

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

SECCIÓN II

APORTACIONES DE ESTA TESIS

CAPÍTULO 4

CLASIFICACIÓN DE LAS METODOLOGÍAS DE CALIBRACIÓN

4.1. Introducción.

En este capítulo 4 se va a clasificar, de manera general, los métodos de calibración en función de dónde se tiene la información y de qué características tiene dicha información. Dicha clasificación está basada en los distintos métodos de calibración puntual y espacial con los que se organiza metodológicamente la calibración de datos de oleaje en cualquier supuesto. A continuación, en este primer apartado, se introducen una serie de conceptos generales que en su mayoría son el resultado del estudio realizado en la Sección I.

Así, generalmente cuando se menciona el término datos de oleaje, en realidad se suele referir a parámetros (estadísticos o espectrales) de oleaje, ver apartado 2.2 del capítulo 2. De entre ellos, el más utilizado para describir la severidad del oleaje de cada estado de mar es la altura de ola significativa (H_s o H_{m0}). Para caracterizar el periodo de oleaje se utiliza el periodo medio (T_m o T_{02}) y el periodo de pico (T_p), aunque se prefiere la caracterización del oleaje a partir del T_m , ya que es menos dependiente de la forma del espectro o de la manera en que ha sido generado. Por ello, en esta tesis, cuando se calibre una base de datos se corregirán los valores de los parámetros del oleaje H_s y T_m , teniendo en cuenta su información direccional siempre y cuando sea posible, mejorando las metodologías existentes actualmente, las cuales se centran solamente en el oleaje escalar

Dado el estado actual de las bases de datos de oleaje es posible siempre caracterizar direccionalmente el oleaje, como mínimo H_s , T_m y θ_m (en adelante la dirección media de procedencia del oleaje se denota por θ); y en cualquier ubicación del ámbito costero español, pues los datos de oleaje provenientes del modelado numérico se simulan con suficiente calidad y resolución espacial para todas las costas españolas. Así mismo, en general también se dispone de datos instrumentales más o menos próximos, escalares o direccionales. Por lo que se parte de la hipótesis, para el desarrollo de esta tesis, de que se calibrarán las bases de datos procedentes de modelado numérico (reanálisis) con información instrumental (fundamentalmente boyas y satélites). La metodología que se va a plantear es suficientemente general como para incorporar otras bases de datos como los visuales, ya que siempre se pueden emplear las técnicas de homogeneización basadas en regresiones (presentadas en el capítulo 3, ver apartado 3.2.3) para poder utilizar conjuntamente datos de varias fuentes de información con calidades distintas.

Para el caso particular de las bases de datos de oleaje en España, la base de datos de reanálisis SIMAR-44 caracteriza el clima marítimo en cada posición con series de estados de mar homogéneos de 44 años de datos, con una resolución espacial temporal de 1-3 horas y una

resolución espacial de entre 0.125° a 0.25° . Se puede comprobar en la figura 4.1, que comparando las series de H_s de SIMAR-44 con los registros de algunas boyas de la Red Exterior de OPPE, SIMAR-44 difiere ligeramente de lo registrado por las redes de medida. Se puede observar que los picos de las series (los temporales) de SIMAR-44 y las boyas coinciden, por lo que SIMAR-44 describe correctamente los instantes de los picos, aunque existen diferencias en la magnitud (OPPE, 2003). Se pone así de manifiesto la necesidad de calibrar la base de datos SIMAR-44 (proveniente del modelado numérico) con datos instrumentales, con el fin de mejorar aún más la calidad de los datos SIMAR-44 en España.

