Revista de Teledetección. 2005. 24: 55-59

Influencia de la distribución espacial de los datos usados para determinar los algoritmos de la temperatura superficial del mar desde satélite

L. Arvelo-Valencia, M. Arbelo y P. A. Hernández-Leal
larvelo@ull.es, marbelo@ull.es, pealhdez@ull.es

Grupo de Observación de la Tierra y la Atmósfera. Departamento de Física. Universidad de La Laguna. 38200 La Laguna, Islas Canarias, España.

RESUMEN

Se examinan los errores en la estimación de la temperatura superficial del mar que tienen lugar cuando usamos datos *in situ* de diferentes bandas latitudinales y longitudinales para calcular los coeficientes del algoritmo global NLSST (Non-Linear Sea Surface Temperature). Se seleccionaron 8 bandas latitudinales y 18 longitudinales. Aplicando los coeficientes obtenidos de la zona ecuatorial y de la zona polar al caso global, se obtiene que la media de los residuos es superior a -0,6 ºC y que la desviación estándar está próxima a 1ºC en ambos casos. Cuando usamos los coeficientes de las bandas longitudinales en el conjunto global, los residuos presentan un sesgo inferior a -0,1ºC en aproximadamente el 80% de los casos. Los errores medios para los algoritmos longitudinales son pequeños cuando se comparan con los latitudinales, pero en ambos casos tienen una dependencia substancial de la temperatura superficial del mar.

PALABRAS CLAVE: temperatura superficial del mar, algoritmo global, AVHRR, NLSST.

ABSTRACT

The aim of this study is to examine the SST accuracy deviations that occur when we use *in situ* data from different latitudinal and longitudinal bands to compute, by linear regression, the NLSST (Non-Linear Sea Surface Temperature) algorithm coefficients. We have restricted the *in situ* SST to 8 latitudinal bands and 18 longitudinal bands. Applying the resulting coefficients of the equatorial region and from the polar zone to the global case, we have found that the mean of the residual is greater than -0,6ºC and the standard deviation near to 1ºC in both cases. When we use the coefficients from longitudinal bands in global data set, the residuals show a bias lower than -0,1 ºC in about the 80% of cases. The mean errors for the longitudinal algorithms are small when compared to the latitudinal, but in both cases have a substantial dependence on the SST *in situ*.

KEY WORDS: sea surface temperatura, global algorithm, AVHRR, NLSST.

INTRODUCCIÓN

La temperatura superficial del mar (TSM) a nivel global tiene una gran importancia para entender los cambios climáticos mundiales. Además, las TSM obtenidas a partir de las observaciones de los satélites se utilizan en una amplia variedad de cálculos, operaciones que van desde la determinación de rutas oceánicas hasta las planificaciones de actividades pesqueras. Uno de los atractivos de las medidas de la TSM mediante los satélites reside en su cobertura global a diferencia de otros tipos de medidas. Los radiómetros infrarrojos múlti-canales operando en condiciones libres de nubes nos proporcionan el conjunto de datos más fiable de la TSM.

Desde 1981, los satélites NOAA llevan instalado el radiómetro AVHRR (Advanced Very High Resolution Radiometer), que va dotado de canales en el infrarrojo térmico. (McMillin et al., 1984) y (McClain et al., 1985) desarrollaron algoritmos para determinar la TSM basándose en las diferencias lineales de las temperaturas de brillo de los canales 4 y 5 del AVHRR. Mejoras ulteriores en la corrección atmosférica incorporaron una formulación no lineal (Walton et al., 1998).

El objetivo de este trabajo consiste en analizar los errores que se producen en la determinación de la temperatura superficial del mar cuando utilizamos datos *in situ*, de diferentes franjas latitudinales y longitudinales, para estimar los coeficientes de un algoritmo global basados en el sensor NOAA-AVHRR.

DATOS

Para el desarrollo y validación de los algoritmos hemos utilizado la base de datos AVHRR Pathfinder Oceans Matchup Database (PFMDB). Estos datos, conocidos como *matchups*, suministran la información tomada *in situ* por boyas fijas y a la deriva y las medidas simultáneas del sensor AVHRR instalado a bordo del satélite NOAA-14, con un ajuste temporal de ± 30 minutos y espacial de ±0,1º de latitud y longitud.

