
NOTAS Y EXPERIENCIAS / NOTES AND EXPERIENCES

Desarrollo de un filtro temático para la delimitación bibliométrica de un área interdisciplinar: el caso de Ciencias del Mar

Development of a thematic filter for the bibliometric delimitation on interdisciplinary area: the case of Marine Science

Rodrigo Costas*, María Bordons*

Resumen: En esta nota metodológica se analiza la precisión y exhaustividad de un filtro temático diseñado para delimitar el área de Ciencias del Mar a través de revistas especializadas y palabras clave, y se compara su eficacia con una delimitación basada sólo en revistas. El estudio analiza la producción científica española en Ciencias del Mar obtenida del SCI, SSCI & AHCI durante los años 1994-2004. La estrategia basada en revistas obtiene una alta precisión (96%) pero baja exhaustividad en la delimitación (próxima al 30%). El filtro mixto presenta una menor precisión (69%) pero una mayor exhaustividad (alrededor del 75%). La delimitación basada en revistas ofrece una imagen incompleta del área e ignora contribuciones que generan alto impacto publicadas en revistas de otras disciplinas. Se pone de manifiesto el interés de desarrollar filtros temáticos más elaborados y el importante papel de los expertos para obtener un equilibrio entre precisión y exhaustividad.

Palabras clave: Ciencias del Mar, exhaustividad, precisión, recuperación de información, delimitación temática, estudios bibliométricos.

Abstract: In this methodological note, the recall and precision of a mixed thematic filter developed for the delimitation of the Marine Science area by means of specialised journals and keywords is analysed, and its efficiency is compared with a delimitation based only on scientific journals. The scientific production of Marine Science in Spain during 1994-2004 as covered by SCI, SSCI & AHCI is analysed. A high precision (96%) but a very low recall (close to 30%) is obtained by the delimitation strategy based only on scientific journals. The mixed filter presents a lower precision (69%) but a higher recall (around 75%). The journal-based delimita-

* CINDOC-CSIC. Madrid. Correo-e: rodrigo.costas@cindoc.csic.es; mbordons@cindoc.csic.es.
Recibido: 11-4-07; 2.^a versión; 5-11-07; 3.^a versión: 22-11-07.

tion provides an incomplete view of the area in which high-impact documents published in journals of other disciplines are ignored. The interest of developing advanced filters for the delimitation of scientific areas and the importance of expert opinions in order to obtain a balance between precision and recall are stressed.

Keywords: Marine Science, recall, precision, information retrieval, thematic delimitation, bibliometric studies.

1. Introducción

Los análisis bibliométricos constituyen una herramienta de gran utilidad para analizar la actividad investigadora de países, centros y disciplinas a través de sus publicaciones científicas. Sin embargo, al abordar el estudio de áreas científicas surge con frecuencia el problema relativo a cómo delimitar y definir de la forma más exhaustiva y precisa posible un campo científico, decisión con indudables repercusiones sobre los resultados de los análisis (Glänzel et al., 1999).

La delimitación de áreas temáticas atendiendo a la clasificación de revistas en disciplinas, como la seguida por la base de datos Web of Science, es una práctica frecuente en los estudios bibliométricos, sobre todo en aquellos a nivel macro que analizan la producción científica de un país en sus distintas áreas científicas (Gómez et al., 2005; Tuzi, 2005, o Bayers, 2005). De hecho, la delimitación temática que ofrecen los informes sobre indicadores de ciencia y tecnología publicados periódicamente en los países más avanzados se basa en la clasificación de revistas en disciplinas, que a su vez son agrupadas en un número reducido de grandes áreas para ofrecer los resultados de una forma sintética (OST, 2004). Sin embargo, esta delimitación dista mucho de ser perfecta, ya que las fronteras entre disciplinas no suelen ser nítidas, especialmente en las áreas emergentes o en aquellas con un importante componente interdisciplinar. Por un lado, los investigadores pueden publicar sus resultados en revistas de distintas disciplinas, y no se puede olvidar que parte de la producción de interés para una disciplina aparece en revistas de carácter multidisciplinar, como *Nature* o *Science* (Rinia et al., 1993; Gómez et al., 1996). Además, el problema de la delimitación temática se acentúa a medida que se desciende en el nivel de análisis (centros, grupos, departamentos o individuos).