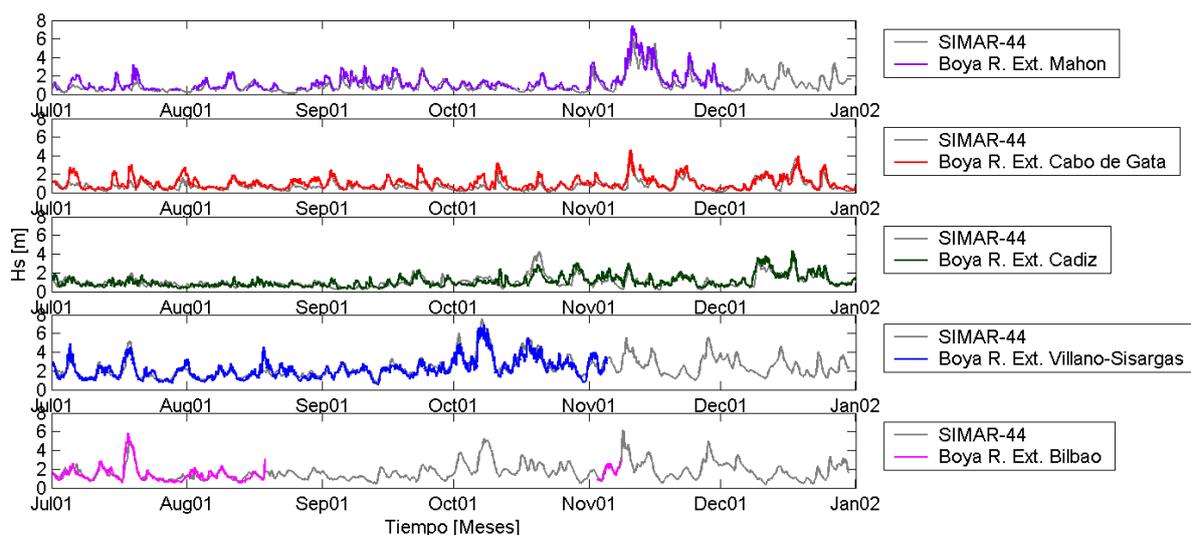


Figura 4.1. Comparación de seis meses de series de altura de ola significativa de SIMAR-44 con 5 boyas de la Red Exterior del OPPE, en la posición de estas boyas.

Históricamente, las metodologías de calibración se han basado en técnicas de regresión fundamentalmente lineales, pero existen otras posibles aproximaciones más sofisticadas (ver más detalles en Gutiérrez *et al.*, 2004) como redes neuronales y probabilísticas para la estimación de regresiones no lineales complejas, modelos autoregresivos y cadenas de Markov para relleno de huecos en series temporales o correlaciones canónicas¹ que son una generalización de las regresiones lineales múltiples para variables espaciales multidimensionales.

En la figura 4.2 se han clasificado las distintas técnicas de calibración, de manera general y atendiendo a las conclusiones y consideraciones del estado del conocimiento actual de la calibración de las bases de datos de oleaje (ver Capítulo 3); pudiendo cada una de estas técnicas estar definidas por sectores direccionales (calibración direccional) o agregando todos

¹ Correlación canónica: el análisis de correlación canónica es similar al de Componentes Principales (ver EOF en Anejo II) pero maximizando la correlación entre las variables en lugar de la varianza, como hacen los EOF (por lo que no se cumple la ortogonalidad entre sus componentes).

los datos (calibración escalar), aunque siempre que haya suficientes datos disponibles se debe utilizar la calibración direccional.

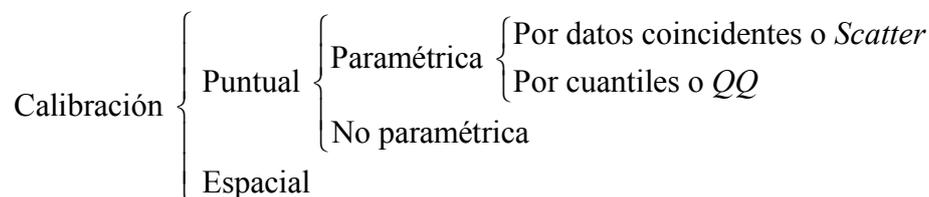


Figura 4.2. Clasificación general de las técnicas de calibración.

En los siguientes capítulos se van a ir describiendo las distintas técnicas de calibración desarrolladas, las puntuales paramétricas en los capítulos 5 y 6 (en el capítulo 6 será con retropropagación, que se explicará posteriormente), las no paramétricas en el capítulo 7, la espacial escalar en el capítulo 8 y la espacial direccional en el capítulo 9. Comparando los resultados obtenidos por cada una de ellas.

Previo a la selección de la mejor manera de calibrar unos datos es necesario conocer las características de la información que se va a utilizar, donde se encuentra ubicada, si sus regímenes son similares,... para poder discriminar, por ejemplo, si es necesario utilizar un método de calibración espacial o si es suficiente con los métodos de calibración puntuales. Por ello, seguidamente se va a hacer una clasificación de los datos de oleaje disponibles en cada caso, en función de su ubicación relativa al punto donde se necesita la información.

