Universitat de València

PhD Thesis

Implementing Ecosystem Approach to Fisheries Management: Advances and New tools

Maria Grazia Pennino

Supervisors:
José María Bellido Millán
David V. Conesa Guillén
Antonio López Quílez

Departamento de Estadística
e Investigación Operativa

Instituto Español de Oceanografía
Centro Oceanográfico de Murcia

Ai miei genitori, sostenitori del mio cammino,
A Angelo, il mio compagno di viaggio.

Acknowledgments

Una persona molto saggia disse *l'importante non è la meta ma il viaggio*, e per me questi quattro anni di dottorato sono stati un cammino molto largo, con molte salite e discese e in cui, senza l'aiuto dei miei compagni di viaggio non avrei mai raggiunto la meta finale.

Un viaggio non può iniziare senza un buon sostegno alle spalle, Mammisia & Peppino, mi avete lasciato partire e sempre mi sostenete, senza voi non avrei mai iniziato neanche il cammino. Se mai dovessi consigliare un' agenzia di viaggi vi suggerirei l'agenzia "Belli-Antonio-David", non solo mi hanno venduto il biglietto facilmente ma sono stati presenti in ogni tappa del percorso, e ognuno, alla sua maniera, è stato fondamentale. Belli mi ha dato il coraggio di intraprendere qualunque viaggio, mi ha sempre indirizzato il cammino pero lasciandomi camminare sola e imparando a essere indipendente; Antonio mi ha dato la sicurezza e la certezza che dopo qualunque salita, c'è sempre una discesa; David è stato letteralmente la luce che ha sempre illuminato il cammino, la persona che ti tende la mano quando inciampi, nessun invito in Sardegna potrà contraccambiarlo. Le insuperabili compagne di viaggio (letteralmente) le mie *pitagorine*: Cris e Gemma che mi hanno sempre rallegrato il cammino, ed in particolare Encarni che oltre a un ottima compagna di viaggio si è dimostrata una insostituibile amica. Facu, per cui non ho parole per descrivere la stima nei suoi confronti e senza cui non avrei mai raggiunto la meta finale cosi velocemente. Hèctor, Anabel, Rufo, Silvia e Guille, un viaggio non è mai piacevole senza una tappa indimenticabile, un posto in cui torneresti sempre perchè è sinonimo di casa, e per me la "tappa valenciana" lo è stata. Antonio, Angel e Javi: l'insostituibile triade, colleghi ma sopratutto amici: in un viaggio non possono mancare le "persone di contatto" sui cui puoi sempre contare in qualunque momento. Raul, compagno di sventure e avventure, che sempre ha reso il cammino più sereno e agibile. Salva, la mia buona entropia nel cammino che ha arrichito il viaggio con innumerevoli piacevoli fuori-pista. Andrea e tutto l'equipaggio SNC che mi hanno dato l'energia necessaria per capire che avevo gia percorso un largo tratto del cammino e non mancava tanto alla meta finale. A tutti i miei colleghi del IEO, e sopratutto a tutti gli *achopijos* che mi sono stati vicini nelle salite per aiutarmi ad affrontarle, e nelle discese per assaporarle: Anita, Mayte, Merche, Stefy& Pablo, Olga, Ana, Rosa, Cris & Migue, José Antonio, Luz, Ursula, Blanca & Juan e Javi & Lidia. Infine, un viaggio non è lo stesso senza una persona al tuo fianco con cui condividerlo, commentarlo, ricordarlo, amarlo e a volte odiarlo. Ancora la nostra meta non è arrivata pero senza te questo viaggio non sarebbe mai stato lo stesso.

A tutti i miei compagni di viaggio, semplicemente GRAZIE e ovviamente, il cammino non finisce qui...esto acaba de empezar.

Contents

List of Figures

List of Tables

Introduction

Since ancient times, fishing has been a major source of food for humanity and a provider of employment and economic benefits to those engaged in this activity. However, with increased knowledge and the dynamic development of fisheries it was realized that aquatic resources, although renewable, were not infinite and needed to be properly managed if their contribution to the nutritional, economic and social well-being of the growing world's population was to be sustained (García et al., 2000).

In recent years, world fisheries have become a dynamically developing sector of the food industry, and coastal states have striven to take advantage of their new opportunities by investing in modern fishing fleets and processing factories in response to growing international demand for fish and fishery products. It became clear, however, that many fishery resources could not sustain an often uncontrolled increase in exploitation.

Concerns have been expressed about the contribution of fisheries to sustainable development and about overfishing, excess catching capacity, the depletion of some stocks, human-induced changes in ecosystems, as well as the increase and globalization of the fish trade with its potential impact on local supplies and equity.

In order to address these problems, the United Nations Food and Agriculture Organization (FAO) has called for the application of an Ecosystem Approach to Fisheries Management (EAFM), which aims to achieve a sustainable exploitation of commercial fisheries, providing specific consideration of the interactions between fishing gears and marine ecosystems (Cotter et al., 2009). EAFM takes into account that fisheries are embedded into the environment and cannot be managed in isolation (Gascuel et al., 2012). It has to be considered as the application of sustainable development principles to the fishing sector, combining ecological sustainability, economic viability and social fairness.

Nevertheless, while it is widely recognized that fishing is important to sustainable development and that its contribution could be improved, the amount of objective scientific information about fishing is limited and what exists is difficult to access.

The reliability of scientific advice for the management of natural resources is highly dependent on the quantity and quality of data that are available for scientific assessment and interpretation (Jennings and Polunin, 1996). Although large amounts of certain types of data about marine ecosystems are readily available, as in the case of satellite-derived remote sensing data or observations based on automatic telemetry, it is far more common to have to deal with limited and irregularly spaced data (e.g. on fish and other marine fauna), and the data may not always be strictly comparable due to variations in environmental conditions between sampling periods.

The main reason is that in fishery research the collection of data is both time-consuming and expensive. Data are difficult to obtain, and the problems increase when the goal of the research is to study long time series on a macro-scale, with the purpose of examining changes in the dynamics of a whole ecosystem. In all these cases that require large databases, with an adequate coverage in space and time of a variety of variables, information is almost non-existent.

In addition, information about the status of fisheries can be derived from different sources such as fishery-independent surveys, fishery-dependent surveys (skipper logbooks and/or observers) and FAO official landings (Vasconcellos and Cochrane, 2005). Each of these sources provides different types of information with additional details, and must be carefully selected according to the type of study and the objectives pursued.

Fishery-independent surveys rely mostly on expensive research campaigns performed over relatively short periods of time. Survey data are considered to be of superior quality because they are independent of management measures, standardized fishing procedures are used, and both sampling statistics and the biological information on target species are taken into consideration during survey design (National Research Council, 2000). However, they generally have a limited coverage in space and time (both in terms of seasonality and the number of years of data), which could lead to biased and imprecise estimations (Hilborn et al., 1992). In brief, there are two main issues regarding fishery-independent surveys: the one of timing in relation to the seasonal cycle and location of the target species (leading to unrepresentative sampling, i.e. bias) and the fact that only a limited amount of data can be collected.

Fishery-dependent surveys can provide a long time series, wide spatial coverage all year round, and information on a large variety of target species, gear types, landing sites and distribution channels (Lunn and Dearden, 2006). Data of this kind sometimes lack particular details such as the location of fishing grounds and species identity (catches are mostly identified to a higher taxonomic level), and then there are issues of bias due to constraints imposed by management and the deliberate misreporting of catches. Some of the disadvantages of fishery-dependent surveys can by overcome by using on-board observers but inevitably only a small fraction of fishing activity can be covered in this way.

FAO official landings are often the only source of data available due to their connection to the economy and business. The main issue in this source of information is that in many fisheries (official) landings and the actual catch are not necessarily the same. Large amounts of unreported or underreported catches, discards and illegal catches are not recorded in landing data (Vasconcellos and Cochrane, 2005), and not all landings enter official statistics (e.g. *black* landings, sales by small-scale fisheres directly to consumers, etc.). Furthermore, while previous data sources imply a spatial reference where species have been captured, the FAO official landings have no georeferencing.

However, the data collected by on-board observers and that from research surveys have relatively recent time series, while official landing data have been available for most FAO countries since 1950 (www.fao.org).

