



*Juan C. Anduckia
janducki@yahoo.com

*Serguei A. Lonin
slonin@costa.net.co

*CIOH

CIOH
www.cioh.org.co

Efecto residual de la nubosidad en la asimilación de temperatura superficial del mar

Resumen

La temperatura superficial del mar, determinada mediante satélites de la NOAA, es tratada mediante la aplicación de un modelo auto-regresivo con el fin de mejorar el proceso de asimilación y utilizarlo cuando el campo satelital de temperaturas, por efecto del vapor de agua y las nubes, presenta una variabilidad temporal con componente aleatoria cuyo ruido no está filtrado completamente en la señal.

Palabras claves: Tratamiento de datos satelitales, temperatura superficial del mar, modelos auto-regresivos

Abstract

Sea-surface temperature determined with NOAA satellites is treated by means of the application of an auto-regressive model. The auto-regressive model should be employed for improving assimilation performance whenever the satellite sea-surface field, because of water vapour and cloudiness effects, presents a temporal variability with a random component whose noise has not been filtered in the signal.

Key words: Satellite data treatment, sea-surface temperature, auto-regressive models

Introducción

Uno de los mayores obstáculos que presenta la asimilación de datos satelitales de temperatura superficial del mar es la presencia de nubes en las imágenes AVHRR, dado que aquellas afectan la medición de los radiómetros, como se muestra en la Figura 1. Aunque la NOAA ha diseñado “tests” para el filtraje de los datos, distinguiendo los píxeles de las imágenes en aceptables

y no, por lo general la calidad de la información de la serie de tiempo para un píxel dado no es satisfactoria debido a los efectos residuales de la atmósfera, los cuales no se remueven plenamente. Estos efectos se deben a los aerosoles, sobre todo las nubes y el vapor de agua. Una solución posible para obtener campos coherentes consiste en producir cada semana mapas compuestos de la temperatura satelital usando cualquier técnica de “valor máximo” en la composición, asumiendo que la máxima temperatura es la más limpia (sin contaminación atmosférica o de las nubes). Las zonas calientes en donde las capas superiores de la columna de agua pueden alcanzar hasta 2 o más grados se originan en el fuerte calentamiento solar en condiciones de calma (no turbulencia); asimismo, las zonas en donde hay una nubosidad marcada pueden resultar en temperaturas hasta de 2 grados por debajo, como se mostró en una contribución anterior (Anduckia *et al.*, 2003). En la producción de campos mensuales basados en estas imágenes promediadas debe nuevamente promediarse para evitar los errores.

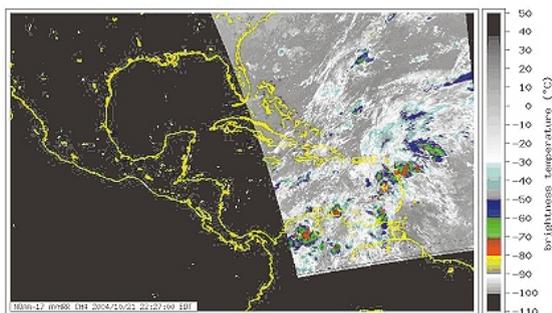


Figura 1. Efecto de las nubes sobre los radiómetros: temperaturas de brillo para un paso típico de satélite en la zona de interés (fuente: Coast Watch)

Procedimientos correctivos de este tipo son ampliamente usados a nivel mundial (NOAA, KNMI). Sin embargo, las imágenes compuestas y promediadas a lo largo de varios días suponen uniformidad en las condiciones del tiempo en la escala de días, condición que

no se cumple. Además, el esquema de pronóstico de las condiciones oceanográficas en operación actualmente en el CIOH (Lonin *et al.*, 2003) es un esquema de tiempo real (lo mismo que la asimilación de datos), por lo tanto la noción de imágenes promediadas debe descartarse para buscar un procedimiento alternativo que permita la asimilación en tiempo real.