4.2. Dónde se tiene la información de oleaje.

Una vez explicadas las necesidades de información para el diseño de obras marítimas (ver apartado 1.2 y tabla 1.1 del capítulo 1), las distintas fuentes de información y bases de datos de oleaje de que se dispone en la actualidad (ver apartado 2.4 del capítulo 2) y la clasificación de las técnicas generales de calibración (ver figura 4.2); a continuación se va a hacer una clasificación de los datos de oleaje en función de su ubicación/profundidad y de su distancia relativa al punto donde se necesita la información. De esta manera se va a poder definir, en función de cada necesidad de información de oleaje y de los datos disponibles de oleaje, el método de calibración más adecuado.

Así, por ejemplo, si se tiene la información necesaria proveniente de datos instrumentales de oleaje justo en la ubicación de la obra (no se necesitan series muy largas ni continuas), no haría falta la calibración de los datos de reanálisis, pues ya se tiene directamente la

caracterización precisa del oleaje a partir de los datos instrumentales. Pero este supuesto pocas veces se presenta, ya que los datos suelen estar más o menos cerca de donde se necesita y casi nunca a la misma profundidad.

La profundidad relativa a la que se encuentran los datos disponibles tiene mucha importancia, puesto que no es lo mismo que los datos de oleaje estén o no afectados por el fondo, es decir que estén o no en profundidades indefinidas. Por eso es habitual utilizar, para caracterizar el clima marítimo de una zona, el oleaje que se define en profundidades indefinidas frente a la zona objetivo, que aún no está modificado por la batimetría local.

La otra cuestión interesante (además de si se está o no en profundidades indefinidas) a la hora de relacionar la ubicación de los datos disponibles con la posición donde realmente se necesitan, es su distancia relativa: si ambas posiciones no coinciden puede que haya que trasladar la información de un punto al otro. A la hora de determinar si están cerca o lejos relativamente, no se va a utilizar un criterio de distancia propiamente dicho, sino que se va a evaluar que sus climas marítimos sean similares. Para ello se comparan ambos climas marítimos en profundidades indefinidas, verificando si se puede definir el clima marítimo donde se necesita a partir de la información disponible.

La forma y criterios para comparar y determinar si dos climas marítimos son similares o distintos vendrán dados en gran medida por las necesidades de caracterización del oleaje, según sean más o menos importantes el régimen medio, distribución por direcciones, persistencias, régimen extremal, tendencias,... Para ello, puede resultar útil ajustar los distintos regímenes a las mismas funciones de distribución, determinar ambas distribuciones de sus flujos medios de energía o sus rosas de oleaje, calcular coeficientes de correlación vectoriales² entre ambas series de oleaje,...

A continuación, en la figura 4.3 se va a clasificar los datos en función de la profundidad a la que están y a la que se van a utilizar. Los datos disponibles se señalan con un punto gris y la ubicación dónde son necesarios con un punto rojo. Cuando los datos están en profundidades indefinidas se denotándose en blanco y en negro cuando no lo están. También se clasifica en función de un segundo criterio, si se parecen o no los climas marítimos en profundidades indefinidas de ambas posiciones (dónde se tiene y dónde se necesita la información de oleaje), denotándose con una flecha verde si son iguales, con una flecha negra si son parecidos y con una flecha roja si son distintos. Con ello se llega a 12 supuestos distintos ($2 \times 2 \times 3 = 12$).

² Coeficiente de Correlación Vectorial, VCC: Existen varios coeficientes que indican la correlación existente entre dos variables vectoriales (ver por ejemplo Crosby *et al.*, 1993).

