

## **Modificación e implementación del método de evaluación espacial MODEMod para su uso operativo en Cuba**

Camilo Fernando Rodríguez-Genó <sup>(1\*)</sup>, Maibys Sierra-Lorenzo <sup>(1)</sup> y Adrián Luis Ferrer-Hernández <sup>(1)</sup>

<sup>(1)</sup> *Centro de Física de la Atmósfera, Instituto de Meteorología, Loma de Casablanca, Regla, La Habana, Apartado 17032, Habana 17. C.P. 11700, Cuba, E-mail: kmi.rodriguez901101@gmail.com*

Recibido: octubre 22, 2014	Acceptado: septiembre 30, 2015
----------------------------	--------------------------------

### **Resumen**

En el presente estudio se realiza una modificación del Método de Evaluación y Diagnóstico Orientado a Objetos Modificado (MODEMod) para evaluar el pronóstico de variables discretas a partir de imágenes provenientes de los satélites meteorológicos. Para comprobar el comportamiento del método se utilizan datos de pronóstico de nubosidad del modelo Weather Forecast and Research (WRF) y las imágenes del satélite GOES-13. El método reconoce las áreas de nubosidad como resultado de un proceso de convolución y aplicación de umbrales para agrupar las áreas de nubes significativas y filtrar los rasgos más débiles. Se utiliza un algoritmo de lógica difusa para agrupar los objetos en un mismo campo y encontrar objetos similares en campos distintos. La herramienta desarrollada permite verificar el pronóstico de elementos como la precipitación, la convección y la nubosidad, cuya variabilidad espacial, estructura y localización específica pueden dificultar el pronóstico.

**Palabras clave:** Evaluación espacial, MODE, MET, WRF.

## **Modification and implementation of the method for spatial evaluation MODEMod for its operative use in Cuba**

### **Abstract**

In the present study a modification of the Modified Method for Object-Based Diagnostic Evaluation (MODEMod) is developed to evaluate the forecast of discrete variables using the processed images in TIFF or ASCII format from meteorological satellites. Integrated cloud fraction forecast from WRF model and images from satellite GOES-13 are used to check the performance of the proposed method. The method recognizes the cloud areas as result of convolution and thresholding processes, which allowed grouping the most significant cloud areas and filtering the weaker features. A fuzzy logic algorithm is used in order of merge objects of the same field, and match similar objects of different fields. The tool developed allows verifying the prediction of elements such as precipitation, convection and clouds, whose spatial variability, structure and specific location may hinder the forecast.

**Keywords:** Spatial evaluation, MODE, MET, WRF.

## 1. Introducción

En la actualidad, una de las principales herramientas para el pronóstico del tiempo son los modelos de pronóstico numérico, los que, basándose en la aplicación de ecuaciones físico-matemáticas que describen el comportamiento de la atmósfera, predicen con bastante exactitud el estado del tiempo. En la mayoría de los casos, estos modelos no son capaces de encontrar soluciones analíticas de las ecuaciones de la física atmosférica, por lo que se obtiene una solución aproximada. Esto se debe fundamentalmente al incompleto entendimiento de los fenómenos que se pretende pronosticar, además de la realización de aproximaciones numéricas para dar solución a los sistemas de ecuaciones.

Debido a este error inherente en los modelos numéricos de pronóstico, es que merece suma importancia la evaluación de sus resultados, que en el caso de los modelos numéricos incluye la determinación sistemática de la exactitud, el error cometido y el significado de los resultados en función de unos criterios determinados respecto a un conjunto de normas con el objetivo de incrementar la calidad. Para realizar la evaluación de un modelo de pronóstico, se comparan los resultados obtenidos con el valor real correspondiente a las observaciones, lo que da una medida del error cometido.

En la actualidad existen tres factores que garantizan un pronóstico correcto:

- el nivel de correspondencia entre el pronóstico numérico y el juicio del pronosticador,
- el grado en que el pronóstico se ajusta a lo ocurrido realmente y
- el valor del pronóstico para la toma de decisiones.

Hasta ahora, los métodos convencionales de verificación no garantizan el cumplimiento de estos factores, ya que no tienen en cuenta la estructura espacial de los fenómenos, por lo que se hace cada vez más urgente la evaluación de los modelos de pronóstico numérico de forma tal que considere la existencia de objetos en los campos espaciales de las observaciones y pronósticos. Dichos objetos pueden tener carácter discreto (áreas de nubes, precipitación), o constituir fenómenos descritos conceptualmente (cyclones tropicales, bajas extratropicales). Ambos casos reúnen un gran grupo de características que los definen, por lo que son mucho más difíciles de evaluar.

Los principales estudios en esta dirección han estado dirigidos fundamentalmente a identificar y delimitar los objetos en ambos campos de forma significativa, y así cuantificar la habilidad del modelo para representar las estructuras espaciales de los fenómenos meteorológicos (Du y Mullen, 2000; Ebert y McBride, 2000; Lakshminarahan *et al*, 2001; Lakshminarahan *et al*, 2002; Mahoney *et al*, 2002; Bullock *et al*, 2004a; Bullock *et al*, 2004b; Brown *et al*, 2004; Basu y Foufoula-Georgiu, 2005; Brown *et al*, 2006a; Brown *et al*, 2006b)

En nuestro país sólo Quintana emplea métodos de evaluación espacial aplicados al pronóstico de precipitación del modelo *Advanced Regional Prediction System* (ARPS) (Quintana, 2012), sin embargo, el método empleado para agrupar los objetos no es eficiente, ya que incluye muchas áreas que no son de lluvia.

Uno de los métodos de evaluación espacial más empleados es el *Method for Object-Based Diagnostic Evaluation* (MODE), el cual constituye una de las técnicas de evaluación de pronóstico espacial pertenecientes al *Model Evaluation Tools* (MET), el cual constituye una de las herramientas fundamentales de verificación de pronósticos numéricos (MET Developers, 2011). Esta herramienta fue modificada con el objetivo de evaluar el pronóstico de variables discretas a partir de las imágenes procesadas en formato TIFF provenientes de los satélites meteorológicos. Dicha modificación, denominada MODE Modificado (MODEMod), posibilitó el reconocimiento de áreas de nubosidad y permite, a partir de un proceso de convolución y aplicación de umbrales, agrupar las áreas de nubes significativas y filtrar los rasgos más débiles. Para la evaluación se empleó una medida de distancia euclidiana que incluye la diferencia de posición del centroide, la diferencia de área y la diferencia de perímetro entre los objetos que están siendo comparados, con el objetivo de encontrar los objetos similares en los campos de observaciones y de pronóstico. Además, el método aplica un umbral de área para discriminar los objetos pequeños que el modelo es incapaz de representar debido a la configuración utilizada (Roque, 2013; Roque *et al*, 2014). Sin embargo, los atributos considerados resultan insuficientes para caracterizar los objetos identificados. Además, el método de agrupamiento puede ser mejorado empleando un algoritmo basado en lógica difusa y fusión de objetos, que tiene la ventaja de aplicar objetivamente los juicios humanos sobre cómo los objetos deben coincidir a partir del cálculo de una gran variedad de atributos que permiten agrupar los objetos de un mismo campo y posteriormente hacer corresponder los objetos de campos diferentes, brindando de esta forma mejores resultados.