The quantity and quality of information varies depending on the country of the fisheries studied. In most countries, detailed information is available and management processes are in place for some important fisheries, while others are poorly documented and hardly managed at all. In particular, but not exclusively, developing countries reflect these data-poor situations, given that the resources and infrastructure needed to collect relevant data, assess system status and implement management requirements have not increased at the same time as the recent rise international trade in fishery products.

Moreover, the need to move towards an EAFM has resulted in an increase in the need for data at different spatio-temporal scales and many areas are currently in a data-poor situation for the purposes of implementing it (Bianchi, 2010). The problems increase when the goal is to study long time series on a macro-scale, with the purpose of examining changes in the dynamics of a whole ecosystem. In all these cases, where sophisticated ecosystem models cannot be applied due to the lack of detailed data, the first step towards applying the EAFM should be to set more realistic goals and use flexible tools that can work effectively despite uncertainty and limited information (Trenkel et al., 2007).

Within this context, the main aim of this thesis is to analyze the different available sources of information on fishery resources, in order to propose new methodologies for efficient fishery management both in terms of data-poor situations, macro-scale/long time series studies, and micro-meso-scale detailed studies.

Several scenarios have been addressed depending on the source of fisheries data used and the objective pursued. For each one of them, a specific methodology has either been developed (when there was no available method) or described (when there was), showing its strengths and limitations, and discussing its implications for the fishery management.

This thesis is organized as follows. After this introduction, Chapter 1 reviews the concept and main aims of the Ecosystem Approach to Fisheries Management (EAFM) in detail, introducing the fishing issues that will be discussed in the thesis, outlining the methodologies described, the different fishing datasets and the types of spatial-temporal scales with which these issues can be addressed.

In Chapter 2 we present two different methodological approaches to studing fishing exploitation through the analysis of selected ecological indicators. These methodologies are very useful in data-poor situations, when the only available source of data are fish market landings, without any additional information, and for macro-scale and long data series studies. The first approach allows us to assess the trends of ecological indicators over time through bootstrap and smoothing tecniques, highlighting shifts, if there are any, in the ecosystem trophic dynamic due to fishing exploitation. The second approach focuses on how the distributions of ecological indicators evolve over time, through the nonparametric estimation of Gaussian kernel density functions and, according to this, show what the probable long run distribution might be. Three different case studies are presented to illustrate these methods.

In Chapter 3 we present a Bayesian hierarchical spatial model to estimate and predict the distribution of fishery non target species using fishery-dependent data. In all these cases, the selection of the sampling locations does not depend on the values of the spatial variable and so these are stochastically independent of the field process. As a result, the species' occurrence is an unbiased indicator of its presence/absence pattern. We present three different types of practical examples. The first two examples use presence/absence data as response variables to estimate and predict the distribution of pelagic and demersal species in specific studied areas. In contrast, the third example is based on a Gaussian response variable in order to address a very topical issue in fisheries: the spatio-temporal distribution of discards.

In Chapter 4, we present a Bayesian hierarchical spatial model to estimate the distribution of fishery target species using fishery-dependent data. In these specific cases, in which the data are collected following a preferential sampling, predicting the distribution of the target species using traditional application of geostatistical methods, could lead to biased predictions. The basic idea is to interpret the data as a marked point pattern, where the sampling locations form a point pattern and the measurements taken in those locations are the marks. As a practical example, we model the abundance of European hake (*Merluccius merluccius*) in the Gulf of Alicante.

In Chapter 5 we propose a Bayesian hierarchical spatial-temporal model to identify fish nurseries for fisheries management purposes. We applied our approach to juvenile European hake (*Merluccius merluccius*) in the central Mediterranean Sea. Time series of fishery-indipendent data were used to map juvenile hake distribution with Bayesian kriging, and to verify density hot-spots persistence.

Finally, in Chapter 6 we present some concluding remarks and future lines of research.

Resumen

Desde la antigüedad, la pesca ha sido una fuente importante de alimentos para la humanidad, así como fuente de empleo y beneficios económicos para quienes se dedican a esta actividad. Sin embargo, con el aumento de los conocimientos científicos y la evolución dinámica de la pesca se hizo evidente de que los recursos acuáticos, aunque renovables, no eran infinitos y era necesario gestionar adecuadamente su contribución al bienestar nutricional, económico y el bienestar social de la población mundial para un crecimiento y desarrollo sostenible (García et al., 2000).

En los últimos años, la pesca mundial se ha convertido en un sector dinámico y de desarrollo de la industria alimentaria. Los estados costeros han procurado aprovechar sus nuevas oportunidades invirtiendo en flotas pesqueras modernas y fábricas de procesamiento en respuesta a la creciente demanda internacional de pescado y productos pesqueros. Sin embargo, se ha constatado que la explotación pesquera no puede mantener un aumento incontrolado sin provocar un perjuicio a la salud de los recursos pesqueros.

En diversos foros se expresa gran preocupación acerca de la contribución de la pesca al desarrollo sostenible, la sobrepesca, el exceso de capacidad de las flotas, el agotamiento de algunas poblaciones, los cambios inducidos por el hombre en los ecosistemas, así como el aumento y la globalización del comercio pesquero con su potencial impacto en el abastecimiento local.

Para hacer frente a estos problemas, las Organización de las Naciones Unidas

para la Agricultura y la Alimentación (FAO) ha pedido la aplicación de un enfoque ecosistémico a la gestión de la pesca (EAFM), cuyo objetivo es lograr una explotación sostenible de la pesca comercial, que tenga en cuenta las interacciones entre las artes de pesca y los ecosistemas marinos (Cotter et al., 2009).

Un EAFM tiene en cuenta que la pesca es una componente del ecosistema y no se puede gestionar de forma aislada (Gascuel et al., 2012). Un EAFM debe considerar la aplicación de los principios del desarrollo sostenible en el sector pesquero, la combinación de la sostenibilidad ecológica, la viabilidad económica y la equidad social.

No obstante, aunque se reconoce que la pesca es importante para el desarrollo sostenible y que su contribución puede ser mejorada, la cantidad de información científica objetiva sobre la pesca es limitada y a menudo la información que existe es de difícil acceso.

La fiabilidad del asesoramiento científico para la gestión de los recursos naturales depende en gran medida de la cantidad y calidad de los datos disponibles para la evaluación científica y su interpretación (Jennings and Polunin, 1996). Es cierto que hay una enorme cantidad de datos sobre los ecosistemas marinos y que además son fácilmente accesibles a través de plataformas online (como por ejemplo los datos de teleobservación obtenidos por satélite y observaciones basadas en telemetría automática). Sin embargo, la gran mayoría de los casos tenemos que lidiar con datos limitados e irregularmente espaciados (por ejemplo, de peces y otra fauna marina). Además, los datos no siempre son directamente comparables debido a las variaciones en las condiciones ambientales entre períodos de muestreo. La razón principal es que en investigación pesquera la recogida de datos es comúnmente costosa tanto en tiempo como en dinero. Además, los problemas aumentan cuando el objetivo de la investigación consiste en examinar los cambios en la dinámica de todo un ecosistema mediante el estudio de largas series temporales en una escala macro-espacial. En estos casos, para los que se requiere grandes bases de datos con una cobertura adecuada en el espacio y tiempo de una variedad de variables, la información es casi inexistente.

Por otra parte, la información sobre el estado de la pesca se puede derivar de muestreos diferentes tales como los muestreos independientes de las pesquerías y los muestreos dependientes de la pesca, así como los datos de desembarcos de la FAO (Vasconcellos and Cochrane, 2005). Cada una de estas fuentes ofrece diferentes tipos de información, que debe ser cuidadosamente seleccionada en función del tipo de estudio y los objetivos perseguidos.

Los muestreos de datos independientes de la pesquería se basan principalmente en campañas de investigación realizadas durante períodos de tiempo relativamente cortos. Estos datos se consideran de calidad superior, ya que son independientes de las limitaciones económicas del muestreo dirigido, utilizan procedimientos normalizados de artes de pesca, y la información biológica sobre las especies se tiene en cuenta durante el diseño de muestreo (National Research Council, 2000). Sin embargo, por lo general, tienen una cobertura limitada en el espacio y el tiempo (tanto en términos de estacionalidad y el número de años de datos), lo que podría dar lugar a estimaciones sesgadas e imprecisas (Hilborn et al., 1992). En resumen, hay dos problemas principales con respecto a los datos independientes de la pesca: la del tiempo en relación con el ciclo de las estaciones y la ubicación de las especies objetivo (que conduce a un muestreo representativo, es decir, sesgo) y el hecho de que sólo una cantidad limitada de datos puede ser recogidos.

