluar el desempeño de los pronósticos de precipitación del Weather and Research Forecasting Model (WRF), con resolución espacial de 4 km, sobre el sur de Sudamérica. Los pronósticos fueron comparados con la estimación de precipitación, Integrated Multi-Satellite Retrievals Final Run (IMERG-F), producto derivado de la misión satelital de medición global de precipitación (GPM, por sus siglas en inglés). Para el período de un mes se evaluó la sensibilidad a la elección de los parámetros de MODE, radio de convolución y umbral de intensidad para la precipitación acumulada a 3 y 24 horas. El período completo de verificación considerado abarcó los años 2017-2018 y además se utilizaron técnicas tradicionales de verificación (ej, Probabilidad de detección, Falsas Alarmas). Adicionalmente, los pronósticos de precipitación acumulada en 24 horas del WRF se compararon con los del modelo global Global Forecast System (GFS). En este estudio se pudo comprobar que los métodos tradicionales de verificación contribuyen a conocer la calidad de un pronóstico de precipitación en forma objetiva. Por otro lado, la verificación por objetos permite evaluar cómo los pronósticos logran representar la posición, tamaño e intensidad del fenómeno de interés. Respecto de los modelos analizados, tanto el WRF como el GFS presentan gran cantidad de eventos sorpresas y falsas alarmas, pero también se destacan porque sus aciertos presentan bajos errores de posición e intensidad de precipitación.

Palabras clave: pronósticos de precipitación, verificación espacial, MODE, WRF, GFS.

SPATIAL VERIFICATION OF HIGH RESOLUTION PRECIPITATION FORECASTS OVER SOUTHERN SOUTH AMERICA

ABSTRACT

Given that there is an increasing demand and use of numerical models to forecast precipitation events, it is essential to advance in the use of different verification methods to measure the quality of the forecasts with the evaluation of errors and biases. The method for object-based diagnostic evaluation (MODE) is a spatial verification method that identifies regions of interest, like precipitation, in the same way that a human would do. This method defines objects in the forecast and observation fields based on user-defined parameters. MODE was used to evaluate the performance of 4-km hourly precipitation forecasts from the Weather and Research Forecasting Model (WRF) over southern South America against the Global Precipitation Measurement (GPM) derived product IMERG Final Run version (IMERG-F). For a one month period, tests were performed to select the values for threshold and the radius of convolution parameters adequate for 3 and 24 hour accumulated precipitation. The whole verification period considered was 2017-2018 and furthermore, traditional verification statistics (eg, Probability of Detection, False Alarm Ratio) were used. Additionally, 24-hour accumulated precipitation forecasts from WRF were compared with those from the Global Forecast System (GFS). This study proved that traditional verification methods allow objectively to know the quality of precipitation forecasts. Conversely, object verification rather than making a pointwise evaluation of hits and misses, identifies precipitation patterns and compares attributes describing position, size and intensity of matched forecasted and observed objects. Regarding the analyzed models, although WRF and GFS present many surprises and false alarms, hit events present low errors associated with location and intensity of precipitation.

Keywords: precipitation forecast, spatial verification, MODE, WRF, GFS.

1. INTRODUCCIÓN

El pronóstico de precipitación, especialmente en los casos de precipitación intensa asociados a tormentas, representa un desafío científico y tecnológico muy complejo. Uno de los métodos para pronosticar su ocurrencia es la utilización de modelos numéricos de la atmósfera en muy alta resolución. Debido a la importancia que conlleva el pronóstico meteorológico, su verificación resulta indispensable para evaluar la calidad de los mismos.

Uno de los métodos para verificar pronósticos de precipitación es la verificación puntual, en donde a partir de una evaluación punto a punto entre el campo pronosticado y observado se comprueba si la precipitación ocurrió o no. Si bien el aumento de la resolución de los modelos meteorológicos mejoró el rendimiento de los pronósticos, a veces este avance no

es captado correctamente por los métodos tradicionales de verificación que no distinguen los errores asociados al desplazamiento del evento o a la extensión de la región precipitante (Brown y otros, 2004; Roebber y otros, 2004; Guilleland y otros, 2020). Es así que en los últimos años se desarrollaron técnicas que contemplan la verificación espacial y entre ellas se encuentra el método de evaluación diagnóstica basado en objetos (MODE, por sus siglas en inglés). MODE está basado en la teoría desarrollada por Davis y otros (2006a) y permite, mediante resultados gráficos y estadísticos, describir y comparar los objetos pronosticados con los observados (Brown y otros, 2009). Esta metodología fue utilizada en diferentes partes del mundo y para diferentes aplicaciones. Mittermaier y Bullock (2013) y Griffin y otros (2017) utilizaron MODE para evaluar pronósticos de cobertura nubosa. El primer estudio demostró que el número de

objetos en los pronósticos de nubosidad sobre el Reino Unido depende de los parámetros elegidos en la verificación espacial, influyendo en el número de aciertos, falsas alarmas y sorpresas. En el segundo estudio se observó que los pronósticos de nubosidad en Estados Unidos son más exactos en verano ya que los objetos pronosticados representan mejor la ubicación y el área de los objetos. Mittermaier y otros (2020) emplearon la verificación basada en objetos para evaluar concentraciones de clorofila sobre el noreste del Océano Atlántico. La metodología les permitió detectar que los modelos oceánicos empleados tienden a sobreestimar la cantidad de objetos, situación que podría mejorarse si se aumentara la resolución espacial de dichos modelos.

Johnson y otros (2013) compararon los resultados de la verificación por MODE entre modelos con distinta resolución espacial: la mejora en la resolución de los pronósticos permite encontrar objetos de menor tamaño y al igual que Kain y otros (2008), sostienen que los errores espaciales de los pronósticos de mayor resolución se deben a la propia resolución del modelo y a sus condiciones de borde. Wolff y otros (2014) demostraron que la verificación espacial en los pronósticos de precipitación en alta resolución, a diferencia de los métodos tradicionales de verificación, permite identificar errores de ubicación, extensión, y desplazamiento. Moser y otros (2015) estudiaron los efectos de la asimilación de datos de radar para mejorar los pronósticos de precipitación en alta resolución. Para ello, MODE fue utilizado en situaciones particulares, permitiendo observar que la exactitud en la predicción de la posición, área e intensidad de la precipitación varía de acuerdo al caso de estudio seleccionado. Los trabajos de Johnson y otros (2013), Wolff y otros (2014) y de Moser y otros (2015), coinciden en que la elección de los parámetros que definen los objetos depende de cada caso de estudio, y por lo tanto la generalización de los resultados a otras aplicaciones se ve limitada. Skinner y otros (2018), y posteriormente Flora y otros (2019), se

basaron en la técnica de MODE para verificar tormentas y mesociclones. Ambos estudios señalaron la habilidad de MODE para evaluar pronósticos y que una incorrecta elección del umbral de intensidad influye en el número de objetos identificados, alterando los valores de los indicadores de verificación. Por su parte, Gilleland (2020) evaluó el desempeño de dos nuevos índices de verificación para intensidades de precipitación de modelos en alta resolución, inspirado en las propiedades de los objetos identificados por MODE.

En particular en Sudamérica se pueden mencionar el trabajo de Charo y otros (2014), donde aplicaron el método de MODE para estudiar un caso de precipitación intensa. En dicho trabajo, MODE permite analizar desde otro punto la calidad del pronóstico de precipitación ya que no solo considera la intensidad de la precipitación pronosticada, sino también el área cubierta, la forma y la orientación del área. Bender y Ynoue (2014) utilizaron índices de la verificación puntual y de MODE para verificar los pronósticos de precipitación del modelo Global Forecast System (GFS) sobre San Pablo y alrededores. Se observó que durante la época lluviosa, el GFS presenta un mejor desempeño y que gracias a la implementación de MODE, pudo detectarse que una de las causas de los errores en los pronósticos se debe a la sobreestimación del área y de la intensidad de precipitación. Posteriormente, Carrasco (2017) avanzó en una metodología basada en MODE para evaluar la precipitación en América del Sur, para lo cual se llevaron a cabo experimentos utilizando las predicciones BRAMS y que fueron comparadas con las predicciones GFS. La mayor proporción de errores estuvo relacionada a eventos no pronosticados y la verificación espacial resultó de utilidad para caracterizar patrones espaciales.

En el presente trabajo se describe y aplica la metodología de identificación de objetos con MODE, con el objetivo de realizar una verificación del desempeño de los pronósticos

de precipitación de alta resolución del modelo Weather Research and Forecasting Model (WRF) operativo en el Servicio Meteorológico Nacional (SMN) para la región sur de Sudamérica, para el periodo 2017-2018. En la sección 2 se muestran los datos y la metodología empleada. En la sección 3, los resultados de la verificación para la precipitación acumulada en 24 y 3 horas, y se analizan las ventajas y limitaciones de la verificación puntual y de MODE. También se compara el desempeño de los pronósticos de precipitación acumulada en 24 horas del WRF con los del modelo GFS. Finalmente, en la sección 4 se presentan las conclusiones.