Además de las limitaciones mencionadas anteriormente, ninguno de estos métodos se encuentra funcionando de forma operativa en nuestro país, evaluando en tiempo real los resultados de los modelos de pronóstico numérico. Por ello se proponen como objetivos:

- Modificar el método de evaluación espacial MODEMod para lograr una mejor caracterización y agrupación de los objetos.
- Comprobar el comportamiento del método en diferentes casos de estudio.
- Implementar de forma operativa las modificaciones realizadas a la herramienta de evaluación espacial MODEMod.

Una vez implementado el MODEMod, se contará con un método de evaluación espacial verificando los resultados de los pronósticos numéricos de forma operativa. Dicho método proporcionará información acerca de las características de las áreas detectadas, así como una evaluación en tiempo casi-real de las salidas de los modelos, lo que permitirá la rápida identificación de los errores de ubicación espacial y temporal de los fenómenos, la magnitud de los mismos, etc.

## 2. Materiales y Métodos

El MODEMod se puede utilizar de manera generalizada para comparar dos campos. El campo 1, considerado como el *pronóstico*, mientras que el campo 2 puede ser considerado como la *observación*. Esta herramienta detecta los objetos en ambos campos; con este propósito realiza un primer proceso llamado convolución y un segundo proceso que se basa en la aplicación de umbrales. Los objetos identificados imitan lo que los humanos llamarían *regiones de interés*. Luego calcula propiedades que caracterizan a los objetos llamadas *atributos* (área, perímetro y posición del centroide). Estos atributos se comparan y se emplea una medida de distancia euclidiana para hacer corresponder los objetos del campo pronóstico con los de la observación (Roque, 2013; Roque *et al.*, 2014). Estos atributos se pueden utilizar para identificar correlaciones y diferencias entre los objetos, lo que conlleva a percepciones sobre las fortalezas y debilidades del pronóstico.

### 2.1 Atributos

Los atributos del objeto se definen tanto para objetos individuales (caracterización) como para los pares de objetos (comparación). Los objetos de un par pueden seleccionarse en un mismo campo para realizar una fusión de objetos, o de campos diferentes para hacer coincidir los objetos similares en campos diferentes, y por lo tanto medir el rendimiento del pronóstico.

Muchos de los atributos son definidos geoméricamente, incluyendo características de los objetos. Los atributos implementados en la aplicación MODEMod (Roque, 2013; Roque *et al.*, 2014) son:

- **Área:** es el número de cuadrículas que ocupa un objeto.
- **Perímetro:** es el número de cuadrículas que conforman el borde de un objeto.
- **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  $x, y$ ).

Los nuevos atributos calculados son:

**Orientación:** Caracteriza la dirección general u orientación del objeto.

Para su cálculo se necesita el centroide del objeto y los puntos del perímetro del mismo. Se halla la recta que pasa por el centro de masa y que cumpla con que la suma de los cuadrados de las distancias desde los puntos del perímetro sea mínima (Chaudhuri y Samal, 2007). La ecuación de dicha recta en función del ángulo es:

$$x \tan \theta - y + \bar{y} \tan \theta = 0 \quad (1)$$

Mientras que la ecuación de la distancia perpendicular de cada punto del perímetro a la recta es:

$$p_i = (x_i - \bar{x}) \sin \theta - (y_i - \bar{y}) \cos \theta \quad (2)$$

Por lo que la suma de los cuadrados de las distancias perpendiculares es:

$$P = \sum_{i=1}^n [(x_i - \bar{x}) \sin \theta - (y_i - \bar{y}) \cos \theta]^2 \quad \frac{\partial P}{\partial \theta} = 0 \quad (3)$$

Aplicando la condición de mínimo la orientación queda expresada de la forma:

$$\tan \theta = \frac{2 \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n [(x_i - \bar{x})^2 - (y_i - \bar{y})^2]} \quad (4)$$

**Razón de aspecto:** Razón entre la longitud del eje menor y el eje mayor del objeto.

Una vez obtenida la orientación del objeto se calculan las rectas que conforman los ejes principales a partir de las ecuaciones de la recta. Posteriormente se clasifican los puntos del perímetro por su posición respecto al cada eje (a la derecha, a la izquierda o sobre el eje), para hallar después los puntos más distantes a cada lado de los ejes principales (Chaudhuri y Samal, 2007).

**Complejidad:** Razón entre el área del objeto y el área de su envoltura convexa.

Para hallar los puntos que conforman la envoltura convexa de un objeto se emplea el método Quickhull (Mount, 2000), que determina la envoltura convexa de forma recursiva. Una vez logrado lo anterior, se recorre el objeto, determinando el área de la envoltura convexa al contar los puntos que se encuentran dentro de los límites de dicho polígono.

A cada objeto se le aplica un umbral de área para discriminar los objetos muy pequeños que la resolución del modelo empleado no permite detectar. El umbral de área mínimo es de dos cuadrículas pues es la cantidad de puntos necesaria para calcular los atributos implementados. Todos estos atributos se definen para objetos individuales. Una vez que éstos se determinan, pueden ser utilizados para calcular los atributos para los pares de objetos:

**Distancia mínima entre los bordes:** Es la distancia entre los puntos más cercanos de ambos objetos. Da una medida de la cercanía de los bordes de los objetos.

**Diferencia de centroide:** Es la diferencia entre los centroides de los dos objetos. Proporciona una idea del desplazamiento que existe entre el centro de masa de ambos.

**Razón de las áreas:** es la razón entre el área del objeto más pequeño entre el área del objeto más grande. Proporciona una medida de la semejanza de los objetos en cuanto a su tamaño.

**Razón de los perímetros:** Es la razón entre el perímetro del objeto más pequeño y el perímetro del objeto más grande. Proporciona una medida de la semejanza de los objetos en cuanto a su forma.