Los datos dependientes de la pesquería suelen proporcionar series temporales mucho más largas, una amplia cobertura espacial durante todo el año, así como información sobre una gran variedad de especies objetivo, artes de pesca, lugares de desembarque y canales de distribución (Lunn and Dearden, 2006). Los datos de este tipo a veces carecen de detalles específicos, como la ubicación de las zonas de pesca y la identidad de las especies (las capturas son en mucho casos identificados a un nivel taxonómico superior), así como problemas de sesgo debido a las limitaciones impuestas por la dirección y la información errónea deliberada de las capturas. Algunas de las desventajas de los datos dependientes de la pesca se pueden superar mediante el uso de observadores a bordo, pero, inevitablemente, sólo una pequeña parte de la actividad pesquera se puede cubrir de esta manera.

Los desembarcos registrados por la FAO son a menudo la única fuente de información disponible debido a su conexión con las actividades económicas. La problemática principal de esta fuente de información es que en muchas pesquerías los desembarcos y las capturas no son necesariamente lo mismo. Grandes cantidades de capturas no declaradas, descartes y capturas ilegales no se registran en los registros de los desembarques (Vasconcellos and Cochrane, 2005), y no todos los desembarques entran en las estadísticas oficiales (por ejemplo, las ventas directas a los consumidores de individuos que son inferiores a las tallas mínimas, etc.). Además esta topología de datos no tiene ninguna referencia espacial de donde las capturas han sido efectuadas, a diferencia de las dos fuentes de datos precedentemente mencionadas.

Sin embargo, los datos recogidos por los observadores a bordo y en campañas científicas disponen de series temporales relativamente recientes, mientras que los datos oficiales de la FAO son disponible (para la mayoría de países de la FAO), desde el 1950 (www.fao.org).

La cantidad y calidad de la información varía según el país y las pesquerías estudiadas. En la mayoría de los países, hay información detallada accesible y que además está bien procesada. Aun así existen otros países donde la información es escasa o inexistente. En particular, pero no exclusivamente, los países en desarrollo reflejan estas situaciones de escasez de datos, dado que los recursos y las infraestructuras necesarias para una buena recogida de los datos y para su posterior evaluación de manera, no han aumentado al mismo tiempo que el reciente aumento del comercio internacional de los productos pesqueros.

Por otra parte, la necesidad de avanzar hacia un EAFM se ha traducido en un aumento de la necesidad de datos a distintas escalas espacio-temporales y muchas áreas se encuentran actualmente en una situación de escasez datos para su aplicación práctica (Bianchi, 2010). Los problemas aumentan cuando el objetivo de los estudios es el análisis de series temporales largas o con una macro escala espacial, con el fin de examinar los cambios en la dinámica de todo un ecosistema. En todos estos casos, donde los modelos ecosistémicos complejos no pueden ser aplicados debido a la falta de datos detallados, el primer paso hacia la aplicación del EAFM es establecer objetivos más realistas y utilizar herramientas flexibles que pueden funcionar con eficacia a pesar de la incertidumbre y la información limitada (Trenkel et al., 2007).

Dentro de este contexto, el objetivo principal de esta tesis es analizar las diferentes fuentes de información disponibles sobre los recursos pesqueros, con el fin de proponer nuevas metodologías para una gestión pesquera eficaz tanto en situaciones de escasez de datos, de macro escalas espaciales, de series temporales largas y estudios más detallados.

Varios escenarios se han tratado según los datos pesqueros utilizados y el objetivo perseguido. Para cada uno de ellos, una metodología específica ha sido bien desarrollada (cuando no había ningún método disponible) o descrita (cuando ya existía una), mostrando sus ventajas y limitaciones, y explicando cuales podrían ser sus implicaciones para la gestión de la pesca.

El resto de esta tesis se organiza de la siguiente manera. Después de esta introducción, el Capítulo 1 revisa en detalle el concepto y los principales objetivos del enfoque ecosistémico de la gestión de la pesca (EAFM), introduce las problemáticas en ecología pesquera que se discutirán en la tesis, describe las metodologías utilizadas, los distintos tipos de datos pesqueros y los tipos de escalas espacio-temporales con los cuales serán abordadas las problemáticas pesqueras mencionadas.

En el Capítulo 2 se presentan dos enfoques metodológicos diferentes para estudiar la explotación pesquera a través del análisis de unos indicadores ecológicos seleccionados. Estas metodologías son muy útiles en situaciones de escasez de datos, cuando la única fuente de datos son los desembarques pesqueros oficiales, sin ningún tipo de información adicional, y cuando el objetivo del estudio es utilizar una macro escala espacial y series temporales largas. El primer enfoque permite evaluar las tendencias de los indicadores ecológicos a través del tiempo, a través técnicas de suavizado y bootstrap. Estas técnicas permiten evidenciar los cambios debido a la explotación pesquera, si los hay algunos, en la dinámica trófica de los ecosistemas. El segundo enfoque se centra en cómo las distribuciones de los indicadores ecológicos evolucionan con el tiempo, a través de la estimación no paramétrica de funciones de densidad kernel y, de acuerdo con esto, para predecir lo que podría ser sus distribuciones a largo plazo. Se presentan tres casos de estudio diferentes para ilustrar estos métodos.

En el Capítulo 3, se presenta un modelo espacial jerárquico Bayesiano para estimar y predecir la distribución de la pesquería de especies no objetivo con datos recogidos por observadores a bordo. En todos estos casos, la selección de las localizaciones del muestreo no depende de los valores de la variable espacial, por lo que estos son estocásticamente independientes del proceso analizado. Como resultado, la ocurrencia de la especie es un indicador parcial de su patrón de ocurrencia. Se presentan tres tipos diferentes de ejemplos prácticos. Los dos primeros ejemplos utilizan los datos de presencia/ausencia como variables respuesta para estimar y predecir la distribución de especies pelágicas y demersales en las zonas estudiadas. En cambio, el tercer ejemplo, se basa en una variable respuesta Gaussiana, con el fin de abordar un tema de gran actualidad en el sector pesquero: la distribución espacio-temporal de los descartes.

En el Capítulo 4, se presenta un modelo espacial jerárquico Bayesiano para estimar la distribución de las especies objetivo de la pesca a partir de datos de observador a bordo. En estos casos específicos, en los que se recogen los datos siguiendo un muestreo preferencial, la predicción de la distribución de las especies objetivo mediante la aplicación de los métodos geoestadísticos tradicionales, podría conducir a resultados sesgados. La idea básica es la de interpretar los datos como un patrón puntual marcado, donde los lugares de muestreo forman un patrón puntual y las mediciones realizadas en estos lugares son las marcas. Como ejemplo práctico, modelamos la abundancia de merluza europea (*Merluccius merluccius*) en el Golfo de Alicante.

En el Capítulo 5 se propone un modelo espacio-temporal jerárquico Bayesiano para identificar áreas de reclutamiento de especies pesqueras. Presentamos como ejemplo práctico de esta metodología la evaluación de las áreas nursery de la merluza europea (*Merluccius merluccius*) en el Mediterráneo central utilizando datos de campañas científicas.

Por último, en el Capítulo 6 se presentan algunas conclusiones y las líneas futuras de investigación.

Chapter 1

The Ecosystem Approach to Fisheries Management

In this chapter we review the concept and main aims of the Ecosystem Approach to Fisheries Management (EAFM) in detail, introducing the fishing issues that will be discussed in the thesis, outlining the methodologies used, the different fishing datasets and the types of spatial-temporal scales with which these issues can be addressed.

1.1 Overview

Many of the world's fish populations are overexploited and the ecosystems that sustain them are degraded (FAO, 2002). The unintended consequences of fishing, including habitat destruction, incidental mortality of non target species, evolutionary shifts in population demographics, and changes in the function and structure of ecosystems are being increasingly recognized. Fishery management to date has often been ineffective; it focuses on maximizing the catch of a single target species and often ignores habitat, predators and prey of the target species, and other ecosystem components and interactions. The indirect social and economic costs of the focus on a single species can be substantial (Pikitch et al., 2004).