El presente trabajo investiga la estructura temporal de las series de temperatura en las imágenes satelitales con el fin de establecer modelos estadísticos que permitan cuantificar el efecto de la nubosidad y el vapor de agua en las mismas. Para ello analiza la dispersión de la señal satelital junto con su patrón de autocorrelación, confirmando la hipótesis (Anduckia *et al.*, 2003) de que la dispersión es mayor, y por lo tanto menor la correlación de los datos en momentos de tiempo cercanos, en las zonas donde se presenta mayor nubosidad. El dato de temperatura cercano al valor real a ser usado en la asimilación se puede restituir con base en un promedio corrido, y con la ayuda de un modelo auto-regresivo se puede obtener la serie completa dentro de la ventana del promedio.

Datos utilizados

Los datos satelitales de temperatura superficial del mar son determinados por radiómetros de alta resolución y fueron descritos en la primera contribución mencionada en la introducción. Varios “tests” se emplean para remover el efecto de las nubes en la señal (Stowe *et al.*, 1991; Jin *et al.*, 2000), pero a pesar de esto, como se aprecia en la Figura 1, la nubosidad tiene un efecto residual importante en la determinación de la temperatura. El “Crucero-2004” efectuado con el fin de revisar la validación de temperaturas realizada en el año 2003 y descrita en la contribución citada (Anduckia *et al.*, 2003), en la que se logró

cuantificar un efecto de sesgo constante independiente de la nubosidad en la temperatura satelital [este efecto es conocido como de *skin layer* (*capa delgada*), y ha sido confirmado por una numerosa literatura, por ejemplo Donlon (2002) y Kurzeja (2002) entre otros,] fue realizado entre el 25 de mayo y el 11 de junio en condiciones de gran nubosidad, como se muestra en la Figura 2.

Aunque la estadística para comparación entre imágenes satelitales y mediciones puntuales no es grande (50 imágenes disponibles y 20 coincidencias imagen-estación), confirma el efecto de “capa delgada”. La diferencia $\Delta T = T_{\text{sat}} - T_{\text{ctd}}$ entre la temperatura determinada por el satélite y la medición puntual presenta ahora un promedio de -0.70 ± 0.94 °C, que concuerda bastante bien con lo obtenido en el 2003 (-0.66 ± 0.65 °C). Se piensa que la diferencia es mayor, así como la dispersión, debido a la persistente nubosidad imperante durante el desarrollo del crucero.

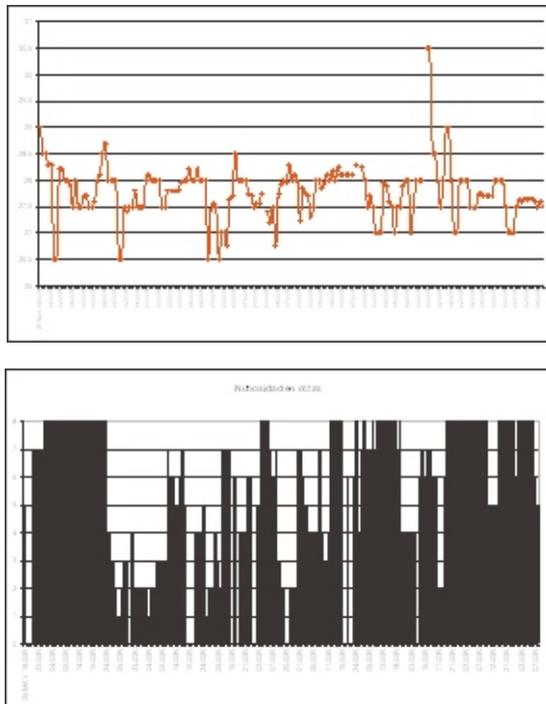


Figura 2. Temperatura y nubosidad durante el Crucero 2004 del CIOH

La pertinaz condición de cielos cubiertos (8 octas en el 90% de las estaciones) creemos es responsable de la mayor diferencia obtenida para ΔT en el día que en la noche:

$$\begin{aligned} \text{Día} \quad \Delta T &= -0.89 \pm 0.92 \text{ }^\circ\text{C} \\ \text{Noche} \quad \Delta T &= -0.59 \pm 1.04 \text{ }^\circ\text{C}. \end{aligned}$$

Este resultado es contrario al obtenido en el Crucero 2003, donde la diferencia era mayor en la noche que en el día (-0.92 ± 0.54 °C, y -0.43 ± 0.67 °C respectivamente). La nubosidad es el único factor que podría explicar la diferencia y el aumento en la dispersión de las mediciones, teniendo en cuenta que bajo un cielo totalmente cubierto cabe esperar en general que la superficie se enfríe menos en las noches (debido al débil calentamiento solar), cuando las nubes constituyen un techo que reemite la radiación hacia la superficie, radiación que en caso de no existir nubes escaparía hacia el espacio.