**Diferencia de ángulo:** Es la diferencia entre los ángulos de orientación de ambos objetos. Proporciona una medida de la diferencia de inclinación de los ejes de los objetos.

**Proporción de aspecto:** Es la proporción existente entre el valor de razón de aspecto más pequeño y el valor de razón de aspecto más grande. Proporciona una medida de la semejanza de los objetos en cuanto a su forma.

**Razón de complejidad:** Es la razón entre el valor de complejidad más pequeño y el valor de complejidad más grande. Proporciona una medida de la semejanza de los objetos en cuanto a su forma.

## 2.2 Correspondencia entre los objetos

Una vez que los atributos  $\alpha_1, \alpha_2, \dots, \alpha_n$  de cada par de objetos son calculados, se emplean en un algoritmo de lógica difusa que realiza los pasos de correspondencia y fusión de los objetos. En el paso de *fusión (merging)* se agrupan los objetos que se encuentran en el mismo campo, mientras que en el paso de *correspondencia (matching)*, se hacen corresponder los objetos similares que se encuentran en campos distintos. El algoritmo de lógica difusa emplea funciones lineales de interés para calcular los valores de interés  $I_i$  de cada atributo, los cuales se encuentran entre cero (sin interés) y uno (interés máximo). Posteriormente se calculan valores de confianza  $C_i$  para cada atributo. Los valores de confianza tienen valor uno para todos los atributos excepto la orientación, la cual depende de la razón de aspecto, ya que mientras más circular sea el objeto, menos confiable será el cálculo de la orientación del mismo. El próximo paso es asignar pesos ( $w_i$ ) a cada atributo en función de la importancia relativa de los mismos. Finalmente se calcula un valor de *interés total* de la siguiente forma:

$$T(\alpha) = \frac{\sum_i w_i C_i(\alpha) I_i(\alpha)}{\sum_i w_i C_i(\alpha)} \quad (5)$$

A este valor de interés total se le aplica un umbral (por defecto 0.7), y los pares de objetos con  $T(\alpha)$  por encima del mismo son fusionados si se encuentran en el mismo campo, o se hacen corresponder si están en campos distintos (MET Developers, 2011).

### 2.3 Técnica de fusión de objetos

Como se explicó anteriormente, todos los pares de objetos que tengan un valor de interés total por encima de un umbral definido por el usuario y se encuentren en el mismo campo (ya sea de pronóstico u observación), son fusionados. Para realizar esta fusión de objetos es necesario realizar los siguientes pasos:

1. Se determinan los puntos más cercanos entre los dos objetos.
2. Se extrae el área que se encuentra entre ambos puntos.
3. En esta área se realiza un suavizamiento de los objetos aplicando un filtro de radio igual a la mitad de la distancia entre los objetos, y empleando el mismo proceso de convolución que se usa para la identificación de objetos.

Al realizar este procedimiento los objetos quedan fusionados de tal manera que se afecta lo menos posible su morfología, y por lo tanto se modifica lo menos posible el campo involucrado.

### Estadígrafos de evaluación orientados a objetos

Para evaluar el rendimiento del modelo se calculó el *Índice de Éxito Crítico* (CSI) orientado a objetos (Davis *et al*, 2005), que en este caso se deriva del número de pares de objetos seleccionados  $N_m$  (aciertos) dividido entre 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).

$$CSI = \frac{N_m}{N_m + N_o + N_p} \quad (6)$$

Además se calcularon otros tres índices orientados a objetos que complementan la información brindada por el CSI.

La *Probabilidad de Detección* es la proporción de ocurrencias de eventos que se pronosticaron correctamente.

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

La *Proporción de Falsas Alarmas* es la proporción de eventos que se pronosticaron y no ocurrieron.

$$FAR = \frac{N_p}{N_m + N_p} \quad (8)$$

El sesgo *Bias* es la relación entre el número de pronósticos de ocurrencia y el número de ocurrencias reales. Nos da una medida de cuál de los dos errores típicos predomina en el pronóstico.

$$Bias = \frac{N_m + N_p}{N_m + N_o} \quad (9)$$

### Herramientas utilizadas para la implementación

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

Para la verificación del pronóstico de nubosidad los datos se obtuvieron de los ficheros de salida del modelo *Weather Research and Forecast- Advanced Research WRF* (WRF-ARW) (Colectivo de autores, 2012), correspondientes a la suma de la variable *cloud fraction* correspondiente a los niveles bajos, medios y altos, obteniendo la variable *integrated cloud fraction* (fracción de cielo cubierto), con una resolución espacial de 3 km y la configuración que se muestra en la Tabla I. Estos datos se extrajeron en formato de texto utilizando la herramienta *grads2ascii*, y posteriormente interpolados a una malla de 4 km de resolución espacial, para de esta forma hacerlo compatible con las observaciones.

**Tabla I** – Configuración del modelo WRF-ARW utilizada.

Configuración WRF-ARW	
Resolución espacial	3 km
Cantidad de nodos (x)	411
Cantidad de nodos (y)	183
Niveles Verticales	27
Tope	100 mb
Duración del Pronóstico	24 horas
Condiciones de Frontera	Cada 3 horas
Microfísica	WSM5
PBL	MYNN 2.5 level TKE
Radiación de Onda Corta	Dudhia
Radiación de Onda Larga	Esquema RRTM
Cúmulos	Grell-Freitas

Los datos del pronóstico se compararon con los datos reales de nubosidad de las imágenes provenientes del sensor GOES-Imager a bordo del satélite GOES-13. Para los horarios seleccionados se utilizaron las imágenes del espectro visible (0.6 $\mu$ m) referidas al canal 1 con una resolución espacial de 4 km. Dichas imágenes fueron acotadas mediante un algoritmo de georreferenciación, desde el ecuador hacia los polos entre los 19.32°N y los 24.28°N, y de Este a Oeste por los 73.75°W y los 85.73°W, y recibidas en formato *csv*.

Para el procesamiento de los ficheros se utiliza GrADS, la cual es una herramienta de escritorio interactiva utilizada para la representación gráfica y el análisis de datos científicos, desarrollada por el Centro para Estudios del sistema Tierra-Océano-Atmósfera de la Universidad de Maryland de los Estados Unidos. Provee un entorno integrado para el acceso, manipulación y representación de datos científicos. Contiene una interfaz programable en forma de lenguaje scripts interpretado. GrADS puede ser corrido en modo *bash*, es decir, a través de estos scripts (Doty, 2003).

Con este programa se ha creado una herramienta que posibilita la evaluación del pronóstico de variables discretas con la información disponible. Está adaptado para la lectura de los datos de las imágenes de satélite y los datos de salida del WRF. No obstante, el método puede ser usado para la evaluación del pronóstico numérico realizado por otros modelos meteorológicos.