In order to address the critical need for a more effective and holistic management approach, a variety of advisory panels have recommended that ecosystem considerations should be considered broadly and consistently in managing fisheries. The Ecosystem Approach to Fishery Management (EAFM) is a new direction for fishery management, essentially reversing the order of management priorities in order to start with the ecosystem rather than the target species.

The overall objective of EAFM is to sustain healthy marine ecosystems and the fisheries they support. In particular, EAFM should (*i*) avoid the degradation of ecosystems, as measured by indicators of environmental quality and system status; (*ii*) minimize the risk of irreversible change to natural assemblages of species and ecosystem processes; (*iii*) obtain and maintain long-term socio-economic benefits without compromising the ecosystem; and (*iv*) generate knowledge of ecosystem processes sufficient to understand the likely consequences of human actions.

Standardized systems, reference points and control rules analogous to single species decision criteria must therefore be derived and developed. The objective must be to ensure that the total biomass removed by all fisheries in an ecosystem does not exceed the total amount of system productivity, after accounting for the requirements of other ecosystem components (e.g. non target species, protected and vulnerable species, habitat considerations, and various trophic interactions). Maintaining system characteristics within certain bounds may protect ecosystem resilience and avoid irreversible changes (Pikitch et al., 2004).

EAFM must delineate all marine habitats utilized by humans in the context of vulnerability to fishing-induced and other human impacts, identify the potential irreversibility of those impacts, and elucidate sensitive habitats for species for vital population preservation processes. Protecting essential habitats for fish and other important ecosystem components from destructive fishing practices increases fish diversity and abundance (Sainsbury et al., 2000). Thus, ocean zoning, in which the type and level of allowable human activity are specified spatially and temporally, will be a critical element of EAFM.

In addition, EAFM devote particular attention to the protection of nursery areas. The implementation of management measures should aim at reducing the effects of fishing on juveniles and their habitats, and consequentely the spatial identification of nurseries is needed.

Another goal of EAFM is to reduce the excessive levels of bycatch and discards (i.e. killing of non target species or undersized individuals of the target species), because juvenile life stages and unmarketable species often play important roles in the ecosystem (Bellido et al., 2011). Single-species management has been successful at reducing the incidental catch of vulnerable species in some cases, but EAFM would also manage the indirect effects. The annual global discards in commercial fisheries have been conservatively estimated at 27 million tonnes, equivalent to one third of the weight of all reported marine landings in commercial fisheries worldwide (Alverson, 1998). Discard problems can be addressed through a depth knowledge of the factors that influence discard rates as well as of their spatio-temporal patterns in order to implement an efficient fishery management.

Finally, EAFM must manage target species in the context of the overall state of the system, sensitive habitat, vulnerable species, and non target species. Single-species target and limit reference points are still appropriate, but they will need to be modified in the context of these other factors.

The EAFM implementation in systems that differ in levels of information and uncertainty can be applied through the judicious use of a precautionary approach. This means selecting the management targets and limits with caution when information is sparse or uncertain. Greater uncertainty would be associated with more stringent management measures. Because ecosystem management involves a wide range of objectives, great ecosystem complexity and a high level of uncertainty in predicting impacts, EAFM inevitably requires robust statistical methods to obtain the best estimates of uncertainty. Ideally, EAFM would shift the burden of proof so that fishing would not take place unless it could be shown not to harm key components of the ecosystem.

In data-poor situations with little or no information about target species status or ecosystem processes and where sophisticated ecosystem models cannot be applied, EAFM may simply involve using flexible tools that can work effectively despite uncertainty and limited information (Trenkel et al., 2007).

Targeted management and improved data collection (ecosystem-based reference points and measures of system status) for high-priority ecosystem interactions could promote more comprehensive EAFM in the future. With the increased richness of data, management evolves toward a system in which performance indicators for each ecosystem-based objective are monitored. With more data, there could be fewer precautionary measures.

Progression from data-poor to data-rich EAFM will be facilitated by adaptive management and a greater understanding of how ecosystems respond to alternative fishing strategies. Moreover, EAFM should be tailored to the management capacity available and allow for sequential improvements.

New analytical models and management tools will be needed as well. Robust models and trophic measures must be refined and expanded to better account for system-level uncertainties, to derive system level reference points and to evaluate the ecosystem-level consequences of proposed EAFM actions. Because EAFM emphasizes habitat and ecosystem function in the context of fluctuations, advanced models for EAFM should incorporate spatial structure and environmental processes.

In a fishery management plan, the impact of management action would be assessed with respect to the ecosystem as well as individual species. It is entirely possible that a fishery could be considered overfished within the ecosystem plan (ecosystem overfishing) when it is not overfished in a single-species context. This can occur when a forage species that serves as a prey resource for marine predators is also the target of a fishery or when overfishing of large predators causes food web shifts. Rebuilding ecosystems from their degraded state, in turn, might inflict short term economic hardship on fishermen. The transition to EAFM might thus involve compensating fishermen and providing incentives to other stakeholders to support EAFM as a long-term strategy.

It will not be easy to make the transition from an established management system based on maximizing individual species yields to an ecosystem-focused approach that acknowledges the uncertainty inherent in marine ecosystems. The difficulties are not insurmountable, however, and should not delay progress. EAFM should move forward now despite current uncertainties about ecosystems and their responses to human actions because the potential benefits of implementation are as large as or greater than the potential risks of inaction.

1.2 Fishery sampling: options for data collections

The reliability of scientific advice for the management of natural resources is highly dependent on the quantity and quality of data that are available for scientific assessment and interpretation (Jennings and Polunin, 1996). In fishery research, several methods are used to collect fishery and biological data. These include fishery observer data, logbooks, official FAO landings and research survey data. Each one of these has positive and negative aspects and the decision to use any particular one usually depends on their availability and the research objectives purpused.

1.2.1 Fishery-dependent data

Fishery observers

Fishery observer programmes are used worldwide to collect fishery data including biological data, species composition, discards, etc. This is the preferred means of gaining accurate and in-depth data, but it is more costly than other fishery-dependent data gathering methods. Observers receive training in collection and sampling techniques from fishery professionals involved with, and often employed by, the fishery organization that manages the fishery. Observer programmes tend to be introduced after a fishery has demonstrated a decline, but their use is a wise monitoring strategy in well-managed or developing fisheries as well. The number of data observers depends on the goals of the management organization.

Under the European Union (EU) Data Collection Framework (DCF), EU members are obliged to collect economic, biological and transversal data. The first DCF ran from 2000 to 2008, while the new framework is covering the period 2009 to 2013.

Data are collected on board a commercial fishing vessel only with the agreement of its skipper, and no financial or other compensatory incentive is provided to encourage vessels to carry an observer.

The selection of vessels for sampling trips is not random in the strict statistical sense. In the early years, the sampling trips were carried out on vessels operating from ports in the locality of the observer's base. Currently, sampling is based on a métier approach (subdivision of a fishery by vessel type), with métiers defined with regard to the DCF (EC Regulation 199/2008), which is based on landing percentage composition, type of gear used, mesh size, and area of operation. Trip selection within a métier is quasi-random, because practical considerations arise (e.g. cooperation of the vessel's skipper) and as a result not all vessels have an equal probability of being selected. Such considerations could compromise the estimation of true variance and bias.

Data-collection protocols and management procedures were established at the outset, but have been refined over time. In 2009 the métier approach involved a shift in the sampling strategy, which changed from a specific species sampling to a concurrent sampling of all commercial species (retained and discarded), prioritizing target species. Fishing trips were defined as the primary sampling units, and hauls within a fishing trip as secondary sampling units.

Biological information (i.e. lengths, weights and otolith samples) are collected from the catch, together with vessel, type of gear used, geographical position and environmental attributes (depth, bottom type). For each observed haul, an estimate of the total catch weight is made by the fishermen and the observer in collaboration. The total catch is then sorted into the retained and discarded components by the commercial fishermen. The total weights of each individual species retained are recorded. If the abundance of a species is small, total numbers and lengths are recorded, otherwise a subsample is taken, numbers and lengths recorded and raised accordingly. The total weight of the discarded portion is approximated, a subsample taken, and then sorted by the observer into species. The total weights and numbers of each discarded species in the subsample are determined and raised based on the total approximated discarded weight.

