

UNIVERSIDAD DE CANTABRIA

**DEPARTAMENTO DE CIENCIAS Y TÉCNICAS
DEL AGUA Y DEL MEDIO AMBIENTE**

TESIS DOCTORAL

**METODOLOGÍAS DE CALIBRACIÓN DE BASES DE
DATOS DE REANÁLISIS DE CLIMA MARÍTIMO**

Presentada por: ANTONIO TOMÁS SAMPEDRO

**Dirigida por: FERNANDO J. MÉNDEZ INCERA
IÑIGO J. LOSADA RODRÍGUEZ**

Mayo, 2009

CAPÍTULO 7

CALIBRACIÓN PUNTUAL NO PARAMÉTRICA

7.1. Introducción.

Como ya se ha definido en el capítulo 4, existen dos tipos fundamentales de metodologías de calibración puntual, las paramétricas y las no paramétricas. Tras haber explicado las metodologías de calibración puntual paramétricas en los capítulos 5 y 6, en este capítulo 7 se va a describir la metodología de calibración puntual no paramétrica.

La calibración no paramétrica consigue corregir una base de datos modificando cada dato sin seguir una ley o relación paramétrica preestablecida. La corrección necesaria de cada dato es determinada a partir de una serie de valores análogos tomados de los datos instrumentales de referencia. Por ese motivo, esta técnica se suele denominar también ‘de análogos’.

El método original de Caires y Ferreira (2005), que se explica en el apartado 7.2, calibra la base de datos de H_s del ERA-40 con información de satélites (TOPEX/POSEIDON). Se definen grandes zonas para conseguir suficientes datos instrumentales para corregir el reanálisis (áreas circulares de 10°), realizando una calibración escalar global de la base de datos, con excelentes resultados al validar con datos de boyas (Caires y Sterl, 2005). Dicho método es ampliamente reconocido en la actualidad, pero para aplicarlo a las zonas costeras españolas es necesario definir ciertas modificaciones.

Como se ha señalado, el método no paramétrico original se define para zonas de mar abierto del Atlántico Norte, pudiendo ser correcto asumir que áreas de 10° tienen climas marítimos similares. Pero esta hipótesis no se cumple¹ para el Cantábrico y fundamentalmente para el Mediterráneo, debiendo rebajar a áreas circulares del orden de 1° . Esto conlleva una serie de modificaciones que se explicarán en el apartado 7.3. Así mismo, la aplicación de este método se realizará calibrando SIMAR-44 escaladamente con datos de satélites y boyas.

En el apartado 7.4 se va a modificar la metodología anterior para poder realizar calibraciones no paramétricas direccionales. Por lo tanto, se va a definir una metodología de calibración que permite calibrar todos los parámetros espectrales del estado de mar de la base de datos SIMAR-44, es decir H_s , T_m y θ , manteniendo las relaciones entre ellas, o lo que es lo mismo, sin modificar la función de distribución conjunta (Tomás *et al.*, 2009).

¹ 10° es aproximadamente la distancia que hay entre el estrecho de Gibraltar y la isla de Menorca, en cambio entre la isla de Ibiza y la de Mallorca hay aproximadamente 1° ; asumir que áreas de 10° tienen el mismo clima marítimo en los mares españoles sería similar a asumir que todas las costas mediterráneas tienen el mismo oleaje, lo que es incorrecto. Es usual asumir zonas homogéneas de oleaje en el Mediterráneo de 1° como máximo, debido a la gran variación espacio-temporal de sus oleajes.

En el apartado 7.5 se estudia la influencia de la cantidad de información instrumental disponible para calibrar direccionalmente el oleaje en una ubicación concreta. Finalmente, en el apartado 7.6, se resume la metodología de calibración puntual no paramétrica presentada y en el apartado 7.7, las conclusiones de este capítulo.

7.2. Metodología original.

La metodología de calibración no paramétrica de Caires y Ferreira (2005) se basa en el estudio de las series temporales (datos de modelado numérico) cuya corrección está condicionada a la información obtenida por otras variables de referencia (datos instrumentales). Así, para cada valor de la serie a calibrar, se busca el conjunto de valores en esa misma serie cuya historia reciente sea similar, de entre todos esos valores se seleccionan los que tienen información instrumental coincidente. Finalmente, con el promedio de las diferencias entre los datos de la serie original seleccionados y sus datos instrumentales simultáneos se corrige cada uno de los valores de la serie.

En concreto, la base de datos a calibrar, H_s , es función del espacio y del tiempo (se denota por $H_s(x,t)$) y la información instrumental de los satélites utilizados se denota por $H_{s_SAT}(x,t)$. Para las aplicaciones de Caires y Sterl (2005) se buscan historias temporales recientes de 3 observaciones ($m=3$), por lo que como el reanálisis utilizado tiene una resolución temporal de 6 horas, para cada secuencia o subserie de cada dato [$H_s(x,t)$, $H_s(x,t-6)$, $H_s(x,t-12)$] se buscan secuencias en el reanálisis [$H_s(x,t_i)$, $H_s(x,t_i-6)$, $H_s(x,t_i-12)$] similares a la original, donde el subíndice i denota el instante de tiempo de la serie temporal del reanálisis. Para buscar las subseries similares a la original se define el parámetro h_n , que acota las máximas diferencias permitidas con los datos buscados (3 observaciones) en la forma siguiente: $|H_s(x,t_i) - H_s(x,t)| < h_n$, $|H_s(x,t_i-6) - H_s(x,t-6)| < h_n$ y $|H_s(x,t_i-12) - H_s(x,t-12)| < h_n$. De las subseries del reanálisis que cumplen este requerimiento de analogía, se identifican las p secuencias que tengan datos coincidentes de satélites en los instantes t_i . A partir de ellos se obtiene el valor calibrado de $H_s(x,t)$, $H_{s_CAL}(x,t)$, dado en la ecuación 7.1:

$$H_{s_CAL}(x,t) = H_s(x,t) + \frac{1}{p} \sum_{i=1}^p [H_{s_SAT}(x,t_i) - H_s(x,t_i)] \quad (7.1)$$

donde t_i es cada uno de los instantes en la serie temporal de reanálisis que cumplen el criterio de analogía definido anteriormente ($|H_s(x,t_i) - H_s(x,t)| < h_n$, $|H_s(x,t_i-6) - H_s(x,t-6)| < h_n$ y $|H_s(x,t_i-12) - H_s(x,t-12)| < h_n$) y que tienen datos de satélites coincidentes.

En la figura 7.1 se presenta un croquis del método de calibración no paramétrico de Caires y Ferreira (2005) en la que se plantea la calibración de un dato, buscando subseries de 3 datos ($m=3$) que disten menos de h_n de la original (zonas sombreadas en amarillo). Para los puntos que cumplan dicha restricción y que tengan información instrumental se evalúa la diferencia $H_{sSAT}(x, t_i) - H_s(x, t_i)$ (marcada con una flecha verde); el promedio de todos esos valores (p) es el residuo o el valor que hay que añadir al dato original para calibrarlo (marcado en azul).

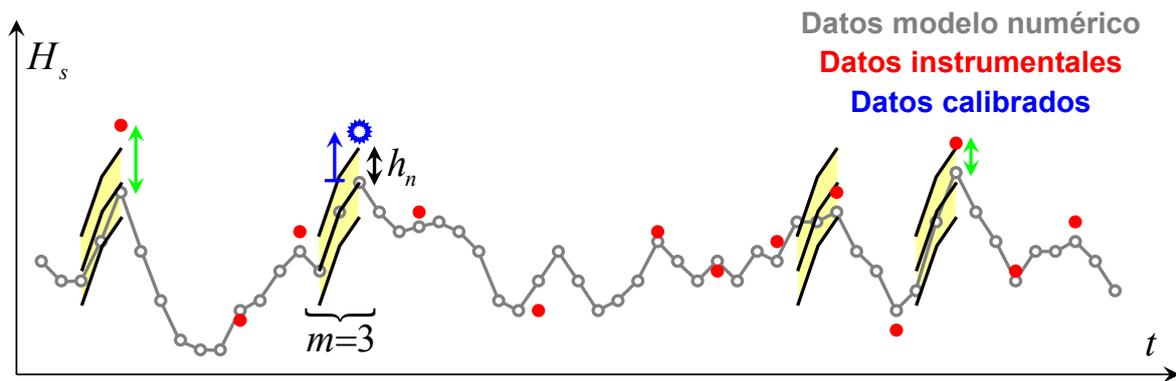


Figura 7.1. Croquis de la calibración no paramétrica definida por Caires y Ferreira (2005).

Para definir completamente el método de calibración aún hay que definir ciertos parámetros como h_n , que depende de m y de la población de datos que se dispone (n). En teoría h_n tiende a 0 cuando n tiende a infinito, siendo la expresión empleada en Caires y Ferreira (2005), la dada en la ecuación 7.2:

$$h_n = (cn^{\alpha-1} \log n^2)^{1/m} \quad (7.2)$$

donde c y α son constantes que deben determinarse empíricamente.

Usualmente este método ha sido implementado suponiendo que $\alpha = 0.2$ y que para una población de 700 datos la máxima diferencia permitida para considerar subseries similares es 0.3m, por lo que $c = 0.3^m / (700^{\alpha-1} \log 700^2)$, que se denotará por $h_{700} = 0.3m$. También deben definirse restricciones para considerar los datos instrumentales coincidentes en posición y en tiempo con los datos de modelado numérico. En la aplicación del método no paramétrico para la base de datos de H_s del ERA-40 en el Atlántico Norte se ha considerado una distancia máxima de 10° para considerar un dato de satélite (en concreto TOPEX/POSEIDON) coincidente con el dato a calibrar (en adelante se denotará dicha distancia por R).

En la figura 7.2 se presenta un ejemplo del área circular de radio $R=10^\circ$ que se considera válida para utilizar datos de satélites en el Atlántico Norte, según Caires y Sterl (2005). Así, para calibrar los datos de modelado numérico de la posición señalada en la figura con un punto rojo, se podría analizar la información de los satélites situada en el interior del círculo marcado en rojo y que sea coincidente en tiempo con los datos de modelado numérico que se van a calibrar. En dicha figura no sólo se presentan los datos del satélite TOPEX/POSEIDON (como la metodología original), sino todos los datos de satélites disponibles para esta tesis (GEOSAT, TOPEX/POSEIDON, GFO, JASON-1 y ENVISAT).

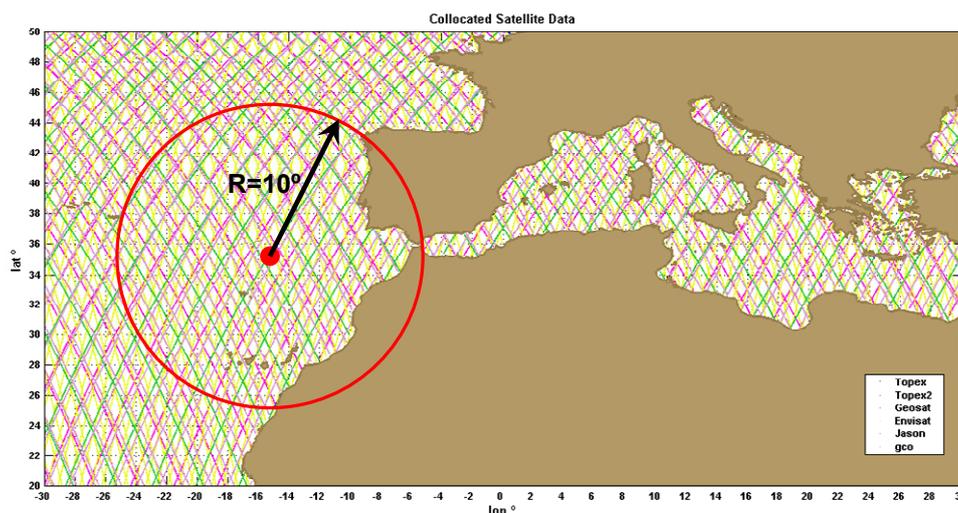


Figura 7.2. Ejemplo del criterio de proximidad para seleccionar los datos de satélites para calibrar, con el método no paramétrico original, los datos de una posición o nodo de modelado numérico.

7.3. Metodología escalar.

En este apartado se va a aplicar, modificando la metodología original de Caires y Ferreira (2005), la metodología de calibración escalar de H_s de SIMAR-44 con datos de satélites. Esta metodología también se aplica de manera escalar a H_s y T_m de SIMAR-44 utilizando datos de boyas.

7.3.1. Calibración con datos de satélites.

Para implementar el método de calibración no paramétrico a las bases de datos de reanálisis de los oleajes costeros españoles ha sido necesario introducir ciertas variaciones respecto del presentado en el apartado 7.2. Dicha metodología calibra escalaramente el valor de H_s en un punto con información de áreas de 10° alrededor suyo ($R=10^\circ$, ver figura 7.2) en el Atlántico Norte; a continuación se irá presentando la calibración de H_s de SIMAR-44 con datos de

satélites de diferentes áreas, distando respectivamente como máximo, $R = 5^\circ$ (figura 7.3), $R = 1^\circ$ (figura 7.4) y $R = 0.5^\circ$ (figura 7.5).

Cabe señalar que el método de calibración se ha aplicado en todas las posiciones donde se tienen boyas de la red exterior de OPPE con datos coincidentes con SIMAR-44 (12 boyas), para poder validar con ellas los resultados de las calibraciones realizadas con satélites. Las conclusiones, definiciones y modificaciones en la metodología de calibración han sido realizadas en base a los resultados obtenidos con todas las posiciones de aplicación, aunque sólo se muestran ejemplos de algunos de ellos, los que se han considerado más representativos. Debido a ello, en este apartado se aplica el método al Mediterráneo (por ser más variable que el Atlántico), en concreto en la posición de dos boyas, Cabo Begur y Cabo de Gata, por ser representativas de distintas zonas del Mediterráneo, y Cabo Begur tener un registro relativamente corto (fondeada en el 2001), en cambio Cabo de Gata cuenta con un registro relativamente largo (se fondeó en 1998).

A la hora de definir la subserie de búsqueda se ha definido (tras diversas pruebas) que las duraciones serán de 6 horas para el Mediterráneo y de 12 horas para el Atlántico, pudiendo variarse en función de la cantidad de información disponible o el grado de excepcionalidad (probabilidad de no excedencia) hasta el que se desea que la calibración tenga efecto; siendo el caso límite el de subseries de un solo elemento ($m = 1$) en el que no se considera la historia reciente de cada dato para buscar datos similares.

Así, en la figura 7.3 se muestra los resultados de la calibración para las dos posiciones propuestas del Mediterráneo, siguiendo la metodología original pero con $R = 5^\circ$ y con $m = 7$, ya que la resolución temporal de SIMAR-44 en el Mediterráneo es de 1 hora y la subserie de búsqueda se define de 6 horas; el resto de parámetros son los del método original, $\alpha = 0.2$ y $h_{700} = 0.3m$.