### Casos de estudio en la región occidental

Como casos de estudio para la aplicación del método propuesto en la región occidental, se emplearon los utilizados para comprobar los resultados de la herramienta MODEMod (Roque, 2013; Roque *et al*, 2014), así como las salidas del modelo y las imágenes de satélite en formato TIFF, con vistas a establecer una comparación entre ambos resultados. Dichos casos son 5 días del mes de noviembre de 2012 y siete días del mes de marzo de 2013, donde hubo un predominio de la nubosidad sobre la región occidental, influenciada fundamentalmente por: las altas presiones migratorias y la presencia de un flujo difluente en la troposfera superior como factor estimulante para los procesos de lluvia, además la presencia de una hondonada en superficie, con representación en la troposfera baja y el paso de sistemas frontales por la región (Ballester y González, 2012; González, 2013).

Los casos de estudio empleados en la región occidental fueron:

1. 17 de noviembre de 2012
2. 18 de noviembre de 2012
3. 19 de noviembre de 2012
4. 24 de noviembre de 2012
5. 25 de noviembre de 2012
6. 1 de marzo de 2013
7. 2 de marzo de 2013
8. 15 de marzo de 2013
9. 17 de marzo de 2013
10. 25 de marzo de 2013
11. 26 de marzo de 2013
12. 27 de marzo de 2013
13. 29 de marzo de 2013

## **Casos de estudio para Cuba**

Como casos de estudio en la región de Cuba se emplearon 4 días del mes de junio de 2014, donde hubo un predominio de nubosidad en todo el país, influenciado fundamentalmente por vaguadas en las cercanías de la región, difluencia en los niveles altos y medios combinada con convergencia del flujo en superficie, alto contenido de humedad en la columna troposférica, circulación de bajo geopotencial en niveles bajos de la atmósfera y presencia de bajas frías (Caymares *et al*, 2014; Hernández *et al*, 2014; Sánchez *et al*, 2014a; Sánchez *et al*, 2014b). Los casos de estudio empleados en la región de Cuba fueron:

1. 2 de junio de 2014 2100Z
2. 4 de junio de 2014 2100Z
3. 5 de junio de 2014 2100Z
4. 9 de junio de 2014 2100Z

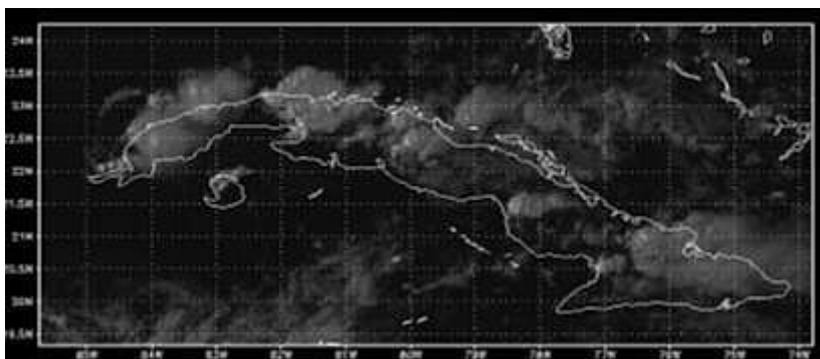
## **3. Análisis y discusión**

### **3.1 Selección de los umbrales para la identificación de los objetos**

Para la identificación de los objetos existentes en ambos campos, es necesario identificar los valores de intensidad mínimos que los conforman. Dichos valores dependen de la escala, la resolución de la malla empleada y el interés del usuario que realiza la evaluación. Además, los umbrales deben estar expresados en las unidades de los campos correspondientes. Las pruebas se realizaron para todos los casos de estudio correspondientes al área de Cuba.

Los experimentos se realizaron para cuatro valores de umbrales distintos en ambos campos (pronóstico y observación) con el objetivo de identificar el valor de intensidad mínimo que permita un mejor ajuste entre las áreas de nubosidad identificadas y las existentes en la imagen original. De forma general se pudo apreciar que a medida que incrementamos el valor del umbral, disminuye el tamaño de los objetos en ambos casos, identificándose en algunos casos un incremento en el número de objetos, aunque esta relación no es directamente proporcional, ya que depende en gran medida de los valores de intensidad existentes.

Para la selección del umbral del campo de observación se escogieron como posibles valores  $T=17$ ,  $T=24$ ,  $T=33$  y  $T=39$ . Este campo, al ser una imagen de satélite, contiene información expresada en niveles digitales de escala de grises que varía en un rango de 0 a 255. En la figura 1 se presenta la imagen original del día 9 de junio de 2014 a las 2100 UTC. Como se aprecia en la figura 2, los umbrales  $T=17$  y  $T=24$  no filtran lo suficiente la información de la imagen, mientras que para el umbral  $T=39$ , se pierde información de la misma. Por ello se concluye que el umbral óptimo para el campo de observación es de  $T=33$ .



*Fig. 1- Imagen original del día 9 de junio de 2014 2100 UTC.*

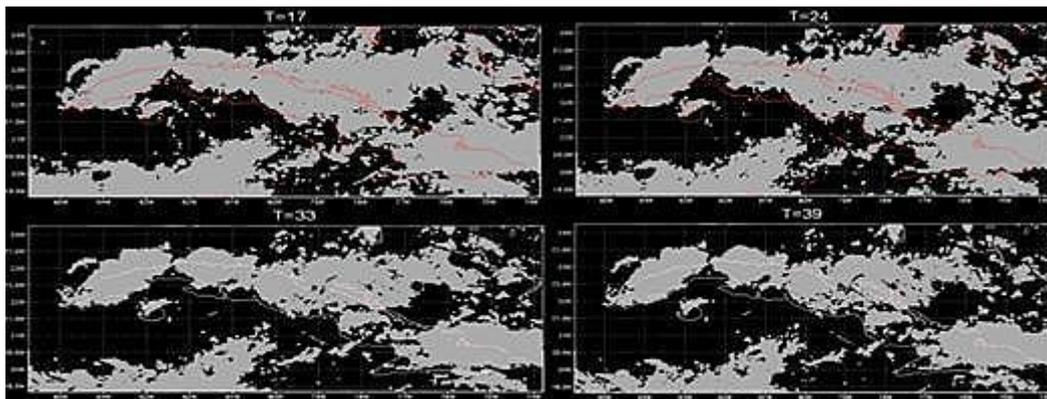


Fig. 2- Aplicación del umbral para el campo de observación.

