

Método de Evaluación Diagnóstico Orientado a Objetos Modificado (MODEMod) utilizando la información de los satélites meteorológicos

Arletis Roque Carrasco¹
Maibys Sierra Lorenzo¹
Nivian Quintana Rodríguez¹

¹Instituto de Meteorología - INSMET
Código Postal 11200 - Regla – La Habana, Cuba
arletis.carrasco@gmail.com, {maibys.lorenzo, nivian.quintana}@insmet.cu

Abstract. A modification of the Method for Object-Based Diagnostic Evaluation (MODE) is developed to evaluate the forecast of discrete variables using satellite images in TIFF format. Cloud forecast from ARPS model and images of GOES-13 geostationary satellite are used to check the performance of the proposed method. The results show that the method recognizes the cloud areas as result of convolution and thresholding processes. Suitable parameters for the objects identification process are convolution radius $R = 4$ and the thresholds $T = 0.1$ and $T = 17$ for the forecast and observation, respectively. The Euclidean distance measure, derived from the centroid, area and perimeter difference, identified objects with similar characteristics in the observation and forecast.

Palavras-chave: Convolution, Objects identification, Numerical weather prediction models, Thresholds, Spatial forecast verification.

1. Introducción

Una de las herramientas fundamentales para el estudio y predicción de los fenómenos atmosféricos son los conocidos modelos meteorológicos. Estos modelos numéricos se basan en la implementación de las ecuaciones matemáticas que describen la física de la atmósfera; cuya resolución conduce usualmente a realizar aproximaciones de la situación de estudio, al no poder encontrar soluciones analíticas para las ecuaciones que se plantean. Es por eso que se debe tener presente que un modelo propone una aproximación a la realidad. La evaluación de esta aproximación se hace indispensable, debido a la importancia que tiene la exactitud de todo el sistema de pronóstico meteorológico y con el propósito de mejorar y perfeccionar la calidad de los mismos.

El proceso de evaluación de la calidad de un sistema de pronósticos basado en una muestra de pronósticos previos y en las observaciones correspondientes se conoce como verificación del pronóstico.

Los estudios realizados sobre la verificación del pronóstico numérico del tiempo han demostrado que el desempeño de los modelos de predicción numérica del tiempo debe ser evaluado dentro de un marco que reconoce la existencia de “objetos” en los campos espaciales de la observación y el pronóstico, debido a que las técnicas estándar de verificación ignoran la estructura espacial de dichos campos y tratan los errores inapropiadamente, por lo que no proporcionan información significativa que pueda ser útil para mejorar los pronósticos u orientar a los usuarios del pronóstico. Los objetos antes mencionados pueden ser parámetros discontinuos tales como la precipitación o un área de nubes, o pueden ser entidades más conceptuales que se definen por una gran conjunto de características, como los ciclones tropicales o extratropicales. Los principales estudios en esta dirección han estado dirigidos, fundamentalmente, a identificar y delimitar los objetos en los dos campos de manera significativa meteorológicamente para cuantificar la habilidad de un modelo de pronóstico (Baldwin et al. 2002, Brown et al. 2002, Brown et al. 2004, Bullock et al. 2004, Chapman et al. 2004, Davis et al. 2006a, Davis et al. 2006b).

Una de las herramientas fundamentales para la verificación del pronóstico numérico del tiempo es el MET (Model Evaluation Tools). El MET cuenta con las metodologías

tradicionales de verificación del pronóstico de variables discretas y continuas como la Raíz del Error Cuadrático Medio y el Índice de Éxito Crítico. Cuenta además, con técnicas avanzadas de verificación del pronóstico espacial destacándose el Método de Evaluación Diagnóstico Orientado a Objetos (MODE, por sus siglas en inglés) que permite la comparación de los campos espaciales, pronóstico y observación, mediante la identificación de objetos con características similares en dichos campos.