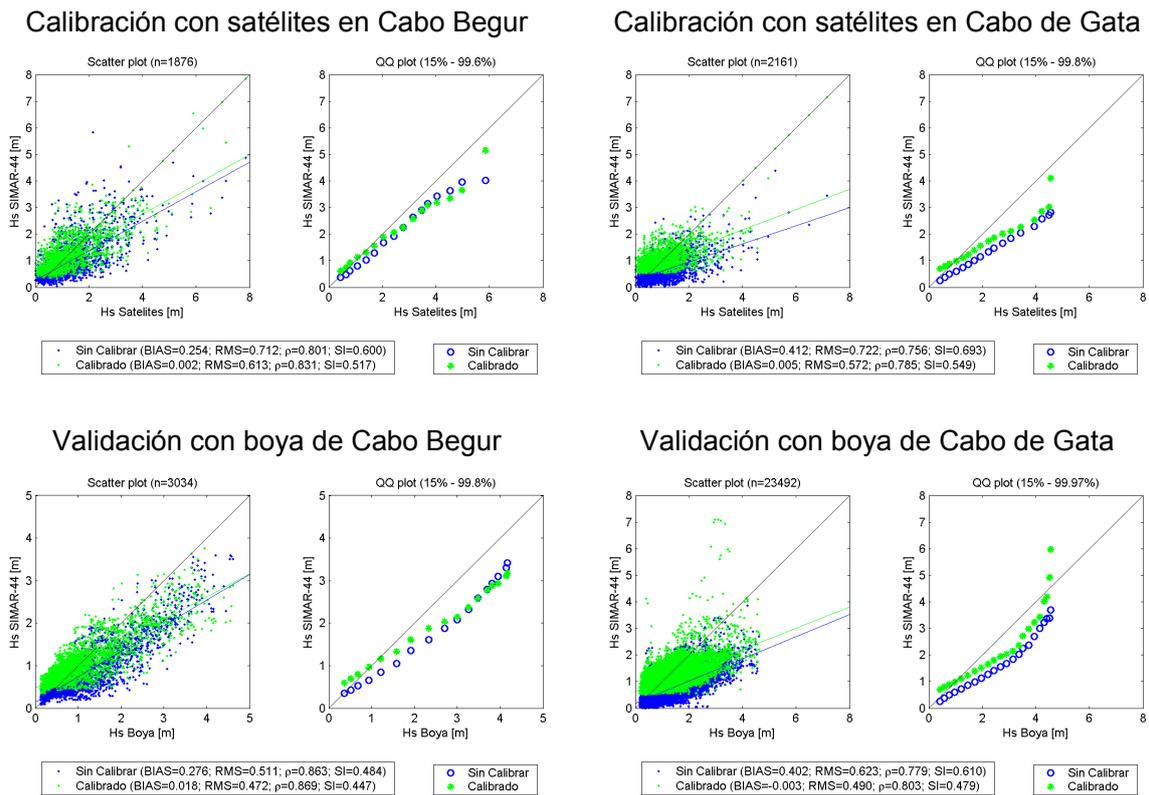


Figura 7.3. Calibración de H_s de SIMAR-44 con satélites (arriba) y validación con boyas (abajo) en las posiciones de las boyas de Cabo Begur (izquierda) y Cabo de Gata (derecha) con $h_{700} = 0.3m$ y $R = 5^\circ$.

Se puede comprobar en la figura 7.3 que todos los parámetros estadísticos mejoran con la calibración, tanto comparando los datos calibrados y sin calibrar de SIMAR-44 con los de los satélites que sirven para calibrar (calibración de la parte superior), como comparándolos con los datos coincidentes de las boyas (validación de la parte inferior). Pero a pesar de ello, cuando se representan los diagramas de dispersión (*scatter plot*) y fundamentalmente los cuantiles que caracterizan el régimen medio (*QQ plot*) no siempre se calibra correctamente. Nótese como con la calibración, los oleajes más energéticos se modifican incorrectamente, pudiendo comprobarse con la validación en las boyas (oleajes de Cabo de Gata que deberían tener H_s de 3-4 m se sobreestiman con H_s prácticamente del doble, 6-7 m). Esto es debido a que se han utilizado para calibrar los datos SIMAR-44, datos de satélites de zonas muy grandes para el Mediterráneo ($R = 5^\circ$), introduciendo datos que recogen lejanos oleajes mucho más energéticos que los de la zona que se calibra. Para evitar ese problema se reduce la distancia de búsqueda, con el inconveniente de que se obtienen menos datos para realizar la calibración.

En las figura 7.4 y 7.5 se reduce la distancia de búsqueda respectivamente a $R = 1^\circ$ y $R = 0.5^\circ$, disminuyendo notablemente el número de datos útiles para calibrar (se pasa aproximadamente de $n \approx 2000$ con $R = 5^\circ$ a $n \approx 500$ con $R = 1^\circ$ y a $n \approx 300$ con $R = 0.5^\circ$). Con la reducción de la información de referencia para calibrar se disminuye la dispersión de

los datos, logrando calibraciones mejores. También se puede observar que los resultados obtenidos mejoran sustancialmente con la disminución de R (sobre todo al pasar de $R = 5^\circ$ a $R = 1^\circ$), tanto con base en las comparaciones cualitativas de los parámetros estadísticos, como en las figuras de *scatter plot* y *QQ plot* de la calibración y de la validación.

No obstante, si se disminuye demasiado el área de búsqueda se obtienen insuficientes datos como para que la calibración sea coherente. Así por ejemplo, con $R = 0.25^\circ$ en la mayoría de las posiciones estudiadas no se selecciona ningún dato de satélite. Es por ello que se recomienda R del orden de $R \approx 0.5^\circ - 1^\circ$ para calibrar los oleajes de las costas españolas con datos de satélites. Estos valores son un compromiso entre que haya suficientes datos para calibrar y que los datos sean representativos de la posición donde se realiza la calibración, ya que si R es muy grande se realizan calibraciones incorrectas (ver ejemplos en figura 7.3), en cambio si R es muy pequeño implica que hay pocos datos de referencia para calibrar, dejando sin calibrar un gran porcentaje de datos, y que generalmente provienen de los temporales más energéticos (que son los menos frecuentes). A partir de los casos estudiados con $R = 5^\circ$ se obtiene un porcentaje de datos sin calibrar del orden del 1%-0.2%, con $R = 1^\circ$ se tiene aproximadamente 2%-0.4% y con $R = 0.5^\circ$ entre 3.9%-0.6%.

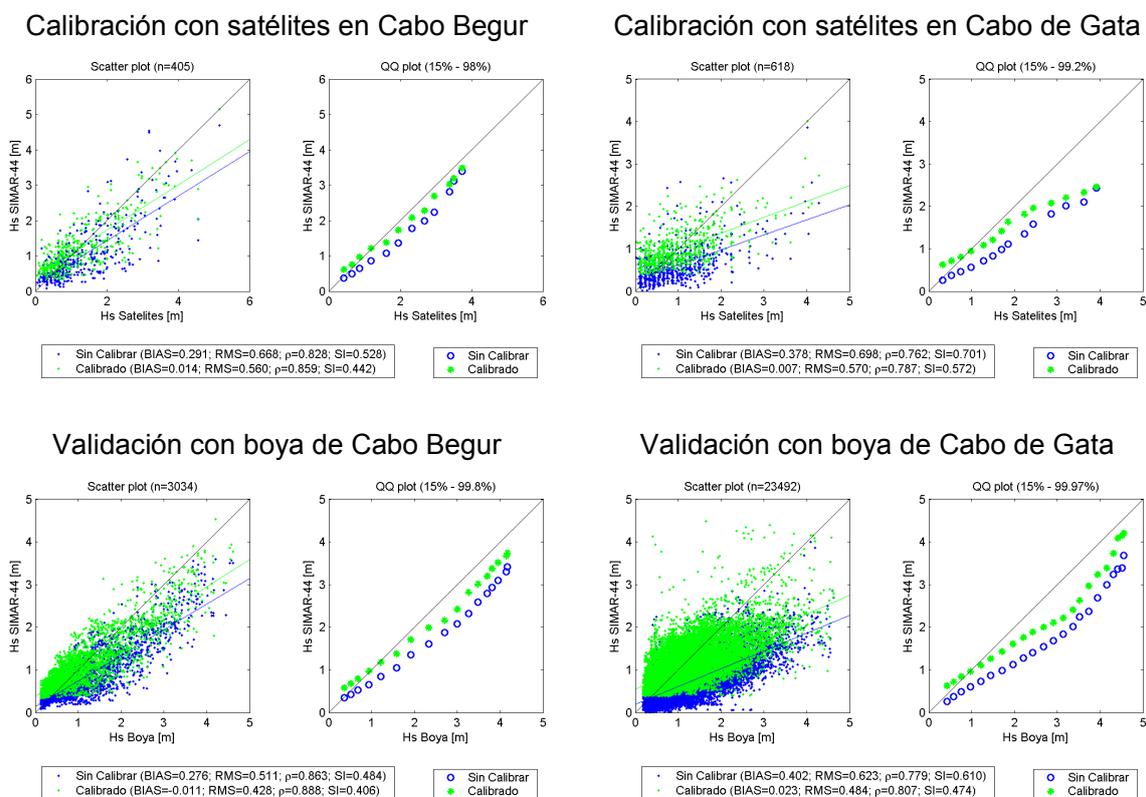
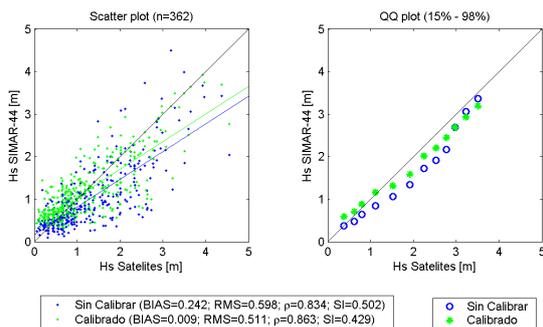
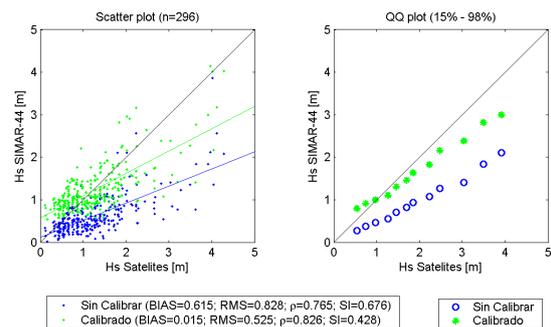


Figura 7.4. Calibración de H_s de SIMAR-44 con satélites (arriba) y validación con boyas (abajo) en las posiciones de las boyas de Cabo Begur (izquierda) y Cabo de Gata (derecha) con $h_{700} = 0.3m$ y $R = 1^\circ$.

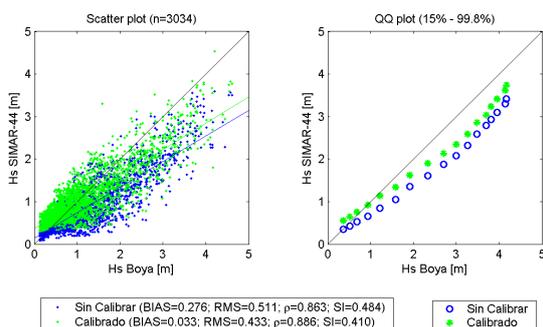
Calibración con satélites en Cabo Begur



Calibración con satélites en Cabo de Gata



Validación con boya de Cabo Begur



Validación con boya de Cabo de Gata

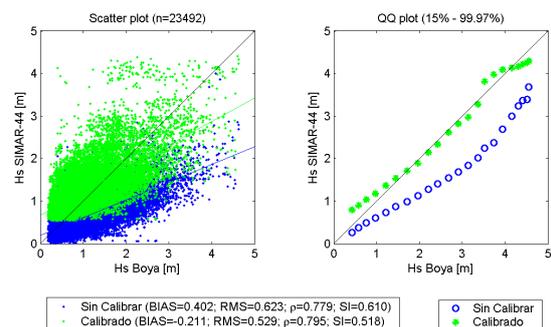


Figura 7.5. Calibración de H_s de SIMAR-44 con satélites (arriba) y validación con boyas (abajo) en las posiciones de las boyas de Cabo Begur (izquierda) y Cabo de Gata (derecha) con $h_{700} = 0.3m$ y $R = 0.5^\circ$.

Pese a todo ello, siempre es preferible tener poca información para calibrar, pero fiable, con lo que se corrigen pocos datos; que tener mucha información, pero más incorrecta, para calibrar, con lo que se corrige un gran porcentaje de datos pero de manera errónea. Por ello se puede concluir que para posiciones cercanas a la costa se recomienda utilizar valores $R \approx 0.5^\circ$.

Por otro lado, a partir de la ecuación 7.1, se comprende que cuantos más datos de referencia para calibrar un dato SIMAR-44 se tengan (es decir existan p subseries que cumpla los requisitos de similitud y que tengan coincidente información instrumental) más consistente y fiable será la calibración de ese dato. A medida que se tengan menos datos se aumenta la incertidumbre de la corrección. El caso límite es que exista únicamente un dato de referencia ($p=1$), con lo que con la calibración de SIMAR-44 se adopta exactamente ese valor instrumental, verificándose para esos casos que los puntos del diagrama de dispersión se sitúan en la bisectriz (ver en las figuras 7.3, 7.4 y 7.5 los *scatter plot* de las calibraciones). Finalmente, cuando no hay ningún dato de referencia para calibrar un dato SIMAR-44 no se modifica dicho dato.

En la metodología de calibración presentada, el requisito de similitud de subseries es quizás muy exigente. La restricción de tomar para una población de 700 datos una diferencia máxima de 0.3m para considerar subseries similares se aproxima a la precisión con la que miden las

boyas y/o satélites las mayores olas (el 5% de 5m es aproximadamente 0.3m), que son las olas en las que más interés se tiene en calibrar correctamente. Para tratar de solucionar este problema se toma $h_{700} = 0.5m$, calibrando de nuevo los casos con $R = 1^\circ$ y $R = 0.5^\circ$; dichas calibraciones se representan en la figuras 7.6 y 7.7, que se pueden comparar con los resultados obtenidos en la figuras 7.4 y 7.5 (con $h_{700} = 0.3m$).

Con $h_{700} = 0.5m$ se obtienen resultados del orden de los obtenidos con $h_{700} = 0.3m$, mejorando las validaciones con datos de boyas, sin distorsionar indebidamente los oleajes más extremos y calibrando un mayor porcentaje de datos (con $R = 1^\circ$ se tiene aproximadamente un porcentaje de datos sin calibrar de 0.8%-0.2% y con $R = 0.5^\circ$ entre 1.5%-0.2%). Por ello, a la vista de estos resultados, se fija para las siguientes aplicaciones de calibración $R = 0.5^\circ$ y $h_{700} = 0.5m$.