Por todo lo anterior, en el análisis de una determinada área temática es recomendable emplear filtros y estrategias de búsqueda que superen las limitaciones impuestas por las clasificaciones de revistas en disciplinas de las bases de datos. El uso combinado de revistas especializadas, palabras clave en el título y/o lugar de trabajo e incluso autores especialistas en la materia ha sido descrito por distintos investigadores (De Bruin y Moed, 1993; Lewison, 1999), con el fin de aumentar la fiabilidad de los estudios bibliométricos al proporcionar una mejor cobertura del área objeto de análisis (Nederhof, 1991; Lundberg et al., 2006).

El área de las Ciencias del Mar es un caso interesante debido, por un lado, a la creciente importancia de esta disciplina en el conjunto de las áreas (Insua y Torto-

sa, 1997; Tapaswi y Maheswarappa, 1999; Gattuso et al., 2005); y por otro, a la gran complejidad que presenta su delimitación temática (Eto, 1999). Existen en la literatura científica distintos trabajos que abordan el estudio de las Ciencias del Mar o sus subespecialidades desde un punto de vista bibliométrico (Pudovkin y Fuseler, 1995; Dastidar y Ramachandran, 2005), sin embargo, no se ha descrito la diferente cobertura del área en función de la delimitación temática utilizada ni de las repercusiones de ésta última en los resultados de los estudios bibliométricos.

2. Objetivos

Este trabajo presenta un filtro de recuperación de información sobre Ciencias del Mar aplicado a las bases de datos SCI, SSCI y AHCI y basado en revistas especializadas y palabras clave. Se analiza la eficacia de este filtro mixto y se cuantifica la aportación de las palabras clave sobre la delimitación por revistas. El filtro aquí descrito ha sido utilizado en un estudio previo para identificar la producción científica española en Ciencias del Mar (Duarte et al, 2006).

3. Metodología

La búsqueda se ha realizado sobre las bases de datos *Science Citation Index*, *Social Science Citation Index* y *Arts & Humanities Citation Index* (versión CD-ROM), elaboradas por Thomson Scientific. Se ha utilizado la versión CD-ROM de estas bases de datos por disponer los autores de la producción científica española durante 1994-2004 descargada a una base de datos relacional, depurada en lo que se refiere a normalización de nombres de revistas y de instituciones, y a la que se han incorporado los descriptores de los documentos procedentes de la versión Web of Science de estas bases (Costas e Iribarren-Maestro, 2007).

Sobre la base de datos mencionada se han identificado los documentos sobre Ciencias del Mar a través de un filtro mixto compuesto de dos estrategias de búsqueda principales: 1) una estrategia basada en revistas especializadas; y 2) una estrategia basada en descriptores y palabras de título.

3.1. Estrategia basada en revistas de publicación

Se ha elaborado una estrategia para identificar los documentos publicados en revistas de interés para el área, incluyéndose: a) las revistas asignadas por la base de datos a las siguientes disciplinas ISI («Subject Categories»): «Oceanography» (41 revistas en JCR 2004), «Engineering, Ocean» (16 revistas), «Engineering, Marine» (6 revistas) y «Fisheries» (40 revistas); b) una selección de 20 revistas, no pertenecientes a ninguna de las cuatro disciplinas anteriores, propuestas por un grupo de expertos en Ciencias del Mar (Duarte et al., 2006).

3.2. Estrategia basada en descriptores

Incluye diferentes descriptores propuestos por los expertos. Se usa aquí el término «descriptor» en sentido amplio, refiriéndose tanto a los descriptores de los autores de las publicaciones como a palabras en el título de los documentos. La estrategia de búsqueda detallada se muestra en el Anexo 1.

Evaluación del filtro. Para el análisis de la efectividad del filtro se calcula su exhaustividad y precisión, definidos de la forma siguiente (Lewison, 1996):

- Exhaustividad: proporción de documentos relevantes recuperados, sobre el total de documentos relevantes existentes en la base de datos.

$$\text{Exhaustividad} = \frac{\text{N. doc. relevantes recuperados}}{\text{N. doc. relevantes}}$$

- Precisión: proporción de documentos relevantes recuperados sobre el total de documentos recuperados.

$$\text{Precisión} = \frac{\text{N. doc. relevantes recuperados}}{\text{N. doc. recuperados}}$$

4. Resultados

El filtro mixto de recuperación obtuvo 9.938 documentos, siendo la estrategia basada en palabras clave la que recuperó mayor número de documentos (81% de los documentos) (tabla I).