Si bien los datos de radiación obtenidos en este crucero son insuficientes para realizar un análisis para mirar la influencia del flujo de calor y el viento en la diferencia de temperaturas y en el efecto de “capa delgada”, salta a la vista que las nubes siguen siendo un desafío para una buena asimilación.

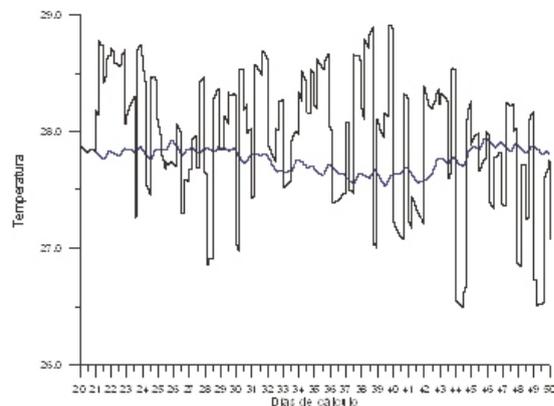


Figura 3. Comportamiento de la temperatura con (línea negra) y sin (azul) asimilación de datos AVHRR, en el punto de control (150,50) de la malla del modelo oceanográfico.

En la Figura 3 se muestra el efecto de la asimilación de la temperatura en un punto representativo de control de la malla del modelo oceanográfico (200 x 100 nodos), ubicado en el Caribe oriental (150,50). Los saltos originados por la asimilación, y cuyo origen ha de atribuirse al efecto residual de la nubosidad, obligan a buscar una ventana de predicción en la que la asimilación tenga un comportamiento “suave”. No hay que olvidar, sin embargo, que en la contribución mencionada (Anduckia *et al.*, 2003) se mostró que la señal satelital, vista a lo largo del tiempo, en promedio reproduce el campo correcto, compensando las deficiencias que presenta en las imágenes instantáneas (Figura 4). Esta hipótesis será puesta a prueba en lo sucesivo.

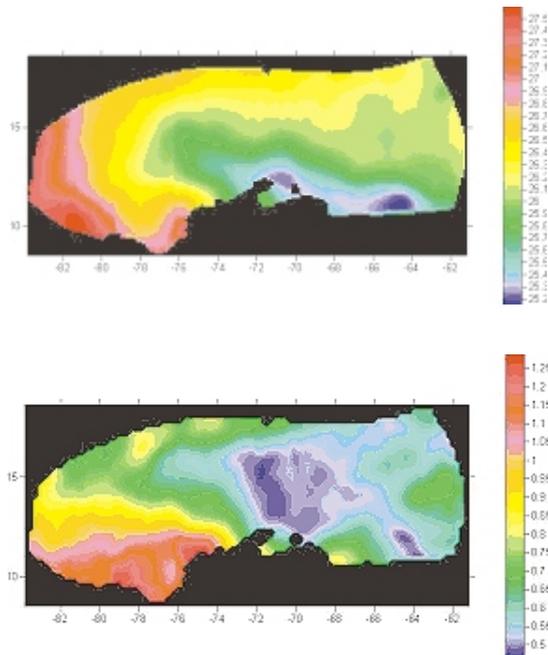


Figura 4. Señal satelital para el mes de marzo de 2003. Arriba: temperatura promedio. Abajo: dispersión (Tomado de: Anduckia *et al.*, 2003)

Metodología

Como una manera de investigar el efecto residual de la nubosidad en las imágenes

satelitales de temperatura, aprovechando el experimento de calibración general del modelo oceánico llevado descrito en Lonin *et al.* (2003), se han construido series de tiempo de la temperatura superficial del mar para seis meses (marzo-septiembre del 2003), interpoladas en los nodos de la malla del modelo.