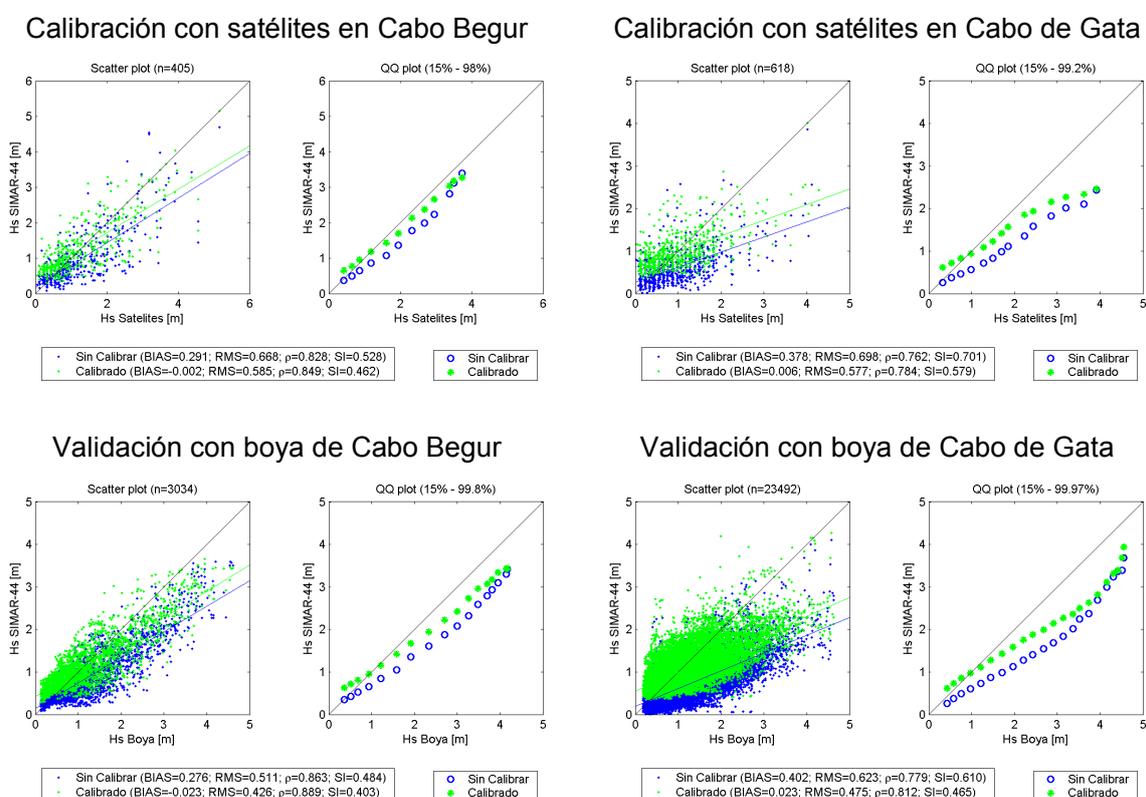
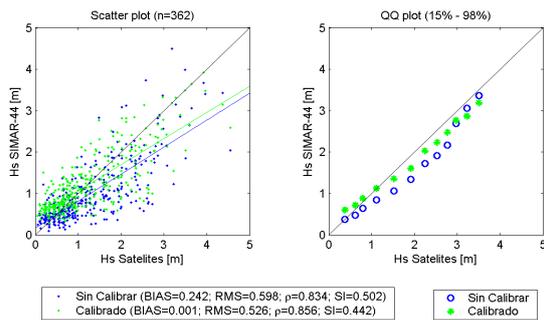
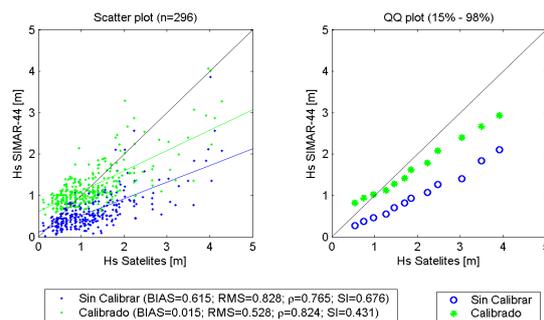


Figura 7.6. Calibración de H_s de SIMAR-44 con satélites (arriba) y validación con boyas (abajo) en las posiciones de las boyas de Cabo Begur (izquierda) y Cabo de Gata (derecha) con $h_{700} = 0.5m$ y $R = 1^\circ$.

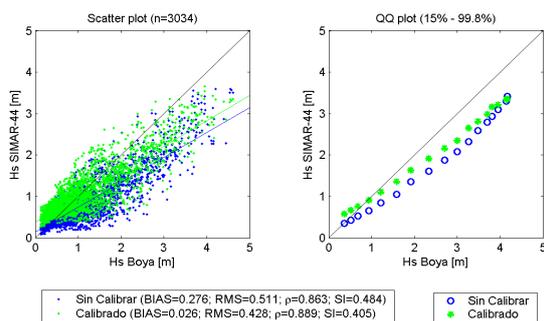
Calibración con satélites en Cabo Begur



Calibración con satélites en Cabo de Gata



Validación con boya de Cabo Begur



Validación con boya de Cabo de Gata

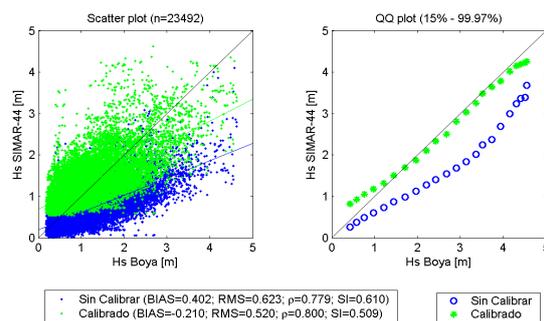


Figura 7.7. Calibración de H_s de SIMAR-44 con satélites (arriba) y validación con boyas (abajo) en las posiciones de las boyas de Cabo Begur (izquierda) y Cabo de Gata (derecha) con $h_{700} = 0.5m$ y $R = 0.5^\circ$.

A pesar de las mejoras producidas al utilizar $R = 0.5^\circ$ y $h_{700} = 0.5m$ aún se tienen algunos problemas para lograr que las series calibradas sean homogéneas. Esto es debido a que al corregir cada dato independientemente, en ciertas ocasiones se obtienen correcciones muy diferentes para oleajes muy similares y contiguos en el tiempo, introduciéndose un cierto ruido en la señal provocada por el procedimiento de calibración. En la figura 7.8 se muestran las series de H_s de SIMAR-44 sin calibrar y calibradas, así como los H_s de satélites de noviembre de 2001, para Cabo Begur y Cabo de Gata calibrados con $R = 0.5^\circ$ y $h_{700} = 0.5m$. En dicha figura también se sombrea en gris el intervalo de confianza² del 95% de los datos calibrados.

Se puede observar en la figura 7.8 que cada dato se corrige de manera distinta, así los picos del temporal, cuando no se tiene información para calibrar, no se modifican; de hecho para esos valores el intervalo de confianza del 95% aumenta considerablemente. Nótese que el caso límite con un solo dato ($p = 1$), el intervalo de confianza tiende a infinito. También se

² Intervalo de confianza de 95%: El cálculo del intervalo de confianza de los datos calibrados se realiza a partir de la cuasidesviación típica muestral, s_x , (ver ecuación 3.3) de la diferencia de los valores originales y de referencia de satélites, siendo $X = H_{sSAT} - H_s$ y también de la función de distribución de la t-Student's, $t-S$ y del número de muestras de X , p . Así el intervalo de confianza del 95% se expresa como $H_s \pm t \cdot S^{-1} \left(1 - \frac{1-0.95}{2}, p-1\right) s_x \sqrt{1/p}$, tomada de Luceño (1989).

puede apreciar (fundamentalmente en la calibración de Cabo Begur) que la serie calibrada tiene un ruido u oscilaciones de alta frecuencia que no tiene la serie original sin calibrar.

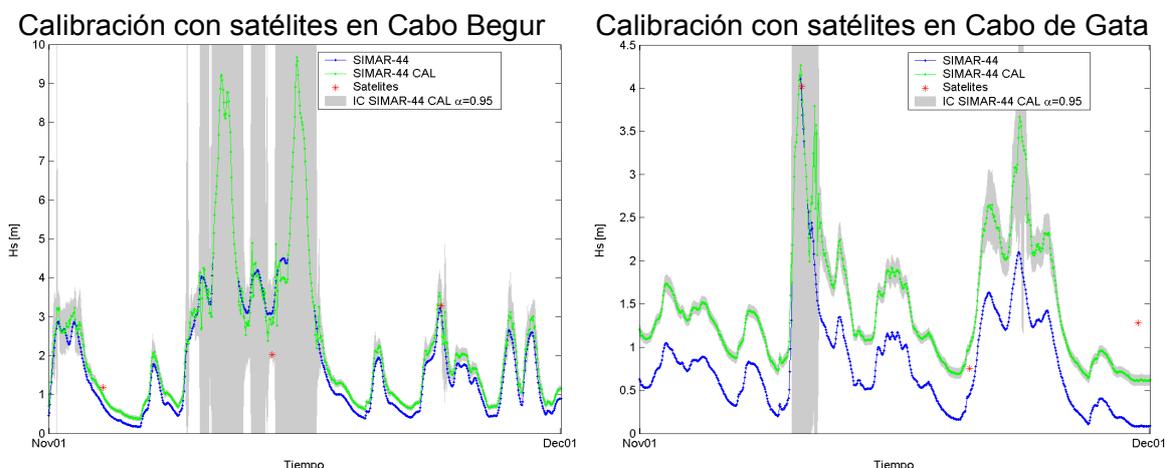


Figura 7.8. Serie temporal (noviembre de 2001) de H_s de SIMAR-44 calibrado y sin calibrar con satélites en las posiciones de las boyas de Cabo Begur (izquierda) y Cabo de Gata (derecha) con $h_{700} = 0.5m$ y $R = 0.5^\circ$.

Para solucionar el problema de las inhomogeneidades o ruido generado con la calibración se han desarrollado varias alternativas, pero la que mejores resultados ha producido es aplicar un filtro que reduzca las oscilaciones de alta frecuencia a los residuos (o promedio de diferencias entre el valor de referencia de satélite y el original de SIMAR-44) que finalmente se agregan a la serie original para calibrarla. Por simplicidad se han probado medias móviles de 3, 5, 7, 9 y 11 horas, aplicándolas tanto una vez como dos veces sucesivas a los residuos, siendo las que mejores resultados generaban las dobles de 5 horas para el Mediterráneo y las dobles de 9 horas para el Atlántico.

En la figura 7.9 se presentan las series temporales de las calibraciones de Cabo Begur y Cabo de Gata, que son las mismas que las de la figura 7.8, pero ahora aplicando los filtros de media móvil. Se puede observar en esta figura, en comparación con la figura 7.8, que con los filtros se elimina el ruido introducido con la calibración, sin distorsionar las oscilaciones de alta frecuencia propias de la señal original. Se produce una homogeneización de la señal que contribuye a reducir el porcentaje de datos sin calibrar, de hecho para los casos de $R = 0.5^\circ$ y $h_{700} = 0.5m$ se pasa de tener el 1.5%-0.2% de los datos sin calibrar antes de aplicar el filtro a porcentajes menores siempre del 1% (0.6%-0.05%) tras el filtrado.

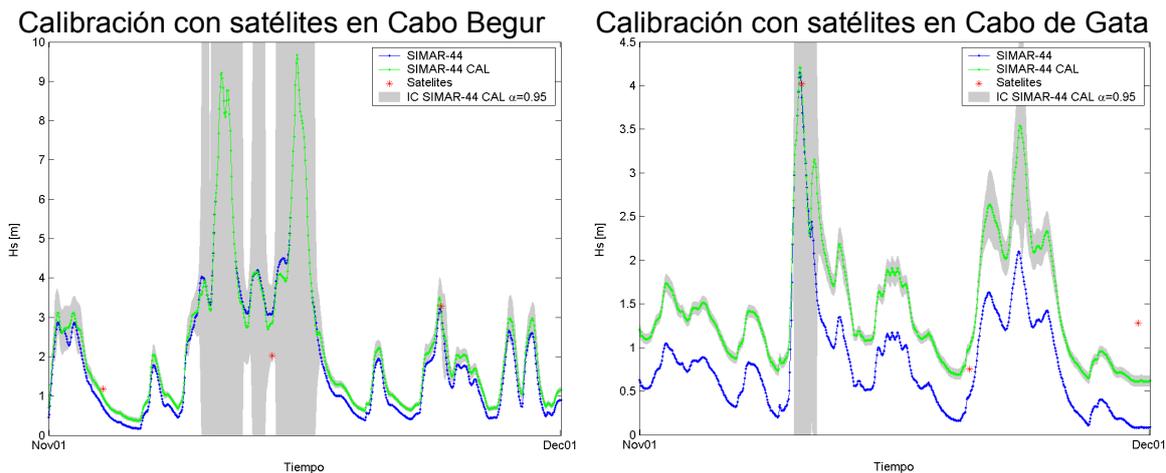
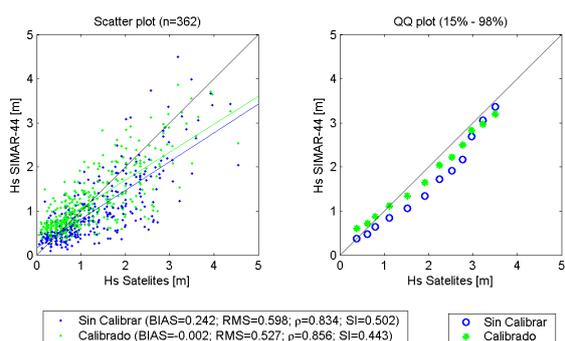


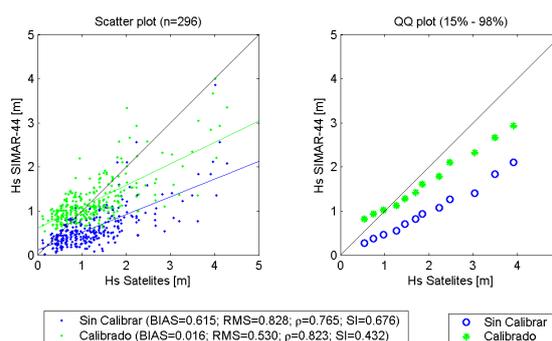
Figura 7.9. Serie temporal (noviembre de 2001) de H_s de SIMAR-44 calibrado y sin calibrar con satélites en las posiciones de las boyas de Cabo Begur (izquierda) y Cabo de Gata (derecha) con $h_{700} = 0.5m$ y $R = 0.5^\circ$. Aplicando doble filtro de media móvil de 5 horas.

Finalmente, en la figura 7.10 se presentan de nuevo los resultados de la calibración escalar con satélites y validando con boyas para Cabo Begur y Cabo de Gata, pero aplicando ahora previamente el doble filtro de media móvil de 5 horas. Comparando dichos resultados con los de la figura 7.7, que son sin filtros, se puede comprobar que prácticamente no varían los resultados de los parámetros cuantitativos de ajuste, pero se consideran mejores estos últimos (con filtros). Se consideran mejores los resultados de las calibraciones introduciendo los dobles filtros de media móvil porque no se generan oscilaciones de alta frecuencia debidas a la escasez de datos instrumentales.