Tabla I
Recuperación de documentos a través del filtro mixto temático

	<i>Doc. recuperados</i>		<i>Doc. relevantes</i>		<i>Doc. no relevantes</i>	
	<i>N. doc.</i>	<i>%</i>	<i>N. doc.</i>	<i>%</i>	<i>N. doc.</i>	<i>%</i>
Revistas especializadas	1912*	16,79	1836	26,61	76	2,50
Palabras clave	8026	80,76	5062	73,38	2964	97,50
Total documentos	9938	100	6898	100	3040	100

* 1669 documentos pertenecían a las 4 disciplinas incluidas en la estrategia de búsqueda y 243 a otras disciplinas.

El filtro se aplicó secuencialmente y de forma excluyente, de modo que los documentos que se recuperaron a través de la estrategia por revistas se excluyeron de la búsqueda por palabras clave.

4.1. Precisión de la búsqueda

Los documentos recuperados fueron revisados por los expertos para asegurar su relevancia para el área de las Ciencias del Mar. Se descartaron un total de 3040 documentos, lo que equivale al 31% de los documentos recuperados originalmente. El 97% de los documentos descartados se habían recuperado en la búsqueda por palabras clave, mientras que el 3% restante respondía a la estrategia de búsqueda en revistas especializadas. De este modo, en la tabla I se observa que la Precisión del filtro mixto se sitúa próxima al 70% (Precisión= $6898/9938=0,69$).

4.2. Exhaustividad

La exhaustividad es el parámetro más difícil de calcular de una forma exacta, dado que sería necesario conocer cuál es el número total de documentos relevantes existentes en toda la base de datos, aspecto que es imposible de obtener ya que equivaldría a analizar todos los documentos españoles recogidos en ISI para determinar si pertenecen o no al área de Ciencias del Mar. Por esta razón, se han planteado dos aproximaciones a la exhaustividad general del filtro a partir de dos subconjuntos de la población estudiada.

a) El análisis de la cobertura obtenida por el filtro de la producción de una serie de centros con actividad nuclear en el área proporciona una aproximación a la exhaustividad de la búsqueda.

Así, en el estudio se han identificado los principales centros activos en Ciencias del Mar (con más de 100 documentos) y con una clara especialización y orientación hacia las ciencias marinas (se han excluido algunas Facultades de Ciencias y centros generalistas debido a que coexisten en ellos grupos con actividad en diferentes áreas). A través de una consulta a la base de datos total de la producción científica española, se han recuperado todas las publicaciones de esos centros, independientemente de su revista o disciplina de publicación. Se ha considerado que para estos centros especializados en Ciencias Marinas, todos sus documentos serían potencialmente relevantes para el área, por lo que el cociente entre el número de documentos recuperados y el número total de documentos del centro ofrece un valor aproximado de la exhaustividad del filtro. Sin embargo, la totalidad de los documentos de un centro puede no ser pertinente para las Ciencias del Mar, por lo que la estimación obtenida es conservadora, estimándose que los valores de exhaustividad reales son superiores a los obtenidos.

La tabla II muestra los centros especializados en Ciencias del Mar con más de 100 documentos, incluyendo su producción total (columna A), su número de documentos en Ciencias del Mar según la estrategia basada en revistas especializadas (Rev) (columna B); y su número de documentos en Ciencias del Mar según el filtro mixto (revistas/descriptores) (Filtro mixto) (columna C). Las últimas dos columnas

indican el porcentaje de la producción total de cada centro recuperado a través del filtro basado en revistas (B/A) y a través del filtro mixto (C/A). De este modo, es posible obtener una aproximación ilustrativa de la exhaustividad del filtro mixto (tabla II).