2. DATOS Y METODOLOGIA

2.1. MODELO WRF, GFS Y ESTIMACIONES SATELITALES DE PRECIPITACIÓN

Los pronósticos numéricos verificados son los generados para Argentina en forma operativa en el SMN con el modelo WRF en su versión con núcleo dinámico Advanced Research WRF (ARW) versión 3.8 (Skamarock y otros, 2005). Los pronósticos tienen una resolución horizontal de 4 km y vertical de 38 niveles con el tope en 50 hPa. Su proyección es Lambert Conformal y su dominio abarca toda la región sur de Sudamérica y los océanos adyacentes tal como se muestra en la Figura 1. La convección se resuelve de manera explícita y se generan cuatro ciclos diarios de pronósticos, inicializados a las 00, 06, 12 y 18 UTC, todos por un plazo de 48 horas, produciendo pronósticos en forma horaria. Los mismos se inicializan a partir de los pronósticos del GFS con una resolución horizontal de 0.25° incorporándose como condiciones de borde en forma horaria (García Skabar y otros, 2018). En este trabajo, se verificaron los pronósticos del WRF inicializados a las 06 UTC por un período de dos años comprendido entre el 1 de enero de 2017 hasta el 31 de diciembre de 2018. Por un lado, se analizó el desempeño del pronóstico de precipitación acumulada a 3 horas desde el inicio de la simulación hasta las 48 horas. Por

otro lado, se analizó la precipitación acumulada en 24 horas (entre las 12 UTC del día y las 12 UTC día siguiente) correspondiente a un plazo de pronóstico entre 6 y 30 horas. El GFS es un modelo meteorológico global producido por el National Centers for Environmental Prediction (NCEP). Los pronósticos del GFS están disponibles con una resolución espacial de 0.25° y temporal de 1 hora, con cobertura global. El GFS cuenta con cuatro ciclos diarios de pronósticos (0, 06, 12 y 18 UTC), de los cuales, al igual que con el WRF, se utilizó el ciclo inicializado a las 06 UTC y se analizó la precipitación acumulada entre las 12 UTC del día y las 12 UTC día siguiente, correspondiente a un plazo de pronóstico entre 6 y 30 horas.

Los pronósticos de precipitación del WRF y del GFS fueron comparados con las estimaciones satelitales de precipitación Integrated Multi-Satellite Retrievals generadas a partir de los datos provenientes de la misión de Global Precipitation Measurement, GPM, en su versión Final Run (Huffman y otros, 2015), denominados IMERG-F de aquí en adelante. Esta estimación cuenta con datos calibrados con estaciones pluviométricas de superficie y presenta proyección regular con dominio casi global (60° N a S), resolución espacial de 0.1° (equivalente a 10 km) y resolución temporal de media hora. En estudios realizados para la región sur de la Cordillera de los Andes, la versión IMERG-F presentó mejores resultados que otras estimaciones para casos de precipitación intensa (Hobouchian y otros, 2017). Asimismo, el procesamiento adicional de IMERG-F mejora el desempeño de la estimación frente a otras versiones de IMERG de menor latencia (Foelsche y otros, 2017; Tan y otros, 2017; Wang y otros 2017).

2.2. MODE

El software Model Evaluation Tools (MET) fue desarrollado por el National Center for Atmospheric Research (NCAR) con el fin de realizar verificaciones de pronósticos numéricos de la atmósfera. Este software es de distribución

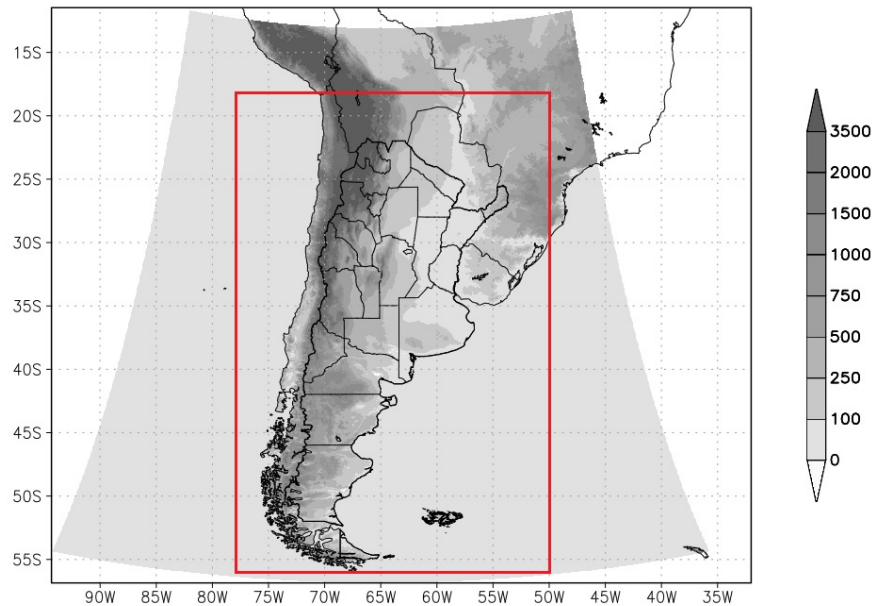


Figura 1: La región sombreada representa el dominio del pronóstico del modelo WRF y la escala de grises indica la topografía en metros. El rectángulo rojo indica el área de verificación.

libre y puede encontrarse en la página web <http://www.dtcenter.org/met/users/>. Entre sus funciones se incluye la verificación punto a punto de las variables meteorológicas, el cálculo de estadísticos utilizados tradicionalmente en meteorología (por ejemplo desvío, raíz cuadrada del error cuadrático medio, critical success index) y módulos de verificación espacial como la verificación basada en objetos MODE, método utilizado en este trabajo (Brown y otros, 2009; Fowler y otros 2017). MODE considera a los objetos como regiones de interés (por ejemplo, zonas de precipitación). Estos objetos se definen a partir del radio de convolución (R) y el umbral de intensidad de precipitación (T). En el proceso de identificación de objetos, primero se realiza un suavizado del campo original de precipitación a partir de la elección de R. Cuanto mayor sea R, mayor será el suavizado y puede ocurrir que haya sistemas de precipitación más pequeños que no sean identificados, prevaleciendo así los objetos más grandes. Luego, se aplica una máscara con el umbral de intensidad asociado a la mínima intensidad de precipitación que se desea estudiar. Para umbrales chicos, se obtienen objetos de gran tamaño y a medida que aumenta

el umbral, el área de los objetos disminuye. Una vez aplicados R y T, se obtiene el campo final en donde se reconstruye el valor original de precipitación en el interior de los objetos. En la Figura 2 se muestra a modo de ejemplo, un campo de precipitación acumulada de 24 horas pronosticada por el modelo correspondiente al día 19 de febrero del 2018 a las 12 UTC y los objetos identificados por MODE utilizando un valor de R de 50 km y T de 10 mm. MODE trabaja con atributos que son propiedades que caracterizan a los objetos y a partir de ellos se puede indicar qué tan parecidos son los objetos identificados en el campo del pronóstico y en el de la observación. Los atributos están definidos tanto para objetos individuales como para pares de objetos (entre el pronóstico y la observación). En el manual de MET (Foelsche y otros, 2017) se describen todos los atributos definidos por MODE. En este trabajo se consideraron los atributos de pares de objetos de: la distancia entre centroides, la diferencia de ángulo, la relación entre áreas y la relación entre los percentiles 50 de intensidad de precipitación.

La diferencia entre centroides indica el error de la posición de los objetos pronosticados (Figura

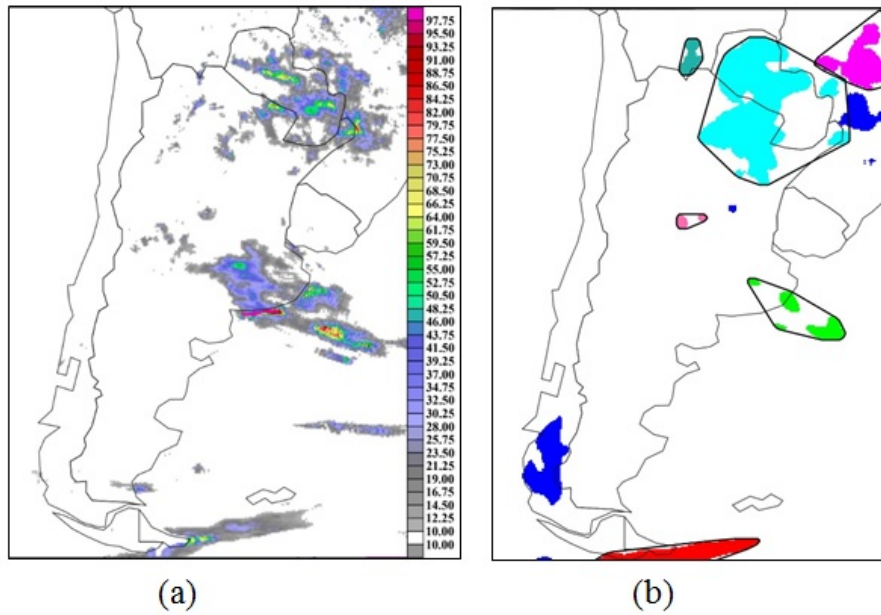


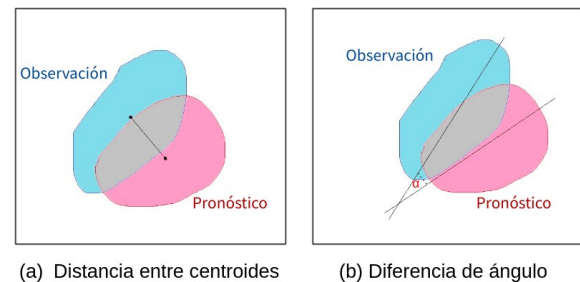
Figura 2: (a) Campo original del pronóstico de precipitación acumulada del WRF entre el 18 y 19 de febrero 2018 a las 12 UTC. (b) Las regiones sombreadas representan los objetos y los objetos agrupados dentro de los bordes negros representan los conjuntos identificados por MODE en el campo (a).

3.a). Si el objeto en el campo del pronóstico tiene su centro en las coordenadas (x_1, y_1) y el objeto de la observación en el punto (x_2, y_2) , entonces la distancia entre centroides se calcula como:

$$distanciaentrecentroides = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (1)$$

La diferencia de ángulo, asociada a la orientación de los objetos, mide las diferencias de inclinación, en grados, entre los ejes de dos objetos (Figura 3.b). Al tener dos objetos, se tienen dos ejes y la diferencia de ángulo es el menor de los ángulos que forman las rectas cuando se cruzan.