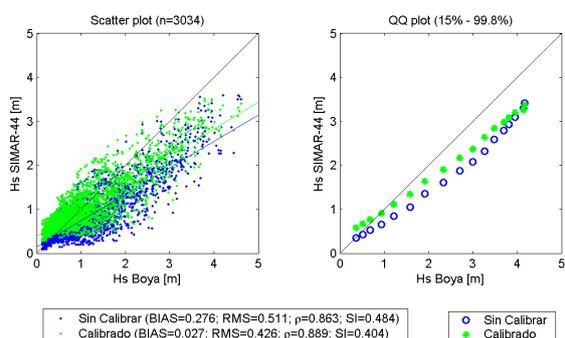
Calibración con satélites en Cabo Begur



Calibración con satélites en Cabo de Gata



Validación con boya de Cabo Begur



Validación con boya de Cabo de Gata

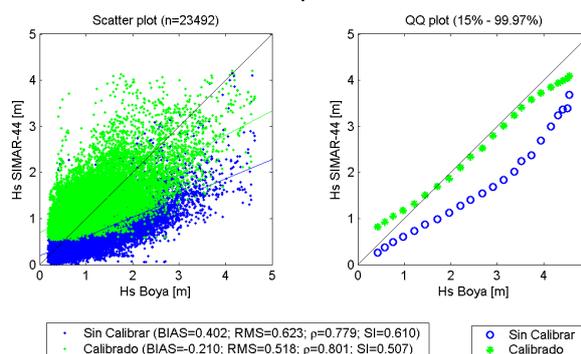


Figura 7.10. Calibración de H_s de SIMAR-44 con satélites (arriba) y validación con boyas (abajo) en las posiciones de las boyas de Cabo Begur (izquierda) y Cabo de Gata (derecha) con $h_{700} = 0.5m$ y $R = 0.5^\circ$.

Aplicando doble filtro de media móvil de 5 horas.

A pesar de que con el filtrado se han solucionado muchos de los problemas inherentes a la metodología de calibración original, como homogeneizar las series, reducir el porcentaje de datos que no se calibran, etc, aún persiste un pequeño problema que afecta a la calibración de los valores próximos a cero de la serie a calibrar. Fijándose en muchas de las figuras presentadas en este apartado se puede observar que con la calibración se introduce un sesgo negativo o aumento sistemático y mayor de lo necesario de los valores más pequeños. Esto es debido a que el método calibra en función la media de los residuos (que son las diferencias entre los datos análogos instrumentales y los datos de referencia a calibrar), pero si al realizar la media uno de los residuos es mucho mayor que el resto, se desvirtúa el resultado de la media. Para solucionarlo se ha tratado de utilizar la mediana en lugar de la media, o hacer el promedio de los cocientes en lugar de las diferencias, o a variar logarítmicamente el valor de h_n en función de H_s para que tenga valores más pequeños con los oleajes próximos a 0; pero las mejoras no han sido significativas, por lo que se ha desechado utilizar estas alternativas. De hecho los pequeños sesgos que se producen no afectan a los oleajes en los que se suele estar más interesado, el régimen medio y alto.

7.3.2. Calibración con datos de boyas.

Otra alternativa a la calibración con datos de satélites es, cuando existen boyas cercanas, la calibración con datos de esas boyas. La ventaja de calibrar con datos de boyas es que proporcionan mucha más información que los satélites en una posición fija, por eso, si se dispone de ellas, es preferible calibrar con boyas. Además, de momento, con los satélites sólo se dispone con suficiente precisión de información escalar de H_s , prescindiendo utilizar para calibrar otros parámetros derivados como T_m o θ .

Así por ejemplo, a continuación en la figura 7.11 se calibra H_s en las posiciones de las boyas de Cabo Begur y Cabo de Gata con la metodología escalar no paramétrica descrita en el apartado anterior, pero en lugar de utilizar datos de satélites se calibra con datos de dichas boyas. Se puede comprobar que es posible calibrar hasta probabilidades de no excedencia mayores que en los casos de calibración con satélites y que se logran mejores ajustes con las boyas (ver figura 7.10). La validación de la calibración realizada se efectúa con los datos de satélites coincidentes, que distan menos de 0.5° , constatándose la mejora de los datos de H_s de SIMAR-44 con la calibración.

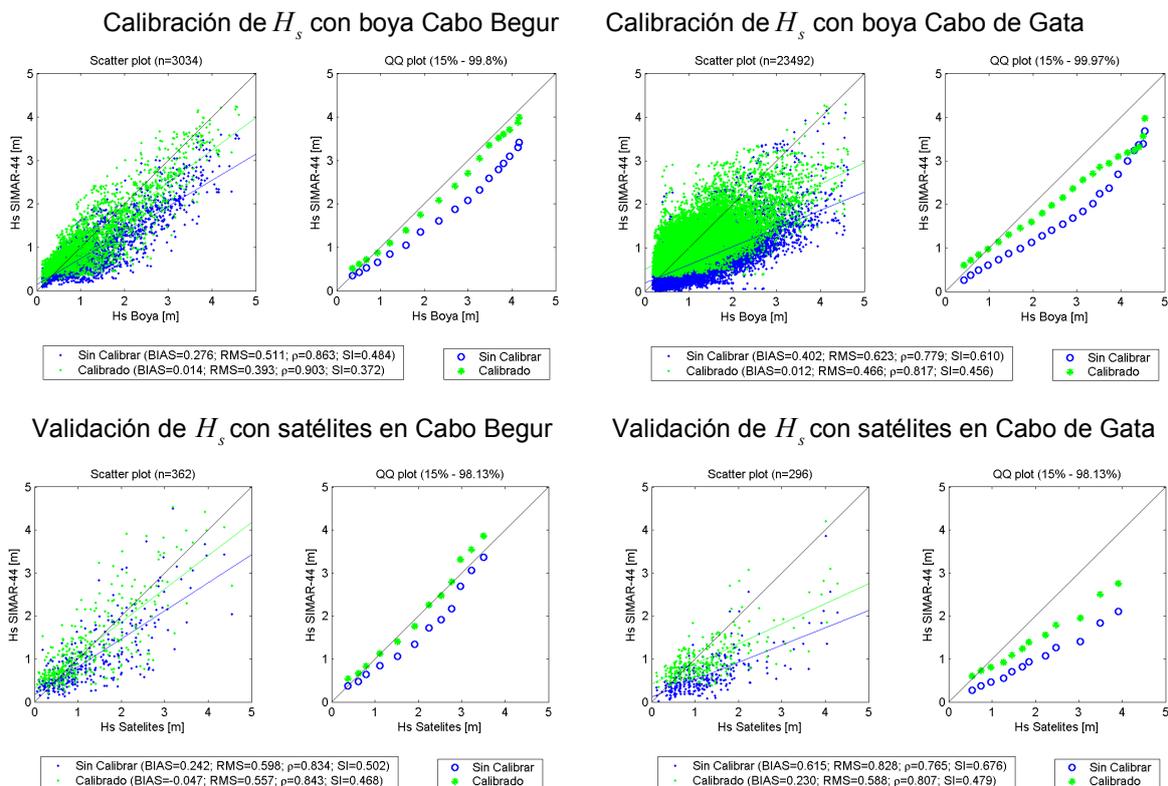


Figura 7.11. Calibración de H_s de SIMAR-44 con boyas (arriba) y validación con satélites (abajo) en las posiciones de las boyas de Cabo Begur (izquierda) y Cabo de Gata (derecha). Aplicando doble filtro de media móvil de 5 horas.

Además de la calibración de H_s de SIMAR-44 con datos de boyas, también se puede calibrar otros parámetros de SIMAR-44 como T_m o θ , pues se dispone de esa información, entre otros parámetros de oleaje, a partir de boyas. Pero para aplicar el método no paramétrico a estas variables hay que determinar ciertos parámetros intrínsecos de la metodología de calibración, como m , c o α . Para m y α se han dejado los mismos parámetros que los determinados para H_s , $m=7$ y $\alpha=0.2$; pero para establecer el equivalente a h_{700} para T_m o θ (denotándose t_{700} y θ_{700} respectivamente), se han realizado distintas pruebas, en función de ellas, de los rangos de variación de las distintas variables y de las precisiones en su determinación se han propuesto:

- Para T_m , con una población de 700 datos, se permite una máxima diferencia para considerar subseries similares de 0.5s, por lo que $c = 0.5^m / (700^{\alpha-1} \log 700^2)$, siendo $t_{700} = 0.5s$.
- Para θ , con una población de 700 datos, se permite una máxima diferencia para considerar subseries similares de 22.5°, por lo que $c = 22.5^m / (700^{\alpha-1} \log 700^2)$, siendo $\theta_{700} = 22.5^\circ$.

En la figura 7.12 se representa la calibración, para Cabo Begur y Cabo de Gata de T_m y de θ únicamente para Cabo Begur pues Cabo de Gata era escalar hasta el año 2003. Se utiliza el método escalar no paramétrico con los parámetros presentados y con un doble filtro de media móvil de 5 horas. Se comprueba que dicho método es perfectamente aplicable a otras variables que no sean H_s , método original. Pero cabe destacar que la representación realizada para verificar la calibración realizada de θ no es cuantitativamente correcta, pues al ser una variable cíclica no tiene sentido, ni los ajustes, ni la representación *QQ-plot*, pero indica cualitativamente que se mejora con la calibración.

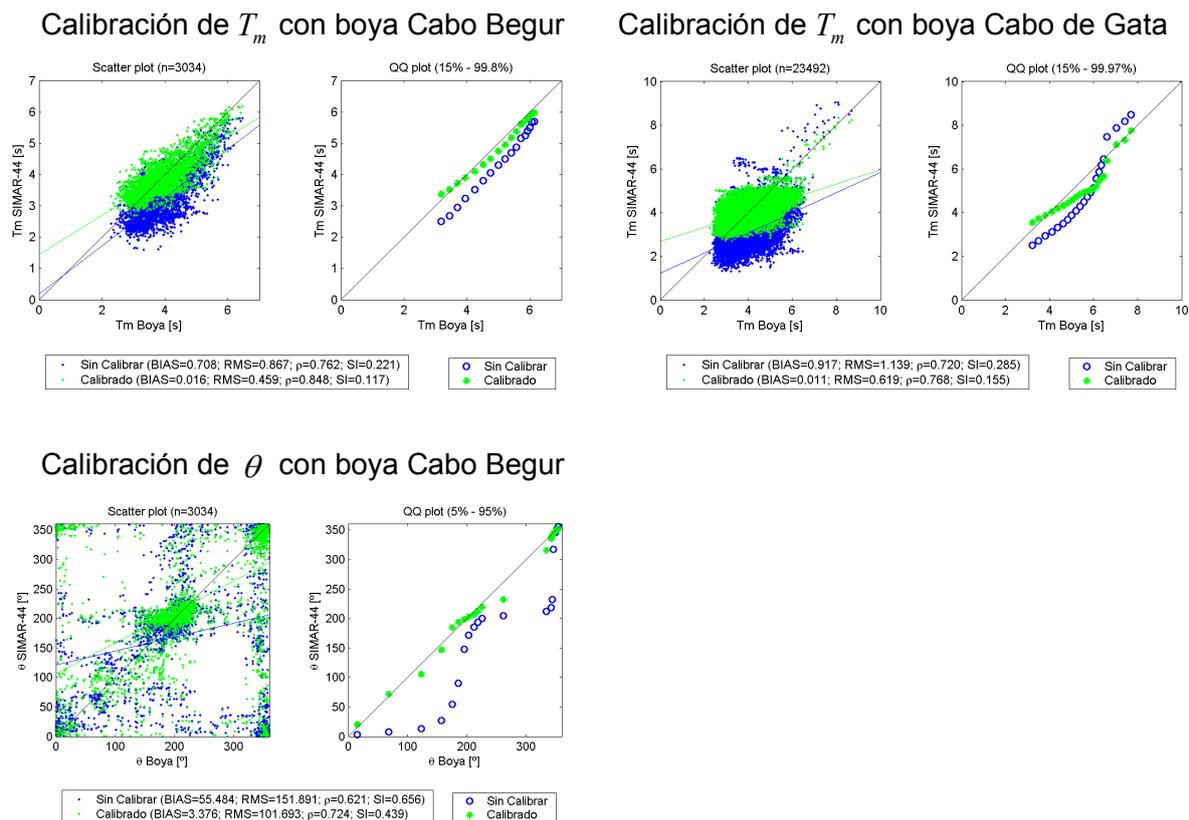


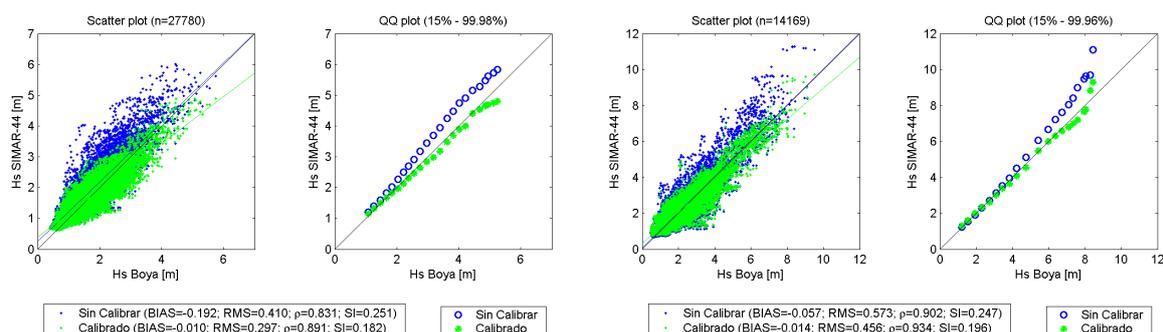
Figura 7.12. Calibración de T_m (arriba) y θ (abajo) de SIMAR-44 con las boyas de Cabo Begur (izquierda) y Cabo de Gata (derecha). Aplicando doble filtro de media móvil de 5 horas.

Hasta ahora las boyas utilizadas para calibrar han sido del Mediterráneo, pero existen algunas diferencias cuando se aplica la metodología al Atlántico, pues la variabilidad espacio-temporal del Atlántico es menor que el Mediterráneo. Así para el Mediterráneo se han utilizado subseries de comparación de 6 horas y como la resolución temporal de SIMAR-44 en el Mediterráneo es de 1 hora, se tiene $m = 7$; en cambio en el Atlántico la resolución temporal de SIMAR-44 es de 3 horas, pues su variación temporal es menor, por eso mismo y debido a las pruebas realizadas se utilizan subseries de comparación de 12 horas para el Atlántico, por lo que se tiene $m = 5$. En la misma línea, los filtros también varían, siendo dobles filtros de media móvil de 5 horas para el Mediterráneo y de 9 horas para el Atlántico. Cabe señalar que introduciendo los dobles filtros de media móvil se evita generar oscilaciones de alta frecuencia debidas a la escasez de datos y al ruido de alta frecuencia de los registros de las boyas.