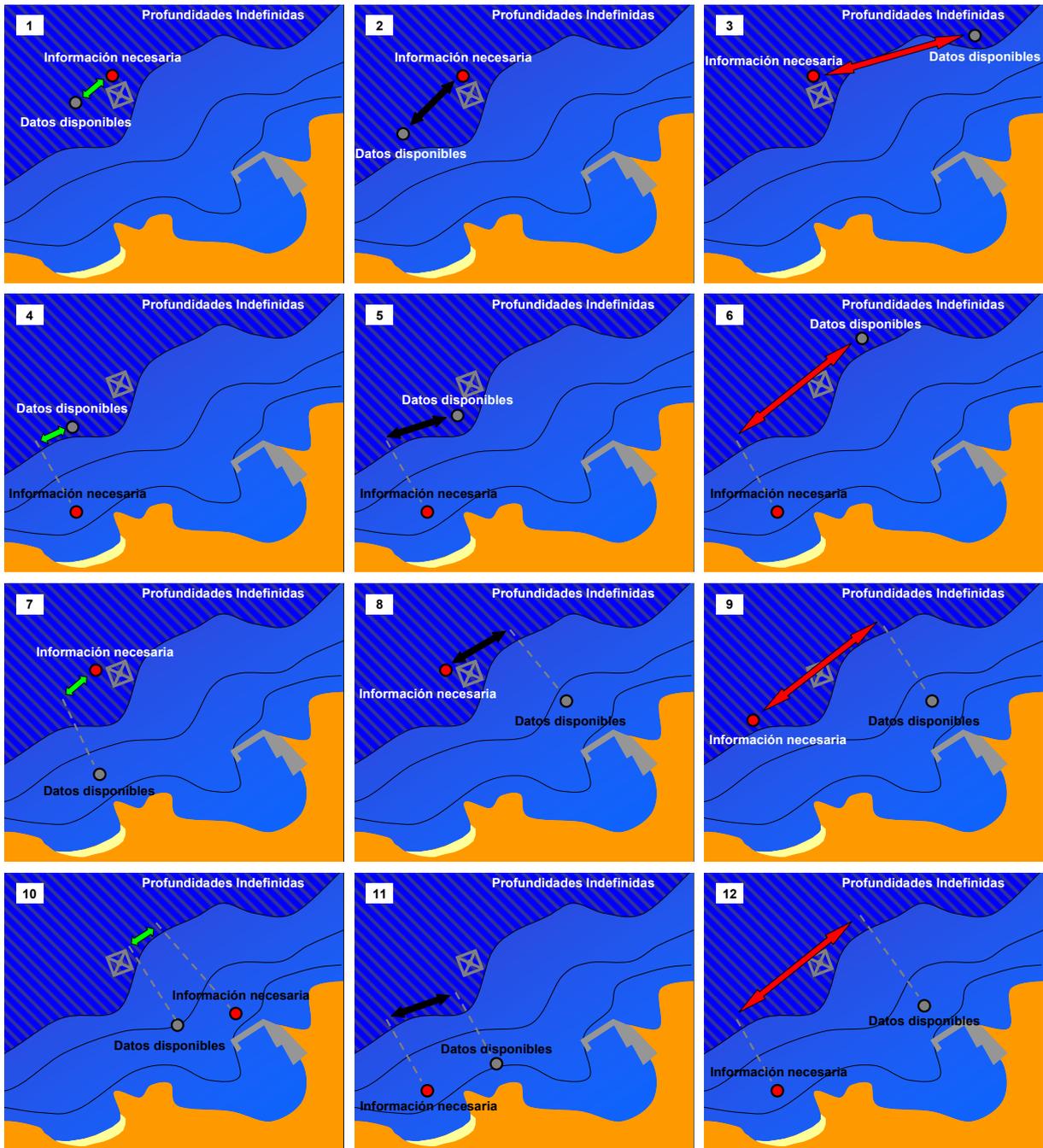


Figura 4.3. Representación de la clasificación de la información de oleaje disponible en función de su ubicación.

En la tabla 4.1 se presenta la misma clasificación de los 12 supuestos que ha sido representada en la figura 4.3. Si los datos disponibles o la información necesaria está en profundidades indefinidas (Prof. indef..) se denota por ✓ y en caso contrario por ✗. Si los climas marítimos en profundidades indefinidas de ambas posiciones (dónde se tiene y dónde se necesita la información de oleaje) son iguales, se denota por ≡; si son parecidos por ≈ y si son distintos por ≠.

	Datos disponibles	Información necesaria	Climas marítimos en Prof. indef.
	¿Prof. indef.?	¿Prof. indef.?	
1	✓	✓	≡
2	✓	✓	≈
3	✓	✓	≠
4	✓	✗	≡
5	✓	✗	≈
6	✓	✗	≠
7	✗	✓	≡
8	✗	✓	≈
9	✗	✓	≠
10	✗	✗	≡
11	✗	✗	≈
12	✗	✗	≠

Tabla 4.1. Clasificación de la información de oleaje disponible en función de su ubicación.

Así, por ejemplo, si se tiene información en profundidades indefinidas, se necesita en profundidades reducidas y el clima marítimo de profundidades indefinidas de ambas posiciones es el mismo (supuesto 4 de la tabla 4.1), simplemente habría que propagar los datos disponibles hasta la ubicación objetivo. En el apartado siguiente se va a presentar cuál es el tratamiento de la información (calibración, propagación,...) necesaria para cada uno de estos 12 supuestos.