De acuerdo con la hipótesis, basada en la Figura 4, la temperatura promedio temporal en un punto dado $T_{real}(x, y)$ puede determinarse con una buena aproximación a través de la observada $T_{obs}(x, y, t)$, si esta última se integra en el tiempo t como:

$$\bar{T}_{real}(x, y) = \lim_{\tau \rightarrow \infty} \frac{1}{2\tau} \int_{-\tau}^{\tau} T_{obs}(x, y, t) dt. \quad (1)$$

Al mismo tiempo, si la autocorrelación de los datos

$$\int_0^{\infty} T'_{obs}(t) T'_{obs}(t + \xi) dt \rightarrow 0, \quad \forall \xi, \quad (2)$$

debido a la calidad de la señal instantánea (los signos “prima” aquí indican los momentos centrales), se puede suponer que el valor esperado en (1) corresponde al comportamiento verdadero de la temperatura, $T_{real}(x, y, t)$, junto con una señal aleatoria de promedio nulo pero con varianza superior a la de $T_{real}(x, y, t)$ (ver Figura 3). A partir de esta consideración resulta que debe existir un valor de τ finito para el cual:

$$\frac{1}{2\tau} \int_{-\tau}^{\tau} T_{obs} dt \rightarrow T_{real}(x, y, t).$$

Supóngase que el momento $t = 0$ es el momento de la asimilación. Entonces, en el intervalo de tiempo $t \in [-\tau, 0]$ la función $T_{obs}(t)$ es conocida para cada punto del espacio a partir de las observaciones o por interpolación con *splines* de las mismas, mientras que en la segunda mitad del intervalo $t \in [0, \tau]$ la función $T_{obs}(t)$ tiene que ser reemplazada por otra: $T_{ARIMA}(t)$, aún desconocida.

En este caso, el valor de temperatura $T_{obs}^{nuevo}(x, y, t=0)$ que reemplaza el valor instantáneo observado $T_{obs}(x, y, t)$ en el momento de asimilación $t=0$, estará compuesto de dos partes en el procedimiento de promedio en la ventana:

$$T_{obs}^{nuevo}(x, y, t=0) = \frac{1}{\tau} \int_{-\tau}^0 T_{obs}(t) dt - \frac{1}{\tau} \int_0^{\tau} T_{ARIMA}(t) dt. \quad (3)$$

La primera parte corresponde al promedio de los datos a lo largo del semi-intervalo de longitud τ , la segunda tiene que ser pronosticada con un modelo auto-regresivo. Las dos siguientes tareas a resolver son entonces: 1) encontrar el intervalo τ satisfactorio para que $\frac{1}{2\tau} \int_{-\tau}^{\tau} T_{obs} dt$ se acerque al valor $T_{real}(x, y, t)$ con un error tolerable y 2) buscar apropiadamente la función $T_{ARIMA}(t)$ para la fórmula (3).

En cuanto al primer punto, por inspección del comportamiento de promedios móviles de distintas series se ha escogido una ventana como la que se muestra en la Figura 5, correspondiente a un promedio móvil de periodo 11 en la serie del punto de control (150,80).

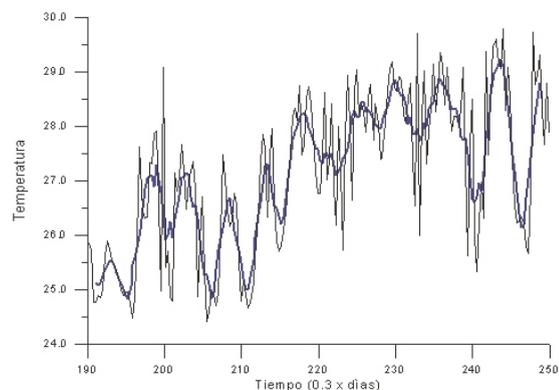


Figura 5. Promedio móvil de orden 11 (línea azul) en la serie de temperatura (línea negra) del punto (150,80)

En cuanto al segundo punto, la investigación del comportamiento de las series de tiempo

consiste en los tres pasos siguientes:

Identificación del modelo $ARIMA(p, d, q)$, donde p corresponde al orden del modelo auto-regresivo (AR); d a la diferenciación y q a la parte estocástica (MA).