La relación entre áreas, que se define como el cociente entre el objeto de menor área y el de mayor área, indica qué tan parecidos son los tamaños de los objetos, y la relación entre el percentil 50 (cociente entre el menor y mayor valor del percentil del par de objetos) mide la relación de dichos percentiles de intensidad de precipitación entre los pares de objetos. Asimismo, MODE define funciones de interés para cada atributo, considerándolos como un



(a) Distancia entre centroides (b) Diferencia de ángulo

Figura 3: Esquema de (a) la distancia entre centroides y de (b) la diferencia de ángulo (α) entre el objeto del pronóstico (rosa) y de la observación (celeste).

argumento y devolviendo un valor de interés que varía entre cero y uno, siendo éste el máximo grado de interés (Figura 4). La función de interés para la distancia entre centroides describe que cuanto más cortas sean las distancias entre los centros de los objetos, más alto será el interés. Respecto de la diferencia de orientación, si los ejes de los objetos difieren entre 0° y 30° , el interés es uno y este decae a medida que la diferencia se incrementa. Tanto para la relación entre áreas como percentiles de intensidad, el

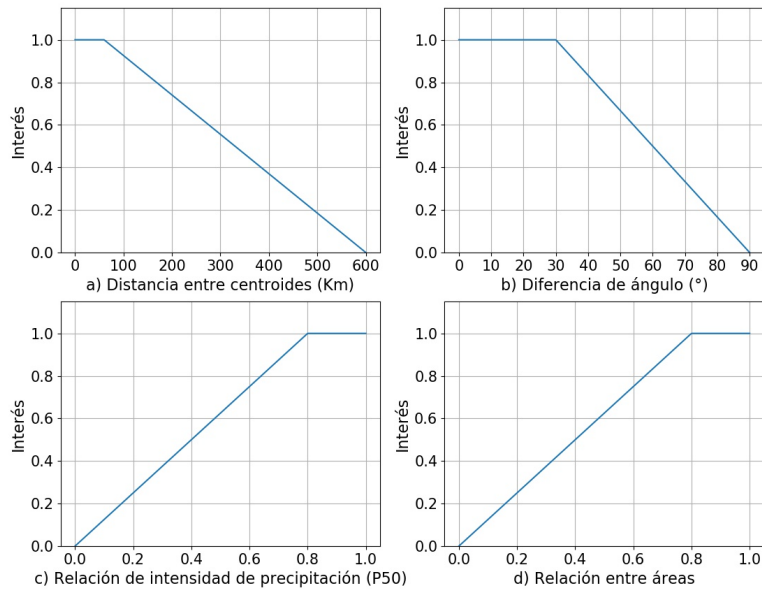


Figura 4: Funciones de interés empleadas en MODE para a) distancia entre centroides, b) diferencia de ángulo, c) relación de intensidad de precipitación y d) relación entre áreas.

interés es alto para los casos donde el cociente está entre 0.8 y 1, siendo 1 el caso de áreas idénticas y de igual intensidad de precipitación, respectivamente.

Adicionalmente, MODE otorga un nivel de confianza a cada atributo a través de los mapas de confianza. Estos mapas toman valores entre cero y uno: el cero indica que no hay ninguna confianza mientras que el valor de uno, hay completa confianza. MODE, para los mapas de confianza de todos los atributos excepto los de la distancia entre centroides y de la diferencia de ángulo, los fija en uno. El ángulo de orientación es muy sensible a la forma del objeto. Por eso, cuanto más parecidos sean a un círculo, más bajo será el nivel de confianza. Respecto de la distancia entre centroides, su mapa de confianza depende de la relación entre las áreas de los objetos y los pares de objetos con áreas diferentes entre sí tienen bajo nivel de confianza.

Luego, a cada atributo se le asigna un peso que puede ser modificado por el usuario. En

este trabajo se utilizaron los pesos por defecto, siendo los atributos asociados a distancias los que presentan mayor importancia. Los pesos asignados fueron: 2 para la distancia entre centroides, 1 para la diferencia de ángulo y para la relación entre áreas, y 0 para la relación entre los percentiles 50 de intensidad de precipitación.

A partir de las funciones de interés, mapas de confianza y pesos relativos de cada atributo, se define el interés total entre pares de objetos (I):

$$I = \frac{\sum_{i=1}^n w_i A_i B_i}{\sum_{i=1}^n w_i A_i} \quad (2)$$

en donde i representa cada atributo, w_i los pesos relativos, A_i los mapas de confianza y B_i , las funciones de interés. El interés total también varía entre cero y uno, siendo uno el valor que indica la máxima relación entre los objetos. En MODE, se debe fijar un umbral para dicho interés ya que los objetos que tengan un valor mayor o igual al umbral fijado serán vinculados si pertenecen a distintos campos (es decir un objeto del pronóstico y uno de la observación) o fusionados si pertenecen al mismo campo. En

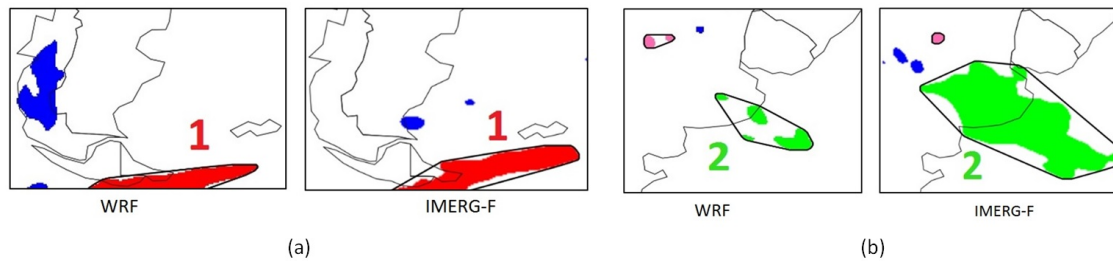


Figura 5: Ejemplo de pares de conjuntos definidos por MODE en el pronóstico y la estimación de precipitación a) par-conjunto 1; b) par-conjunto 2.

este trabajo, se fijó el umbral de interés total en 0.7 al igual que en la mayor parte de los trabajos que utilizan esta metodología (Davis y otros, 2009; Clark y otros, 2012; Bender y Ynoue, 2014; Yan y Gallus, 2016).

También puede definirse la máxima distancia posible que debe existir entre los centroides del par de objetos. Eligiendo un adecuado valor para esta distancia, el tiempo de ejecución de MODE se reduce significativamente, evitando las comparaciones entre objetos muy distantes, como por ejemplo un objeto del pronóstico de la región del norte argentino con uno de la observación de la región patagónica. En este trabajo, la máxima distancia permitida es de 400 km.

Finalmente, se definen los conjuntos en cada campo, formados por la agrupación de objetos vinculados y/o fusionados. En la lógica de MODE, primero se identifican los objetos vinculados entre los campos observados y pronosticados, y a partir de esta selección se fusionan los objetos dentro de un mismo campo, observado o pronosticado. Es así que un conjunto puede estar formado por un objeto en cada campo, o por varios objetos. En la Figura 5 se muestran dos ejemplos. El par-conjunto 1 está formado por un objeto perteneciente al pronóstico y otro a la observación, en cambio el par-conjunto 2 está formado por 4 objetos en el campo del pronóstico y 1 objeto en el campo de la observación. En el par-conjunto 2, los 4 objetos del pronóstico se fusionaron porque cada uno de ellos fue vinculado con el único

objeto del campo de IMERG-F. Al igual que los objetos, los conjuntos son caracterizados por sus atributos.

2.3. ELECCIÓN DE RADIOS DE CONVOLUCIÓN Y UMBRALES DE INTENSIDAD

Para determinar los valores del suavizado y de los umbrales de precipitación, que fueran adecuados para la verificación de precipitación, se realizaron varios experimentos con diferentes combinaciones de R y T para los pronósticos del WRF inicializados a las 06 UTC para el mes de febrero del 2018.

Para la precipitación acumulada en 24 horas, en base a una evaluación cualitativa de los resultados y a la experiencia de trabajos previos en otras regiones (Davis y otros, 2006a; Davis y otros, 2009; Johnson y Wang, 2012a; Bender y Ynoue, 2014; Clark y otros, 2014; Yang y Gallus, 2016), se analizaron en mayor detalle las combinaciones de radios de 50 y 100 km, con umbrales de precipitación mayores a 1 y 10 mm.