El MET necesita la información en formato RAW de los satélites meteorológicos. En Cuba solo se cuenta con la imagen procesada en formato TIFF, desventaja que limita su uso y es la causa fundamental de la poca utilización de las técnicas de verificación espacial del pronóstico de variables discretas. Quintana (2012) utiliza un método de evaluación orientado a objetos dirigido a verificar la habilidad del modelo ARPS para pronosticar fenómenos de escala local como la lluvia, sin embargo el algoritmo seleccionado para agrupar los objetos tiene la desventaja que incluye muchas áreas que no son de lluvias en los objetos encontrados.

En el presente trabajo se realiza una modificación del MODE con el objetivo de evaluar el pronóstico de variables discretas a partir de la información en formato TIFF de los satélites meteorológicos.

2. Materiales y métodos

2.1. Método de Evaluación Diagnóstico Orientado a Objetos Modificado (MODEMod)

El Método de Evaluación Diagnóstico Orientado a Objetos Modificado (MODEMod), como lo indica su nombre, representa una variación del MODE desarrollado por Davis (2009).

El MODE se utiliza de manera generalizada para comparar dos campos: “el pronóstico” y “la observación”. Esta comparación se realiza mediante la identificación de objetos o regiones de interés en ambos campos mediante dos procesos. El primer proceso realiza una convolución con el propósito de suavizar o interpolar los datos originales y agrupar las áreas significativas de nubosidad o precipitación mediante una función de filtro circular:

$$C(x, y) = \sum_{u, v} \varphi(u, v) f(x - u, y - v) \quad (1)$$

donde f es el campo que contiene los datos sin procesar, φ es la función de filtro y C es el campo que se obtiene como resultado. Las variables (x, y) y (u, v) son las coordenadas de la rejilla. La función de filtro φ es un filtro circular simple determinado por el radio de influencia R (o radio de convolución) y la altura H :

$$\varphi(x, y) = H \quad \text{si} \quad x^2 + y^2 \leq R^2, \quad \varphi(x, y) = 0 \quad (2)$$

Los parámetros R y H no son independientes. Están relacionados por:

$$\pi R^2 H = 1 \quad (3)$$

El radio de influencia R debe ser ajustado por el usuario. Una vez que se seleccione R , se calcula H por la ecuación (3).

El segundo proceso se basa en la aplicación de umbrales para filtrar los rasgos más pequeños que no son de interés para el usuario, obteniéndose así un nuevo campo M definido por una matriz de ceros y unos, siendo los objetos las regiones conectadas donde $M=1$:

$$M(x, y) = 1 \quad \text{si} \quad C(x, y) \geq T, \quad M(x, y) = 0 \quad (4)$$

Posteriormente se aplica la máscara al campo original para obtener el campo objeto F :

$$F(x,y) = M(x,y)f(x,y) \quad (5)$$

Una vez concluido el proceso de identificación, se obtiene información de cada uno de los objetos identificados mediante el cálculo de una serie de propiedades que los caracterizan y que son llamadas atributos. Esta información es utilizada en un algoritmo de lógica difusa para asociar los objetos dentro de un mismo campo, así como para hacer corresponder los objetos del campo pronóstico con los de la observación.

Para cada uno de los objetos identificados se calcularon los siguientes atributos:

El área es el número de cuadrículas que ocupa un objeto.

El perímetro es el número de cuadrículas que conforman el borde de un objeto.

La posición o centroide es el centro de masa del objeto. Se utiliza para tener una medida de la distancia entre los objetos. Es caracterizado por dos valores escalares (latitud-longitud, o las coordenadas de la rejilla x,y).

La variación introducida al algoritmo que utiliza el MODE parte de la lectura de las imágenes de satélite en formato TIFF. En segundo lugar se sustituye el proceso de asociación y correspondencia de los objetos mediante la lógica difusa, por el cálculo de una medida de distancia euclidiana D en función de la diferencia entre los atributos área, perímetro y centroide entre los objetos que están siendo comparados.

$$D = \sqrt{\frac{(Cp - Co)^2}{N1} + \frac{(Ap - Ao)^2}{N2} + \frac{(Pp - Po)^2}{N3}} \quad (6)$$

donde Cp, Co, Ap, Ao, Pp y Po son el centroide, el área y el perímetro correspondientes a los objetos encontrados en el pronóstico y en la observación, respectivamente. Estos valores fueron normalizados con el propósito de garantizar que todos los atributos tuvieran la misma importancia relativa para seleccionar cuáles objetos se corresponden. Los pares de objetos seleccionados son aquellos para los cuales se obtuvo el menor valor de D.