Tabla II
Exhaustividad del filtro temático sobre la producción de una selección de centros

Centro	A	B	C	B/A (%)	C/A (%)
	<i>N. doc. total áreas</i>	<i>N. doc. C. Mar rev.</i>	<i>N. doc. C. Mar filtro mixto</i>	<i>% doc. rev.</i>	<i>% doc. filtro mixto</i>
I. Cienc. del Mar CSIC, Barcelona	590	207	545	35,08	92,37
I. Inv. Marinas CSIC, Vigo	521	200	391	38,39	75,05
C. Andaluz Superior Estud. Mar, U. Cádiz	461	69	306	14,97	66,38
C. Est. Avanz. Blanes CSIC, Girona	405	66	257	16,30	63,46
Fac. CC. Mar. U. Las Palmas	258	82	192	31,78	74,42
I. Cienc. Marinas CSIC, Cádiz	194	42	168	21,65	86,60
I. Acuic. T. Sal CSIC, Castellón	200	63	167	31,50	83,50
I. Esp. Oceanografía, La Coruña	104	56	101	53,85	97,12
Media				28,12	74,40

En la tabla II se observa que la estrategia basada únicamente en revistas especializadas se acompaña de una baja cobertura del área, ya que sólo se recupera en torno al 28% de la producción de los centros.

A través del filtro mixto se recupera entre el 63% y el 97% (promedio del 74%) de la producción de los centros. Las cifras más bajas corresponden a centros que tienen alguna línea de trabajo ajena al mundo de las ciencias marinas (p. ej. el Instituto de Investigaciones Marinas de Vigo incluye una línea de investigación de tecnología de los alimentos).

b) Dadas las limitaciones del cálculo anterior de la exhaustividad de la búsqueda, se presenta otra aproximación a la misma. La estimación de la exhaustividad de un filtro más utilizada en la literatura consiste en comparar los resultados de una búsqueda manual realizada por expertos («gold standard set») con los obtenidos por la estrategia de búsqueda que se desea evaluar (ver por ej. Sladek et al, 2006).

En este estudio los expertos revisaron los documentos publicados por los investigadores españoles en el subconjunto formado por todas las revistas especializadas incluidas en la estrategia de búsqueda, y señalaron aquellos relevantes para las Ciencias del Mar, conformando esta muestra el «gold standard set» o búsqueda de referencia. Se planteó la siguiente pregunta: ¿qué porcentaje de los documentos publicados en revistas especializadas se recuperan a través de la estrategia basada en palabras clave? En la tabla III se presenta un resumen de este análisis.

Tabla III
Evaluación de la estrategia de palabras clave

	<i>Doc. relevantes</i>	<i>Doc. no relevantes</i>	<i>Total</i>
Doc. recuperados	1275 (67%)	45 (2%)	1320
Doc. no recuperados	561 (29%)	31 (2%)	592
Total	1836	76	1912

$$\text{Exhaustividad} = 1275/1836 = 0,69$$

Como se puede observar, el filtro de palabras clave presenta una exhaustividad del 69%, que es moderadamente alta, si se tiene en cuenta que este filtro se diseñó con la intención de complementar la delimitación basada en revistas especializadas y disciplinas ISI, que aporta el núcleo básico de actividad en el campo. Dado que el filtro mixto incluye una búsqueda por revistas especializadas que recupera el 100% de los documentos relevantes publicados en dichas revistas y una búsqueda por palabras clave (exhaustividad estimada del 69%), y que la búsqueda por revistas aporta el 27% de los documentos relevantes mientras que la búsqueda por palabras clave el 73% restante (véase la tabla I), se estima la exhaustividad del filtro mixto en un 77% ($0,27*1+0,73*0,69$) (tabla IV).

Tabla IV
Evaluación del filtro mixto

	<i>Revistas</i>	<i>Filtro mixto (Revistas+Palabras clave)</i>
Precisión	0,96 (1836/1912)	0,69 (6898/9938)
Exhaustividad	0,27 (1836/6898)	0,77 (*)

(*) Valor estimado ($0,27*1+0,73*0,69$).

4.3. Interdisciplinariedad de la investigación en Ciencias del Mar

La creciente interdisciplinariedad de la investigación es uno de los factores que dificulta la delimitación de las áreas temáticas. Una muestra del carácter interdisciplinar de la investigación en ciencias marinas es la distribución de las publicaciones en más de 130 disciplinas diferentes, mostrándose las principales en la Tabla V. Hay que tener en cuenta que una revista puede estar clasificada hasta en seis disciplinas diferentes, por lo que hay solapamiento entre ellas. Se observa que el 40% de los documentos están en la disciplina «Biología Marina y de Aguas Continentales». Las disciplinas que se presentan en la tabla V recogen cerca del 80% del total de la producción.