La Figura 6 muestra el ejemplo de un pronóstico de precipitación acumulada de 24 horas entre las 12 UTC del 18 y las 12 UTC del 19 de febrero de 2018. Se observa que al aumentar el umbral de precipitación disminuye el tamaño de los objetos, independientemente del suavizado elegido. Un umbral superior a 10 mm en 24 horas está asociado a sistemas que perduran en el tiempo o sistemas de precipitación intensa. En las áreas de precipitación diaria mayores a 1 mm, al ser un umbral muy bajo para

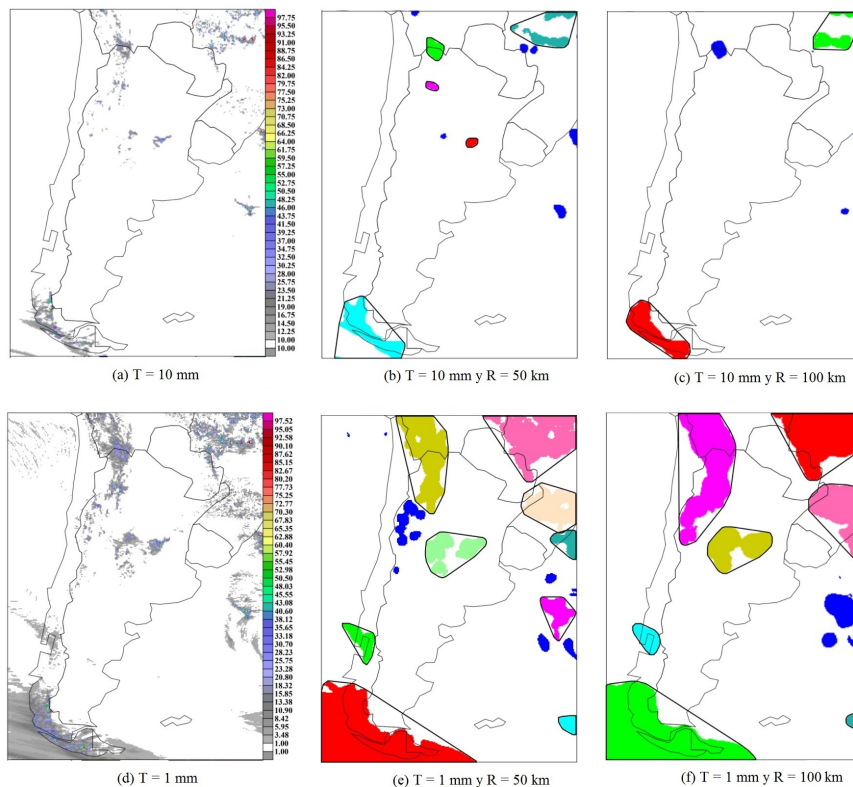


Figura 6: Precipitación pronosticada acumulada entre el 18 y 19 de febrero 2018 a las 12 UTC con un umbral de a) 10 mm y d) 1mm. Objetos y conjuntos identificados por MODE con: b y c) umbral de 10 mm y R de 50 y 100 km respectivamente, e) y f) umbral de 1 mm y R de 50 y 100 km respectivamente. Los colores identifican objetos y los bordes negros, los conjuntos.

ese intervalo de tiempo, los objetos, y por consiguiente los conjuntos, están asociados a la presencia o no de precipitación. Por otro lado, si bien IMERG-F obtiene mejores resultados que otras estimaciones, la detección de precipitación con umbrales tan bajos sigue siendo sensible y varía de acuerdo al umbral seleccionado (Tan y otros, 2017), impactando en el resultado de la verificación de los modelos. También se observa que los conjuntos con bajos valores de R son más parecidos al campo original y a medida que aumenta el radio de convolución, estos se suavizan más, es decir tienen menos detalle, o incluso no se identifican como ocurre con los conjuntos para una tasa de precipitación igual o superior a 10 mm, en el centro y norte del país.

Además, la Tabla I muestra la cantidad de objetos vinculados (los que forman parte de algún conjunto) para el mes de febrero del 2018, según los parámetros elegidos. Se puede observar que a medida que aumenta el umbral de precipitación, disminuye la cantidad de objetos vinculados y al aumentar el radio de convolución, se produce el mismo efecto.

Al analizar los resultados de las distintas combinaciones entre los parámetros de R y T se pudo ver que por un lado, cuanto mayor es el suavizado menor es la definición o el detalle de los conjuntos; por otro lado, al aumentar el umbral de precipitación menor es el número de conjuntos ya que la búsqueda de objetos está centrada en encontrar áreas con precipitación intensa. También, a medida que el umbral

	$T \geq 1 \text{ mm}$	$T \geq 10 \text{ mm}$
$R = 50 \text{ km}$	172 (176)	147 (150)
$R = 100 \text{ km}$	102 (101)	89 (93)

Tabla I: Cantidad de objetos vinculados para diferentes valores de T y R, para precipitación acumulada en 24 horas durante el mes de febrero 2018 para el pronóstico del modelo WRF y entre paréntesis para la estimación IMERG-F.

aumenta, el tamaño de los conjuntos disminuye. Por ende si se eligieran bajos valores de R y T, MODE encontrará mayor cantidad de objetos y el área de los conjuntos será grande, mientras que para altos valores de R y T, se obtendrán menos objetos y de menor tamaño.

Debido a la importancia de los pronósticos de precipitación intensa asociada a tormentas y por qué una buena predicción puede ser significativa para los tomadores de decisiones o para usuarios en general, se tomó la decisión de aplicar la identificación de objetos para sistemas precipitantes más desarrollados. Por dicha razón se eligió trabajar con un suavizado de 50 km y con precipitación acumulada mayor o igual a 10 mm cada 24 horas. Para la precipitación acumulada cada 3 horas, luego de un análisis cualitativo y tomando como referencia los trabajos de Davis y otros (2006a), Wolff y otros (2014), y de Yan y Gallus (2016), se consideró el radio de convolución de 50 km y el umbral de intensidad de 3 mm.

2.4. ESQUEMA DE VERIFICACIÓN

Dado que los pronósticos y las estimaciones de precipitación tienen retículas distintas y de diferente resolución espacial, para realizar la comparación se decidió interpolar la información de los pronósticos del WRF y del GFS a la retícula de IMERG-F. MET tiene definido algunos métodos para realizar la interpolación horizontal y para este caso se utilizó el método bilineal, el cual considera los cuatro puntos de retícula más cercanos al punto que se desea

interpolar (Davis y otros, 2009; Johnson y otros, 2013; Carrasco, 2017). De esta forma, al interpolar el campo del WRF al de la estimación, se degradó su resolución horizontal de 4 km a 10 km, mientras que para el caso del GFS, este pasó de 25 km (0.25°) a 10 km.

Para la verificación espacial se utilizaron los atributos de los pares de conjuntos de distancia entre centroides, diferencia de ángulo, relación entre el percentil 50 de intensidad, relación entre áreas y el interés total. Respecto de las medidas de verificación puntual, se realizaron tablas de contingencia para un determinado umbral de precipitación a partir de las cuales se calcularon los índices BIAS Score, la probabilidad de detección (POD), la relación de falsas alarmas (FAR) y el índice Equitable Threat Score (ETS).

$$BIASScore = \frac{H + F}{H + M} \quad (3)$$

$$POD = \frac{H}{H + M} \quad (4)$$

$$FAR = \frac{F}{F + H} \quad (5)$$

$$ETS = \frac{H - C}{H + F + M - C}, \quad \text{con} \quad (6)$$

$$C = \frac{(H + F) \cdot (H + M)}{H + F + M + CN}$$

en donde H son los aciertos, F las falsas alarmas, M sorpresas y CN los correctos negativos. Valores de BIAS Score cercanos a 1 indican un buen desempeño del modelo, valores por encima de 1 indican una sobreestimación de la cantidad de eventos de precipitación pronosticados y valores menores a 1 indican una subestimación. POD varía de 0 a 1, siendo un pronóstico perfecto aquel cuyo valor sea 1. El FAR que representa las falsas alarmas también varía entre 0 y 1, siendo un pronóstico perfecto cuando vale 0. El ETS está asociado a los eventos correctamente pronosticados, considerando los aciertos aleatorios y sus valores varían entre $-1/3$ y 1: un pronóstico perfecto tendrá un valor de ETS igual a 1. Un mayor detalle de estos estadísticos y su definición se puede encontrar en <http://www.cawcr.>

gov.au/projects/verification/ y en Nurmi (2003).

3. RESULTADOS

3.1. PRECIPITACIÓN ACUMULADA EN 24 HORAS

3.1.1. Verificación de pronósticos de precipitación del WRF

Para la verificación puntual y considerando el umbral de 10 mm, se calcularon los valores medios de los índices BIAS Score, POD, FAR y ETS de los pronósticos de precipitación del WRF (Tabla II) para el período 2017-2018. Si bien el BIAS Score es cercano a 1, los valores de POD y ETS son bajos por lo que el modelo tiene un bajo desempeño para pronosticar correctamente los eventos de precipitación para un umbral de 10mm. Además, el índice FAR evidencia la presencia de falsas alarmas. Estos resultados son similares a los encontrados en la verificación del sistema de pronóstico operativo implementado en el SMN para el año 2020 frente a observaciones de estaciones convencionales (Matsudo y otros, 2021). Asimismo, los resultados de los índices POD, ETS y FAR pueden estar influenciados por la doble penalización de la verificación punto a punto (Zingerle, C., y Nurmi, P., 2008). Primero, porque en la comparación puntual, en el campo de la observación se observa precipitación pero no así en el campo del pronóstico (evento sorpresa) que detecta la precipitación con un pequeño desplazamiento en otra posición. Segundo, porque esa precipitación en el campo del pronóstico resulta a la vez ser una falsa alarma ya que en la retícula de la observación, no se observaron valores de precipitación acumulada. Esta doble penalización es más notable cuanto mayor es la resolución espacial del pronóstico, y por estos motivos resulta conveniente recurrir a metodologías de verificación espacial para validar pronósticos en alta resolución (Roebber y otros, 2004; Rossa y otros, 2008).

	BIAS SCORE	POD	ETS	FAR
WRF	1,08	0,45	0,27	0,54
GFS	1,13	0,48	0,28	0,53

Tabla II: Valores medios de estadísticos de verificación para el pronóstico de precipitación acumulada en 24 horas del WRF y del GFS considerando intensidades de precipitación mayores o iguales a 10 mm, para el periodo 2017-2018.