2.2. Herramientas utilizadas para la implementación

La implementación del método propuesto se realizó en un programa escrito en el lenguaje de programación C, dicho programa contiene todos los pasos que sigue el algoritmo para la identificación de los objetos, tanto en el campo del pronóstico como en la observación.

Los datos del pronóstico de la variable nubosidad se obtuvieron de los ficheros de salida del modelo ARPS, correspondientes a las razones de mezcla para el agua de nube (qc), agua de lluvia (qr), hielo de nube (qi), nieve (qs) y granizo o graupel (qh); estos datos se extrajeron en formato de texto usando el módulo arpsread. Estos ficheros están basados en un pronóstico de 24 horas inicializado a las 0400 h (0900 UTC) con una resolución espacial de 3 km. Para el procesamiento de dichos ficheros y su representación gráfica se utilizó la herramienta GrADS.

Los datos del pronóstico se compararon con los datos reales de las imágenes provenientes del sensor GOES-Imager a bordo del satélite geoestacionario GOES-13. Para los horarios seleccionados de los casos de estudio se utilizaron las imágenes del espectro visible (0.6 μ m) referidas al canal 1 con una resolución espacial de 1 km. Dichas imágenes fueron acotadas mediante un algoritmo de georreferenciación, desde el ecuador hacia los polos entre los 21.02°N y los 23.72°N y de Este a Oeste por los -85.37°W y los -79.42°W.

Para el procesamiento de las imágenes de satélite se utilizó la biblioteca Libtiff. El uso de esta biblioteca permitió generar las imágenes correspondientes a cada uno de los procesos incluidos en el método desarrollado, facilitando el análisis y la comprobación del comportamiento de dicho método.

3. Resultados y discusión

3.1. Selección de los parámetros utilizados para la identificación de los objetos

El tamaño del filtro y el radio de influencia R fueron especificados en términos de unidades de rejilla. El umbral se especificó en términos de niveles digitales para el caso de las observaciones y en términos de las unidades de la razón de mezcla (g/Kg) para los datos del pronóstico. Con el objetivo de seleccionar los valores adecuados para dichos parámetros se llevaron a cabo varias pruebas. Las primeras pruebas se realizaron fijando un tamaño para el filtro circular y variando el radio; se pudo observar que para un radio $R=1$ el campo suavizado resultó ser idéntico al campo original, obteniéndose un mayor suavizado a medida que se incrementó el valor del radio. Se seleccionó para este parámetro un valor de $R=4$, pues para valores mayores se pierde mucha información en el proceso de suavizado.

Una vez seleccionado el radio de influencia, se hicieron comparaciones para un mismo valor del radio variando el tamaño del filtro. Se pudo comprobar que el nivel de suavizado no depende del valor definido para este parámetro, obteniéndose el mismo producto resultante del proceso de suavizado independientemente de cuál fuese el tamaño del filtro. Se pudo concluir entonces, que es el radio de influencia el de mayor importancia en el proceso de convolución, y por lo tanto el que determina cuánto se desea suavizar la imagen. Estas pruebas se realizaron tanto para el pronóstico como para la observación.