Estimación de los parámetros p, d y q .

Verificaciones diagnósticas sobre la adecuación del modelo.

La búsqueda de los parámetros del modelo primero requiere establecer su orden, es decir, los valores respectivos de p, d y q en el modelo auto-regresivo. El primer paso en la identificación del modelo auto-regresivo adecuado es el examen de las funciones de autocorrelación de las series de temperatura.

Puesto que las series de temperatura presentan a simple vista una tendencia (Figura 6a), para removerla fue suficiente con construir las series de diferencias de orden 1 para volverlas estacionarias. Previamente dichas series fueron restituidas con un procedimiento de *spline* para llenar los vacíos (por ausencia de valores en el punto correspondiente en las imágenes) y para producir una serie distribuida uniformemente en el tiempo (Figura 6b). Un ejemplo de estas series de diferencias restituidas que se forman tomando $z_t = \gamma_{t+1} - \gamma_t$ se muestra en la Figura 6c; la función de autocorrelación de esta serie de diferencias estacionarias (Figura 7) revela que en el rezago 1 ya hay un corte abrupto. No presenta estacionalidad, es decir es “ruido blanco”, sin estructura ni relaciones con el ciclo diario. No hay correlación apreciable en ningún valor de rezago excepto el primero (mayor que 0.2), indicando que el modelo auto-regresivo que se puede adaptar es de los más simples, un modelo auto-regresivo de promedios móviles de orden 1, $ARIMA(1,1,1)$, de acuerdo con la metodología de Box-Jenkins. Estos resultados son para el punto de control (150,50) ubicado

en el Caribe oriental, y corresponden a 911 imágenes satelitales, aproximadamente 210 días entre el mes de marzo y septiembre del 2003.

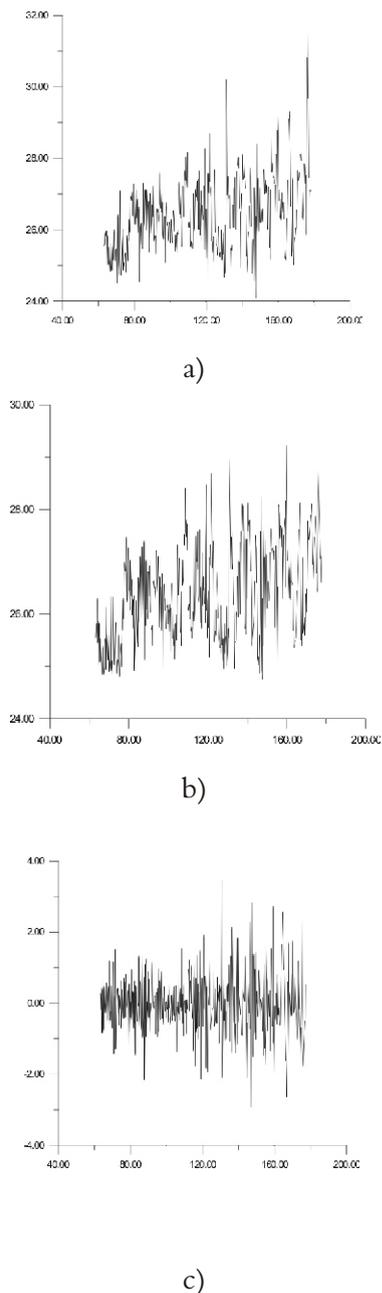


Figura 6. (a) Serie de temperaturas en un punto de control (150,50); (b) serie uniforme de temperaturas con spline cúbico aplicado; (c) serie de diferencias con $d = 1$.