Para analizar el desempeño del WRF en los casos de aciertos, a partir de la verificación espacial, se realizaron histogramas de frecuencia relativa de algunos atributos de los pares de conjuntos, como se muestra en la Figura 7. En este análisis no fueron considerados los objetos sorpresas y falsas alarmas, es decir que se descartaron aquellos objetos con interés total del par pronóstico-observación menor a 0.7 (que representan más del 50% del total de los objetos identificados por MODE). De esta forma, solo se analizaron los objetos que cumplen con un radio de convolución de 50 km, un umbral de intensidad mayor o igual a 10 mm y con el interés total entre pares mayor o igual a 0.7. En la Figura 7 se observa que para el par conjunto pronóstico-observación, las distancias más frecuentes son las que se hallan entre 40 y 120 km. También se observa que dominan los casos con diferencia de orientación menor a 10°. Las funciones de interés de estos atributos (Figura 4) muestran que cuanto más bajos sean sus valores, más alto será el interés, lo que implica que el WRF presenta un buen desempeño para pronosticar la posición de la precipitación. El histograma del percentil 50 de intensidad de precipitación muestra frecuencias altas para valores de relación mayor a 0.8, indicando que el WRF presenta pocos errores en los pronósticos de intensidad de precipitación. Respecto de la relación entre áreas, la distribución de frecuencias es homogénea por lo que la habilidad del WRF para pronosticar el área de la precipitación es baja. Por otro lado, pudo corroborarse que dominan los pares

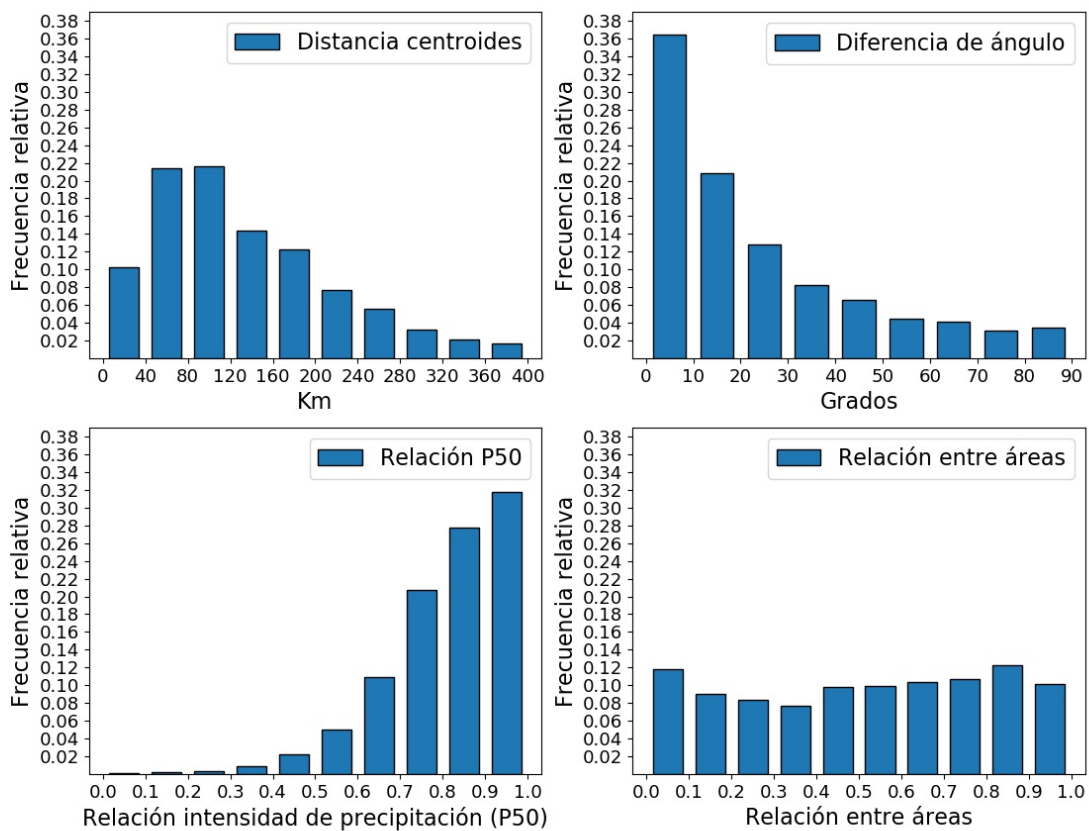


Figura 7: Frecuencias relativas de atributos entre pares de conjuntos detectados para precipitación acumulada en 24 horas del WRF para el período 2017-2018. Se muestran la distancia entre centroides, la diferencia de ángulo, la relación entre el percentil 50 de intensidad de precipitación y la relación entre las áreas de los conjuntos entre el campo del WRF y del IMERG-F.

de conjuntos, entre el campo del WRF y de IMERG-F, que presentan intersección entre sus áreas. Sin embargo, existen algunos pocos casos (que representan sólo el 5%) en los cuales el interés total resulta mayor a 0.7 y no existe área de intersección, es decir que la posición de la región precipitante pronosticada no concuerda con la observada. En estos casos, el WRF pronostica la precipitación con un leve desplazamiento respecto de la precipitación observada, pero la forma y orientación de la misma es similar a la observada.

El histograma de interés total (Figura 8) condensa la información de los atributos vistos anteriormente. Se observa que el modelo presenta altos valores de interés con los conjuntos asociados al campo de IMERG-F, dominando los pares de conjuntos con un interés total superior a 0.95. Esto se debe principalmente a que gran parte de los conjuntos identificados por MODE presentan distancias cortas entre sus centroides, con poca diferencia entre sus ejes de inclinación y un buen pronóstico de intensidad de precipitación.

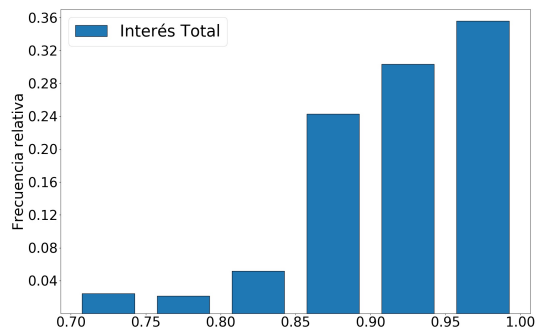


Figura 8: Frecuencia relativa del interés total entre pares de conjuntos para precipitación acumulada en 24 horas del WRF para el período 2017-2018.

3.1.2. Verificación de pronósticos de precipitación del WRF

Comparando los resultados entre el WRF y GFS de la verificación puntual de los pronósticos de precipitación acumulada mayor o igual a 10 mm en 24 horas, la Tabla II muestra bajos valores de POD y ETS, indicando que ambos modelos presentan dificultades para pronosticar correctamente los eventos de precipitación. Respecto del BIAS Score, se observa que el GFS tiende a sobreestimar los eventos de precipitación. La Tabla III presenta los estadísticos para verano e invierno, tomando el verano como los meses de diciembre, enero y febrero y el invierno como junio, julio y agosto. Comparando los valores medios de POD, FAR y ETS entre estaciones, se observa que los pronósticos de eventos de precipitación de ambos modelos mejoran levemente durante el invierno (que es la época menos lluviosa), aunque los índices POD y ETS permanecen bajos y los valores del FAR indican un número importante de falsas alarmas. Como resultado de la verificación espacial y a partir del análisis diario, se obtuvieron los diagramas de cajas de la frecuencia diaria de objetos que cumplen con un radio de 50 km y un umbral de intensidad de 10 mm en los campos del WRF, GFS e IMERG-F, y de los objetos vinculados con un interés total mayor o igual a 0.7 (Vin.WRF y Vin.GFS) para

	VERANO		INVIERNO	
	WRF	GFS	WRF	GFS
BIAS SCORE	0,94	1,07	1,27	1,22
POD	0,36	0,41	0,52	0,53
ETS	0,20	0,22	0,31	0,34
FAR	0,60	0,60	0,59	0,49

Tabla III: Valores medios de estadísticos de verificación para el pronóstico de precipitación acumulada del WRF y del GFS considerando intensidades de precipitación mayores o iguales a 10 mm acumulada en 24 horas, para los meses de verano y de invierno del periodo 2017-2018.

los meses de verano e invierno (Figura 9.a y 9.b, respectivamente). Se observa que en verano se encuentra un mayor número de objetos que en invierno, tanto en los campos pronosticados como en los observados, coincidiendo con la climatología que indica que en la mayor parte del país, en verano se registra la mayor frecuencia de eventos de precipitación (Marino, 2007; Salio y otros, 2007). En verano (Figura 9.a), la mediana de los objetos encontrados diariamente en los campos del WRF y de la estimación satelital tiene un valor de 13 objetos mientras que para el GFS presenta 10 objetos, indicando una subestimación de los eventos de precipitación pronosticados por este modelo. En invierno (Figura 9.b) también se observa una subestimación de eventos en el GFS. Los objetos vinculados son los objetos asociados entre el campo del pronóstico y de la estimación satelital, con un interés total igual o superior a 0.7, y que forman parte de algún conjunto. Respecto de los vinculados diariamente por MODE en el periodo de estudio, se observa que la proporción de objetos vinculados disminuye respecto del número total de objetos (Figura 9 y Tabla IV). Esto se debe a que MODE compara un objeto del campo de pronóstico con un objeto de IMERG-F, siempre que la distancia entre centroides sea menor o igual a 400 km, ocurren varios casos en donde los atributos del par son bajos y en consecuencia se obtiene un interés total menor a 0.7. Esta situación ocurre por 3 posibles causas: 1)