El siguiente paso fue la selección del umbral; las pruebas se realizaron para diferentes valores de umbrales y con el propósito de identificar el área de nubosidad mínima de interés y optimizar la cantidad de objetos a detectar. Se pudo comprobar que a medida que se incrementa el umbral, el tamaño de los objetos disminuye, propiciando además una variación en el número de objetos detectados, aunque en este caso no se pudo encontrar una relación directamente proporcional a la variación del umbral. Para el caso de las imágenes de satélite, o sea, para el campo de la observación, se escogió un umbral $T=17$ pues el área que ocupan los objetos identificados es muy parecida al área que ocupan los objetos de interés en la imagen original; para valores mayores el valor de este atributo disminuye. De igual manera se seleccionó para el campo del pronóstico un umbral $T=0.1$, siendo el que mejor identificó las áreas de nubes de interés.

En la Figura 1 se muestra un ejemplo de las primeras 3 etapas implicadas en el proceso de identificación de objetos a partir de los parámetros definidos. La Figura 1d muestra el campo resultante una vez que la máscara obtenida después del umbral fue aplicada a los valores de nubosidad originales para obtener el campo objeto.

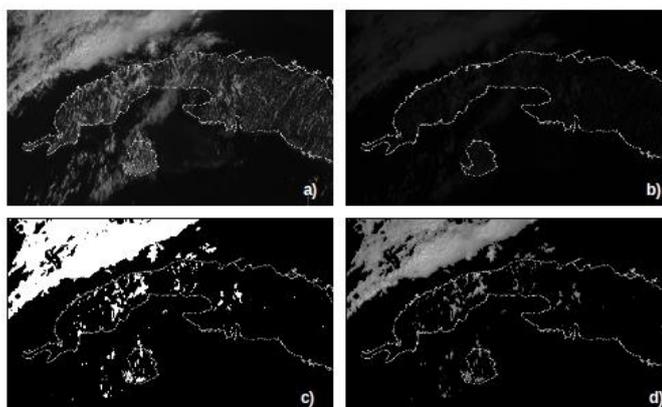


Figura 1. Convolución y aplicación del umbral para la identificación de objetos en el campo de la observación: a) campo de nubosidad, b) campo de convolución c) campo después que se aplica el umbral, d) campo objeto

3.2. Aplicación del método de evaluación

Una vez seleccionados los parámetros adecuados, el método fue aplicado al pronóstico de nubosidad realizado por el modelo ARPS. El método desarrollado detectó correctamente las áreas de nubosidad en ambos campos como resultado de la convolución y la aplicación de umbrales, procesos que permitieron agrupar las áreas de nubes significativas y filtrar los rasgos más débiles que no eran de interés para el usuario. Los atributos calculados permitieron encontrar diferencias y semejanzas entre los objetos y, por lo tanto, determinar para un objeto en el campo pronóstico cuál le correspondía en el campo de la observación. No obstante, no se pudo encontrar una correspondencia para todos los objetos, quedando objetos sin pareja en cada uno de los campos. Los objetos no observados en el pronóstico pudieran ser consecuencia de la resolución con que se realizaron las corridas del ARPS la cual no es suficiente para describir las áreas muy pequeñas de nubosidad, sobre todo aquellas que son resultado de fenómenos locales; las condiciones de inicialización del ARPS que están asociadas a los errores que puede tener el WRF (modelo del cual se alimenta el ARPS); a los datos del terreno y a la parametrización de la microfísica utilizada para representar, fundamentalmente, los procesos de convección. Por estas razones se consideró importante añadir una condición que desprecia los objetos con áreas pequeñas que no son de interés para el usuario.

Se utilizó como indicador del comportamiento del pronóstico realizado por el modelo el Índice de Éxito Crítico (CSI por sus siglas en inglés) orientado a objetos, que en este caso se deriva del número de pares de objetos seleccionados N_m (aciertos) dividido por la suma del número de pares de objetos seleccionados, el número de objetos detectados en la observación a los que no se les encontró una correspondencia en el pronóstico N_o (fallos) y el número de objetos detectados en el pronóstico a los que no se les encontró una correspondencia en la observación N_p (falsas alarmas).

$$CSI = \frac{N_m}{N_m + N_o + N_p} \quad (7)$$

Esta aplicación del CSI difiere de la aplicación tradicional en la verificación del pronóstico numérico, porque aquí la evaluación se lleva a cabo para objetos en vez de para celdas individuales de la rejilla, (Jolliffe y Stephenson, 2003). Calculamos el CSI orientado a objetos, ya que tiene una relación estadística con las métricas de verificación tradicionales, pero conserva la ventaja de la perspectiva orientada a objetos.