Logbooks

Logbooks are used in many fisheries but logbook data can be highly variable and therefore suspicious. Despite this, logbook data are commonly used in stock assessments and are one of the major data collection sources in numerous fisheries. Fishermen are required to fill out logbooks while at sea. The following data can be recorded in logbooks: species identification, number caught, type of gear used and amount used, location, time of set and haul back, depth and water temperature. It is widely recognized that fishermen do not always record data about their catches accurately and frequently identify species incorrectly. Fishermen busy bringing in and working on their catch are unlikely to record accurate data at the expense of fishing productivity. Many fishermen do not fill in their data at the time of fishing and recreate data from memory at later dates. Fishermen's illiteracy is a problem in some regions. The correct identification of species is a major issue, because most fishermen are not scientifically trained in proper identification techniques. In addition, many fishermen dislike any type of management planning and are unwilling to go out of their way to collect data (Lordan et al., 2011). Finally, there may be no quality control of logbook data, with no on-board monitoring of logbook entries. However, this type of data collection is inexpensive and is often the only method available.

FAO official landings

In the 1990s, the FAO completely revised the fishery production statistics time series available, computerizing them back to 1950, disaggregating data by fishing areas, taking account of political changes (e.g. the emergence of new countries), adjusting species identification (as taxonomy evolves), and improving the discrimination between aquaculture and capture fisheries production. In cases where data are missing or are considered unreliable, the FAO includes estimates based on the best available information from any source, such as regional fishery organizations, project documents, industry magazines, or statistical interpolations. The resulting data sets have been made widely available on the FAO web (www.fao.org). These data sets consist of recorded nominal catches and do not include discards. In the areas in which the FAO has not yet had the means to work effectively e.g. production from illegal fishing, there is no information at all at global level. The great advantage of the FAO's landings statistics is that they are global in coverage, have complete time series since 1950 and are regularly updated, so they can be used to provide overview trends in fisheries by region and to provide resource status indicators when other data are lacking.

1.2.2 Fishery-independent data

One of the major problems with using commercial catch data for estimating stock distribution and abundance is that fishermen go where the fish are. Fishing effort is normally concentrated on the highest densities of fish and attempts to assess the range or total abundance of fish from commercial catch and effort data can be expected to be biased. Management agencies normally attempt to avoid the biases of commercial catch data by using research surveys, either by chartering commercial fishermen or using government research vessels.

Indeed, fisheries independent data include surveys that are not directly influenced by harvesting activities and provide critical information on the status of fish and shellfish stocks (Morgan and Burgess, 2005). Fishery surveys are specifically designed to satisfy multiple objectives and are often cross disciplinary in nature. In addition to collecting fishery data, the surveys provide important information on the wider marine environment.

Fishery-independent surveys provide valuable measures of relative abundance, rates of population change and size and sex composition for a wide range of species. As these measures are obtained from scientific sampling or within an experimental design, they are less subject to the unknown and often confounding factors that complicate the interpretation of fishery-dependent indices of stock status. The major problem with surveys is cost. Operating vessels are so expensive that the number of samples is not usually nearly as large as would be desired by the scientific staff using the data.

Fishery surveys can been divided into two main types: monitoring surveys for pelagic species and monitoring surveys for demersal species. The main difference between these species is that pelagic fish live near the surface or in the water column, in contrast to demersal fish, which live on or near the bottom.

In order to determine the relative abundance of the target pelagic species, acoustic surveys are usually carried out. This information is then used to determine catch rates and management advice for the following year. They are generally carried out on spawning and pre-spawning aggregations of fish. Outside the spawning season many pelagic species are generally very scattered over a large geographical area and difficult to detect using acoustic methods. These surveys use sound waves emitted from a *transducer* to estimate the density of plankton and fish shoals. The survey vessel tows the transducer under water, and it is linked to an echo sounder in the vessel which records the shoals of fish as *marks* on a screen or paper trace. The species composition of each mark is then identified by taking samples. The density and number of marks are then converted into the biomass (weight) of the different species.

Systematic bottom trawl survey surveys intend to produce basic information on benthic and demersal species in terms of population distribution as well as demographic structure, on the continental shelves and along the upper slopes at a global scale.

During the survey, the scientists use a trawl net to fish at different depths. The species caught in the trawl are brought on board, identified and measured. For commercial species, as well as species of specific scientific interest, other biological information such as weight, age, sex and maturity is also collected.

The major advantage of survey use of fishing gear over commercial data is that the survey can use a planned sampling design rather than relying on fishermen's choices of where to fish. Additionally, conditions of gear use (tow speed, soak time, etc.) can be better controlled. The use of research survey fishing is well established, particularly for trawl gear. The major problem with fish gear surveys is estimating the proportion of fish captured by the gear. Trawl surveys normally take the area swept by the gear, and expand the catches by the proportion of the area surveyed.

1.2.3 Fishing gear and methods

Fishing gear is defined as the tools used to capture marine/aquatic resources, whereas how the gear is used is the fishing method. Additionally, a single type of gear may also be used in multiple ways. Different target species require different fishing gear to effectively catch the target species. Fishing gear falls into two general categories, active gear and passive gear. Active gear is designed to chase and capture target species, while passive gear generally sits in one place, allowing the target species to approach the capture device.

A common way to classify fishing gear and methods is based on the principles of how the fish or other preys are captured and, to a lesser extent, on the construction of the gear or the gear materials used. Following the FAO's definition and classification, the main categories of fishing gear are as follows:

- *Surrounding nets*

  The net is roughly rectangular in shape without a separate bag. It is set vertically in water to surround the school of fish, generally of a pelagic nature. The nets are subdivided into three categories: a one-boat seine, a two-boat seine; and a surrounding net without a purse line.

- *Seine nets*

  Seine nets are cone-shaped nets with two wings where in the wings are normally larger than those of trawl nets. The net is pulled towards a stationary boat or onto a beach.

- *Trawl nets*

  A conical bag-shaped net with two or more wings, pulled by one or two boats for a period of time, to catch mainly demersal fish or other aquatic animals that live directly on, or stay near, the seabed. The trawl is subdivided into three major types: bottom trawl, pair trawl and beam trawl.

- *Lift nets*

  A sheet of net, usually square, but which may sometimes be conical, mounted either by several rods and ropes or on a frame and is either at the bottom or in mid-water for some time and then lifted to trap the fish lying above it.

- *Gill nets*

  Gill nets are curtain-like nets that are fitted with sinkers on the lower end and floats on the upper end and are set transversely to the path of migrating fish. Fish trying to make their way through the net wall are entangled, gilled or enmeshed in the mesh.

- *Scoop nets*

  A bag-shaped net with a fixed or variable opening and is usually operated in shallow waters. The gear catching mechanism is operated by filtering a certain volume of water and trapping the fish into it in a scooping manner.

- *Traps*

  The gear is set or stationed in the water for a certain period of time to trap moving fish in the water. Trapping is done with the use of a non-return valve fitted in the entrance of the gear. The gear may or may not include a netting material.

- *Hooks and Lines*

  The gear generally consists of line(s) and hook(s) to which artificial or edible baits are attached to lure and catch fish or other aquatic animals.

- *Other Fishing Gear And Methods*

  This group covers a variety of fishing gear and methods not classified into the above groups, such as the use of gleaning along the shore for shellfish, seaweed or fish.

Fishing methods have continuously evolved throughout recorded history. Fishermen are inventive and not afraid of trying new ideas. The opportunities for innovation have been especially good in recent decades with advances in fibre technology, the mechanization of gear handling, improved performance of vessels and motorization, computer processing for gear design, navigation of aids, fish detection, to mention only a few technologies (FAO, 2002).

In the past the technological development of fishing gear and methods was aimed at increasing production. On the contrary, nowdays, gear development is focussed on selective fishing and gear with less impact on the environment, due to the present situation of overfished stocks and concerns about the environmental impact of the fishing operation.