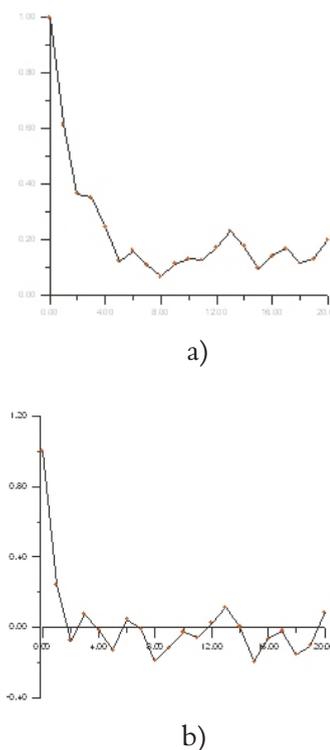


Figura 7. (a) Función de autocorrelación parcial de la serie original con *spline*; (b) función de autocorrelación parcial de la serie de diferencias con *spline*

Para un modelo ARIMA(1, d , 1) con término constante (δ), si z representa los valores de la serie de diferencias de orden 1 ($d = 1$), se tiene:

$$z_t = \delta + \phi_1 z_{t-1} + a_t - \theta_1 a_{t-1} \quad (4)$$

donde a_t es la señal aleatoria en el momento de tiempo t ; y ϕ_1 y θ_1 son los coeficientes del modelo mixto, por definir. Para un modelo de promedios móviles MA(1) se tiene:

$$z_t = a_t - \theta_1 a_{t-1} \quad (5)$$

En el caso de un modelo puramente autorregresivo, las ecuaciones de Yule-Walker (Box y Jenkins, 1970)

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p}, \quad k > 0 \quad (6)$$

se utilizan para encontrar los coeficientes de un modelo auto-regresivo de orden p . ρ_k en (6) es la función de auto-correlación del proceso generador de la serie. En el caso de modelos ARIMA(1,d,1) y MA(1) estos coeficientes están dados respectivamente por (Bowerman, 1993):

$$\rho_1 = \frac{(1-\phi_1\theta_1)(\phi_1-\theta_1)}{1+\theta_1^2-2\phi_1\theta_1} \quad (7)$$

$$\rho_k = \phi_1 \rho_{k-1}, \quad k \geq 2$$

y

$$\rho_1 = \frac{-\theta_1}{1+\theta_1^2} \quad (8)$$

$$\rho_k = 0, \quad k > 1$$

Es de notar que los parámetros del modelo serán en cada caso función de las coordenadas espaciales (x,y), y que por ahora el modelo correspondiente no dependerá del tiempo.

Resultados y discusión

La Figura 8 muestra la estructura de los rezagos de orden 1 y 2 para todo el campo satelital, confirmando la dependencia espacial del modelo auto-regresivo a emplear y de sus respectivos coeficientes.

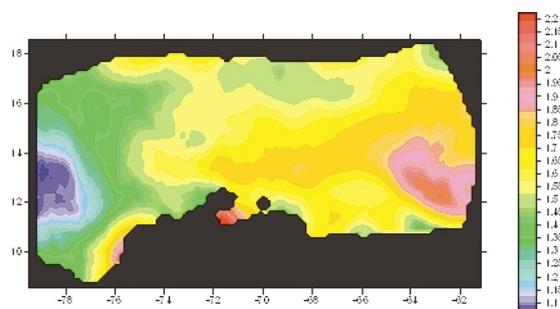
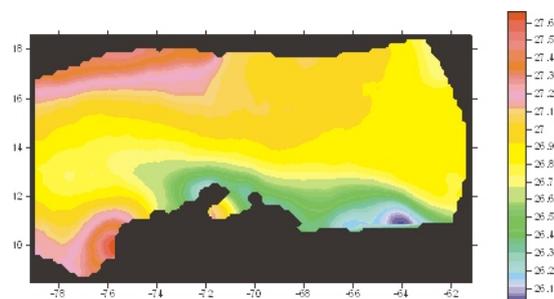


Figura 8. Rezagos de orden 1 (arriba) y 2 (abajo) para el campo satelital de temperatura

La estimación de parámetros del modelo fue realizada con la ayuda de rutinas programadas por los autores. Así, por ejemplo, el modelo auto-regresivo obtenido para el punto de control (150,80) fue un AR(2) con orden 1 de diferencias:

$$z_t = 0 + 26.44z_{t-1} + a_t - 0.12a_{t-1},$$

es decir:

$$y_t - y_{t-1} = 0 + 26.44(y_{t-1} - y_{t-2}) + a_t - 0.12a_{t-1},$$

siendo que δ , el término constante en el modelo auto-regresivo, no resultó significativamente diferente de cero. Para este modelo en particular, la Figura 9 ilustra los resultados obtenidos en la ventana de periodo 11, con $\tau = 5$. En la figura se observa que la diferencia entre la temperatura promedio observada y la restituida para la asimilación es mínima, no supera un valor de 0.15°C , es decir, el error es comparable con la precisión de las mediciones de contacto (termómetros de cazoleta) y es unas diez veces menor que el error de los sensores remotos.