Tabla V
Distribución de documentos por disciplinas ISI
(sólo disciplinas con más de 200 documentos)

<i>Disciplinas ISI</i>	<i>N. Docs</i>	<i>%</i>
Marine & Freshwater Biology	2817	40,84
Oceanography	879	12,74
Ecology	751	10,89
Fisheries	726	10,52
Environmental Sciences	589	8,54
Geosciences, Multidisciplinary	436	6,32
Plant Sciences	355	5,15
Biotechnology & Applied Microbiology	306	4,44
Zoology	300	4,35
Chemistry, Analytical	299	4,33
Biochemistry & Molecular Biology	280	4,06
Food, Science & Technology	249	3,61
Microbiology	228	3,31
Toxicology	206	2,99

4.4. Impacto de la investigación en Ciencias del Mar

Es interesante señalar que las publicaciones de los investigadores españoles identificadas a través de descriptores se publicaron en revistas de mayor factor de impacto que las identificadas a través de revistas especializadas, lo que en parte se debe a la inclusión en la primera delimitación de revistas multidisciplinares como *Nature* o *Science*, que presentan un factor de impacto muy elevado (superior a 20 en 2005). Es decir, que una delimitación basada sólo en revistas especializadas no permite identificar los documentos publicados en revistas de muy alto prestigio pero de carácter multidisciplinar. ¿Significa esto que tampoco se identifican los documentos más citados, ya que no necesariamente se publican en las revistas especializadas seleccionadas en nuestra estrategia? En nuestro estudio, no se encontraron diferencias en el número medio de citas por documento de las publicaciones en función de estar publicadas en revistas especializadas (incluidas en la búsqueda) o de otras disciplinas (identificadas por descriptores). No obstante, entre los documentos muy citados («highly cited papers» [HCP] ó 1% de documentos más citados) se detectó una alta presencia relativa de documentos identificados por descriptores. Así, el 84% de los HCP había sido identificado por descriptores, mientras que esta estrategia recuperó el 73% del total de los documentos objeto de estudio (tabla VI). Es decir, la delimitación por revistas no necesariamente infravalora el impacto medio de la actividad científica en el área, pero deja de percibir algunos trabajos muy destacados que han producido un impacto extraordinario en el área.

Tabla VI
Distribución de HCP según su estrategia de identificación
(revistas o descriptores)

Tipo delimitación	HCP (>52 citas/doc.)	Resto de (0-52 citas/doc.)	Total doc.
Revistas	11 (15,71%)	1825 (26,73%)	1836 (26,21%)
Descriptores	59 (84,29%)	5003 (73,27%)	5062 (73,38%)
Total	70	6828	6898

Nota: HCP son aquellos documentos con más de 52 citas en el periodo.

Conclusiones

Una delimitación del área de Ciencias del Mar basada únicamente en un conjunto de revistas especializadas presenta una gran precisión (96% de los documentos recuperados son relevantes para el estudio) pero una baja exhaustividad, ya que se pierden documentos relevantes publicados en revistas no especializadas en ciencias marinas. Por otro lado, una delimitación únicamente basada en palabras clave se caracteriza por una mayor exhaustividad, pero un mayor ruido en la recuperación de información. Estos datos ponen de manifiesto el interés de combinar ambas metodologías. El filtro mixto que se presenta en este documento combina una estrategia de búsqueda basada en revistas especializadas (recupera el 27% de los documentos relevantes) y palabras clave (recupera el 73% restante) y se acompaña de valores de precisión y exhaustividad del 69% y 77% respectivamente.

Resulta llamativo que «Biología Marina y de Aguas Continentales», disciplina que aporta el mayor número de documentos, no estuviera incluida en la estrategia de búsqueda por disciplinas ISI. La razón es que los expertos aconsejaron no incluirla por contener documentos relacionados con las aguas continentales, que no eran objeto del estudio. No obstante, los resultados de este trabajo muestran que el 86% de los documentos españoles en revistas de «Biología Marina y de Aguas Continentales» se recuperaron a través del filtro mixto, y que sólo el 14% fue considerado no relevante. Ello aconseja la inclusión de esta disciplina en la estrategia de búsqueda en futuros estudios.