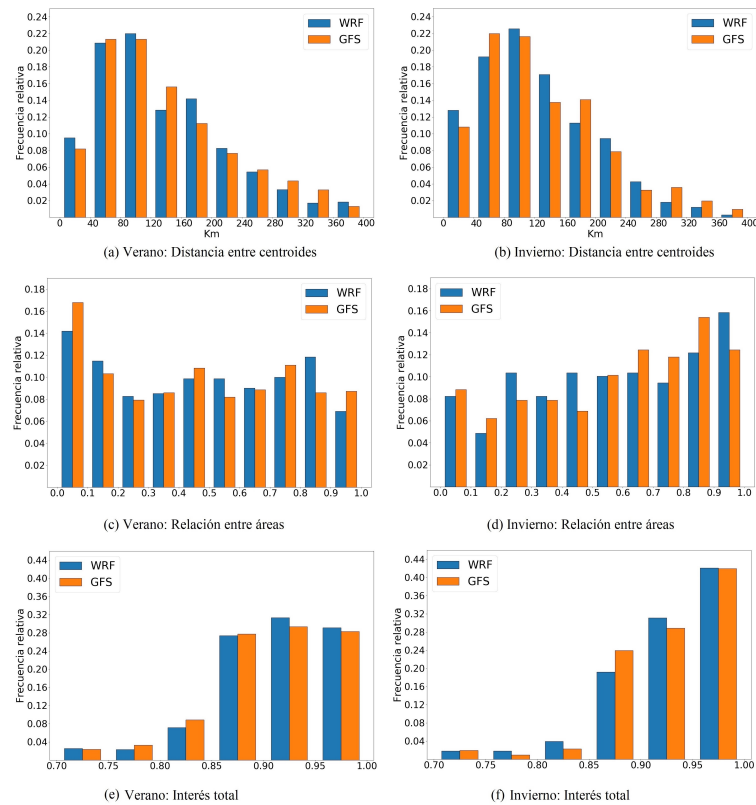


Figura 9: Diagrama de cajas de la frecuencia de objetos encontrados diariamente en los pronósticos del WRF y del GFS, y en la estimación IMERG-F para precipitación acumulada de 24 horas en el período 2017-2018. Para el total de objetos (WRF, GFS, IMERG-F) y objetos vinculados (Vin.WRF, Vin.GFS) en el verano (a) y en el invierno (b). La caja de los diagramas está formada por el intervalo intercuartil, entre el percentil 75 (P75) y el percentil 25 (P25). La parte angosta corresponde a la mediana. El bigote superior se extiende hasta el último dato menor que $P75 + 1.5 \cdot (P75 - P25)$, y el bigote inferior se extiende hasta el primer dato mayor que $P25 - 1.5 \cdot (P75 - P25)$. Los puntos representan los eventos fuera de rango.

cuando MODE busca comparar un objeto del campo del pronóstico que no está presente en el campo observado (Falsa Alarma), 2) cuando MODE busca comparar un objeto del campo observado que no está presente en el campo pronosticado (Sorpresa), y 3) cuando los centros de los objetos están muy separados o sus atributos no son lo suficientemente similares. Es así que se obtiene una disminución en la frecuencia de ocurrencia de los objetos vinculados con respecto a las ocurrencias de objetos individuales del WRF y GFS: en verano este valor desciende a 5 eventos diarios y en invierno, se detectan entre 2 y 3. Además, tanto en invierno como en verano, son más los objetos

vinculados en promedio en el caso del WRF que en el GFS, indicando un mejor desempeño del WRF en la verificación espacial de los sistemas precipitantes. Para continuar con el análisis y comparación de la verificación espacial de ambos modelos, se realizaron los histogramas de frecuencia relativa de la distancia entre centroides, la relación entre áreas y el interés total, para verano e invierno, de los pares de conjuntos con un interés mayor o igual a 0.7. Al igual que en la sección 3.1.1, los conjuntos se formaron a partir de los objetos vinculados, representando entre el 43% y 52% de los objetos totales, según el modelo y la época del año (Tabla IV). En los histogramas de la

	VERANO		INVIERNO	
	WRF	GFS	WRF	GFS
Objetos totales	2370	1958	822	659
Objetos vinculados	1021 (43%)	899 (46%)	379 (46%)	343 (52%)

Tabla IV: Cantidad de objetos totales y vinculados con radio de convolución de 50 km y umbral de intensidad mayor o igual a 10 mm, para precipitación acumulada en 24 horas durante el verano e invierno del período 2017-2018, para los pronósticos de los modelos WRF y GFS. Entre paréntesis se muestra el porcentaje de objetos vinculados respecto del total de objetos.

distancia entre los centros de los conjuntos (Figura 10.a y 10.b), se observa el dominio de casos en donde la distancia entre los centros se encuentra entre 40 y 120 km y que la frecuencia de pares de conjuntos con distancias menores a 40 km es mayor en el WRF que en el GFS. Los histogramas de la relación entre áreas (Figura 10.c y 10.d) se caracterizan por su homogeneidad, demostrando la poca eficacia de los modelos para pronosticar la extensión de la precipitación. Por último, se observa que tanto el WRF como el GFS presentan altos valores de interés con los conjuntos asociados al campo de IMERG-F (Figura 10.e y 10.f). Esto se debe principalmente a que gran parte de los conjuntos identificados por MODE presentan distancias cortas entre sus centroides. Además, se observa que la cantidad de casos con interés total entre 0.95 y 1 es superior durante el invierno, resultado influenciado por la mejora de los modelos al pronosticar la relación entre áreas de los sistemas precipitantes.

3.2. PRECIPITACIÓN ACUMULADA EN 3 HORAS

Los valores medios de los índices estadísticos del BIAS Score, POD, FAR y ETS con un umbral de intensidad de 3 mm, para la verificación de los pronósticos de precipitación acumulada

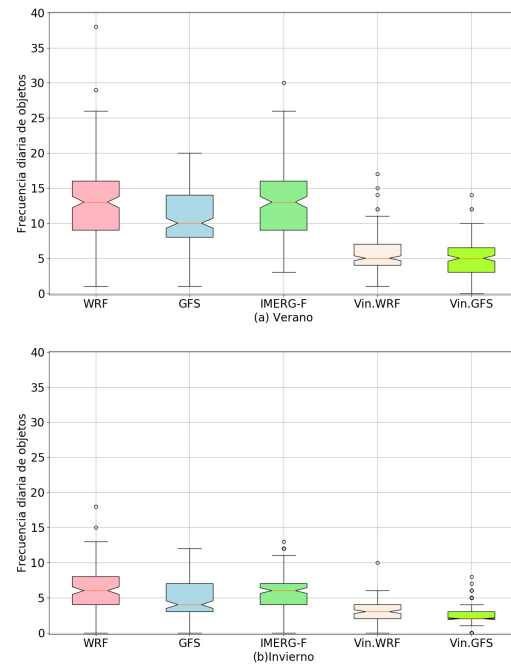


Figura 10: Frecuencias relativas de atributos entre pares de conjuntos detectados para precipitación acumulada en 24 horas del WRF (azul) y del GFS (naranja) para el verano e invierno del período 2017-2018. Distancia entre centroides (a y b), relación entre áreas (c y d), interés total (e y f).

de 3 horas del WRF desde su inicio hasta las 48 horas, se muestran en la Figura 11. Se puede ver que el pronóstico no tiene un buen desempeño en las primeras 3 horas a excepción de las falsas alarmas que presenta una reducción. Esto es producto del período de spin up del pronóstico dado que la convección se resuelve en forma explícita requiriendo más tiempo para el desarrollo de la precipitación (Done y otros, 2004). Se observa también que los valores de POD y ETS tienden a disminuir de acuerdo avanza el plazo del pronóstico, y en el caso del FAR, estos empiezan a aumentar. En otras palabras, como es de esperar a medida que aumenta el plazo del pronóstico, este se vuelve menos preciso. Además, los valores de

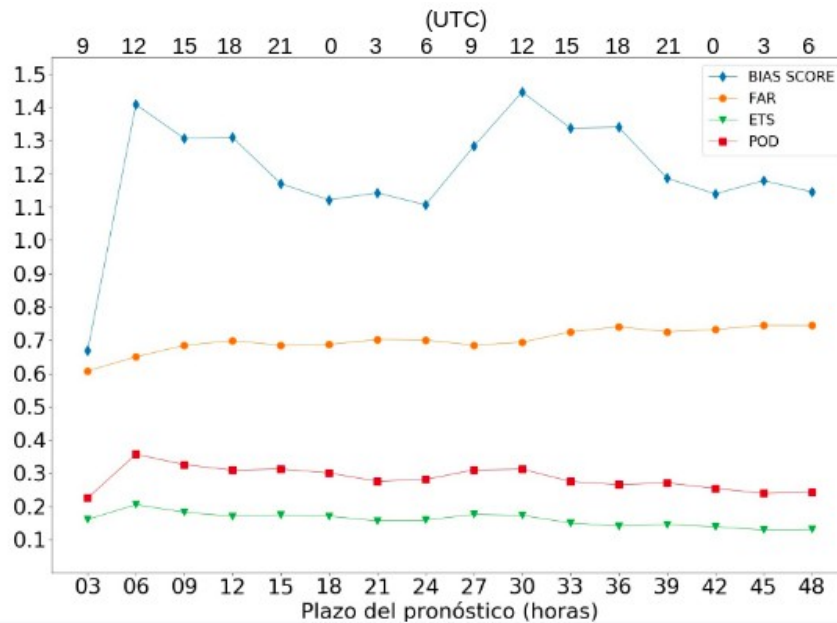


Figura 11: Valores medios de estadísticos de verificación para el pronóstico de precipitación acumulada en 3 horas del WRF considerando intensidades de precipitación mayores o iguales a 3 mm, para el periodo 2017-2018. El eje x superior indica la hora de validez del pronóstico en horas UTC.

los estadísticos en el análisis de la precipitación acumulada en 24 horas, resultaron mejores. En los pronósticos de precipitación acumulada cada 3 horas, al aumentar la resolución temporal se acentúa el efecto de la doble penalización en los índices de verificación puntual, ya que a los errores en los desplazamientos espaciales se les suma los errores asociados a pequeños desfases en el tiempo, por ejemplo, diferencias en el tiempo de inicio y finalización del evento.