1.3 The importance of the spatial and temporal scales

With the development of an EAFM, research has increasingly focused on the environmental effects of fishing in marine habitats as well as on management scenarios or options (e.g. effort reduction, closed areas, technical measures) that take these impacts into account (Sainte-Marie et al., 2009). Therefore, more than in traditional fishery management, this produces the need for a reappraisal of the issue of scale in space and time when estimating the impact of fishing, which, in most cases, boils down to one question: what is the appropriate scale for the study of a particular fishery problem?

Many studies that estimated the impact of fishing emphasized the importance of the spatial and temporal component even though the scale on which they studied the relationship differed. Problems range in scale from local to global. Moreover, local effects might scale-up cumulatively to be more-than-additive (Leuven and Poudevigne, 2002) and some phenomena apparent on one scale can either disappear or be subsumed at other scales. Therefore, conclusions about fishery phenomena from one scale might be modified partly or completely on other scales (Thompson and McGarigal, 2002).

Dinmore et al. (2003) studied the microscale distribution of fishing on the scales of 1 x 1, 2 x 2, 4 x 4, and 8 x 8 nautical miles (*n.mi.*) and observed that the spatial scale may have a critical effect on any interpretation of the impact of fishing.

Ideally, diagnosis and response require explicit recognition of scale hierarchies, but this is challenging for several reasons.

First, fish biology is scale-dependent across life stages or activities. Examples are daily feeding, which varies between habitats, and seasonal migration which extends within or beyond the whole ecosystem. These scale dependent activities vary within and between fish species.

Secondly, although fishery research increasingly addresses scale, studies relating fish (and many other organisms) to their environment overwhelmingly involve small spatio-temporal extents. At the other extreme, true large-scale patterns in marine systems are often obscured by non-independent sampling, weak inference and models that are poorly tested or over-extrapolated (Vaughan and Ormerod, 2005). These challenges prevent fishery managers from quantifying the importance of large-scale, anthropogenic disturbances.

Thirdly, the restoration and conservation of fish communities are limited by scale and complexity. Actions often involve small segments, thereby excluding larger-scale environmental processes that often control fish distribution (Ormerod, 2004). This is because the proper selection of the spatial-temporal scale of a research should be detected before data collection. Ideally, this would be accomplished beforehand via a benchmark survey, which would determine whether the pattern or process studied was on a local, regional or global scale. This is rare in real field research, however, due to the lack of time and funding and consequentely the availability of data driving the scale selection.

In observation, scale determines the relative fineness and coarseness of different details as well as the selectivity among the patterns that these data may form. Thus scale becomes a filter, or a window of perception through which analysis, observation, knowledge, and information can be considered and/or defined.

Obviously, aggregated global syntheses necessarily leave out local details, because the quality and quantity of data available for a local scale are usually higher than a global scale. In addition, there is a positive correlation between the spatial extent of a study and the time window that it addressed. This pattern is consistent with the expectation that processes with relatively coarse spatial extents will also have relatively long temporal windows. Consequentely, long time series, such as macro-scale studies, do not provide the same detailed information that recent time series ensure.

As a result, each type of scenario provides a different approximation. Large-scale and long time series studies give a broader and more complete view of an ecosystem's status, an outcome which is usually not achievable in the case of smaller-scale studies. On the contrary, specific studies in space and time offer very detailed information on a particular habitat or species, while a meso-scale analysis indicates a scale size that falls between the micro and macro-scale, and may also reveal connections between them. Different scales contribute an understanding of different processes, or alternative views of the same process.

In this context a multiscale approach is the best option for improving our understanding of the fishery dynamics of fisheries and marine resources. The use of a multiscale approach would provide information benefits by improving the assessment findings, and their applicability, on all scales.

1.4 Can we implement the Ecosystem Approach to Fishery Management in data-poor fisheries?

Many fisheries worldwide have limited data, particularly but not exclusively in developing countries. In some circumstances even catch statistics might not be reliable and effort statistics may not be available. Many of these fisheries often have only very general or no clear management objectives, and infrastructure and resources are insufficient to support comprehensive and continuous data collection, scientific research, and fishery management.

Moreover, data-poor problems increase in long time series and macro-scale studies, which have the main purpose of examining changes in the dynamics of a whole ecosystem as the principles of EAFM require.

Finally, poor communication and coordination in policy formulation and the implementation of management plans among agencies and stakeholders are common problems that result in data-poor fishery situations (Hilborn et al., 1992).

Implementing an EAFM in these fisheries should still be feasible and needs first to set the objectives realistically and to develop management strategies that can operate effectively with limited information in the face of uncertainty and to monitor the fishery cost-effectively. In all these cases a shift is from require model-based to qualitative common-sense, and data- and knowledge-based indicators corresponding to management objectives are required (Trenkel et al., 2007).

One approach, for this purpose, is the development and selection of a suitable set of ecological indicators that can provide a readily understood, cost-effective tool for assessing trends regarding sustainable development objectives without requiring extensive and costly data collection and complex modelling for describing the current state of fishery resources and fishery activity.

Indicators are succinct measures that aim to describe as much of a system as possible in as few points as possible, in order to understand, evaluate and improve it. In order to be useful, ecosystem indicators should be sensitive to changes in ecosystem integrity through space and time, easily measured, understandable, informative and based on accessible data. Indicators do not necessarily need to be precise; they just need to give a reliable picture of the ecosystem they represent.

At the macro-scale level, countries can use indicators to produce a holistic picture of the fisheries sector and its environment, while at the micro-scale level, indicators provide an operational tool in fisheries management, as a bridge between objectives and management action.

Like any reductionist approach, an indicator must be understood in context. An indicator rarely captures the complete richness and complexity of an ecosystem, but a set of indicators could greatly improve the situation, especially in all the data-poor cases in which they are the only tools that can be used.

1.5 Towards an Ecosystem Approach to Fishery Management: Marine Spatial Planning

Marine Spatial Planning (MSP) is an emerging tool to support the implementation of an EAFM. By means of managing current and future sea uses, marine plans assist in solving potential conflicts between multiple uses of the marine environment (Douvere, 2008).

Despite its broad acceptance and wide range of principles, definitions and guidelines, the EAFM is still more a concept which is widely discussed at scientific fora, but with few examples of actual practice. It is increasingly clear that governments and stakeholders lack the necessary tools to make an ecosystem approach operational in marine environment.

A range of tools and measures will be needed to implement the multiple objectives of an EAFM, but a focus on the spatial and temporal aspects of EAFM is one way to make this approach more tangible. MSP can do this because it:

- *Addresses the heterogeneity of marine ecosystems in a practical manner.*

  MSP takes into account that some things only occur in certain places. Important ecological areas, for example, are located in areas of high diversity, sensitive habitats of vulnerable species, spawning and nursery areas, and migration stopover points. At the same time, economic activity will (and can) only take place where the marine resources are located. Acquiring spatio-temporal knowledge of marine resources could allow the management of this overlap.

- *Focuses on influencing the behaviour of humans and their activities over time.*

  Although goals and objectives for a certain area are usually set for both ecosystem/natural processes and human activities, it is only the human component (human activities and resource use) that can be managed (not the ecosystem itself), e.g. through management measures (incentives) that change the behaviour of fisheries activities over time.

- *Makes conflicts and compatibilities among human uses visible, and therefore tangible.*

  Through the mapping of ecosystems, their characteristics and the human activities affecting it, one can see where conflicts are or will be located.

The place-based characteristics of ecosystems, natural resources, and human activities affecting them, increases the need to look at the “system” from a spatial and temporal perspective and implies that all policies and management strategies (e.g. fisheries management, marine transportation management, and marine protected area management) directed toward influencing the human use of ecosystems and their resources will inherently have a spatial and temporal dimension.

In order to achieve these aims, the MSP requires an integrated assessment of multiple marine plan objectives and potential conflicts of use, and the risk of the combined effects of multiple uses on the marine environment. Such an integrated assessment requires practical tools to both visualise spatially complex relationships and assess marine spatial planning scenarios under given planning objectives. In turn, the analysis of spatial relationships and planning options requires sound methods by which to assess interconnectivity between pressures and the response of the marine environment within the study area of interest.

To date, a number of studies have assessed empirically the spatial and temporal patterns of multiple human uses and their combined pressures and/or impacts on ecosystem components to support an EAFM (Dalton et al., 2010).