Para la selección de umbral del campo de pronóstico se escogieron como posibles valores  $T=0.2$ ,  $T=0.25$ ,  $T=0.3$  y  $T=0.35$ . Este campo contiene información de la variable *integrated cloud fraction* (fracción de cielo cubierto), que es adimensional y varía en un rango de 0 a 3.0. En la figura 3 se presenta el campo pronóstico original del día 2 de junio de 2014 a las 2100 UTC. Como se aprecia en la figura 4, los umbrales  $T=0.2$  y  $T=0.25$  no filtran lo suficiente la información del mismo, mientras que para el umbral  $T=0.35$ , se pierde demasiada información. Por ello se concluye que el umbral óptimo para el campo de pronóstico es  $T=0.3$ .

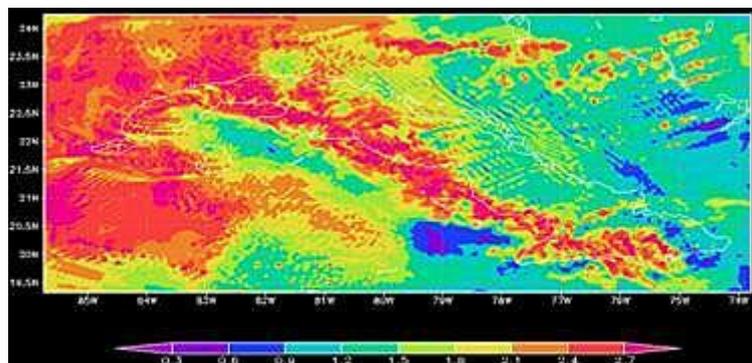


Fig. 3 - Campo pronóstico original del día 2 de junio de 2014 2100 UTC.

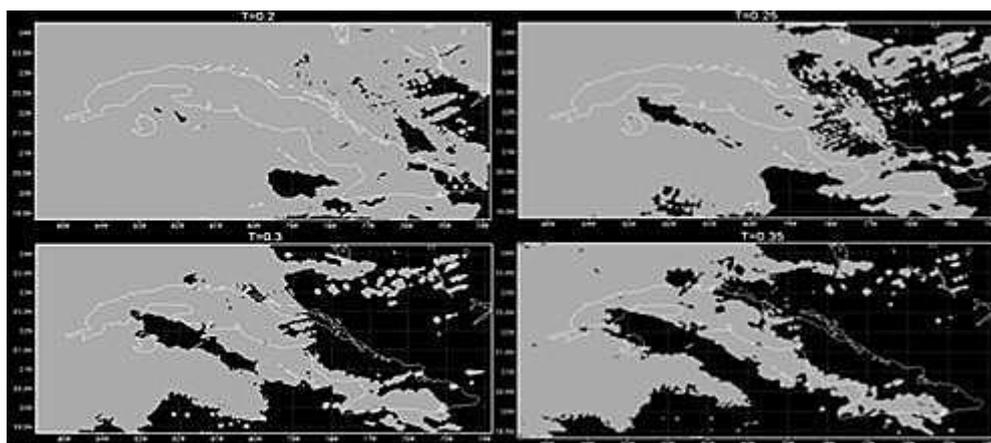


Fig. 4- Aplicación del umbral para el campo de pronóstico.

### 3.2 Selección de los límites de las funciones de interés

En los procesos de comparación y asociación, las funciones de interés que conforman el algoritmo de lógica difusa juegan un papel fundamental. Por ello dichas funciones deben estar dentro de los límites óptimos que garanticen un resultado correcto de la evaluación realizada. Con este objetivo se experimentó con cuatro configuraciones distintas para todos los casos de estudio, las cuales se enumeran a continuación:

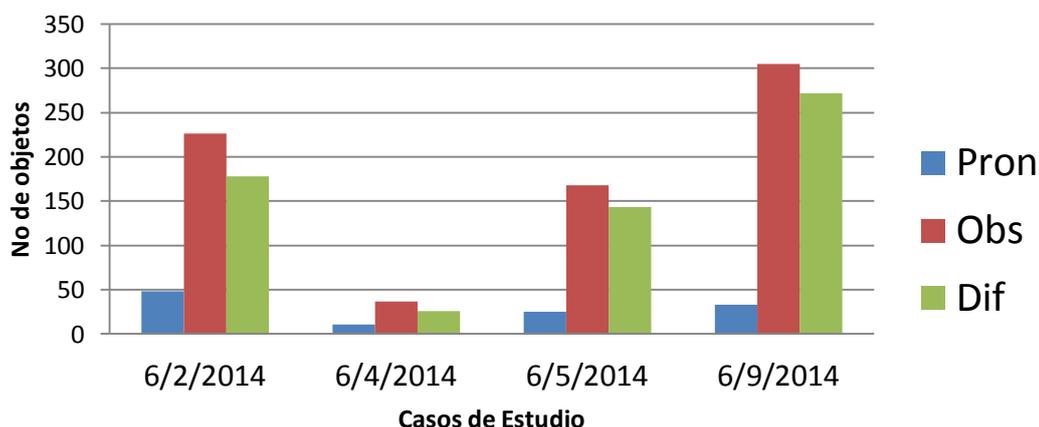
- a) Distancia mínima entre los bordes: 150 pts/res  
Diferencia de centroide: 300 pts/res
- b) Distancia mínima entre los bordes: 200 pts/res  
Diferencia de centroide: 400 pts/res
- c) Distancia mínima entre los bordes: 800 pts/res  
Diferencia de centroide: 400 pts/res
- d) Distancia mínima entre los bordes: 400 pts/res  
Diferencia de centroide: 600 pts/res

En los experimentos realizados se obtuvo que las configuraciones **a** y **b**, tienden a subestimar el resultado de la evaluación, mientras que la configuración **c** tanto subestima como sobrestima la evaluación realizada, por lo que se concluye que la configuración más adecuada es la **d**, que además es la que emplea por defecto en la versión 4.1 del *Model Evaluation Tools* (MET), herramienta de evaluación a la que pertenece el MODE (MET Developers, 2011).

### 3.3 Proceso de fusión de objetos

Uno de las principales limitaciones de la evaluación espacial en Cuba es que no se han aplicado técnicas de fusión de objetos que permitan equilibrar el número de áreas detectadas en ambos campos. Esto trae como consecuencia una desproporción entre la cantidad de objetos a comparar en los campos de pronóstico y observación, lo que afecta el resultado final de la evaluación. Como se observa en la figura 5, la diferencia entre la cantidad de objetos detectados originalmente en cada campo es desproporcionada, encontrándose muchas más áreas en el campo de observación que en el de pronóstico para cada uno de los casos de estudio escogidos en el área de Cuba.