3.2.1 Caso de estudio 25 de marzo de 2013

El 25 de marzo de 2013 afectó al occidente cubano el frente frío número 15 de la temporada invernal 2012-2013, dicho sistema frontal estuvo precedido por una línea de inestabilidad que ocasionó nublados con algunas lluvias, chubascos y tormentas eléctricas, principalmente en zonas de la costa norte. Este sistema fue clasificado como un frente frío débil por su intensidad y clásico por su tipo (González y Ortega, 2013).

La Figura 2a corresponde a la imagen de satélite de las 1102 h (1602 UTC) del 25 de marzo de 2013. En dicha imagen se puede observar la proximidad del sistema frontal a la costa norte occidental del territorio cubano y los nublados asociados a dicho sistema. Las áreas más pequeñas de nubosidad están relacionadas a la inestabilidad presente por delante del frente frío, producto de los vientos de componente sur en esa región. El pronóstico realizado por el modelo para ese horario manifestó un adelanto del frente con respecto a la observación, indicando que para esa hora, el territorio cubano ya se encontraba bajo la influencia del sistema frontal, (Figura 2b).

En la Figura 2c y Figura 2d se observan los objetos que fueron detectados en cada uno de los campos, observación y pronóstico, respectivamente. Los pares de objetos identificados están representados con igual color en ambas imágenes. En este caso el objeto 1 en el

pronóstico se asoció con el objeto 2 en la observación, representado por el color amarillo, coincidiendo con la nubosidad correspondiente al frente frío.

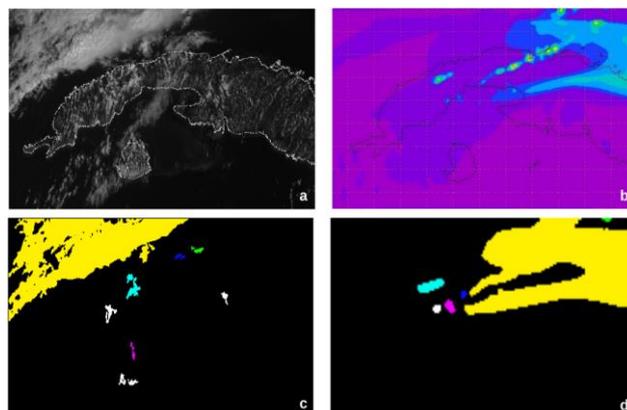


Figura 2. Identificación de objetos en los campos observación y pronóstico (25 de marzo de 2013): a) campo observación, b) campo pronóstico, c) y d) máscaras correspondientes a la observación y el pronóstico respectivamente con los pares de objetos pintados con igual color

Para este día el método de evaluación desarrollado, detectó 6 objetos en el pronóstico y 9 en la observación. De ellos encontró una correspondencia para 5 objetos (aciertos), identificando 4 fallos y 1 falsa alarma, por lo tanto el CSI calculado fue de 0.5, valor que está influenciado por la cantidad de fallos encontrados en el pronóstico.

El pronóstico realizado por el modelo coincidió en gran medida con la observación en cuanto a la forma y el área de las zonas de nubosidad detectadas, pero manifestó un corrimiento en el tiempo en cuanto a la posición de la nubosidad correspondiente al sistema frontal presente.

3.2.2 Caso de estudio 26 de marzo de 2013

El 26 de marzo de 2013, el frente frío número 15 de la temporada invernal 2012-2013, al cual se hizo referencia en el epígrafe anterior, se encontraba transitando sobre las provincias occidentales y centrales del territorio cubano. La nubosidad asociada a dicho sistema se observa en la Figura 3a.