El grado de cobertura de un área ofrecida por una delimitación basada en revistas puede ser muy variable según las bases de datos y las características de cada disciplina. En este caso, la delimitación basada en revistas aportó el 30% de los documentos de Ciencias del Mar (o el 50% si se incluye «Biología Marina y de Aguas Continentales»). La disciplina más productiva fue Biología Marina y de Aguas Continentales, que aportó el 40% de los documentos. Este resultado es compatible con los señalados por van Leeuwen (2007), Aksnes (2000), Lewison (1999) o Bradford (1934), quienes indican que 2/3 partes de las publicaciones de un área no aparecen en las revistas de su especialidad.

Una delimitación temática basada en revistas especializadas es adecuada para identificar los principales actores activos en un tema (Gómez et al, 1996), pero para realizar un análisis a nivel meso o micro, es recomendable complementar la estrategia basada en revistas con otras más elaboradas. A través de una delimitación temática basada en revistas especializadas se obtiene una visión incompleta, y no siempre representativa del total de la actividad realizada en un área (Lewison, 1996). En nuestro estudio no se observan diferencias en el impacto medio del área (medido a través de citas) en función de su delimitación por revistas o descriptores, pero la primera delimitación no logra identificar algunos de los documentos que más repercusión científica han producido, muchos de ellos publicados en revistas multidisciplinares.

Por último, señalar que este estudio contó con una importante implicación de un grupo de expertos, que colaboraron en el diseño de la estrategia y en la posterior revisión de los documentos recuperados. La participación de los expertos fue esencial para descartar documentos no relevantes y optimizar la búsqueda.

Agradecimientos

Este trabajo ha sido posible gracias a una beca I3P concedida a Rodrigo Costas para desarrollar su investigación en el CSIC. Se agradece la participación del grupo de expertos en Ciencias del Mar en la delimitación temática del área y en el análisis de la relevancia de los documentos.

Referencias

- AKSNES, D.W.; OLSEN, T.B.; SEGLEN, P.O. (2000). Validation of bibliometric indicators in the field of microbiology: a Norwegian case study. *Scientometrics*, 49(1): 7-22.
- BAYERS, N. (2005). Using ISI data in the analysis of German national and international research output. *Scientometrics*, 62(1), 155-163.
- BRADFORD, S.C. (1934). Sources of information on specific subjects. *Engineering*, 137: 85-86
- COSTAS, R.; IRIBARREN-MAESTRO, I. (2007). Variations in content and format of ISI database in their different versions: the case of the Science Citation Index in CD-ROM and the Web of Science. *Scientometrics*, vol. 72 (2):167-183.
- DASTIDAR, P.G.; RAMACHANDRAN, S. (2005). Engineering research in ocean sector: An international profile. *Scientometrics*, vol. 65 (2), 199-213.
- DE BRUIN, RE, MOED, H.F.(1993). Delimitation of scientific subfields using cognitive words from corporate address in scientific publications. *Scientometrics* vol. 26 (1), 65-80.
- DUARTE, C.M.; ACUÑA, J.L.; ÁLVAREZ-SALGADO, X.A.; BLASCO, D.; BORDONS, M.; COSTAS, R.; DAÑOBEITIA, J.J.; HERNANDEZ, S.; LOSADA, I.; MORALES-NIN, B.; NOMBELA, M.A.; RUIZ, J.; ZANUY, S. (2006). *Las ciencias y tecnologías marinas en España*. Madrid: CSIC.