Para ilustrar la versatilidad de la metodología de calibración escalar desarrollada, también se van a mostrar los resultados de la calibración con dos boyas del Atlántico. Las dos boyas seleccionadas para aplicar el método de calibración son Gran Canaria y Villano-Sisargas, estando la primera en una zona donde existen ciertas carencias en el modelado numérico, y la segunda (Villano-Sisargas) en una zona donde el reanálisis simula con mucha más precisión

lo registrado por fuentes instrumentales. Así en las figuras 7.13 y 7.14 se presenta respectivamente la calibración de H_s y T_m de SIMAR-44 en ambas boyas, así como la validación para ambos parámetros. Pero en estas ocasiones se muestra otra alternativa de validación que no depende de los datos de satélites y que permite validar T_m . En este caso no se ha calibrado con todos los datos disponibles de las boyas, sino que se han reservado aproximadamente el 25% de los datos para validar, en concreto se ha calibrado con los datos hasta el año 2000 inclusive y se ha validado con los datos del año 2001.

Calibración de H_s con boya Gran Canaria Calibración de H_s con boya Villano-Sisargas



Validación de H_s con boya Gran Canaria Validación de H_s con boya Villano-Sisargas

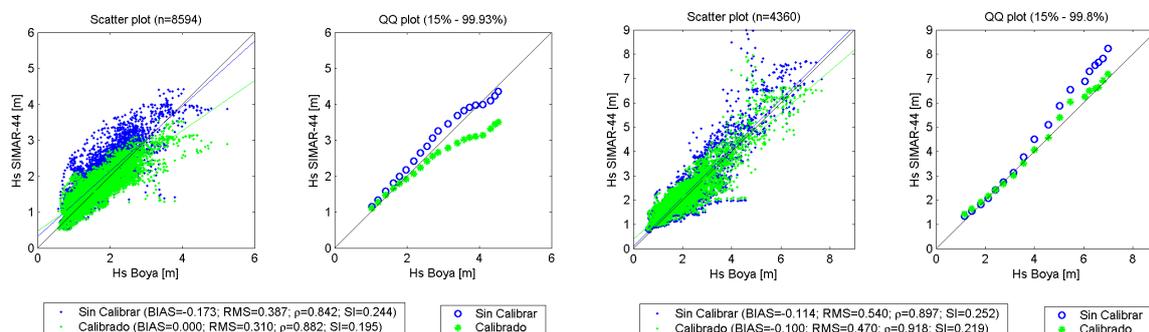


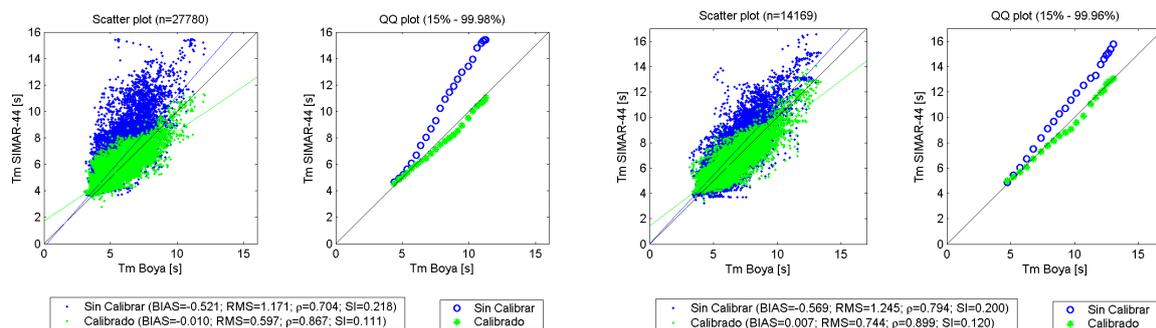
Figura 7.13. Calibración de H_s de SIMAR-44 con boyas, hasta el año 2000 (arriba) y validación con datos de boyas del año 2001 (abajo); las boyas son las de Gran Canaria (izquierda) y Villano-Sisargas (derecha).

Aplicando doble filtro de media móvil de 9 horas.

En la figura 7.13 se puede comprobar que tanto en la calibración como en la validación de Villano-Sisargas se reduce la dispersión de los datos y se mejora tanto el régimen medio de H_s como el medio-alto, logrando unos muy buenos resultados con la calibración. En cambio en Gran Canaria, en la verificación de la calibración se obtienen resultados satisfactorios, mejorando bastante los parámetros de ajuste, pero en la validación, a pesar de que los parámetros de ajuste son buenos, se observa que en la rama medio-alta del régimen se subestima el oleaje con la calibración, esto es debido a que existen pocos datos disponibles para calibrar y que dichos datos no contienen las correcciones necesarias para los datos más extremos del año 2001.

Calibración de T_m con boya Gran Canaria

Calibración de T_m con boya Villano-Sisargas



Validación de T_m con boya Gran Canaria

Validación de T_m con boya Villano-Sisargas

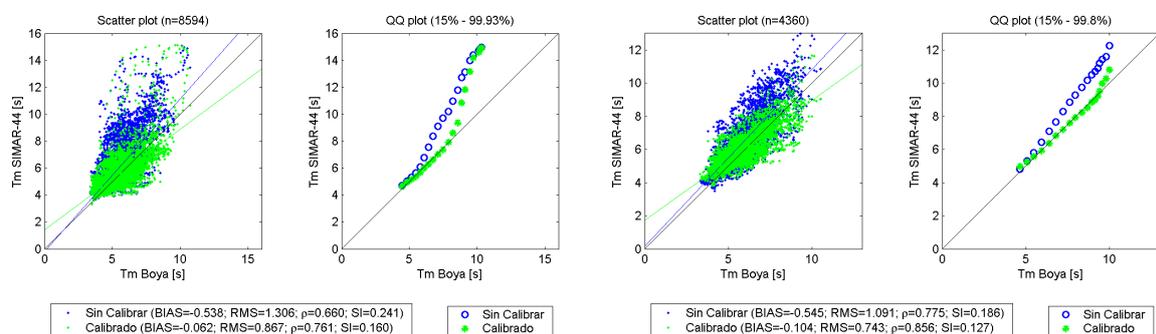


Figura 7.14. Calibración de T_m de SIMAR-44 con boyas, hasta el año 2000 (arriba) y validación con datos de boyas del año 2001 (abajo); las boyas son las de Gran Canaria (izquierda) y Villano-Sisargas (derecha). Aplicando doble filtro de media móvil de 9 horas.

De manera similar a la figura 7.13, en la figura 7.14 se comprueba que tanto en la calibración como en la validación de Villano-Sisargas se reduce la dispersión de los datos y se mejora tanto el régimen medio de T_m como el medio-alto, logrando unos muy buenos resultados con la calibración. En Gran Canaria, con la verificación de la calibración se obtienen unos buenos resultados, mejorando bastante los parámetros de ajuste, pero en la validación, a pesar de que los parámetros de ajuste son buenos, no se consigue corregir la rama medio-alta del régimen, dejando prácticamente sin calibrar estos datos; esto es debido (como en la figura anterior) a que existen pocos datos disponibles para calibrar y que dichos datos no contienen las correcciones necesarias para los datos más extremos de T_m del año 2001.

Se pone de manifiesto la importancia de validar las calibraciones realizadas, debiendo dejar siempre disponibles datos para validar, pues en cada caso y según las necesidades de calibración, pueden ser necesarios más o menos datos para corregir una determinada variable correctamente. Por ello, con la validación se puede determinar si una calibración es correcta o el grado de validez que tiene.

Finalmente, en la figura 7.15 se representan varias series temporales de H_s (arriba) y T_m (abajo) para las posiciones de la boya de Gran Canaria (izquierda) y Villano-Sisargas (derecha). Se muestran las series de las boyas y de SIMAR-44 sin calibrar y calibrado con la metodología escalar no paramétrica con datos de las boyas hasta el año 2000. Como las series presentadas son del año 2001, esta gráfica valida la calibración realizada, así como los intervalos de confianza del 95% mostrados, que han sido determinados sin la información de estos datos de las boyas del año 2001.

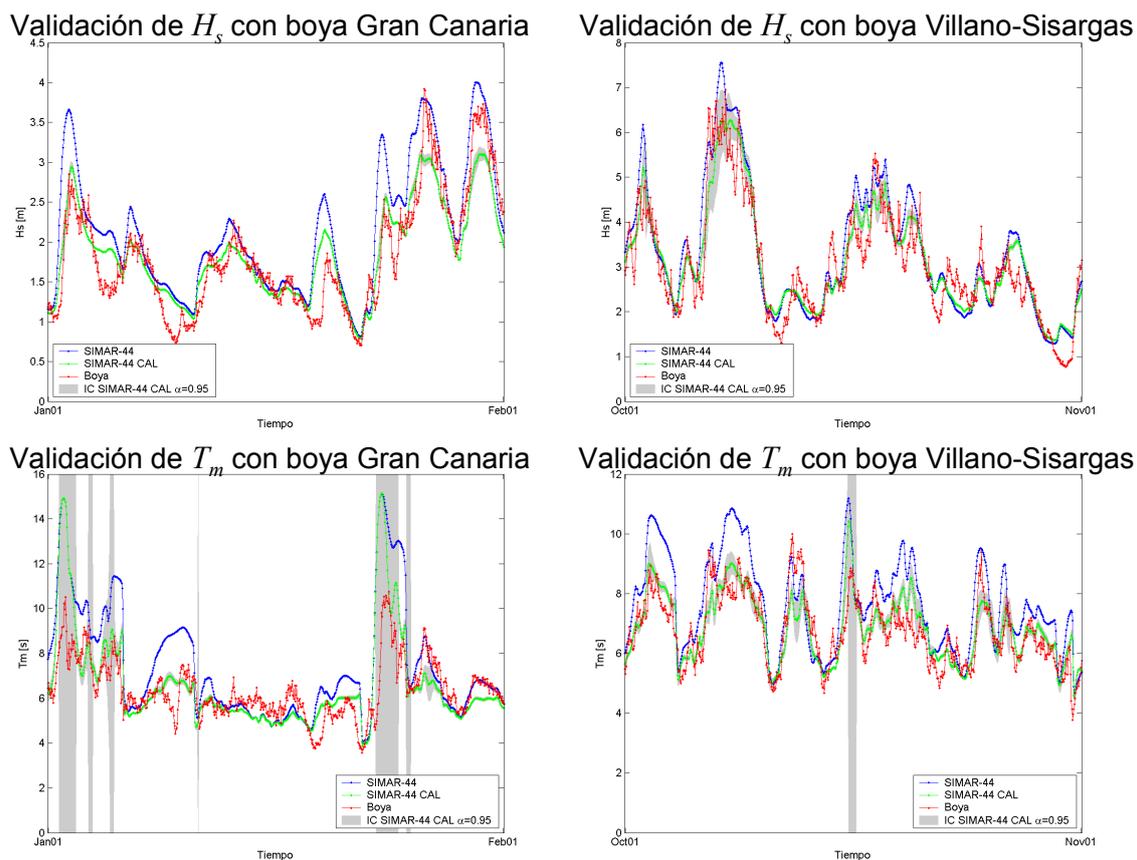


Figura 7.15. Series temporales de H_s (arriba) y de T_m (abajo) de SIMAR-44 sin calibrar y calibrado con la metodología escalar, utilizando las boyas de Gran Canaria (izquierda) y Villano Sisargas (derecha) hasta el año 2000, con el intervalo de confianza del 95% de los datos calibrados y la serie temporal de las boyas.

Puede comprobarse, en las validaciones mostradas de las series de H_s y T_m de la figura 7.15 y conjuntamente con las validaciones de H_s de la figura 7.13 y de T_m de la figura 7.14, que las series calibradas reproducen mejor lo registrado por las boyas que antes de la calibración. Por ejemplo, la H_s calibrada de SIMAR-44 en la posición de la boya de Villano-Sisargas se superpone, prácticamente en la totalidad de los datos, a los registros de la boya. En cambio, para H_s de Gran Canaria se subestiman algunos de los eventos más energéticos, habida cuenta que el caso de la calibración de Gran Canaria es el que peores resultados presenta. Los T_m de ambas boyas se corrigen correctamente, a pesar de que existe la tendencia de que los mayores T_m debieran disminuirse en mayor medida, ya que los datos calibrados se sitúan

entre los datos de SIMAR-44 sin calibrar y los de las boyas. También puede observarse que para los datos de T_m que requieren una calibración mayor, sus intervalos de confianza aumentan considerablemente, siendo esto un indicativo de que no existen suficientes datos análogos para corregir estos eventos.

7.4. Metodología direccional.

En la metodología escalar de calibración no paramétrica se calibra una variable (por ejemplo H_s) sin tener en cuenta las relaciones con otras variables (por ejemplo θ), por eso se ha denominado escalar. Pero a la hora de buscar subseries temporales análogas a una dada se puede comparar para cada instante de tiempo los valores de la propia variable u otras variables del oleaje. Así, por ejemplo para una subserie dada $[H_s(x,t), H_s(x,t-6), H_s(x,t-12)]$ se buscan secuencias $[H_s(x,t_i), H_s(x,t_i-6), H_s(x,t_i-12)]$ similares a la original comparando no sólo H_s a través del parámetro h_n sino también otras variables como θ mediante θ_n , debiendo verificarse simultáneamente $|H_s(x,t_i) - H_s(x,t)| < h_n$, $|\theta(x,t_i) - \theta(x,t)| < \theta_n$, $|H_s(x,t_i-6) - H_s(x,t-6)| < h_n$, $|\theta(x,t_i-6) - \theta(x,t-6)| < \theta_n$, $|H_s(x,t_i-12) - H_s(x,t-12)| < h_n$ y $|\theta(x,t_i-12) - \theta(x,t-12)| < \theta_n$, siendo en este caso $m=6$ (3 instantes de tiempo de cada subserie por 2 variables a comparar) las condiciones que deben cumplirse para considerar las subseries similares.

Al aumentar el número de condiciones o restricciones que se deben satisfacer para que dos subseries sean similares se reduce el número p de secuencias con datos instrumentales coincidentes que cumplen dichas restricciones, lo que implica que se reduzca el porcentaje de datos calibrados. Pero los datos que se consiguen calibrar con esta metodología tienen una gran calidad y fiabilidad, pues distintos oleajes de variadas características direccionales pueden requerir calibraciones diferentes. Para solucionar el problema de calibrar el mayor número de datos que se pueda con la mayor precisión posible, aprovechando la información direccional que se tiene con el modelado numérico, se define la calibración direccional en dos fases. En una primera fase se calibra escalarmente cada una de las variables del oleaje por separado (como se ha explicado en el apartado 7.3), corrigiendo de esta forma la práctica totalidad de los datos. Y a partir de estos datos calibrados, en una segunda fase, se recalibra imponiendo más restricciones a las distintas variables del oleaje, de forma que cuando haya información suficiente como para calibrar direccionalmente un dato se corrige, y sino, se mantiene la calibración escalar.