Por otro lado, también se evidencia que las curvas del BIAS Score, POD y ETS presentan dos máximos (Figura 11), vinculados a la precipitación acumulada entre las 9 y 12 UTC. Analizando el ciclo diario de la tasa de precipitación acumulada cada 3 horas estimada por IMERG-F, se puede observar una relación entre la tasa de precipitación y los valores de los índices estadísticos. En horas de la mañana se produce la menor tasa de precipitación en la región de estudio, por lo que los máximos del BIAS se relacionan con la sobreestimación de los eventos de precipitación por parte del WRF. Asimismo, la mejora en los valores del

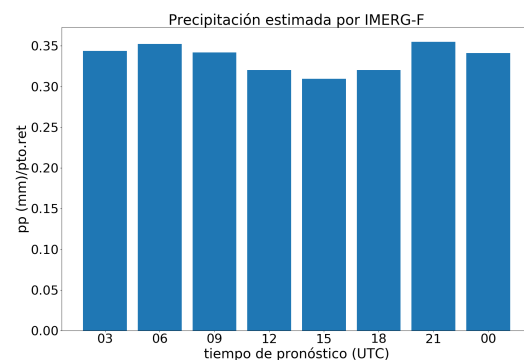


Figura 12: Tasa de precipitación acumulada cada 3 horas por punto de retícula estimada por IMERG-F, para el periodo 2017-2018 y para el dominio de verificación.

POD y ETS, se observan también en las horas de mínima tasa de precipitación estimada por IMERG-F

Considerando los pares de conjuntos identificados con un radio de convolución

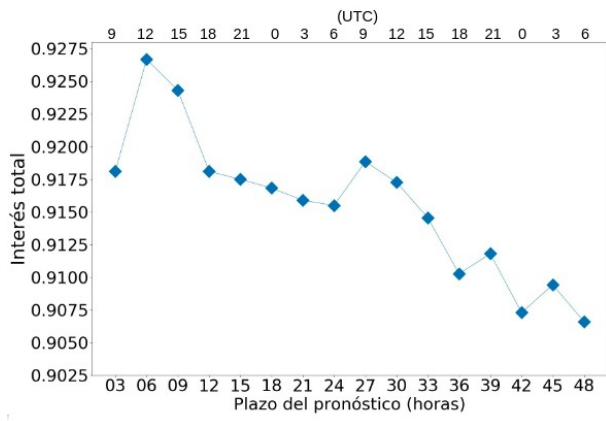


Figura 13: Valor medio del interés total entre pares de conjuntos para la precipitación acumulada en 3 horas del WRF en función del plazo de pronóstico para el período 2017-2018. El eje x superior indica la hora de validez del pronóstico en horas UTC.

de 50 km, un umbral de 3 mm y un interés total de 0.7, se muestra la evolución del interés total medio (Figura 13). Al igual que los índices de la verificación puntual, se observa el spin up en las primeras 3 horas del plazo del pronóstico y máximos durante la mañana, cuando disminuye la tasa de precipitación estimada. Si bien el interés total tiende a decaer con el tiempo, sus valores se mantienen superiores a 0.9. Este alto interés entre el par pronosticado y observado se debe a que los valores de distancias entre centroides y diferencia de ángulo permanecen bajos para todos los plazos del pronóstico, la relación entre la intensidad pronosticada por el WRF y la estimada por IMERG-F sigue siendo alta, y por último, porque el tamaño de los conjuntos pronosticados es similar a los observados (no se muestra).

El gráfico del interés total está basado sobre los objetos vinculados, es decir que estos objetos fueron agrupados y dieron origen a los conjuntos. En la Figura 14 se observa que MODE encuentra mayor cantidad de conjuntos durante la tarde y la noche (entre las 15 y 00 UTC) y comparando este resultado con la Figura 12, el aumento del número de conjuntos

está relacionado con el aumento de la tasa de precipitación. También puede observarse que de la gran cantidad de objetos identificados en el campo del WRF fijando un radio de convolución de 50 km y un umbral de 3 mm, solo un poco más del 40 % corresponden a los objetos vinculados. Es decir que si bien el pronóstico del WRF obtuvo un buen resultado en la verificación espacial de los objetos vinculados, no se están contemplando el 60 % de los casos que resultan de falsas alarmas en el campo pronosticado y sorpresas en el campo de IMERG-F.

4. CONCLUSIONES

En este trabajo se aplicó por primera vez la metodología de verificación de pronósticos por objetos empleada por el módulo MODE para los pronósticos de precipitación del modelo WRF-SMN y GFS para todo el territorio argentino y para un período de dos años. Además se compararon los resultados con metodologías tradicionales puntuales. El módulo MODE está disponible en el software MET. El mismo requiere la elección de los valores de radio de convolución y umbral, que dependen de los sistemas precipitantes que se quieran estudiar. En este caso, el estudio estuvo focalizado en los sistemas de precipitación moderada o intensa y en sistemas precipitantes que perduran en el tiempo. Para ello se utilizaron un radio de convolución de 50 km y un umbral de 10 mm para los pronósticos de precipitación acumulada a 24 horas. Para los pronósticos con valores acumulados a 3 horas se empleó el mismo suavizado con un umbral de 3 mm. Estos parámetros permanecieron fijos durante los dos años de análisis.

Los modelos verificados presentan falsas alarmas y sorpresas que se ven reflejados en los índices de la verificación puntual y en los porcentajes de objetos vinculados de la verificación espacial. Sin embargo, el porcentaje de objetos vinculados es sensible a los valores de R y T ya que como se muestra en la Tabla I, la cantidad de objetos identificados por MODE varía de acuerdo a la elección de dichos parámetros. La ventaja

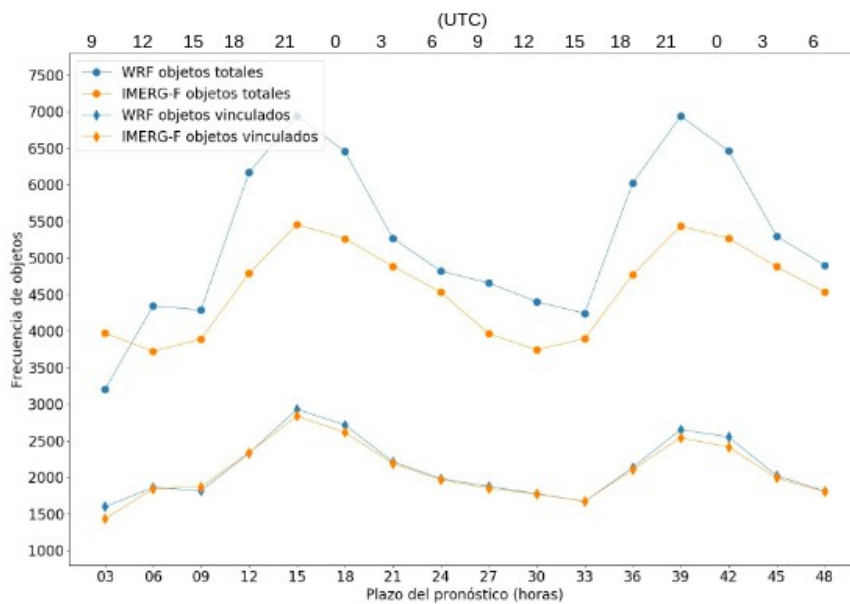


Figura 14: Frecuencia absoluta de objetos en el campo de precipitación pronosticada del WRF (azul) y en el campo de IMERG-F (naranja), para la precipitación acumulada cada 3 horas en función del plazo de pronóstico para el período 2017-2018. Se muestran objetos totales y objetos vinculados. El eje x superior indica la hora de validez del pronóstico en horas UTC.

que propone MODE es analizar en qué medida los modelos logran pronosticar la precipitación, tanto en intensidad, ubicación y extensión.

Tanto el análisis de precipitación acumulada de 24 horas como de 3 horas demostraron que la verificación espacial y puntual se complementan. La verificación puntual permite tener una noción general del comportamiento de los modelos, aunque carga con la doble penalización de las falsas alarmas y las sorpresas ya que el evento que no fue pronosticado resulta luego en un evento sorpresa si es detectado en otra posición, aunque haya sido leve su desplazamiento. Por su parte, la metodología por objetos permite verificar la forma, tamaño, posición e intensidad de la precipitación pronosticada, lo cual le agrega valor a los pronósticos que resultan de utilidad. Es importante destacar que en el análisis de las características de la precipitación, al considerar sólo los objetos vinculados, no se contemplaron las falsas alarmas y sorpresas.

Se encontró que el pronóstico de precipitación del WRF tiende a sobreestimar la cantidad

de eventos de precipitación y presenta falsas alarmas, pero tiene pocos errores asociados a la posición, intensidad y extensión de la precipitación, cuando se realiza una verificación espacial de los conjuntos identificados. Por otro lado, comparando los resultados del WRF con los del GFS, la verificación punto a punto mostró que los pronósticos de precipitación acumulada en 24 horas del GFS resultaron ser levemente mejores que los del WRF. Mientras que a partir de la verificación espacial, se observó que el WRF representa con mayor precisión la posición de la precipitación. Tanto la verificación espacial como la verificación puntual demostraron que el desempeño de ambos modelos mejora durante el invierno.

Este trabajo es un primer paso hacia la inclusión de técnicas de verificación espacial en un esquema operativo. A partir de los resultados obtenidos se considera necesario continuar evaluando los diferentes parámetros de MODE. Al variar el umbral de intensidad de precipitación o el suavizado de los campos, la cantidad de eventos y objetos que representan

falsas alarmas y sorpresas se verá modificado. Si bien los valores de radio y umbral de intensidad elegidos para la verificación espacial resultaron apropiados para un análisis general, se plantea la necesidad de considerar una variación de dichos valores en función de la región y de la época del año, considerando la utilización de percentiles de intensidad de precipitación (Skinner y otros, 2018).

Agradecimientos: Agradecemos al equipo de soporte técnico del MET por su asesoramiento en el uso del software. Además agradecemos al Servicio Meteorológico Nacional de Argentina por facilitar los pronósticos generados con el modelo WRF y también al NCEP por proveer generosamente los análisis y pronósticos del GFS.