- ETO, H. (1999). The interest of scientific communities in sea-related research topics. *Scientometrics*, vol. 45 (2), 167-183.
- GATTUSO, J.P.; DAWSON, N.A.; DUARTE, C.M.; MIDDELBURG, J.J. (2005). Patterns of publication effort in coastal biogeochemistry: a bibliometric survey (1971 to 2003). *Marine Ecology Progress Series*, vol. 264 (9-22), 9-22.
- GLÄNZEL, W.; SCHUBERT, A.; CZERWON, H.J. (1999). An item-by-item subject classification of papers published in multidisciplinary and general journals using reference analysis. *Scientometrics*, vol. 44 (3), 427-439.
- GÓMEZ, I.; BORDONS, M.; MORILLO, F. (2005). Regionalisation of science and technology data in Spain. *Research Evaluation*, vol. 14 (2), 137-148.
- GÓMEZ, I.; BORDONS, M.; FERNANDEZ, M.T.; MENDEZ, A. (1996). Coping with the problem of subject classification diversity. *Scientometrics*, vol. 35 (2), 223-235.
- ÍNSUA, A.; TORTOSA, E. (1999). *Análisis bibliométrico de la producción científica española en ciencias marinas*. Informe de la Acción Especial Contribución del buque oceanográfico Hespérides a la producción científica española financiada por el Plan Nacional de I+D CICYT. Programa Nacional CYTMAR. Ref.:MAR96-1563. 1996-1997.
- LEWISON, G. (1996). The definition of biomedical research subfields with title keywords and application to the analysis of research outputs. *Research Evaluation*, vol. 6 (1), p.25-36.
- LEWISON, G. (1999). The definition and calibration of biomedical subfields. *Scientometrics*, vol. 46 (3), 529-537.
- LUNDBERG, J.; FRANSSON, A.; BROMMELS, M.; SKAR, J.; LUNDKVIST, I. (2006). Is it better or just the same? Article identification strategies impact bibliometric assessments. *Scientometrics*, vol. 66 (1), 183-197.
- NEDERHOF, A.J. (1991). Delimitation of a medical research topic: interaction with experts in selecting a database and constructing a search strategy. *Research Evaluation*, vol. 1 (3), 149-154.
- OST. (2004). *Indicateurs de Sciences et de Technologies: rapport de l'Observatoire des Sciences et des Techniques*. France: OST<<http://www.obs-ost.fr/rapport2004.pdf>> [Consulta 11-10-2006]
- PUDOVKIN, A.I.; FUSELER, E.A. (1995). Indices of journal citation relatedness and citation relationships among aquatic biology journals. *Scientometrics*, vol. 32 (3), 227-236.
- RINIA, E.J.; DE LANGE, C.; MOED, H.F. (1993). Measuring national output in physics: Delimitation problems. *Scientometrics*, vol. 28 (1), 89-110.
- SLADEK, R.; TIEMAN, J.; FAZEKAS, B.S.; ABERNETHY, A.P.; CURROW, D.C. (2006). Development of a subject search filter to find information relevant to palliative care in the general medical literature. *Journal of the Medical Library Association*. 94(4): 394-401.
- TAPASWI, M.P.; MAHESWARAPPA, B.S. (1999). Ranking serials in oceanography: an analysis based on the contributions and their citations. *Scientometrics*, vol. 44 (1), 93-127.
- TUZI, F. (2005). The scientific specialisation of the Italian regions. *Scientometrics*, vol. 62 (1), 87-111.
- VAN LEEUWEN, T.N. (2007). Modelling of bibliometric approaches and importance of output verification in research performance assessment. *Research Evaluation*, vol. 2 (1): 93-105.

Anexo 1
Estrategia de búsqueda basada en descriptores

(*marine* or *ocean* or *estuar* or *beach* or *crustacean* or *teleost* or
mangrove or *salt*marsh* or *plankton*)
OR
(*cont* margin* or *cont*-margin*)
OR
(*coast* and not *pancoast*)
OR
(((alga* or * alga* or *-alga*) and not *algaas* and not *algan* and not
algaSB) or *microalga* or *macroalga*)
OR
((*fish* and not *fisher * and not *fisher,*) or *kingfish* or *shellfish*)
OR
((sea* or * sea* or *-sea*) and not *search* and not *season*)
AND NOT
(*lake* or *tinca* or *trout*)

Notas: 1) El asterisco indica truncamiento en la búsqueda. 2) Términos negados por introducir ruido: pancoast, fisher, search, season. Se niegan lake, tinca, trout por recuperar documentos relativos a temas de agua dulce.

3) El término sea* no se trunca por la izquierda para evitar ruido (ej. disease). 4) Se niega fisher como final de palabra, pero se recupera fisheries y términos relacionados.