En la ecuación 7.3 se resume la calibración escalar de los tres parámetros seleccionados para caracterizar los estados de mar SIMAR-44 (H_s , T_m y θ) calibrándolos con boyas, o lo que es lo mismo, se presenta la primera fase de la calibración direccional. En dicha primera fase se

utilizan las parametrizaciones y filtros explicados en la metodología escalar, por ejemplo subseries de 6 horas o dobles filtros de media móvil de 5 horas para el Mediterráneo.

$$\left. \begin{array}{l} |H_s(x, t_i) - H_s(x, t)| < h_n \\ \vdots \\ |H_s(x, t_i - 6) - H_s(x, t - 6)| < h_n \\ |T_m(x, t_i) - T_m(x, t)| < t_n \\ \vdots \\ |T_m(x, t_i - 6) - T_m(x, t - 6)| < t_n \\ |\theta(x, t_i) - \theta(x, t)| < \theta_n \\ \vdots \\ |\theta(x, t_i - 6) - \theta(x, t - 6)| < \theta_n \end{array} \right\} \begin{array}{l} H_{s \text{ CAL } 1}(x, t) = H_s(x, t) + \frac{1}{p_h} \sum_{i=1}^{p_h} [H_{s \text{ BOYA}}(x, t_i) - H_s(x, t_i)] \\ T_{m \text{ CAL } 1}(x, t) = T_m(x, t) + \frac{1}{p_t} \sum_{i=1}^{p_t} [T_{m \text{ BOYA}}(x, t_i) - T_m(x, t_i)] \\ \theta_{\text{CAL } 1}(x, t) = \theta(x, t) + \frac{1}{p_\theta} \sum_{i=1}^{p_\theta} [\theta_{\text{BOYA}}(x, t_i) - \theta(x, t_i)] \end{array}$$

p_h secuencias coincidentes en H_s con datos de boya
 p_t secuencias coincidentes en T_m con datos de boya
 p_θ secuencias coincidentes en θ con datos de boya

(7.3)

Tras la primera calibración escalar de la fase 1 ($CAL 1$) de la calibración direccional, ver ecuación 7.3, se realiza la fase 2 ($CAL 2$) de la calibración direccional de H_s , que se describe en la ecuación 7.4:

$$\left. \begin{array}{l} |H_{s \text{ CAL } 1}(x, t_i) - H_{s \text{ CAL } 1}(x, t)| < h_n \\ |T_{m \text{ CAL } 1}(x, t_i) - T_{m \text{ CAL } 1}(x, t)| < t_n \\ |\theta_{\text{CAL } 1}(x, t_i) - \theta_{\text{CAL } 1}(x, t)| < \theta_n \\ \vdots \\ |H_{s \text{ CAL } 1}(x, t_i - 6) - H_{s \text{ CAL } 1}(x, t - 6)| < h_n \\ |T_{m \text{ CAL } 1}(x, t_i - 6) - T_{m \text{ CAL } 1}(x, t - 6)| < t_n \\ |\theta_{\text{CAL } 1}(x, t_i - 6) - \theta_{\text{CAL } 1}(x, t - 6)| < \theta_n \end{array} \right\} \begin{array}{l} H_{s \text{ CAL } 2}(x, t) = H_{s \text{ CAL } 1}(x, t) + \dots \\ \dots + \frac{1}{p_{ht\theta}} \sum_{i=1}^{p_{ht\theta}} [H_{s \text{ BOYA}}(x, t_i) - H_{s \text{ CAL } 1}(x, t_i)] \end{array}$$

$p_{ht\theta}$ secuencias coincidentes en H_s , T_m y θ con datos de boya

(7.4)

La segunda fase es propiamente la calibración direccional, pues en los criterios de búsqueda de subseries temporales se introduce la comparación con θ . De hecho se buscan coincidencias con las tres variables ya calibradas en la fase 1 ($H_{s \text{ CAL } 1}$, $T_{m \text{ CAL } 1}$ y $\theta_{\text{CAL } 1}$). También se ha estudiado comparar únicamente la variable a calibrar junto con su dirección, pero los mejores resultados se han obtenido introduciendo siempre las tres variables, sea cual sea la que se finalmente calibre, pues se respetan las correlaciones entre H_s y T_m .

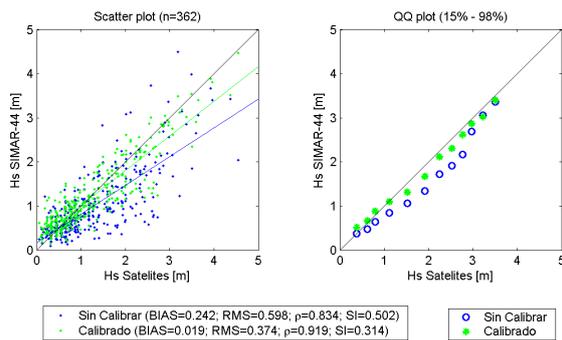
Todos los ajustes de los parámetros, variaciones de la metodología original, introducción de filtros desarrollados para la metodología escalar se mantienen para la direccional. Por ejemplo, tras la calibración de cada fase se pasa un doble filtro media móvil de 5 horas para el Mediterráneo y de 9 horas para el Atlántico; se buscan subseries de 6 horas o de 12 horas según se esté en el Mediterráneo o en el Atlántico, lo que implica que en la fase 1 sean respectivamente $m=7$ y $m=5$, en cambio en la fase 2 (con tres variables) son respectivamente $m=21$ (7x3) y $m=15$ (5x3) las condiciones a verificar por cada subserie; se parametrizan las máximas diferencias permitidas de cada variable a partir de $h_{700} = 0.5m$, $t_{700} = 0.5s$ y $\theta_{700} = 22.5^\circ$ con $\alpha = 0.2$; y también se conserva, para la calibración con satélites, $R = 0.5^\circ$.

A continuación, al igual que se ha realizado para la metodología de calibración escalar no paramétrica, se va a aplicar la calibración direccional con datos de satélites y con datos de boyas.

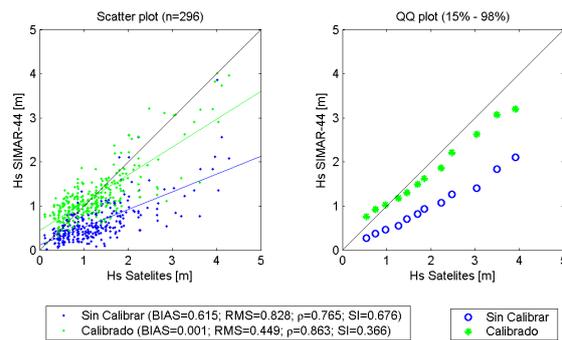
7.4.1. Calibración con datos de satélites.

Como ya se ha justificado anteriormente con datos de satélites sólo se calibra H_s , por lo que en la aplicación de la calibración direccional con dichos datos, en la primera fase, sólo se corrige H_s , dejando T_m y θ sin calibrar. En la figura 7.16 se presentan los resultados de la calibración direccional con satélites y validando con boyas para Cabo Begur y Cabo de Gata aplicando el doble filtro de media móvil de 5 horas. Comparando dichos resultados con los de la figura 7.10, que son el resultado de la calibración escalar, se puede comprobar que mejoran notablemente los resultados de los parámetros cuantitativos de ajuste con la metodología direccional, fundamentalmente en la verificación de la calibración; pero también se mejora significativamente al validar con los datos de las boyas.

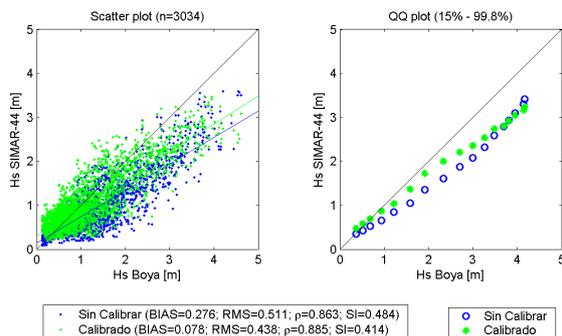
Calibración con satélites en Cabo Begur



Calibración con satélites en Cabo de Gata



Validación con boya de Cabo Begur



Validación con boya de Cabo de Gata

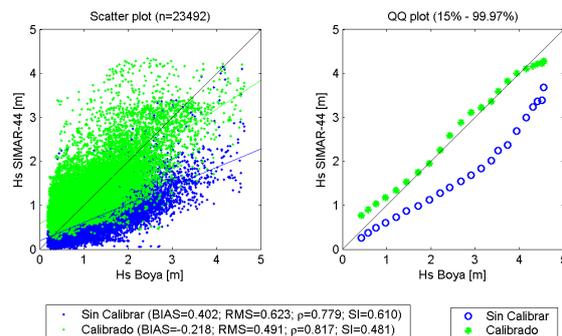


Figura 7.16. Calibración direccional de H_s de SIMAR-44 con satélites (arriba) y validación con boyas (abajo) en las posiciones de las boyas de Cabo Begur (izquierda) y Cabo de Gata (derecha) con $h_{700} = 0.5m$ y $R = 0.5^\circ$.

7.4.2. Calibración con datos de boyas.

Al igual que con la aplicación de la metodología direccional de calibración no paramétrica con datos de satélites, que se han calibrado de nuevo los casos que se calibraron inicialmente con la metodología escalar, para comparar y verificar las mejoras que se obtienen; con la aplicación de la metodología de calibración con datos de boyas también se van a volver a calibrar los casos inicialmente calibrados con la metodología escalar, en sus mismos términos pero ahora direccionalmente, para poder comparar los resultados obtenidos con ambas metodologías.

Así por ejemplo, en la figura 7.17 se calibra H_s en las posiciones de las boyas de Cabo Begur y Cabo de Gata (Mediterráneo) con la metodología direccional no paramétrica descrita anteriormente, calibrando con datos de dichas boyas y validando con información de satélites. Se puede comprobar que es posible calibrar hasta probabilidades de no excedencia mayores que en los casos de calibración con satélites y que se logran mejores ajustes calibrando con los datos de boyas (ver figura 7.16). Comparando con los resultados de la calibración escalar de la figura 7.11 se puede constatar que también se iguala o mejora con la calibración

direccional, tanto en la verificación (con notorias mejoras) como en la validación con datos de satélites.

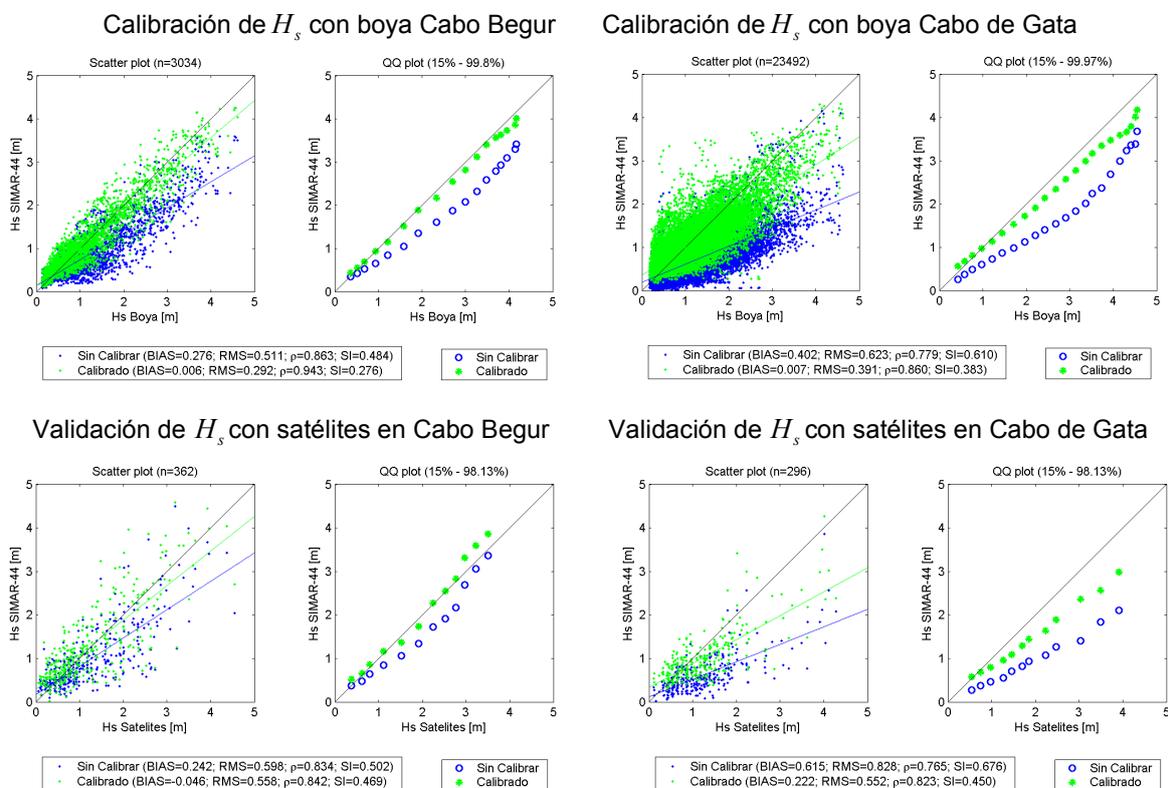


Figura 7.17. Calibración direccional de H_s de SIMAR-44 con boyas (arriba) y validación con satélites (abajo) en las posiciones de las boyas de Cabo Begur (izquierda) y Cabo de Gata (derecha).

Cuando no se utilizan datos de boyas para calibrar, como ya se ha presentado, también se pueden calibrar otros parámetros de SIMAR-44 (además de H_s) como T_m o θ , pues se dispone de esa información proveniente de boyas. En la figura 7.18 se representa la calibración direccional de T_m para Cabo Begur y Cabo de Gata y de θ únicamente para Cabo Begur. Comparando los resultados obtenidos con los de la figura 7.12 de la calibración escalar se puede verificar que aplicando la metodología direccional a variables distintas de H_s también se logran ajustes y calibraciones mucho mejores que con la escalar. Pese a que la representación realizada para verificar la calibración de θ no es cuantitativamente correcta, pues al ser una variable cíclica no tiene sentido ni los ajustes ni la representación *QQ-plot*, pero indica cualitativamente que se mejora con la calibración direccional.