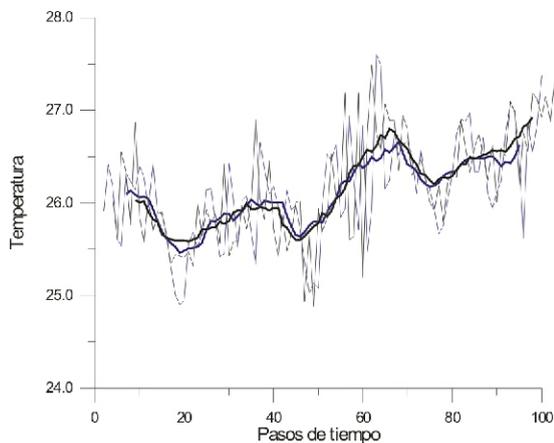


Figura 9. Temperatura satelital restituída (líneas azules) y pronosticada (líneas negras) por un modelo AR(2) en el punto de control (150,80). (Los trazos delgados indican la temperatura instantánea, los gruesos promedios móviles)

Los parámetros δ , ϕ_1 y θ_1 , en consecuencia, son funciones del espacio (x, y) y del tiempo t , y tienen que definirse desde un modelo ARIMA generalizado, atendiendo a la estadística temporal del campo de la temperatura satelital.

Se sabe que la variabilidad del campo de temperatura superficial producido por el modelo oceánico del CIOH, con asimilación de datos, logra resolver bien fenómenos como la surgencia en aguas de la Guajira o el calentamiento de las aguas en la cuenca de Panamá y El Darién. La mayor variabilidad de la temperatura, de acuerdo con la estación correspondiente, se presenta en zonas donde la nubosidad es abundante, como la cuenca de Panamá (debido a la interacción de la Zona de Convergencia Intertropical con otros sistemas sinópticos), y es mucho menor en zonas donde la nubosidad tiende a ser menos persistente, como en la Guajira y en el centro y este del Caribe, como se mostró en Anduckia *et al.* (2003) para el final de la época seca (marzo de 2003).

Para los seis meses considerados ahora la

dispersión de la señal de los satélites se muestra en la Figura 10. A pesar de que la variabilidad en este periodo, que podemos llamar época húmeda, sigue siendo elevada en la cuenca de Panamá, donde las autocorrelaciones son menores y justificarían un modelo de promedios móviles, resulta ahora mayor en las zonas de surgencia de la época seca y en general en el sector central del Caribe, donde la autocorrelación en la época seca era mayor, como para justificar un modelo auto-regresivo. Lo que esto quiere decir es que los parámetros del modelo auto-regresivo en cada punto deben depender del tiempo, como se aprecia comparando las Figuras 4 y 10, pero la variabilidad temporal de estos parámetros tiene que ocurrir en una escala lenta de tiempo, así que para períodos relativamente cortos el modelo estadístico puede considerarse cuasi-estacionario.