Fundamental to any spatially-explicit risk assessment is information on the sensitivity of the environment to a pressure and a measure of the frequency of the occurrence of a pressure. While for most physical pressures empirical data on their temporal and/or spatial pattern exist, a measure of ecosystem sensitivity is often expressed as expert knowledge or another qualitative measure (Stelzenmüller et al., 2010). Both components of such a risk-assessment model introduce uncertainty into the analysis of the relationship between human pressures and adverse effects on the ecosystem. As a MSP describes the spatial and temporal allocation of resource use, it is crucial to assess the uncertainty associated with the data used, and to visualise the uncertainty associated with the outcomes of possible spatial management scenarios.

The three dimensions of uncertainty in a model-based decision support tool are described by Walker et al. (2003) as location (where the uncertainty occurs in the model complex), level (where the uncertainty occurs on the gradient between knowledge and ignorance), and nature (whether uncertainty is due to knowledge gaps or to variability inherent in the system). Thus the development of spatial management scenarios to support marine planning requires a spatially-explicit framework that incorporates various sources of uncertainty.

By utilizing an appropiate statistical approach that accounts for uncertainties in such settings, one is able to obtain accurate estimations and predictions (in space and or time) and in-depth knowledge of marine resources.

1.6 Spatial analysis: need for Marine Spatial Planning

In order to be applicable, the Marine Spatial Planning process requires the use of software tools and well-defined spatially-explicit methodologies.

The first step in understanding ecosystem processes is to identify spatial patterns. Fishery data are usually characterized by spatial structures due to spatial autocorrelation. Spatial autocorrelation refers to the pattern in which observations from nearby locations are more likely to have a similar magnitude than by chance alone. The magnitude, intensity as well as extent of spatial autocorrelation can be quantified using spatial statistics. Most fishery data exhibit some degree of spatial autocorrelation, and this influences the interpretation of statistical models by affecting tests of signficance of the association between distributions of species and environmental factors as well as calculating correlations among such variables (Babcock et al., 2005).

In addition, geographically contagious biotic processes (such as population growth, geographic dispersal and competition dynamics) can also promote spatial autocorrelation in the distributions of species. In many instances, these two influences operate simultaneously, inducing spatial heterogeneity in ecological communities (Babcock et al., 2005).

The goal of employing spatial statistical models is to account explicitly for the effects of these two sources of influence. Taken together, space can be seen as a predictor, when the goal is to exploit the mechanisms (e.g. shared common factors, geographic diffusion) that generate spatial autocorrelation in distributions of species in space, and/or as a covariable when the goal is to adjust for spatial variation when testing for associations between distributions of species and environmental factors.

Once the spatial structures of the studied variable are assessed, and the spatial statistical model (such as linear, generalized or additive models) is used to relate the observed structure to the hypothesized generating processes, prediction tecniques (such as kriging) are applied to obtain spatial estimation of the observed variable at any point in the studied area. In this way, predicted spatial maps of marine resources are generated as efficient tools for fishery management.

Nevertheless, different issues arise in spatial prediction. First, there are problems related to sampling. In general, the spatial area covered by the resources is vastly larger than the area sampled and, correspondingly, the spatial unit of prediction is usually larger than the sites sampled on the ground. An additional sampling problem relates to the heterogeneity of sampling intensity: while large parts of a domain may be unsampled, other parts may be relatively heavily sampled. There may also be sample bias in the available data, so that sampling locations and/or intensity are unevenly distributed with respect to relevant characteristics of the region sampled (Latimer et al., 2006). These problems are often ignored in spatial modeling, or are addressed indirectly by attempting to minimize bias through stratified sampling across major environmental gradients (Elith et al., 2006).

Second, as we mentioned above, there is the problem of spatial dependence or spatial autocorrelation. Using models that ignore this dependence can lead to inaccurate parameter estimates and inadequate quantification of uncertainty. Equally important, to ignore this spatial dependence is to throw out meaningful information (Wikle 2003). Ecological prediction often ignores this problem or deals with it in a less than satisfactory way (see Guisan and Zimmermann (2000)). One might, for example, include latitude and longitude through a trend surface in the mean to improve prediction, but this approach may still miss spatial dependence which explicit modeling of spatial association can capture. Another alternative provided by the generalized regression analysis and spatial prediction (GRASP) modeling package treats spatial autocorrelation at the data stage, by feeding into the model a new data layer that reflects neighborhood values in model predictions (software available online). This approach does deal with autocorrelation, but fails to quantify explicitly the strength of spatial patterns in the residuals and its associated uncertainty.

Third, spatial modeling presents problems in quantifying uncertainty. Predictions frequently involve extrapolation to unobserved parts of the study region and to larger-scale areal units. Assessment of uncertainty in such predictions is crucial when they are used to set conservation policy or to evaluate the impact of climate change on species.

In this way, Bayesian spatial models can be used to deal with these three kinds of problems in a straightforward, transparent way and as practical tools to allow the implementation of MSP.

1.7 The Bayesian framework as a practical tool to support Marine Spatial Planning

Mathematical models are commonly used to describe fishery data (Hilborn et al., 1992). To relate a model to data observed in a fishery, an appropriate method is required to estimate parameters in the model. In general, there are two statistical approaches that can be used for parameter estimation: frequentist and Bayesian approaches. The statistical problem is similar for these two approaches: both are used to make statistical inferences about unknown parameters in the model (Gelman et al., 2004).

Frequentist inference is commonly used in fishery studies (Hilborn et al., 1992). It assumes that the parameters being estimated are a fixed constant and that data are random observations from some unknown statistical population. An objective function needs to be defined based on assumptions made on random variables. Parameters and their confidence intervals can then be estimated by optimizing the objective function.

Bayesian inference has been used increasingly in fisheries (Zhang and Holmes, 2010; Cao et al., 2011; Viana et al., 2011). This approach assumes that parameters are random as opposed to constant for frequentist inference. Instead of estimating the "true" values of the parameters as in frequentist inference, it only looks at the statistical distributions of the parameters values.

Bayesian modelling uses Bayes' theorem to combine the information in the data with additional, independently available information (the prior) to produce a full probability distribution (posterior distribution) for all parameters (Gelman et al., 2004). The Bayesian perspective allows us to ask directly how probable the hypotheses are, given the data.

The posterior probability distribution provides a full picture of what is known about each parameter based on the model and the data, together with any prior information. For a particular parameter, this posterior distribution, unlike the mean and confidence interval produced by classical analyses, enables explicit probability statements about the parameter. Thus the region bounded by the 0.025 and 0.975 quantiles of the posterior distribution has an intuitive interpretation: under the model, the unknown parameter is 95% likely to fall within this range of values.

One feature of Bayesian models, and also a source of much debate within the statistics community, is their ability to incorporate already-known or "prior" information about parameters into the models. In some applications, this can be a critical advantage, particularly when data are sparse, decisions must be made, and expert opinion is available (Gelman et al., 2004). However, when the primary interest in a particular study is to learn what information is contained in the data, the use of vague or uninformative prior distributions allows posterior distributions for all parameters to be driven by the data.

In the spatial context, the Bayesian approach is particularly appropriate because it allows both the observed data and model parameters to be random variables (Banerjee et al., 2004), resulting in a more realistic and accurate estimation of uncertainty. Spatial autocorrelation can be incorporated into a regression model through random effects that capture spatial dependence in the data. Since the random effects are model parameters, they also emerge with a full posterior distribution that allows the quantification of uncertainty. By adding hierarchical levels to a regression model, issues of sampling intensity and gaps in the data can be dealt with explicitly (Gelfand et al., 2006).

Hierarchical models are statistical models in which data can enter at various levels and thus model parameters or unknowns are themselves functions of other model parameters and data. These hierarchical stages can describe conceptual but unobservable latent processes that are ecologically important, such as error in the observation process. In this way, uncertainty attached to unknowns at different model stages is propagated across model levels to more accurately reflect overall inferential uncertainty.

Bayesian spatial models may also aid analyses of data with geographically uneven levels of survey effort because such a bias can be incorporated within the spatial random-effect term, thereby reducing its influence on estimates of the effects of environmental variables. By treating spatial effects as a variable of interest, hierarchical Bayesian spatial models can suggest the identity of additional environmental covariates that may improve model fit or the existence of area effects that may limit population viability.