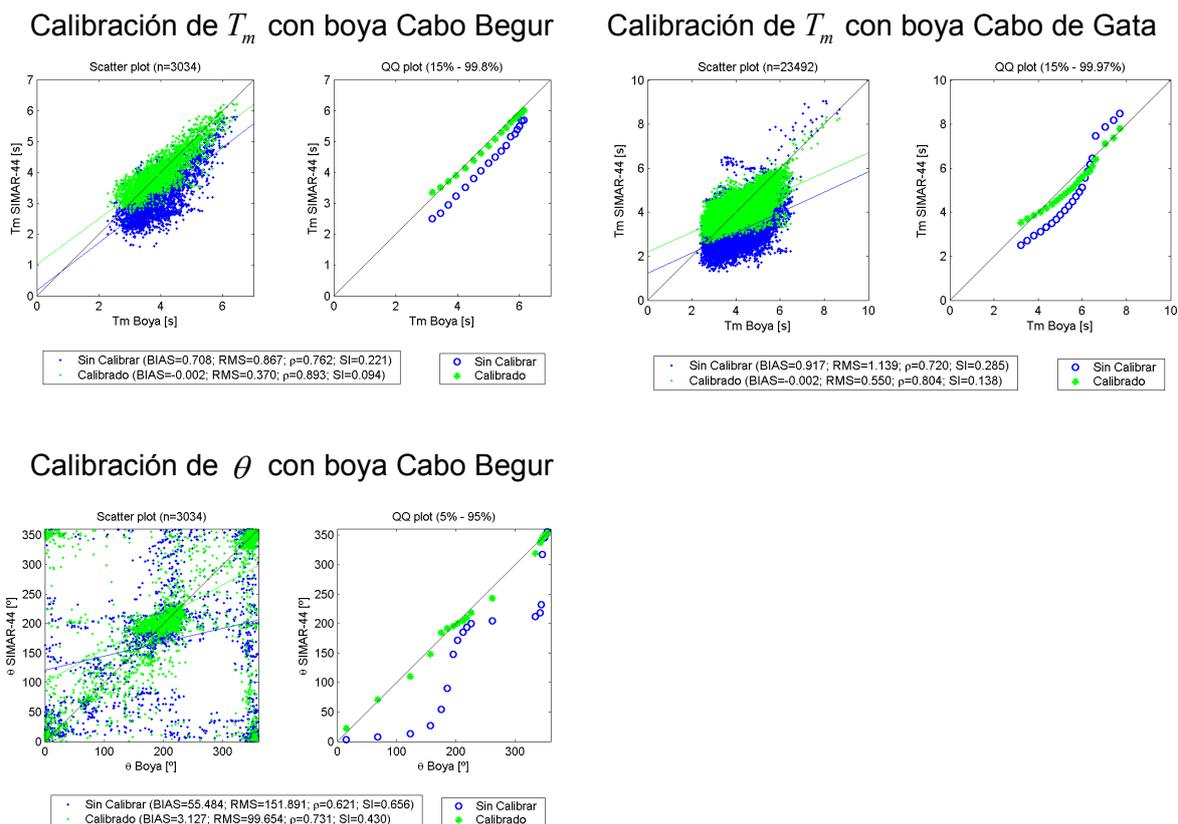
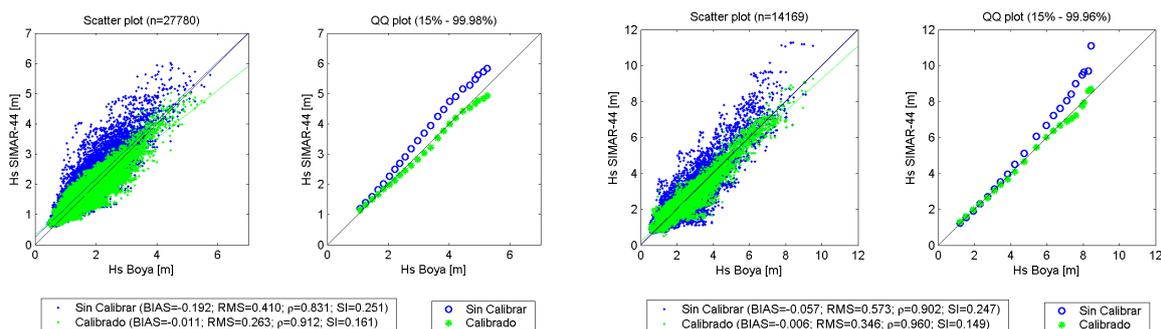


Figura 7.18. Calibración direccional de T_m (arriba) y θ (abajo) de SIMAR-44 con las boyas de Cabo Begur (izquierda) y Cabo de Gata (derecha).

Al igual que para la metodología escalar, también se va a aplicar la metodología direccional a dos boyas del Atlántico, Gran Canaria y Villano-Sisargas, sin utilizar información de satélites para validar, por ello se calibra con datos hasta el año 2000 inclusive y se valida con datos del año 2001 (aproximadamente el 25% de la información total disponible). Así en las figuras 7.19 y 7.20 se presentan respectivamente la calibración de H_s y T_m de SIMAR-44 en ambas boyas, mostrando las validaciones para ambos parámetros con los datos del año 2001.

Calibración de H_s con boya Gran Canaria

Calibración de H_s con boya Villano-Sisargas



Validación de H_s con boya Gran Canaria

Validación de H_s con boya Villano-Sisargas

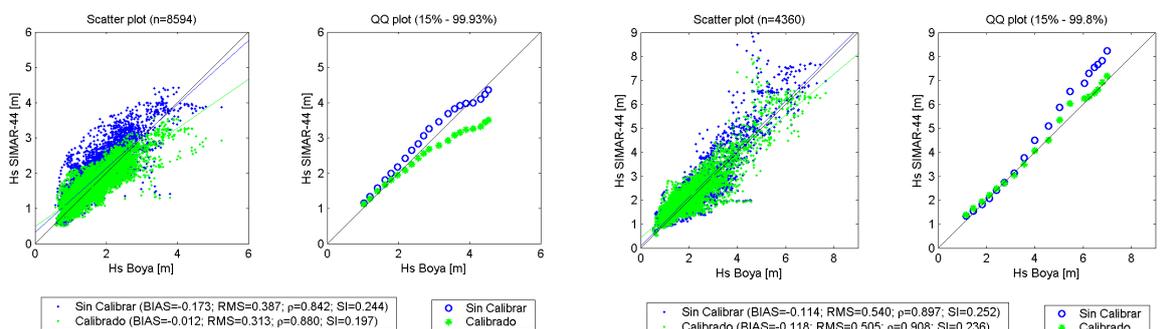
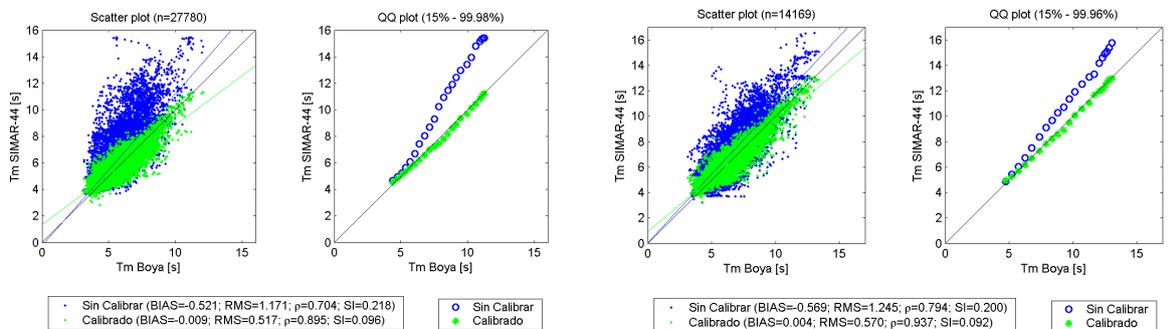


Figura 7.19. Calibración direccional de H_s de SIMAR-44 con boyas, hasta el año 2000 (arriba) y validación con datos de boyas del año 2001 (abajo); las boyas son las de Gran Canaria (izquierda) y Villano-Sisargas (derecha).

Calibración de T_m con boya Gran Canaria

Calibración de T_m con boya Villano-Sisargas



Validación de T_m con boya Gran Canaria

Validación de T_m con boya Villano-Sisargas

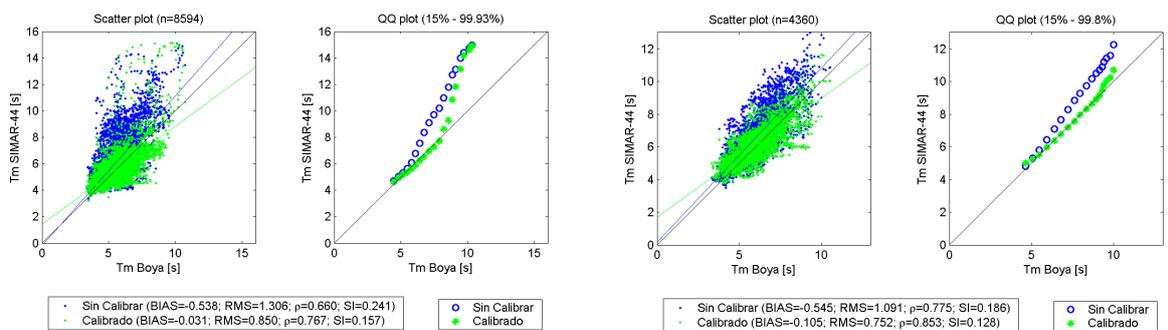


Figura 7.20. Calibración direccional de T_m de SIMAR-44 con boyas, hasta el año 2000 (arriba) y validación con datos de boyas del año 2001 (abajo); las boyas son las de Gran Canaria (izquierda) y Villano-Sisargas (derecha).

De manera similar a lo mostrado en las figuras 7.13 y 7.14 para la calibración escalar de H_s y T_m , con la calibración direccional de H_s y T_m (figuras 7.19 y 7.20) se logran unos excelentes resultados en la calibración de Villano-Sisargas (mejores que los logrados con la calibración escalar) y en la validación (similares a los de la calibración escalar); en cambio, a pesar de que en la verificación de las calibraciones de Gran Canaria se obtienen unos buenos resultados (mejorando los de las calibraciones escalares) en la validación se evidencian los mismos problemas (aunque atenuados) que presentaba los resultados de la calibración escalar, se observa que la rama medio-alta del régimen de H_s se subestima con la calibración y la rama medio-alta del régimen de T_m no se modifica, dejando prácticamente sin calibrar estos datos. Esto es debido a que existen pocos datos disponibles para calibrar la posición de Gran Canaria y que dichos datos no contienen las correcciones necesarias para los datos más extremos del año 2001.

Estas mismas conclusiones se pueden obtener en base a la representación de las series temporales de H_s (arriba) y T_m (abajo) para las posiciones de la boya de Gran Canaria (izquierda) y Villano-Sisargas (derecha) de la figura 7.21. En ella se muestran las series de las boyas y de SIMAR-44 sin calibrar y calibrado con la metodología direccional no paramétrica con datos de las boyas hasta el año 2000 y se representa en gris el intervalo de confianza³ del 95%. Dicha figura valida los resultados de la calibración, pues las series presentadas son del año 2001 y mejora los resultados respecto a los de la metodología escalar (ver figura 7.15). Cabe señalar que los intervalos de confianza aumentan en relación a los mostrados para la metodología escalar, pues se acumula la incertidumbre en la calibración de las dos fases que constituyen la metodología direccional. De hecho la segunda fase, al tener mayores restricciones para encontrar datos análogos, suele tener muestras de datos poco numerosas, lo que hace que aumente significativamente la incertidumbre de la calibración; a pesar de ello los resultados obtenidos con la calibración direccional son de una gran calidad.

³ Intervalo de confianza del 95%: En el cálculo del intervalo de confianza de los datos calibrados cuando la calibración se hace en dos fases sucesivas, a partir de la suma de los residuos de cada una, se asume que las varianzas estimadas de los residuos son iguales. Definiéndose dichos residuos como $X = H_{s\ BOYA} - H_s$ e $Y = H_{s\ BOYA} - (H_s + X)$, siendo el número de muestras de cada uno respectivamente p_X y p_Y y sus cuasidesviaciones típicas muestrales, s_X y s_Y respectivamente. Así el intervalo de confianza del 95% se expresa mediante $H_s \pm t_{S^{-1}(1-\frac{1-0.95}{2}, p_X + p_Y - 2)} s \sqrt{1/p_X + 1/p_Y}$, tomada de Luceño (1989), definiéndose s como $s = \sqrt{(s_X^2(p_X - 1) + s_Y^2(p_Y - 1)) / (p_X + p_Y - 2)}$.

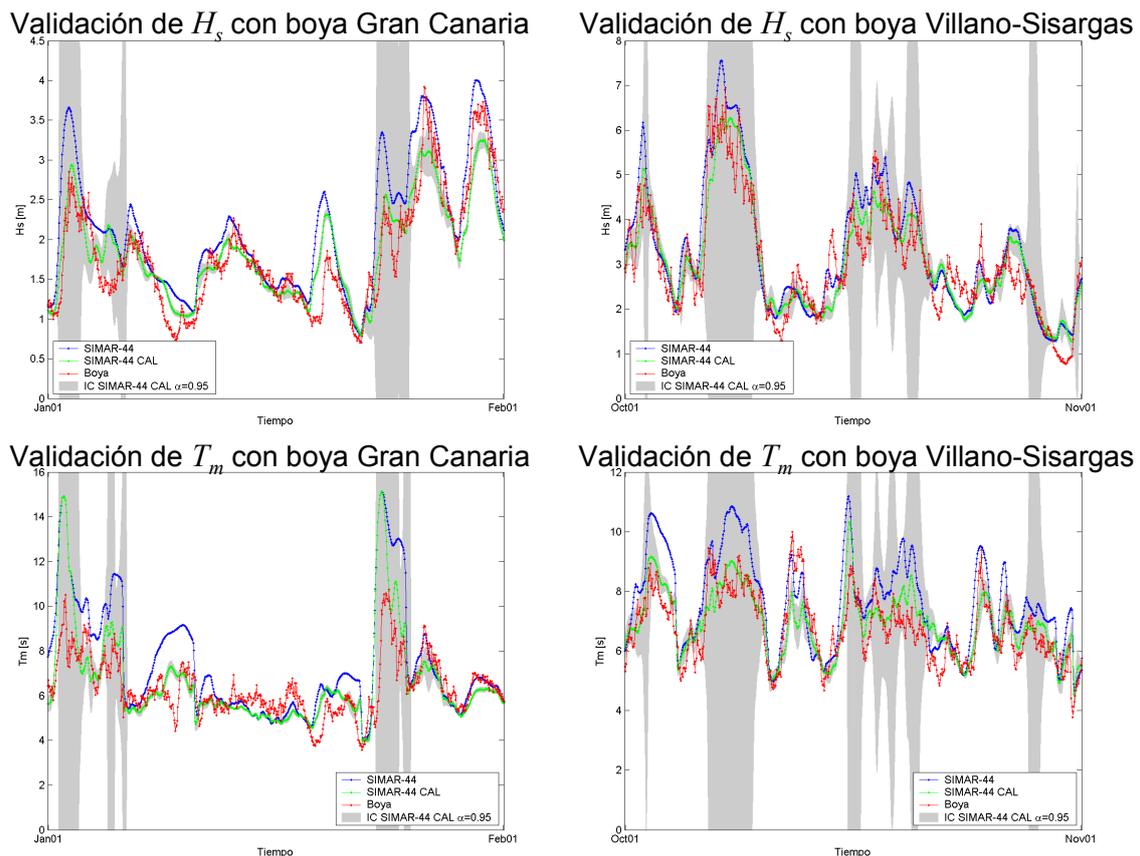


Figura 7.21. Series temporales de H_s (arriba) y de T_m (abajo) de SIMAR-44 sin calibrar y calibrado con la metodología direccional, utilizando las boyas de Gran Canaria (izquierda) y Villano Sisargas (derecha) hasta el año 2000, con el intervalo de confianza del 95% de los datos calibrados y la serie temporal de las boyas.