4.3. Tipos generales de tratamiento de los datos.

Una vez clasificada la información del oleaje disponible en función de su ubicación relativa al lugar donde se necesita la información para el diseño de la obra marítima (los 12 supuestos de la tabla 4.1), se define cómo tratar la información para trasladarla hasta la zona objetivo.

En la tabla 4.2 se presenta de manera general el tratamiento que se debe realizar a los datos disponibles. En dicha tabla se presentan varias acciones a realizar, usar directamente, propagar, retropropagar, calibración puntual y calibración espacial, que a continuación se pasan a explicar:

- Usar directamente (UD): aunque no resulta muy usual, cuando se tiene información instrumental suficiente justo en la posición donde se necesita (supuestos 1 y 10), no hace falta tratar esos datos; se usan directamente.
- Propagar (PR): se denomina usualmente como propagación al proceso de trasladar el oleaje desde profundidades indefinidas hasta una posición en profundidades intermedias o reducidas (zona de interés), de manera que se resuelven los procesos que

afectan al oleaje (refracción, difracción, asomeramiento, reflexión, rotura,...). Así, por ejemplo, en el supuesto 4 únicamente se debe propagar el oleaje desde profundidades indefinidas. En cambio, en los supuestos 5, 6, 11 y 12, la propagación es la parte final del tratamiento de la información necesaria (a veces ésta última propagación también es útil para validar el tratamiento realizado en profundidades indefinidas).

- Retropropagar (RePR): se va a denominar retropropagación al proceso propagación del oleaje desde una posición en profundidades intermedias o reducidas hasta profundidades indefinidas. En el supuesto 7 únicamente se debe retropropagar el oleaje hasta profundidades indefinidas, en cambio, en los supuestos 8, 9, 11 y 12 la retropropagación es parte del tratamiento que hay que realizar a la información disponible previamente a su calibración.
- Calibración puntual (CaP): dentro de los tipos de calibración (ver figura 4.2), que van a ser siempre realizados en profundidades indefinidas, en la calibración puntual los oleajes en profundidades indefinidas deben ser similares (\approx), ver supuestos 2, 5, 8 y 11.
- Calibración espacial (CaE): cuando los oleajes disponibles y necesarios están muy alejados, generalmente sus climas marítimos no son parecidos. Con la calibración espacial se pueden utilizar posiciones en las que sus climas marítimos en profundidades indefinidas no sean similares (pueden ser distintos, \neq), ver supuestos 3, 6, 9 y 12.

	Datos disponibles	Información necesaria	Climas marítimos en Prof. indef.	Tratamiento de la información (datos disponibles) para llevarla hasta el punto objetivo.
	¿Prof. indef.?	¿Prof. indef.?		
1	✓	✓	≡	UD
2	✓	✓	≈	CaP
3	✓	✓	≠	CaE
4	✓	✗	≡	PR
5	✓	✗	≈	CaP + PR
6	✓	✗	≠	CaE + PR
7	✗	✓	≡	RePR
8	✗	✓	≈	RePR + CaP
9	✗	✓	≠	RePR + CaE
10	✗	✗	≡	UD
11	✗	✗	≈	RePR + CaP + PR
12	✗	✗	≠	RePR + CaE + PR

Tabla 4.2. Tratamiento de los datos disponibles en función de la clasificación de la información de oleaje.

4.4. Calibración con datos de oleaje en España.

Particularizando las metodologías de calibración para el caso de España, a partir de los supuestos de la tabla 4.2, se va a clasificar las metodologías de calibración en función de los distintos tipos de datos existentes en las costas españolas. No se realizarán calibraciones cuando se tenga información instrumental suficiente en la posición de interés, casos excepcionalmente escasos (denotado por \equiv en las tablas 4.1 y 4.2), por lo que no se incluyen los supuestos 1, 4, 7 y 10 en la clasificación que se va a realizar a continuación.

Se parte de la premisa de que siempre se tienen datos de modelado numérico a una resolución adecuada (SIMAR-44 o en su defecto WANA) pero que deben ser calibrados con datos instrumentales. No se van a considerar los datos visuales, pues siempre se van a tener datos instrumentales de mejor calidad, pero por su distribución espacial podrían considerarse con características similares a los de los satélites. Dentro de los datos instrumentales para calibrar se utilizarán boyas y satélites, aunque podrían asimilarse a ellos datos de otras fuentes de información como radares HF, ADCP, ... si se dispone de ellos.