El periodo de datos analizado cubre los años 1995-96. Todos los datos han sido filtrados para garantizar que estén libres de nubes y fallos de calibración, aplicando los tests propuestos por (Kilpatrick et al., 2001). Un total de 62.765 *matchups* pasaron estos tests.

METODOLOGÍA

El algoritmo utilizado para obtener los campos de la TSM a partir del sensor AVHRR fue el NLSST (Walton et al., 1998). Este algoritmo presenta la forma siguiente:

$$TSM_{sat}=a+bT_4+c(T_4-T_5)T_{sfc}+d(T_4-T_5)(\sec\theta-1) \quad (1)$$

donde TSM_{sat} es la temperatura superficial estimada por el satélite, T_4 y T_5 son las temperaturas de brillo de los canales 4 y 5 del AVHRR, T_{sfc} es un primer valor supuesto de la TSM, y θ es el ángulo cenital del satélite. Los coeficientes constantes a, b, c y d se obtienen por regresión lineal entre las TSM medidas por las boyas y las medidas por el satélite en el mismo instante. Para calcular los coeficientes, hemos utilizado las temperaturas de las boyas como T_{sfc}, mientras que para validar los resultados elegimos la OISST (Optimal Interpolated Sea Surface Temperature) (Reynolds et al., 1994) por ser esta una información fiable de la que disponemos en la base de datos. La OISST se produce semanalmente con una resolución espacial de un grado.

Para estudiar la influencia de la distribución espacial de los datos usados para determinar los algoritmos de la TSM desde satélite, hemos dividido las TSM medidas por las boyas durante los años 95 y 96, independientemente de la época del año, en diferentes bandas latitudinales y longitudinales (Fig. 1). Los datos de cada una de estas bandas fueron divididos aleatoriamente en dos subconjuntos totalmente separados y de igual tamaño, el subconjunto de entrenamiento, para la estimación de los coeficientes del algoritmo NLSST, y el subconjunto de validación, para determinar y analizar el comportamiento del algoritmo. Con el conjunto global se realizó la misma partición, incluyéndose al mismo en todos los cálculos y análisis posteriores como una "banda global".

Bandas latitudinales

Se seleccionaron 8 bandas latitudinales que hemos agrupado de la siguiente forma: 20N a 20S, 30N a 30S, 40N a 40S, 50N a 50S, 50NP (Norte al polo) más 50SP (Sur al Polo), 40NP más 40SP, 30NP más 30SP, y 20NP más 20SP. Las primeras cuatro bandas fueron obtenidas ampliando la latitud desde el Ecuador hacia los Polos y las cuatro restantes de manera inversa. De esta forma las primeras bandas contienen siempre temperaturas de la franja ecuatorial mientras que las segundas siempre incluyen temperaturas frías de las regiones próximas a los polos. En la Tabla 1, se muestran las estadísticas de las bandas seleccionadas que contienen sólo los matchups que han pasado los test de calidad sugeridos por (Kilpatrick et al., 2001).

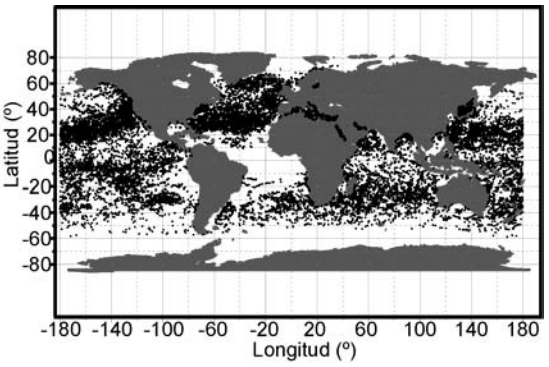

Figura 1. Bandas longitudinales y latitudinales en que se dividió el conjunto total de matchups disponibles para el sensor NOAA/AVHRR para los años 1995-1996.