A pesar de que la cantidad de datos disponibles para calibrar las dos posiciones propuestas de Gran Canaria y Villano-Sisargas son similares, para Villano-Sisargas son suficientes pues el reanálisis en esta zona simula con mucha precisión lo registrado por la boya, en cambio para Gran Canaria no son suficientes para corregir correctamente el régimen medio-alto de oleaje. Se pone de manifiesto la importancia de validar las calibraciones realizadas, debiendo dejar siempre disponibles datos para validar, pues en cada caso y según las necesidades de calibración, pueden ser necesarios más o menos datos para corregir una determinada variable correctamente.

7.5. Influencia de la información disponible en la calibración.

Para estudiar la influencia de la cantidad de información disponible para calibrar en los resultados y validación de los mismos, así como poder analizar los valores o probabilidades hasta los que se mantiene correcta y fiable una determinada calibración, en este apartado se van a realizar diferentes pruebas de calibración direccional en la ubicación con mayor

cantidad de información instrumental disponible, la boya de la red exterior de OPPE de Mahón, pues la serie de datos de dicha boya comienza en el año 1993 y la disposición de las trazas de los satélites permite tener 670 datos de H_s de satélites hasta finales del 2001 a menos de 0.5° de distancia.

Como en la zona del Mediterráneo próxima a Mahón los datos del reanálisis SIMAR-44 se aproximan bastante a los registrados por la boya, las conclusiones a las que se lleguen en este apartado serán tomadas como mínimos que deberán superarse para las zonas que requieran mayores distorsiones de la base de datos original.

Se van a realizar cuatro calibraciones diferentes a partir de los datos de la boya de Mahón, siempre utilizando la metodología no paramétrica de calibración direccional. Las diferencias consisten en la cantidad de datos utilizados para calibrar, se va a calibrar utilizando respectivamente los datos de los años 1993-1994, 1993-1996, 1993-1999 y 1993-2001, que suponen aproximadamente el 25%, 50%, 75% y 100% de los datos de la boya. Se validarán los resultados de dos formas. La primera consiste en comparar los resultados con los datos de la boya de los años 2000 y 2001 (aproximadamente en 25% del total) cuando no hayan sido utilizados para calibrar; esto se muestra en la tabla 7.1, que resume los parámetros cuantitativos medios del diagrama de dispersión entre datos de SIMAR-44 y boya, usando los datos de SIMAR-44 sin calibrar y calibrados.

	Sin Calibrar				Calibrado				n
	BIAS	RMS	ρ	SI	BIAS	RMS	ρ	SI	
Calibración con los datos de la boya de Mahón de los años 1993-1994	0.286	0.531	0.864	0.404	0.028	0.452	0.887	0.343	2476
Calibración con los datos de la boya de Mahón de los años 1993-1996	0.286	0.531	0.864	0.404	0.027	0.437	0.893	0.332	2476
Calibración con los datos de la boya de Mahón de los años 1993-1999	0.286	0.531	0.864	0.404	0.028	0.436	0.897	0.324	2476

Tabla 7.1. Validación de las calibraciones direccionales no paramétricas de H_s de SIMAR-44 con datos de la boya de Mahón, se validan las calibraciones con los datos de la boya de Mahón de los años 2000 y 2001.

La segunda forma de validar se realiza comparando las calibraciones con los datos coincidentes de satélites, ver la tabla 7.2. En ella se resume los parámetros cuantitativos medios del diagrama de dispersión entre datos de SIMAR-44 y satélites, usando los datos de SIMAR-44 sin calibrar y calibrados.

	Sin Calibrar				Calibrado				n
	BIAS	RMS	ρ	SI	BIAS	RMS	ρ	SI	
Calibración con los datos de la boya de Mahón de los años 1993-1994	0.151	0.575	0.863	0.461	-0.103	0.570	0.864	0.457	670
Calibración con los datos de la boya de Mahón de los años 1993-1996	0.151	0.575	0.863	0.461	-0.116	0.571	0.864	0.458	670
Calibración con los datos de la boya de Mahón de los años 1993-1999	0.151	0.575	0.863	0.461	-0.114	0.565	0.866	0.453	670
Calibración con los datos de la boya de Mahón de los años 1993-2001	0.151	0.575	0.863	0.461	-0.111	0.564	0.866	0.452	670

Tabla 7.2. Validación de las calibraciones direccionales no paramétricas de H_s de SIMAR-44 con datos de la boya de Mahón, se validan las calibraciones con los datos de satélites con $R = 0.5^\circ$.

Se puede comprobar, a la vista de las validaciones presentadas en las tablas 7.1 y 7.2, que siempre se mejora con la calibración, sea cual sea el porcentaje de datos utilizados y que a mayor cantidad de datos, mejores resultados se obtienen. Las validaciones con datos de satélites muestran mejoras no muy acusadas (debido a que los datos originales son bastante correctos) aunque con la utilización de más de cuatro años de datos (1993-1996) las mejoras son más apreciables. Con las validaciones con los datos de la boya de los años 2000 y 2001 se evidencia mucho más las mejoras logradas con el aumento de los datos para calibrar, fundamentalmente a partir de los dos años de datos (1993-1994).

Por lo tanto, se concluye que a partir de 3-4 años de datos de boyas se puede calibrar correctamente las zonas de SIMAR-44 que no precisan de una corrección o calibración muy fuerte, como por ejemplo Villano-Sisargas o Mahón, en cambio otras zonas necesitan más información para calibrar correctamente los eventos más extremos, por ejemplo Gran Canaria.

Una vez validado y verificado, con distintos porcentajes de datos de la boya, que los resultados son cada vez mejores aunque muy similares, en este caso parece más apropiado valerse de la mayor cantidad de datos para calibrar. Por ello, se calibra utilizando todos los datos de la boya para calibrar, dejando únicamente los datos de satélites para validar. En la figura 7.22 se presenta la verificación de esta calibración (izquierda) y la validación con satélites (derecha) comprobándose la mejora lograda con la calibración, en el ajuste de los datos y disminución de su dispersión, así como en la similitud de los regímenes de oleaje. También puede compararse con los resultados obtenidos para este mismo caso de calibración por la mejor técnica de calibración puntual no paramétrica direccional (QQ ponderada y $MVRE_{Ext}$, ver figura 5.22 del capítulo 5), siendo los resultados de la metodología no

paramétrica notablemente mejores que los de la paramétrica, tanto en la verificación como en la validación con datos de satélites.

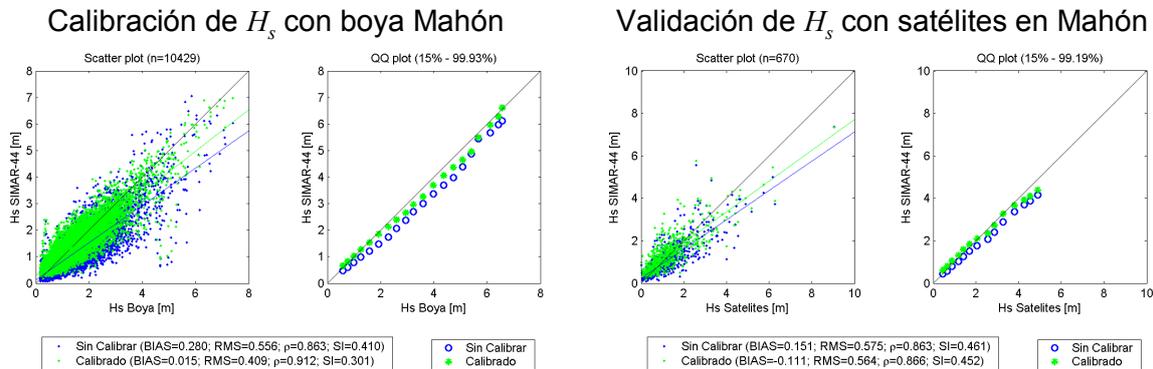


Figura 7.22. Calibración direccional no paramétrica de H_s de SIMAR-44 con todos los datos de la boya de Mahón, verificación de la calibración con todos los datos de la boya (izquierda) y validación con los datos coincidentes de satélites con $R = 0.5^\circ$.

En la figura 7.23 se representa la serie temporal de los mismos dos meses que los presentados en la calibración paramétrica direccional de la figura 5.23, obteniendo resultados considerablemente mejores con la metodología no paramétrica. Se verifica como con la calibración unas veces se aumenta el valor de H_s y otras se disminuye (incluso con valores originales sin calibrar muy similares) para aproximarse a lo registrado por la boya; también se puede observar que a pesar de que para los mayores valores de H_s el intervalo de confianza aumenta considerablemente, los valores calibrados se ajustan muy bien a los datos de la boya.

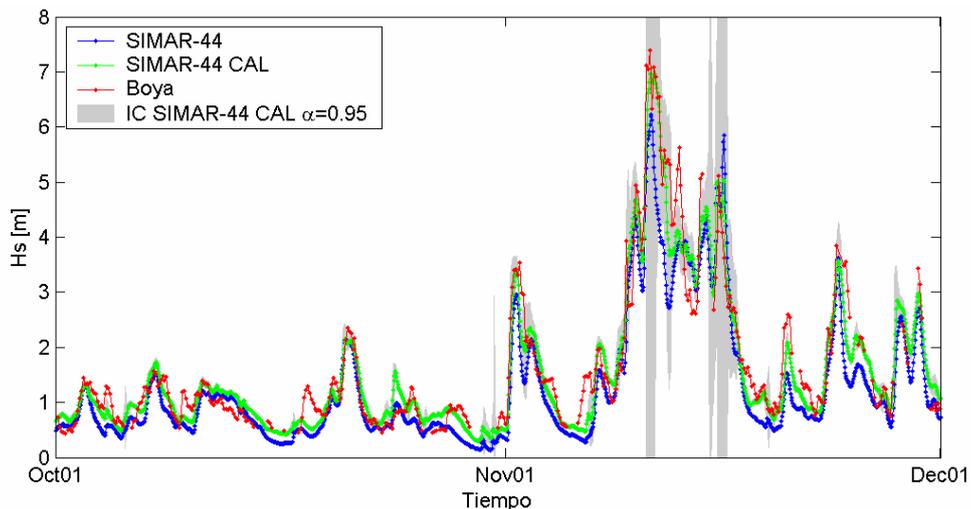


Figura 7.23. Serie temporal de H_s de SIMAR-44 sin calibrar y calibrado con la metodología direccional no paramétrica, utilizando todos los datos de la boya de Mahón, con el intervalo de confianza del 95% de los datos calibrados y la serie temporal de las boya.

7.6. Resumen de la calibración puntual no paramétrica.

A continuación se van a resumir esquemáticamente la metodología de calibración puntual no paramétrica presentada en este capítulo 7:

- Inicialmente, tras la introducción de este capítulo, en el apartado 7.2 se explica la metodología de calibración original de Caires y Ferreira (2005), que es una técnica de calibración escalar que ha sido aplicada para $R = 10^\circ$ y sin filtros, obteniendo buenos resultados para zonas de mar abierto del Atlántico Norte cuando se utilizan únicamente datos de H_s de la misión del satélite TOPEX/POSEIDON.
- En el apartado 7.3 se ha explicado la metodología de calibración escalar, basada en la original de Caires y Ferreira (2005), pero introduciendo una serie de variaciones para implementarla en ámbito costero español.
- En el apartado 7.3.1 se ha aplicado la metodología de calibración puntual escalar no paramétrica con datos de H_s de satélites (GEOSAT, TOPEX/POSEIDON, GFO, JASON-1 y ENVISAT), validando con datos de boyas. Se ha definido $R = 0.5^\circ$, $h_{700} = 0.5m$ y dobles filtros de media móvil, aplicados a los residuos, de 5 horas para el Mediterráneo y de 9 horas para el Atlántico.
- En el apartado 7.3.2 se ha aplicado la metodología de calibración puntual escalar no paramétrica, ahora con datos de H_s , T_m y θ de boyas, validando con datos de satélites o reservando un porcentaje de datos de boyas del orden del 25% para validar. Se ha definido $t_{700} = 0.5s$ y $\theta_{700} = 22.5^\circ$ respectivamente para calibrar T_m y θ .
- En el apartado 7.4 se ha explicado la metodología de calibración direccional, que se puede descomponer en dos fases. En la primera se calibra escalarmente y de forma independiente, cada una de las tres variables H_s , T_m y θ que se utilizan en esta técnica de calibración. En la segunda fase (propriadamente la calibración direccional) se buscan subseries o datos análogos imponiendo restricciones en las tres variables que definen el estado de mar.
- La metodología de calibración puntual direccional no paramétrica se ha aplicado con datos de H_s de satélites, validando con datos de boyas (apartado 7.4.1). También se ha aplicado con datos de H_s , T_m y θ de boyas, validando con datos de satélites o reservando un porcentaje de datos de boyas del orden del 25% para validar (apartado 7.4.2).

- Finalmente, en el apartado 7.5, se ha estudiado la influencia de la información disponible en los resultados de la calibración puntual direccional no paramétrica. Utilizando distintos porcentajes de datos de la boya de Mahón (25%, 50%, 75% y 100%).

7.7. Conclusiones.

A continuación se resume una serie de conclusiones a las que se ha llegado en el presente capítulo 7:

- La calibración puntual no paramétrica es la metodología con la que se obtienen los mejores resultados de calibración del régimen medio y medio-alto, calibrando cada dato con la corrección media de los datos instrumentales análogos. Por lo que los datos más energéticos, si no tienen datos instrumentales análogos, no se corrigen, en cuyo caso el régimen extremal no se calibra. Con esta metodología de calibración se puede definir el intervalo de confianza de cada dato calibrado y no es necesario que las series de oleaje estén perfectamente en fase, aunque si es necesario que exista una suficiente cantidad de datos coincidentes.
- A partir del método de calibración no paramétrico escalar original de Caires y Ferreira (2005) se ha desarrollado la metodología de calibración puntual escalar, implementándola para el ámbito costero español; desarrollando a partir de ella una nueva metodología de calibración puntual direccional no paramétrica, que permite calibrar el oleaje manteniendo la distribución conjunta direccional (H_s , T_m y θ).
- El mayor problema que presentan los métodos de calibración puntual no paramétricos es que si la información instrumental no es muy homogénea y presenta saltos o quiebros en la serie temporal, pueden ser introducidos en la serie temporal calibrada. Este problema ha sido resuelto, con la introducción de filtros de media móvil en el residuo que debe sumarse a la serie original para calibrarla.
- La metodología de calibración puntual no paramétrica (escalar y direccional) se ha aplicado exhaustivamente a los datos H_s , T_m y θ de SIMAR-44 del Atlántico y Mediterráneo para calibrarlos con datos de boyas y/o satélites, validando los resultados de distintas formas. También se ha estudiado la influencia en los resultados de la cantidad de datos utilizados para calibrar, definiendo unos criterios o recomendaciones para utilizar $R = 0.5^\circ$ (distancia entre cada dato de satélite y la posición SIMAR-44 a calibrar) para satélites y un mínimo de 4 años de datos para

CAPÍTULO 7

boyas en zonas donde no deba corregirse considerablemente los datos de reanálisis, aunque siempre se debe verificar estos supuestos en función de los requisitos concretos de cada calibración.