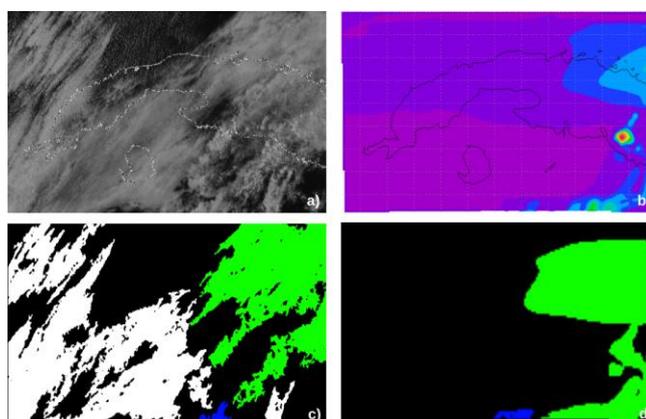


Figura 3. Identificación de objetos en los campos observación y pronóstico (26 de marzo de 2013): a) campo observación, b) campo pronóstico, c) y d) máscaras correspondientes a la observación y el pronóstico respectivamente con los pares de objetos pintados con igual color

Para este día el modelo solo representó correctamente el área de nublados localizada sobre las provincias centrales, identificadas en la parte oriental del dominio, Figura 3b. El método de evaluación encontró en la observación (Figura 3c), las áreas de nubosidad con características similares a las áreas detectadas en el pronóstico (Figura 3d), identificando 2 aciertos y 4 fallos, resultados para los cuales se obtuvo un valor de CSI de 0,333333.

3.2.3 Caso de estudio 2 de marzo de 2013

El 1ro de marzo de 2013 se denotó que desde un complejo sistema de bajas presiones extratropicales situado en los 40.0°N y los 68.0°W, se extendía al suroeste sobre el Atlántico occidental y muy próximo a la costa oriental de los Estados Unidos un frente frío que cruzaba las Bahamas y continuaba sobre el canal de la Florida y los mares adyacentes a la costa norte de las provincias occidentales de Cuba. Este frente avanzó al este desde el final de la mañana de ese día afectando al occidente cubano, produciendo además un incremento de los nublados e imposición de vientos del cuarto cuadrante. Ese sistema frontal se convirtió en el frente frío número 11 de la temporada invernal 2012-2013. Y al continuar su avance hacia el este sobre el territorio cubano su banda frontal se unió a un sistema casi estacionario que se hallaba extendido desde días anteriores sobre las provincias desde Pinar del Río hasta Matanzas. La fusión de esos sistemas frontales ocasionó lluvias y chubascos desde la madrugada hasta la mañana del día 2 de marzo en la región occidental, situación que se trasladó hacia la región central en esa tarde y después continuó sobre el oriente en la noche y madrugada del día 3 (González y Ortega, 2013).

En la Figura 4a se observa la nubosidad correspondiente a este frente frío, sistema que fue representado correctamente por el modelo (Figura 4b). El método de evaluación identificó el área de nublados relacionados con el frente frío como un solo objeto en la observación (Figura 4c) y el pronóstico (Figura 4d), que fue considerado como acierto; no encontró fallos ni falsas alarmas, por lo que el valor del CSI=1 indicó que el pronóstico realizado por el modelo para ese día, coincidió perfectamente con lo observado en la imagen del satélite meteorológico.

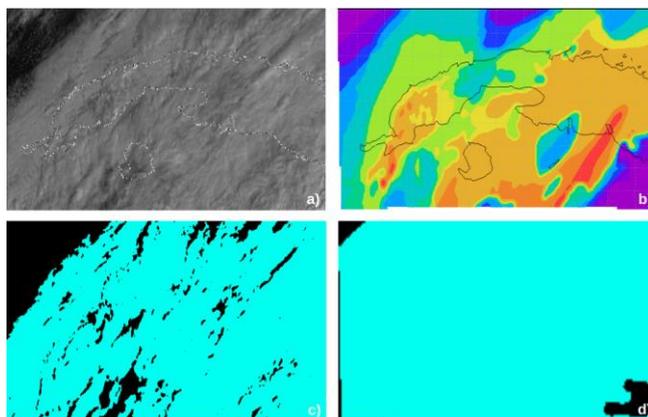


Figura 4. Identificación de objetos en los campos observación y pronóstico (2 de marzo de 2013): a) campo observación, b) campo pronóstico, c) y d) máscaras correspondientes a la observación y el pronóstico respectivamente con los pares de objetos pintados con igual color