Banda Latitudinal	Número de matchups	Temp. promedio (°C)	Temp. Mínima (°C)	Temp. Máxima (°C)
Global	62,765	23,73	2,84	33,3
20N-20S	18,731	27,30	15,64	32,4
30N-30S	34,845	26,41	10,60	33,3
40N-40S	51,432	25,10	3,20	33,3
50N-50S	59,572	24,20	2,84	33,3
50NP-50SP	3,193	8,85	3,07	19,9
40NP-40SP	11,333	12,10	2,84	27,0
30NP-30SP	27,920	16,83	2,84	32,9
20NP-20SP	44,034	20,58	2,84	33,3

Tabla 1. Número de matchups y estadísticas para cada banda longitudinal y el conjunto global.

Bandas longitudinales

Por otro lado se han dividido los datos globales en 18 bandas longitudinales de 20º cada una, también en este caso independientemente de la época del año.

En la Tabla 2, se presentan las estadísticas de cada banda, siguiendo el mismo procedimiento de filtrado anterior que garantice la calidad de los *matchups* seleccionados. Podemos observar que algunas franjas contienen pocos datos y una pobre distribución espacial (Fig. 1), lo que a priori debería provocar grandes errores en la aplicación del algoritmo al conjunto global.

Banda Longitudinal	Número de matchups	Temp. promedio (°C)	Temp. Mínima (°C)	Temp. Máxima (°C)
Global	62,765	23,73	2,84	33,3
160W-180W	4,867	25,58	3,20	31,70
140W-160W	6,996	23,74	3,10	31,10
120W-140W	8,521	18,54	4,00	31,31
100W-120W	3,173	22,34	3,85	31,17
80W-100W	5,146	24,65	5,20	33,3,
60W-80W	5,912	18,25	3,10	30,50
40W-60W	3,102	20,88	3,22	30,32
20W-40W	3,414	18,13	3,37	29,35
0-20W	2,332	16,27	2,84	28,40
0-20E	2,568	20,01	3,50	29,72
20E-40E	1,712	21,56	3,30	32,30
40E-60E	2,378	23,89	3,80	32,50
60E-80E	1,771	24,15	3,71	32,40
80E-100E	1,573	23,90	3,11	31,85
100E-120E	885	23,03	3,53	31,70
120E-140E	2,661	23,32	3,36	32,20
140E-160E	2,871	25,21	3,38	31,90
160E-180E	2,883	24,34	3,80	31,62

Tabla 2. Número de matchups y estadísticas para cada banda longitudinal y conjunto global.

Las bandas 20E-40E, 60E-80E, 80E-100E y 100E-120E se caracterizan por tener el menor número de *matchups*, todas ellas con menos de 2.000 datos, desde los 885 de la franja 100E-120E a los 1.771 de la región delimitada por las longitudes 60E-80E. Si observamos la Fig. 1, encontramos que estas bandas cubren en su mayoría áreas del océano Indico, donde el número de sistemas de medida *in situ* no es muy elevado. Y a su vez, por encima de 20N no existen medidas porque nos hallamos en regiones terrestres.

ANÁLISIS Y RESULTADOS

Para cada una de las bandas anteriores se estimaron los coeficientes de los algoritmos mediante regresión lineal múltiple con cada uno de los conjuntos de entrenamiento. Los algoritmos de cada banda fueron aplicados al conjunto de validación global para la estimación de los residuos (diferencia entre la TSM *in situ* y la TSM estimada con los datos de satélite por el algoritmo).

Análisis latitudinal

En la Tabla 3 se muestran las estadísticas de los residuos de los ocho algoritmos latitudinales determinados, donde podemos destacar que los mayores errores en la media y desviación estándar (D.E.) se refieren a la zona ecuatorial (20N a 20S), con un valor medio de –0,63 ºC y una de D.E. de 0,98 ºC, y la región próxima a los polos (50NP-50SP) con un error medio de –0,59 ºC y una D.E. de 1,05 ºC.

Banda Latitudinal	Número de matchups	Temp. promedio (°C)	Temp. Mínima (°C)	Temp. Máxima (°C)
Global	0,6479	0,0036	-3,763	3,956
20N-20S	0,9834	-0,6368	-5,284	3,822
30N-30S	0,7272	-0,2375	-4,051	3,949
40N-40S	0,6510	-0,0344	-3,748	3,971
50N-50S	0,6477	0,0061	-3,756	3,961
50NP-50SP	1,0543	0,5902	-6,666	3,800
40NP-40SP	0,7355	-0,1047	-4,823	3,633
30NP-30SP	0,6515	0,0474	-3,829	3,847
20NP-20SP	0,6482	0,0452	-3,702	3,989

Tabla 3. Estadísticas de los residuos después de aplicar los algoritmos latitudinales al conjunto de validación global (31.383 matchps).