**Objetos Detectados**



*Fig. 5- Objetos detectados.*

En la figura 6 se muestra la cantidad de objetos detectados en los mismos campos después de aplicar la técnica de fusión de objetos. Se aprecia la notable reducción del número de objetos y la existencia de un mayor equilibrio al disminuir la diferencia entre los dos campos, lo que permite una mejor comparación y evaluación espacial de dicho pronóstico.

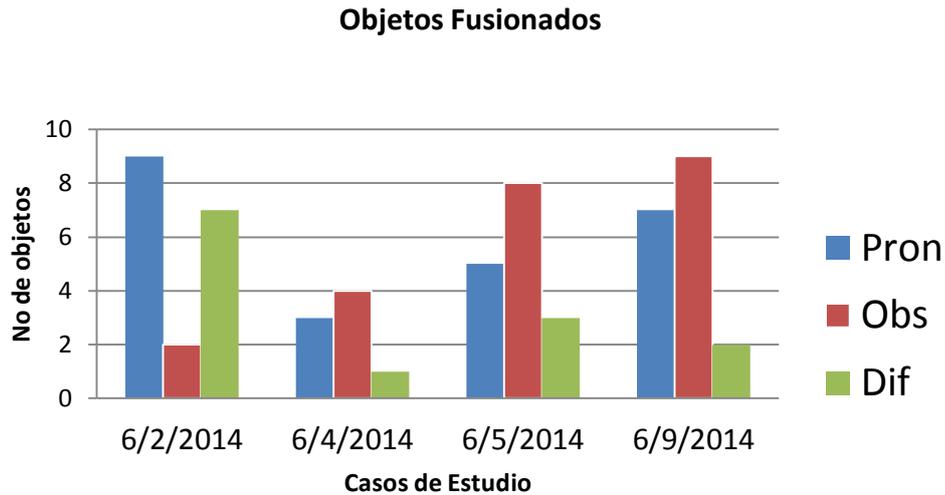


Fig. 6- Objetos detectados al aplicar la técnica de fusión de objetos.

### Comparación entre los resultados obtenidos por el MODEMod y el MODEMod\_1.0

Una vez configurado el método, se procede a la comparación entre los resultados obtenidos por (Roque, 2013; Roque *et al*, 2014) empleando el MODEMod y los obtenidos empleando la metodología implementada en la aplicación MODEMod\_v1.0. Como se aprecia en la figura 7, los valores de CSI obtenidos por la aplicación desarrollada son más representativos y cercanos a la evaluación real de los pronósticos de los casos de estudio escogidos. Esto se debe al empleo de un algoritmo de lógica difusa y una técnica de fusión de objetos que permite reducir la diferencia entre el número de objetos de ambos campos, los que favorecen y mejoran el proceso de comparación y asociación.

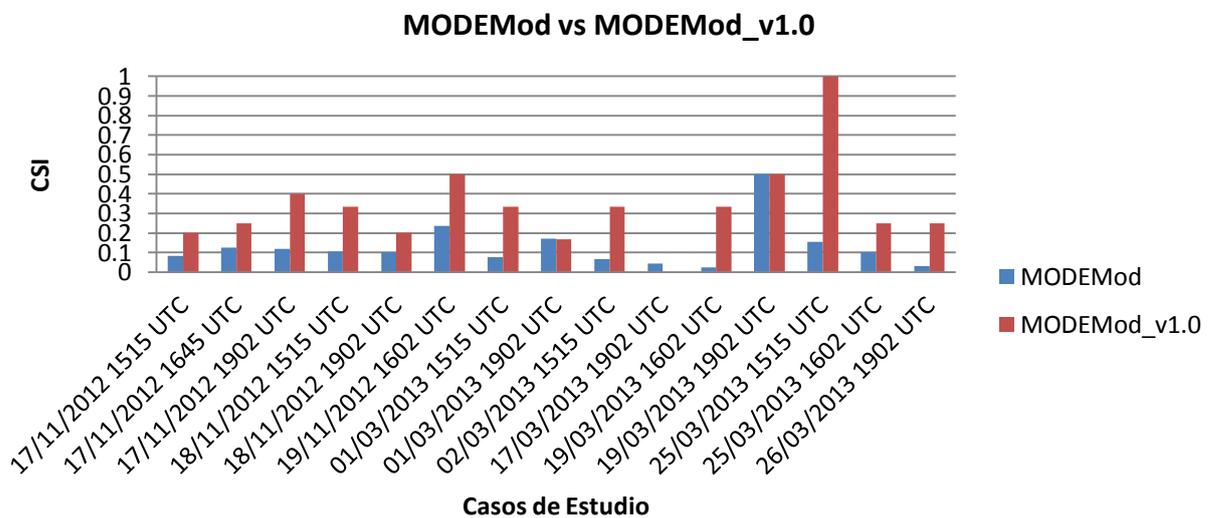


Fig. 7- Comparación entre MODEMod y MODEMod\_v1.0.

### Resultados obtenidos con la configuración escogida

Una vez escogidos los umbrales para los campos de pronóstico y observación, y elegida la configuración para los límites de las funciones de interés, se procede a evaluar los resultados de los casos de estudio escogidos de ambas áreas de estudio.

#### Región occidental

Para la región occidental los resultados fueron evaluados entre valores de CSI entre 0.2 y 0.7, excepto el caso del día 17 de marzo de 2013 a las 1902 UTC, donde el pronóstico fue erróneo. Además se aprecia en la figura 8 una evaluación perfecta para los días 25 de noviembre de 2012 a las 1902 UTC y 25 de marzo de 2013 a las 1515 UTC, donde predominan celdas aisladas de nubosidad, por lo que una conclusión preliminar sería que el método falla en estos casos, pero aún no se cuenta con la suficiente información para realizar una afirmación en este aspecto.