Las bases de datos de reanálisis que se van a utilizar están en profundidades indefinidas, por lo que las calibraciones van a ser definidas para estas profundidades. En caso de tener la posibilidad de definir distintos tratamientos de datos o calibraciones se preferirá la utilización de datos en profundidades indefinidas que no sea necesario propagar o retropropagar, para distorsionar lo menos posible la información disponible. De hecho este criterio primará frente al de elegir métodos de calibración puntuales en lugar de los espaciales (cuando sea posible).

En esta misma línea, la de distorsionar lo menos posible los datos originales, cuando se dispone de datos de fuentes de información distintas se prefiere calibrar con unos y validar con otro conjunto de datos. La otra posibilidad sería homogeneizar inicialmente todos los datos disponibles para posteriormente calibrar con un porcentaje de dichos datos y validar con el resto; pero esta alternativa modifica más la información disponible, que *a priori* es correcta.

En la tabla 4.3 se presenta la clasificación de los distintos tipos de calibraciones (o tratamientos de la información disponible) de la base de datos de modelado numérico (SIMAR-44, marcado en negro) con información instrumental (Boyas y Satélites), marcando en verde los datos utilizados en el proceso de calibración, en amarillo los reservados para la validación (del orden del 25% como mínimo) y en rojo los que no es necesario utilizar. En las boyas se especifica si están fondeadas en profundidades indefinidas o no y si su clima marítimo es similar (\approx) o diferente (\neq) al del punto de interés. Los datos de satélites, gracias a su cobertura espacial, se considera que tienen siempre oleaje similar y en profundidades indefinidas.

CLASIFICACIÓN DE LAS METODOLOGÍAS DE CALIBRACIÓN

Boyas				Satélites	SIMAR-44		
Prof. indef.		No prof. indef.					
≈	≠	≈	≠				
				x / x	x	2	CaP
x				x	x	2	CaP
	x			x / x	x	3	CaE
		x		x	x	8	RePR + CaP
			x	x / x	x	9	RePR + CaE
x	x			x	x	2	CaP
x		x		x	x	5	CaP + PR
x			x	x	x	2	CaP
	x	x		x / x	x	6	CaE + PR
	x		x	x / x	x	6	CaE + PR
		x	x	x	x	8	RePR + CaP
	x	x	x	x / x	x	12	RePR + CaE + PR
x		x	x	x	x	5	CaP + PR
x	x		x	x	x	2	CaP
x	x	x		x	x	5	CaP + PR
x	x	x	x	x	x	5	CaP + PR

Tabla 4.3. Metodologías de calibración para los datos de oleaje españoles.

Cuando sea posible, es preferible realizar la validación propagando los datos calibrados en profundidades indefinidas hasta boyas situadas cerca de la zona de interés, o comparando con los datos de satélites próximos. Los satélites son una excelente fuente de información para validar y verificar que la calibración está correctamente realizada, pero en caso de que no se puedan adquirir dichos datos, la tabla anterior quedaría de la forma mostrada en la tabla 4.4.

Boyas				SIMAR-44		
Prof. indef.		No prof. indef.				
≈	≠	≈	≠			
x / x				x	2	CaP
	x / x			x	3	CaE
		x / x		x	8	RePR + CaP
			x / x	x	9	RePR + CaE
x / x	x			x	2	CaP
x		x		x	5	CaP + PR
x / x			x	x	2	CaP
	x	x		x	6	CaE + PR
	x		x	x	6	CaE + PR
		x / x	x	x	8	RePR + CaP
	x	x	x	x	12	RePR + CaE + PR
x		x	x	x	5	CaP + PR
x / x	x		x	x	2	CaP
x	x	x		x	5	CaP + PR
x	x	x	x	x	5	CaP + PR

Tabla 4.4. Metodologías de calibración para los datos de oleaje españoles, salvo los provenientes de satélites.

Como comentario final, cabe destacar la importancia de la validación de los resultados, pues si los resultados no son satisfactorios deberá rehacerse la calibración, verificando los datos utilizados. Si aún así no se encuentra la fuente de los errores (poca información disponible, mucha incertidumbre de las técnicas de calibración,...) se podría aumentar el porcentaje de datos para calibrar en detrimento de los reservados para validar, para tratar de determinar el origen de los errores en la calibración o verificar con técnicas de calibración menos sofisticadas si se obtienen aproximadamente los mismos resultados.