REFERENCIAS

- Bender F.D. y Ynoue R., 2014: Application of a spatial verification method to GFS precipitation forecasts. *Revista Brasileira de Geofísica*. 32. 561-571.
- Brown, B. G., Bullock, R. R., Davis, C. A., Gotway, J. H., Chapman, M. B., Takacs, A., Gilleland, E., Manning, K. y Mahoney, J. L., 2004: New verification approaches for convective weather forecasts. In *Proceedings of the 22nd Conference on Severe Local Storms*.
- Brown, B. G., Brown, B. G., Gotway, J. H., Bullock, R., Gilleland, E., Fowler, T., Ahijevych, D. y Jensen, T., 2009: The model evaluation tools (MET): community tools for forecast evaluation. In *Preprints, 25th Conf. on International Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology*, Phoenix, AZ, *Amer. Meteor. Soc. A*, 9(6).
- Carrasco A.R., 2017: Método de avaliação orientada a objeto aplicado às previsões de precipitação sobre a América do Sul. *Instituto Nacional de Pesquisas Espaciais (INPE)*, São José dos Campos, 116p.
- Charó, G. D., Collini, E. A. y Dillon, M. E., 2014: La utilización del met (model evaluation tool) para la verificación de los pronósticos del modelo wrf-arw/shn-smn durante la primavera de 2011. *Meteorologica*, 39(2), 49-68.
- Clark, A. J., Weiss, S., Kain, J., Jirak, I., Coniglio M., Melik, C., Sietwet C., Sobash, R., Marsh, P., Dean, A., Xue, F., Thomas, K., Wang, Y., Brewster, K., Gao, J., Wang, X., Du, J., Novak, D. y Correia, J.J., 2012: An overview of the 2010 hazardous weather testbed experimental forecast program spring experiment. *Bulletin of the American Meteorological Society*, 93(1), 55-74.
- Davis, C., Brown, B. y Bullock, R., 2006.a: Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Monthly Weather Review*, 134(7), 1772-1784.
- Davis, C., Brown, B. y Bullock, R. 2006.b: Object-based verification of precipitation forecasts. Part II: Application to convective rain systems. *Monthly Weather Review*, 134(7), 1785-1795.
- Davis, C. A., Brown, B. G., Bullock, R. y Halley-Gotway, J., 2009: The Method for Object-Based Diagnostic Evaluation (MODE) applied to numerical forecasts from the 2005 NSSL/SPC Spring Program. *Weather and Forecasting*, 24 (5), 1252-1267.
- Done, J., Davis, C. A. y Weisman, M., 2004: The next generation of NWP: Explicit forecasts of convection using the Weather Research and Forecasting (WRF) model. *Atmospheric Science Letters*, 5(6), 110-117.
- Flora, M. L., Skinner, P. S., Potvin, C. K., Reinhart, A. E., Jones, T. A., Yussouf, N. y Knopfmeier, K. H., 2019: Object-based verification of short-term, storm-scale probabilistic mesocyclone guidance from an experimental Warn-on-Forecast system. *Weather and Forecasting*, 34(6), 1721-1739.
- Foelsche, U., Kirchengast, G., Fuchsberger, J., Tan, J. y Petersen, W. A., 2017: Evaluation of GPM IMERG Early, Late, and Final rainfall estimates using WegenerNet gauge data in southeastern Austria. *Hydrology and Earth*

- System Sciences, 21(12), 6559-6572.
- Fowler, T., Halley Gotway J., Newman, K., Jensen, T., Brown, B. y Bullock, R., 2017: The Model Evaluation Tools v7.0 (METv7.0) User's Guide. Developmental Testbed Center, 407 pp.
- García Skabar, Y., Matsudo, C. M., Sacco, M., Ruiz, J. J. y Righetti, S.A., 2018: Implementación modelo de pronóstico numérico WRF. Nota Técnica SMN, 2018-45.
- Gilleland, E., Ahijevych, D. A., Brown, B. G. y Ebert, E. E., 2010: Verifying forecasts spatially. *Bulletin of the American Meteorological Society*, 91(10), 1365-1376.
- Gilleland, E., 2020: Novel measures for summarizing high-resolution forecast performance. Submitted to *Advances in Statistical Climatology, Meteorology and Oceanography* on, 19.
- Gilleland, E., Skok, G., Brown, B. G., Casati, B., Dorninger, M., Mittermaier, M. P., Roberys N. y Wilson, L. J., 2020: A Novel Set of Geometric Verification Test Fields with Application to Distance Measures. *Monthly Weather Review*, 148(4), 1653-1673.
- Griffin, S. M., Otkin, J. A., Rozoff, C. M., Sieglaff, J. M., Counce, L. M., Alexander, C. R., Jenssen, T. L. y Wolff, J. K., 2017: Seasonal analysis of cloud objects in the High-Resolution Rapid Refresh (HRRR) model using object-based verification. *Journal of Applied Meteorology and Climatology*, 56(8), 2317-2334.
- Hobouchian, M. P., 2015: Estimaciones de precipitación por satélite en el área Subtropical de Sudamérica: análisis y validación. Tesis de Licenciatura. Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales.
- Hobouchian, M. P., Salio, P., Skabar, Y. G., Vila, D. y Garreaud, R., 2017: Assessment of satellite precipitation estimates over the slopes of the subtropical Andes. *Atmospheric research*, 190, 43-54.
- Huffman G.J., Bolvin D.T. y Nelkin E.J., 2015: Integrated Multi-satellite Retrievals for GPM (IMERG) Technical Documentation. NASA/GSFC Code 612, 47.
- Johnson, A. y Wang, X., 2012a: Verification and calibration of neighborhood and object-based probabilistic precipitation forecasts from a multimodel convection-allowing ensemble. *Monthly Weather Review*, 140(9), 3054-3077.
- Johnson, A., Wang, X., Kong, F. y Xue, M., 2013: Object-based evaluation of the impact of horizontal grid spacing on convection-allowing forecasts. *Monthly Weather Review*, 141(10), 3413-3425.
- Kain, J. S., Weiss, S. J., Bright, D. R., Baldwin, M. E., Levit, J. J., Carbin, G. W., Schawartz, C.R., Weisman, M.L., Droegemeier, K.K., Weber, D. B. y Thomas, K. W., 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Weather and Forecasting*, 23(5), 931-952.
- Marino, M. B., 2007: Variabilidad de la precipitación en Argentina en diferentes escalas temporales, relacionada con actividad convectiva observada. Documentación de la metodología para el tratamiento de la información. Verificación del desempeño de un modelo regional de pronóstico de precipitación. Tesis Doctoral, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires.
- Matsudo, C., Salles, M. A., y García Skabar, Y., 2021: Verificación de los pronósticos del esquema determinístico del modelo WRF para el año 2020. Nota Técnica SMN, 2021-95.
- Moser, B. A., Gallus Jr, W. A. y Mantilla, R., 2015: An initial assessment of radar data assimilation on warm season rainfall forecasts for use in hydrologic models. *Weather and Forecasting*, 30(6), 1491-1520.
- Mittermaier, M. P. y Bullock, R., 2013: Using MODE to explore the spatial and temporal characteristics of cloud cover forecasts from high-resolution NWP models. *Meteorological Applications*, 20(2), 187-196.
- Mittermaier, M., North, R., Maksymczuk, J., Pequignet, C. y Ford, D., 2020: Using feature-based verification methods to explore

- the spatial and temporal characteristics of forecasts of the 2019 Chlorophyll-a bloom season over the European North-West Shelf. *Ocean Science Discussions*, 1-47.
- Nurmi, P., 2003: Recommendations on the verification of local weather forecasts. ECMWF Tech. Memo. N. 430 (19pp).
- Roebber, P. J., Schultz, D. M., Colle, B. A. y Stensrud, D. J., 2004: Toward improved prediction: High-resolution and ensemble modeling systems in operations. *Weather and Forecasting*, 19(5), 936-949.
- Rossa A, Nurmi P. y Ebert E., 2008: Overview of methods for the verification of quantitative precipitation forecasts. *Precipitation: Advances in Measurement, Estimation and Prediction*, Michaelides S (ed.). Springer Verlag; 419-452.
- 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. *Monthly Weather Review*, 135(4), 1290-130
- Skinner, P. S., Wheatley, D. M., Knopfmeier, K. H., Reinhart, A. E., Choate, J. J., Jones, T. A., Creager, G., Dowell, D., Alexander, C., Ladwig, T., Wicker, L. Heinselman, P., Minnis, P. y Palikonda, R., 2018: Object-based verification of a Prototype Warn-on-Forecast System. *Weather and Forecasting*, 33(5), 1225-1250.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W. y Powers, J. G., 2005: A description of the advanced research WRF version 2 (No. NCAR/TN-468+ STR). Tech. Note TN-475+STR, NCAR/TN-468+STR. Tan, J., Petersen, W. A., Kirstetter, P. E. y Tian, Y., 2017: Performance of IMERG as a function of spatiotemporal scale. *Journal of hydrometeorology*, 18(2), 307-319.
- Wang, Z., Zhong, R., Lai, C. y Chen, J., 2017: Evaluation of the GPM IMERG satellite-based precipitation products and the hydrological utility. *Atmospheric Research*, 196, 151-163.
- Wolff, J. K., Harrold, M., Fowler, T., Gotway, J. H., Nance, L. y Brown, B. G., 2014: Beyond the basics: Evaluating model-based precipitation forecasts using traditional, spatial, and object-based methods. *Weather and Forecasting*, 29(6), 1451-1472.
- Yan, H. y Gallus Jr, W. A., 2016: An evaluation of QPF from the WRF, NAM, and GFS models using multiple verification methods over a small domain. *Weather and Forecasting*, 31(4), 1363-1379.
- Zingerle, C. y Nurmi, P., 2008: Monitoring and verifying cloud forecasts originating from operational numerical models. *Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling*, 15(3), 325-330.