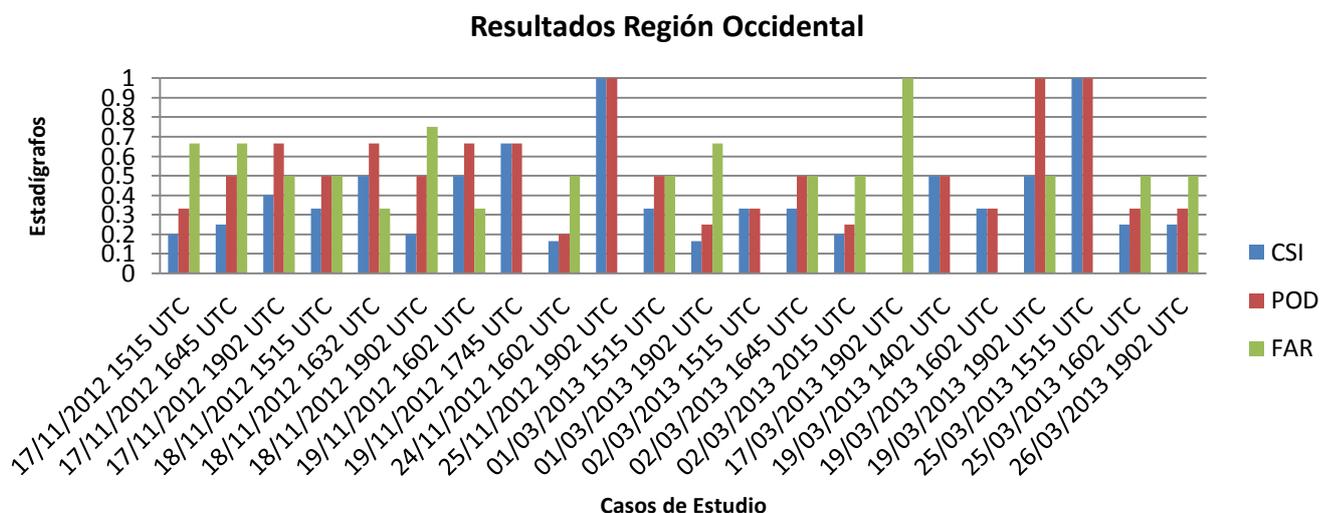


Fig. 8- Resultados del programa en la región occidental.

#### Área de Cuba

En los cuatro casos para el área de Cuba antes mencionada, los valores de CSI se encontraron entre 0.05 y 0.2. En el caso correspondiente al 4 de junio de 2014 a las 2100 UTC el programa subestima la evaluación del pronóstico, mientras que en los casos restantes lo evalúa de forma correcta. Estos resultados se pueden apreciar gráficamente en la figura 9.



Fig. 9- Resultados del programa en el área de Cuba.

## Automatización del MODEMod\_v1.0

En el presente trabajo se modificó el código fuente de la aplicación MODEMod para llevar a cabo los pasos explicados en el epígrafe 2. Dicho programa fue dividido en módulos para lograr una mayor organización y optimización del código.

**nruil.h:** En este módulo se encuentran implementadas algunas funciones útiles que facilitan el trabajo con vectores y matrices, y son empleadas en todos los módulos del programa (Colectivo de autores, 1988-1992).

**preprocessing.h:** Este módulo se encarga de leer la salida del modelo en formato *ascii* con resolución de 3 km, e interpola la misma a una malla de resolución 4 km, usando el método del vecino más cercano. Además se encarga de extraer de la imagen de satélite original en formato *csv*, que se encuentra de sur a norte entre los  $9.02^{\circ}N$  y los  $33.275^{\circ}N$ , y de este a oeste entre los  $58.355^{\circ}W$  y los  $101.674^{\circ}W$ , una subimagen de la región de interés que se ubica de sur a norte entre los  $19.3149^{\circ}N$  y los  $24.2737^{\circ}N$ , y de este a oeste entre los  $73.74768499^{\circ}W$  y los  $85.732608184^{\circ}W$ .

**read\_data.h:** Se encarga de leer los datos necesarios para ejecutar el programa de evaluación. Puede leer los datos de imágenes de satélite, tanto en formato *csv* como *TIFF* y la salida de los modelos tanto en formato *ascii* como *csv*. Además se encarga de leer los ficheros de configuración que contiene las opciones y parámetros definidos por el usuario.

**structs\_convolution.h:** Contiene la declaración de todas las estructuras empleadas por el programa. Lleva a cabo los procesos de convolución y aplicación de umbrales para definir los objetos en los campos de pronóstico y observación.

**single\_atrib.h:** Realiza la caracterización individual de todos los objetos identificados a partir del cálculo de todos sus atributos.

**pair\_atrib.h:** Contiene las funciones que calculan los atributos para los pares de objetos en ambos campos. Determina el valor de interés de cada atributo, y a partir del mismo, determina el interés total para cada par de objetos. Si el interés total de un par de objetos que se encuentren en un mismo campo excede el umbral definido por el usuario, este módulo se encarga de fusionarlos.

Para la automatización de este programa se crearon tres scripts de *Shell*.

**download\_VIS4km.sh:** Se encarga de descargar las imágenes de satélite de la red interna del INSMET y ubicarlas en una carpeta para su posterior uso.

**copy\_cloud.sh:** Extrae los datos de fracción de cielo cubierto de las salidas del modelo WRF-ARW y los ubica en una carpeta para su uso posterior.

**run\_eval.sh:** Ejecuta el programa de evaluación en los horarios correspondientes y guarda los resultados y salidas gráficas para su almacenamiento en la base de datos.

## Conclusiones

Analizando los resultados obtenidos se llegó a las siguientes conclusiones:

1. Se logró implementar de forma operativa el método de evaluación MODEMod\_v1.0.
2. Se demostró que los atributos añadidos permiten mejorar la comparación y asociación de los objetos.
3. La técnica de fusión de objetos disminuye la diferencia entre la cantidad de objetos detectados en ambos campos.
4. El empleo del algoritmo de lógica difusa y la técnica de fusión de objetos mejora el método de evaluación utilizado.
5. Los umbrales en los casos de estudio empleados es de 33 en niveles digitales de la escala de grises para las imágenes de satélite, mientras que para la salida del modelo WRF es 0.3 para la variable *integrated cloud fraction* (fracción de cielo cubierto).
6. La mejor configuración de los límites de las funciones de interés para los casos de estudio utilizados son 400 pts/resolución para la distancia mínima entre los bordes y 600 pts/resolución para la diferencia de centroide.