4. Conclusiones

A partir del análisis de los resultados obtenidos en este trabajo y presentados en los epígrafes anteriores se arribó a las siguientes conclusiones:

1. Se elaboró un método que permite evaluar el pronóstico numérico de variables discretas a partir de la información en formato TIFF de los satélites meteorológicos.

2. Se encontró que los parámetros adecuados para el proceso de identificación de objetos son un radio de convolución $R=4$ y los umbrales $T=0.1$ para el pronóstico y $T=17$ para la observación.
3. La distancia euclidiana derivada de la diferencia de centroide, la diferencia de área y la diferencia de perímetro, permitió identificar la correspondencia entre los objetos de la observación y el pronóstico.

Referencias Bibliográficas

Baldwin, M. E., S. Lakshmivarahan, and J. S. Kain, 2002. **Development of an events-oriented approach to forecast verification**. 15th Conf. Numerical Weather Prediction, San Antonio, TC., August 12-16, 2002.

Brown, B.G., J.L. Mahoney, C.A. Davis, R. Bullock, and C.K. Mueller, 2002. **Improved approaches for measuring the quality of convective weather forecasts**. Preprints, 16th Conference on Probability and Statistics in the Atmospheric Sciences. Orlando, 13-17 May, American Meteorological Society (Boston), 20-25.

Brown, B.G., R. Bullock, C.A. Davis, J.H. Gotway, M. Chapman, A. Takacs, E. Gilleland, J. L. Mahoney, and K. Manning, 2004. **New verification approaches for convective weather forecasts**. Preprints, 11th Conference on Aviation, Range, and Aerospace, Hyannis, MA, 3-8 October.

Bullock, R., B.G. Brown, C.A. Davis, K.W. Manning, and M. Chapman, 2004. **An Object-oriented Approach to Quantitative Precipitation Forecasts**. Preprints, 17th Conference on Probability and Statistics in the Atmospheric Sciences. Seattle, 11-15 January, American Meteorological Society (Boston).

Chapman, M., R. Bullock, B. G. Brown, C. A. Davis, K. W. Manning, R. Morss, and A. Takacs, 2004. **An Object Oriented Approach to the Verification of Quantitative Precipitation Forecasts: Part II - Examples**. Preprint, AMS 2004.

Davis, C., B. Brown, and R. Bullock, 2006a. **Object-based verification of precipitation forecasts. Part I: Methods and application to mesoscale rain areas**. Mon. Wea. Rev., 134, 1772-1784.

-----, 2006b. **Object-based verification of precipitation forecasts, Part II: Application to convective rain systems**. Mon. Wea. Rev. 134, 1785-1795.

Davis, C. A., B. G. Brown, R. G. Bullock, and J. Halley Gotway, 2009. **The Method for Object-based Diagnostic Evaluation (MODE) applied to WRF forecasts from the 2005 Spring Program**. Wea. Forecasting, 24, 1252-1267.

González C. y A. Ortega, 2013. **Resumen sinóptico mensual**. Grupo de investigaciones del centro de pronóstico del Instituto de Meteorología. Disponible en: <http://www.insmet.cu>.

Jolliffe, I. T., and D. B. Stephenson, 2003. **Forecast Verification. A Practitioner's Guide in Atmospheric Science**. Wiley and Sons, 240 pp.

Quintana, N., 2012. **Evaluación espacial del pronóstico de precipitación del modelo ARPS para la campaña de lluvia en el occidente de Cuba**. Revista Cubana de Meteorología, 18, No.2, Noviembre 2012.