Asimismo, cabe señalar la disminución en las medias y las desviaciones estándares conforme aumentamos el número de datos tanto en dirección polar como ecuatorial. El mayor residuo en valor absoluto, próximo a -7 ºC lo encontramos en la aplicación del algoritmo desarrollado para la zona 50NP-50SP, y se corresponde con una temperatura *in situ* de 29,87 ºC, situada muy próxima a la costa oeste de USA, en la posición 122,41º W y 36,75º N.

Análisis longitudinal

Cuando se aplican los coeficientes de las 18 bandas longitudinales al conjunto de datos global, las estadísticas de los residuos evolucionan tal como podemos ver en la Tabla 4.

Los algoritmos que peor responden ante el conjunto de validación global son aquellos generados por los datos de las franjas 60W-80W, 60E-80E, 20W-40W, 20E-40E, 160W-180W, como era de esperar tal como se comentó anteriormente.

Fijémonos por ejemplo en la distribución espacial de la franja 60W-80W (Fig. 1), con la mayor parte de los datos situados frente a la costa este de USA, y que presenta los peores resultados, con un valor medio de los residuos de 0,23 ºC y una dispersión de 0,65 ºC. Las 13 bandas restantes presentan siempre un error medio absoluto inferior a 0,1 ºC, lo que a priori podría considerarse aceptable.

Banda Longitudinal	Número de matchups	Temp. promedio (°C)	Temp. Mínima (°C)	Temp. Máxima (°C)
Global	0,6479	0,0036	-3,763	3,956
160W-180W	0,6664	-0,1048	-3,740	3,950
140W-160W	0,6490	-0,0740	-3,824	3,870
140W-120W	0,6744	-0,0739	-4,271	3,599
100W-120W	0,6666	-0,0684	-4,215	3,641
80W-100W	0,6889	0,0192	-3,588	4,084
60W-80W	0,6548	0,2338	-3,607	4,212
40W-60W	0,6510	-0,0253	-3,786	4,007
20W-40W	0,6484	-0,1243	-3,884	3,804
0-20W	0,6564	-0,0913	-4,157	3,748
0-20E	0,6867	0,0343	-3,743	4,081
20E-40E	0,7008	0,1117	-4,169	4,104
40E-60E	0,6535	-0,0900	-3,845	3,879
60E-80E	1,0163	-0,1848	-4,987	5,435
80E-100E	0,6564	-0,0888	-3,741	3,921
100E-120E	0,6635	-0,0291	-3,661	4,091
120E-140E	0,6581	-0,0490	-3,760	3,869
140E-160E	0,6652	-0,0216	-3,681	4,095
160E-180E	0,6663	-0,0175	-3,677	4,037

Tabla 4. Estadísticas de los residuos por bandas longitudinales (conjunto de validación: 31.383 matchps).

DISCUSIÓN Y CONCLUSIONES

La aplicación de los algoritmos determinados para las bandas latitudinales 20N-20S y 50NP-50SP al conjunto de validación global, presenta, en ambos casos, un error medio negativo, -0.63 ºC para la banda tropical y -0.59 ºC para la franja cercana a los polos. Resulta sorprendente que siendo conjuntos de datos tan dispares muestren un comportamiento promedio similar, esto es, del mismo signo. En un estudio mas detallado de estos errores encontramos que en realidad, los dos algoritmos se comportan de forma contraria. En la Fig. 2 hemos representado la tendencia de los residuos frente a la temperatura *in situ*, donde podemos observar que el error medio para la zona 20N-20S es casi nulo para temperaturas superiores a 23 ºC, mientras que para temperaturas inferiores a este umbral aparece un error negativo dependiente de la temperatura, aumentando a medida que esta disminuye. Para el caso de la banda 50NP-50SP, la tendencia parece totalmente contraria, el error promedio para valores < 23 ºC es casi nulo y para valores > 23 ºC el error tiende a hacerse negativo cuando aumenta la temperatura, provocando un error promedio total negativo.