## Referencias bibliográficas

- Ballester, M. y González, A. O. noviembre 2012. Resumen sinóptico mensual [En línea]. Grupo de investigaciones del centro de pronóstico del Instituto de Meteorología. Disponible: [www.insmet.cu](http://www.insmet.cu).
- Basu, S., Venugopal, V. y Foufoula-Georgiou, E. 2005. A new metric for comparing precipitation patterns with an application to ensemble forecasts. *J. Geophys. Res.*, 110: D8, D08111 10.1029/2004JD005395.
- Brown, B.G., Davis, C.A., Manning, K.W., Bullock, R. 2004. An Object-oriented Approach to Quantitative Precipitation Forecasts. *Preprints, 17th Conference on Probability and Statistics in the Atmospheric Sciences. Seattle, WA.*
- Brown, B., Davis, C. A., y Bullock, R. 2006a. Object-based verification of precipitation forecasts. Part I: Methods and application to mesoscale rain areas. *Monthly Weather Review*, 134, 1772-1784.
- Brown, B., Davis, C. A., y Bullock, R. 2006b. Object-based verification of precipitation forecasts, Part II: Application to convective rain systems. *Monthly Weather Review*, 134:1785-179.
- Bullock, R., Davis, C.A., Gotway, J.H., Chapman, M., Takacs, A., Gilleland, E., Mahoney J., Brown, L. B.G. y Manning, K. 2004. New verification approaches for convective weather forecasts. *Preprints, 11th Conference on Aviation, Range, and Aerospace*, Hyannis, MA.
- Bullock, R., Brown, B. G., Davis, C., A., Manning, K.W., Morss, R., Chapman, M., y Takacs, A. 2004. An Object Oriented Approach to the Verification of Quantitative Precipitation Forecasts: Part II - Examples. *Preprint, AMS 2004.*
- Caymares, A., Cedeño, Y. y Lemes, M. 2014. Estado General del Tiempo 3/6/2014 0000 UTC. Centro de Pronósticos del Tiempo, INSMET.
- Chaudhuri, D. y Samal, A. 2007. A simple method for fitting of bounding rectangle to closed regions. *Pattern Recognition*, 40, 1981-1989.
- Colectivo de autores. 2012. WRF-ARW V3: User's Guide. Disponible: [http://www.mmm.ucar.edu/wrf/users/docs/arw\\_v3.pdf](http://www.mmm.ucar.edu/wrf/users/docs/arw_v3.pdf)
- Doty, B. 2003. The Grid Analysis and Display System GrADS V1.5.1.12 [En línea], Disponible: <ftp://grads.iges.org/grads/beta/doc/gadoc151.ps>.
- Du, J. y Mullen, S. L. 2000. Removal of Distortion Error from an Ensemble Forecast. *Monthly Weather Review*, 128:3347-3351.
- Ebert, E. E. y McBride, J. L. 2000. Verification of precipitation in weather systems: Determination of systematic errors. *Journal Hydrol*, 239:179-202.
- González, C. y González, A. O. marzo 2013. Resumen sinóptico mensual [En línea]. Grupo de investigaciones del centro de pronóstico del Instituto de Meteorología, Disponible: [www.insmet.cu](http://www.insmet.cu).
- Hernández, M. A., Vázquez, E., Otero, A. y Estévez, G. 2014. Estado General del Tiempo 5/6/2014 0000 UTC. Centro de Pronósticos del Tiempo, INSMET.
- Lakshminarahan S., Baldwin, M. E. y Kain, J. S. 2001. Verification of mesoscale features in NWP models. *Preprints, 9th Conf. on Mesoscale Processes*, Ft. Lauderdale, FL.
- Lakshminarahan S., Baldwin, M. E. y Kain, J. S. 2002. Development of an events-oriented approach to forecast verification. *15th Conf. Numerical Weather Prediction*, San Antonio, TC.
- Mahoney J.L., Davis C.A., Bullock R., Brown, B.G. y Mueller, C.K. 2002. Improved approaches for measuring the quality of convective weather forecasts. *Preprints, 16th Conference on Probability and Statistics in the Atmospheric Sciences*. Orlando, FL.
- MET Developers. 2011. The MODE Tool. Model Evaluation Tools Version 3.0.1 (METv3.0.1). User's Guide 3.0.3. Developmental Testbed Center. pages 98-122.
- Mount, D. M. 2000. Lecture Notes for Computational Geometry Course. Dpt. Of Computer Science, University of Maryland.
- 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.
- Roque, A. 2013. *Método de Evaluación Diagnóstico Orientado a Objetos Modificado (MODEMod) utilizando la información de los satélites meteorológicos*. Trabajo de Diploma. Instituto Superior de Tecnologías y Ciencias Aplicadas.
- Roque Carrasco, A., Sierra Lorenzo, M., Quintana Rodríguez, N. 2014. Método de Evaluación Diagnóstico Orientado a Objetos Modificado (MODEMod) utilizando la información de los satélites meteorológicos. *Ciencia de la Tierra y el Espacio*, Vol. 15, No. 2, 139-150.
- Sánchez, A., Bermúdez, Y. y Wong, A. 2014. Estado General del Tiempo 6/6/2014 0000 UTC. Centro de Pronósticos del Tiempo, INSMET.
- Sánchez, A., Bermúdez, Y. y Wong, A. 2014. Estado General del Tiempo 10/6/2014 0000 UTC. Centro de Pronósticos del Tiempo, INSMET.

**Acerca de los autores:**

**Camilo Fernando Rodríguez Genó:** Graduado de Licenciatura en Meteorología en el Instituto Superior de Tecnologías y Ciencias Aplicadas (InSTEC), 2014. Se desempeña como Reserva Científica en el Centro de Física de la Atmósfera ubicado en el Instituto de Meteorología de Cuba. Las líneas de investigación en que se desarrolla abarcan el pronóstico numérico del tiempo (Weather Research & Forecast, WRF) y los métodos de evaluación espacial del pronóstico numérico. Domina los idiomas español e inglés.

**Maibys Sierra Lorenzo:** Graduada de Licenciatura en Ciencias de la Computación en Universidad de la Habana, 2008. Obtuvo el Máster en Ciencias Matemáticas en Universidad de la Habana, 2011. Se desempeña como Especialista Superior en Meteorología Aplicada en Instituto de Meteorología y como profesora de la carrera de Licenciatura en Meteorología con Categoría Docente Instructor. Las líneas de investigación en que se desarrolla abarcan el Pronóstico Numérico (modelos MM5V3, ARPS y WRF), Wavelets y solución de Ecuaciones Diferenciales en Derivadas Parciales y Tormentas Locales Severas. Domina los idiomas español, inglés y portugués. Se desempeña además como Árbitro de la Revista Cubana de Meteorología.

**Adrián Luis Ferrer Hernández:** Graduado de Licenciatura en Meteorología en el Instituto Superior de Tecnologías y Ciencias Aplicadas (InSTEC), 2013. Se desempeña como Reserva Científica en el Centro de Física de la Atmósfera ubicado en el Instituto de Meteorología. Las líneas de investigación en que se desarrolla abarcan la modelación numérica del tiempo (Weather Research & Forecast, WRF y Advanced Regional Prediction System, ARPS), software de visualización y estudio de fenómenos a mesoescala. Domina los idiomas español e inglés.