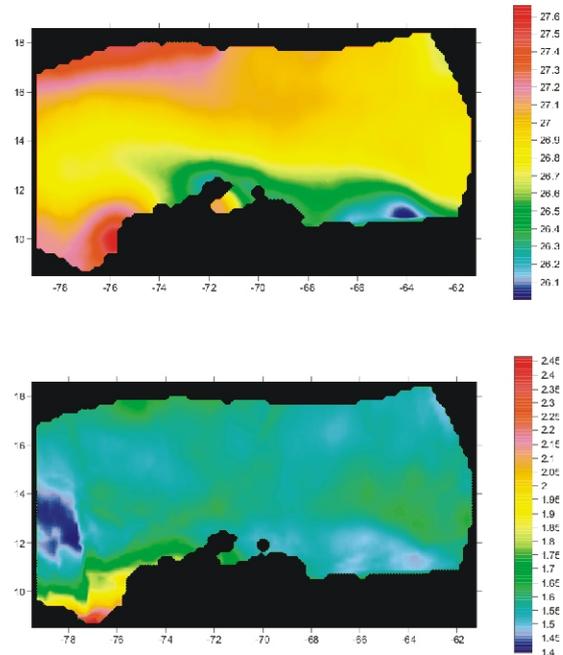


Figura 10. Campo promedio satelital de la temperatura superficial del mar (arriba, en $^{\circ}\text{C}$) y variabilidad temporal (abajo, en $^{\circ}\text{C}$) para seis meses, marzo-septiembre de 2003.

Conclusión

La motivación principal para el desarrollo de un modelo auto-regresivo de la temperatura superficial del mar, determinada satelitalmente, es su empleo de manera continua en el modelo de pronóstico oceanográfico del CIOH. Los datos de la temperatura de la superficie del mar se asimilan en el modelo con el propósito de actualizar los campos simulados para el diagnóstico y pronóstico operativo. Si los errores en los datos de entrada son pronunciados, la asimilación de estos datos producirá, como regla general, errores mayores. El procedimiento desarrollado permite suavizar en el tiempo el proceso de asimilación de la temperatura en cada punto y constituye un modelo de predicción de la temperatura. Gracias a la metodología propuesta, el error satelital en los datos asimilados se reduce en un orden de magnitud.

Habiendo mostrado aquí resultados para algunos puntos de control, el siguiente paso será establecer un modelo independiente en cada punto del espacio, que además dependa del tiempo y tenga en cuenta la respectiva época del año.

Con la inclusión de un radiómetro en futuros cruceros oceanográficos se tendrían elementos para poder diferenciar el efecto de las nubes y los distintos tipos de nubes, vapor de agua y polvo atmosférico en las mediciones satelitales, complementando el análisis estadístico aquí desarrollado.

Agradecimientos

Este trabajo fue realizado bajo el aporte de la Dirección General Marítima dentro del

proyecto de investigación “Seguimiento de las condiciones meteorológicas en el Caribe colombiano”. Los autores agradecen a la tripulación del ARC “Malpelo” que participó en el crucero “Caribe-2004”, así como al Dr. Joaquín Triñanes del nodo Caribe de CoastWatch/NOAA.

Referencias bibliográficas

- Anduckia J., Lonin S., Parra C., Molaes R.** (2003), “Validación de imágenes satelitales de temperatura superficial del mar mediante observaciones *in situ*”, *Boletín Científico No. 21*, pp. CIOH2003.
- Bowerman B., O'Connell R.** (1993), *Forecasting and time series: an applied approach*, Wadsworth Pub. Co, CA. ISBN 0-534-93251-7.
- Box G., Jenkins G.** (1970), *Time-series analysis: forecasting and control*. San Francisco, CA: Holden-Day, 1970.
- Donlon C., Minnett P., Gentemann C., Nightingale T., Barton I., Ward B., Murray M.** (2002), “Toward improved validation of satellite sea surface skin temperature measurements for climate research”, *Journal of Climate* **15** (feb. 2002), pp. 353-368.
- Kurzeja R. J.** (2002), “Skin temperature measurements on small bodies of water”, WSRC-MS-2002-00875, 2002.
- Jin M., Dickinson R.** (2000), “A generalized algorithm for retrieving cloudy sky skin temperature from satellite thermal infrared radiances”, *Journal of Geophysical Research* **105** D22, pp. 27037-27047 (nov. 2000)
- Lonin S., Anduckia J., Parra C., Molaes R.** (2003). Sistema de Pronóstico de las Condiciones Oceanográficas del Mar Caribe para Operaciones Navales. *Boletín Científico No. 21*, pp. 8-27, CIOH2003.
- Stowe L., Mc Clain E., Carey R., Pellegrino P., Gutman G., Davis P., Long C., Hart S.** (1991), “Global distribution of cloud cover derived from NOAA/AVHRR operational satellite data”, *Advances in Space Research* **11** (3), p. 351-354.