Se aprecia que al restringir el conjunto de datos utilizado en la regresión de la TSM a zonas con latitudes altas o bajas, los errores presentan un notable incremento que disminuye a medida que se amplia la banda y por consiguiente la variabilidad climática.

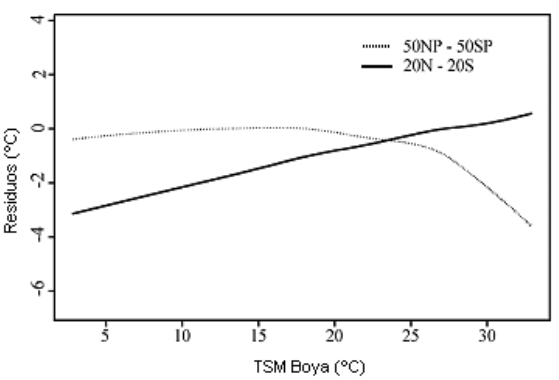

Figura 2. Tendencia de los residuos frente a la TSM "in situ" para los algoritmos latitudinales 20N-20S y 50NP-50SP.

Con respecto a las bandas longitudinales, encontramos que aquellas con más datos y mejor distribuidos espacialmente presentan errores pequeños

en los residuos (Tabla 4). A pesar de estos, aparentemente, buenos resultados, la tendencia de los residuos dependiente de la TSM *in situ*, entre otros parámetros físicos, como se muestra en la Fig. 3. En definitiva, se observan desviaciones significativas en los residuos para ciertos rangos de la TSM *in situ* aunque los errores promedios no lo acusen.

De los resultados obtenidos podemos concluir que los errores promedio para los algoritmos longitudinales son pequeños cuando se los compara con los latitudinales. Sin embargo, debe tomarse un cuidado especial cuando medidas *in situ* son elegidas para desarrollar algoritmos globales ya que en ambos casos (latitudinal y longitudinal) existe una dependencia considerable de la TSM medida por las boyas.

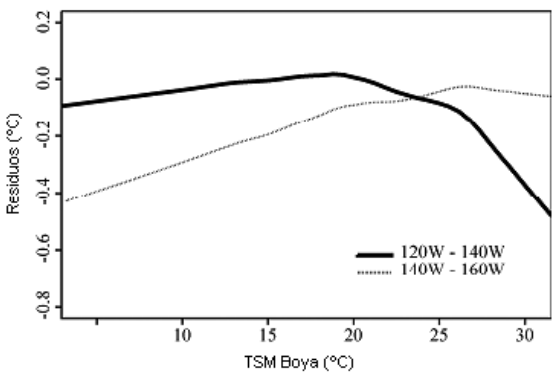

Figura 3. Tendencia de los residuos globales frente a la TSM "in situ" para los algoritmos longitudinales 120W-140WS y 140W-160W.

AGRADECIMIENTOS

Este estudio ha sido financiado por el Ministerio de Educación y Ciencia (Proyecto de Investigación CGL2004-06099-C03-02/CLI).

BIBLIOGRAFÍA

MCMILLIN, L.M. y D.S. CROSBY 1984. Theory and validation of multiple window sea surface temperature technique. *Journal Geophys. Res*. 80: 5113-5117.

MCCLAIN, E.P., W.G. PICHEL. y C.C. WALTON 1985. Comparative performance of AVHRR-based multichannel sea surface temepratures. *J. Geophys. Res.* 90: 11587-11601.

WALTON C.C., W.G. PICHEL, J.F. SAPPER, y D.A. MAY 1998. The development and operational aplication of nonlinear algorithms for the measurement of sea surface temperatures with NOAA Polar-Orbiting Enviromental Satellites. *Journal of Geophys. Res.* 103: 27999-28012.

KILPATRICK, K., G. P. PODESTA y R.H. EVANS 2001. Overview of the NOAA/NASA advanced very high resolution radiometer Pathfinder algorithm for sea surface temperature and associated matchup database. *Journal de Geophys. Res.* 106: 9179-9198.

REYNOLDS, R.W. y T.M. SMITH 1994. Improved global sea surface temperature analyses using optimun interpolation. *J. Clim.* 7: 929-948.