Crucial to the Bayesian approach is the notion of transparency. One reaction of ecological modelers to the problems inherent in spatial analysis is to adopt more flexible methods like neural networks, random forest, and discriminant analysis (Moisen and Frescino, 2002). These methods offer the advantages of responding flexibly to interactions and nonlinearities in data relationships, but often at the expense of interpretability or mechanistic insights.

Implementing an effective MSP in order to establish an EAFM requires an accurate knowledge of the spatial-temporal distribution of marine resources in relation to the environment, and Bayesian methods could be the appropriate tools.

The Bayesian framework can support the decision making process by helping to provide informed decisions, through the assessment of potential outcomes and related uncertainty from management measures in a spatial context, and by offering a visualisation tool that facilitates the engagement of different stakeholders in such a process.

Chapter 2

Ecological indicators for data-limited fisheries

In this chapter, we present two different methodological approaches to study fishing exploitation through the analysis of selected ecological indicators. These methodologies are very useful in data-poor situations, when the only available source of data are official landings without any additional information, and for macro-scale and long data series studies.

The first approach allows us to assess the trends of the ecological indicators over time through bootstrap and smoothing tecniques, highlighting shifts, if there are any, in the ecosystem trophic dynamic due to fishing exploitation.

The second approach focuses on how distributions of ecological indicators evolve over time, through the nonparametric estimation of Gaussian kernel density functions and, according to this, to show how the probable long term distribution might be. Three different cases study are presented to illustrate these methods.

2.1 Introduction

An Ecosystem Approach to Fisheries Management (EAFM) extends fishery management to all the aspects of fisheries. This extension increases information needs and the complexity of the system under consideration and therefore also increases model complexity and the multiplicity in stakeholders and objectives as well as the uncertainty involved in assessing ecosystem states and the outcomes of management scenarios. On the contrary, data-poor fisheries do not have sufficient data to support complex assessments and models. To overcome the data limitation and inability to make use of sophisticated modelling, ecological indicators may prove the most realistic approach.

Many structural and functional properties of ecosystems are difficult to measure directly, and therefore easier-to-measure indicators can be used as proxies for the difficult-to-measure properties. For data-limited fisheries, the indicators should (1) be observable and understandable by all stakeholders, (2) be based on easily obtainable and reliable data, and (3) adequately reflect the condition of the resource.

In order to assess the fishing exploitation on a whole ecosystem, fishery ecological indicators can be used to monitor trends in condition over time, and to provide an early warning signal of changes in the ecosystem.

In particular, fishery-based trophic indicators can describe the stages of exploitation of resources and analyze the state of the ecosystem relative to past periods for which there are no other data available except the amount of landings from fish markets (Pauly et al., 1998; Caddy et al., 1998; Stergiou and Karpouzi, 2002).

According to Pauly and Watson (2005), in marine ecosystems, it would be possible to use the Mean Trophic Level (MTL) of landings as an index of sustainability of the level of exploitation of fish resources. Pauly et al. (1998) used aggregated landing statistics from the United Nations (UN) Food and Agriculture Organization (FAO) and estimates of Trophic Level (TL) for individual species derived from food web models, to demonstrate that the MTL of global fisheries has declined significantly since the late 1950s.

This phenomenon, known as *Fishing Down the marine Food Webs* (FDFW), occurs because the species most susceptible to collapse and generally more often targeted are those of greater size with long life cycles; once these stocks become depleted, exploitation is directed toward smaller sized species with a faster growth rate, and thus the MTL of landings decreases (Pauly et al., 1998).

Nevertheless, several criticisms have been made on the interpretation of the MTL indicator and, in particular, Branch (2010), demonstrated the potential danger to interpret the trend of a single indicator to assess the ecological health of an ecosystem, which could lead to misleading conclusions. Indeed, the interpretation of an indicator could be excessively subjective. Points of reference or limit values with which to unequivocally assess the results obtained with an indicator are very difficult to establish due to the complexity of ecosystems. Secondly, the indicators, by definition, must be simple and general, and as a result might overlook some key ecological features in the assessment of the ecological balance of an ecosystem.

In order to obtain a comprehensive view of ecosystem dynamics it is essential to assess a number of indicator trends jointly.

In this chapter we propose a set of selected ecological indicators derived from Trophic Level (TL) information to assess the ecological balance of marine ecosystems in data-poor situations and for long time-series. Two different methodological approaches to study the fishing exploitation through the analysis of these indicators are presented.

Firstly, a combination of smoothing techniques and bootstrapping is shown in order to highlight the changes in the ecosystem trophic dinamic. Smoothers are non-parametric estimators that produce smooth estimates of regression functions. The bootstrap (Efron, 1979) is a computer-intensive approach that can provide measures of uncertainty (confidence intervals, standard errors, etc.) for a wide range of problems. It is based on the basic idea of repeated resampling with replacement from an original sample of data in order to create replicate datasets from which inferences can be made on the quantities of interest. Thus for the trophic level indicators, bootstrapping allows resampling of the data to create multiple series which have similar stochastic structure to the original series.

Moreover, both discards and Illegal, Unreported and Unregulated fishing (IUU) may represent a significant source of error in assessments of fishery dynamics (Fernández, 2009) and they could also influence the performance of indicators based on the trophic levels of landed fish. In the absence of data on IUU and discards, these effects could be investigated using simulations.

The use of a combination of ecological indicators together with bootstrap and smoothing techniques and discard simulations, could be a first step in the EAFM implementation in data-poor fisheries situations. We applied this first approach to two different Large Marine Ecosystems (LMEs) with a macro-scale study level to obtain a wider view of these ecosystems dynamics.

LMEs are regions of the world's oceans, encompassing coastal areas from river basins and estuaries to the seaward boundaries of continental shelves and the outer margins of the major ocean current systems. They are relatively large regions on the order of 200,000 km^2 or greater, characterized by distinct bathymetry, hydrography, productivity, and trophically dependent populations. The system of LMEs has been developed by the US National Oceanic and Atmospheric Administration (NOAA) to identify areas of the oceans for conservation purposes. The objective is to use the LME concept as a tool for enabling ecosystem approach to provide a collaborative way to management of resources within ecologically-bounded transnational areas.

LME-based conservation is based on recognition that the world's coastal ocean waters are degraded by unsustainable fishing practices, habitat degradation, eutrophication, toxic pollution, aerosol contamination, and emerging diseases, and that positive actions to mitigate these threats require coordinated actions by governments and civil society to recover depleted fish populations, restore degraded habitats and reduce coastal pollution. Although the LMEs cover only the continental margins and not the deep oceans and oceanic islands, the 64 LMEs produce 95% of the world's annual marine fishery biomass yields. Most of the global ocean pollution, overexploitation, and coastal habitat alteration occur within their waters.

Secondly, a meso-scale level study is proposed to accurately capture the fishing exploitation in the Mediterranean ecosystem. Nonparametric density estimation, stochastic kernel distributions and transition probability matrices are used in a complementary way. This methodology allows to ascertain whether there is a tendency for Mediterranean countries to converge in their characteristics, either towards the best or worst practice fisheries, and consequently, on ecosystems impacts.

2.1.1 Fishery ecological indicators:

Fishery ecological indicators quantify the magnitude of stress, degree of exposure to the stresses, or degree of ecological response to the exposure and are intended to provide a simple and efficient method to examine the ecological composition, structure, and function of complex ecosystems (Dale and Beyeler, 2001).

In order to be useful, fishery ecological indicators should be sensitive to changes in ecosystem integrity through space and time, easily measured, understandable, informative and based on accessible data. Indicators do not necessarily need to be precise; they just need to give a reliable picture of the ecosystem they represent.

Following these criteria, we have selected three different fishery ecological indicators:

The Marine Trophic Index

The Marine Trophic Index (MTI) was developed based on the assumption that a decline of the mean trophic level of fisheries catches (mTL=MTI) is generally due to a fishery-induced reduction of the biomass and hence reduced biodiversity of vulnerable predators (Pauly et al., 1998). The MTI tracks changes in mTL and is calculated from a combination of fisheries landings and diet composition data for the landed fish species. It is computed, for each year k from:

$$mTL_k = \frac{\sum(TL_i)(Y_{ik})}{\sum(Y_{ik})} \tag{2.1}$$

where mTL is the mean trophic level of all landings in year k, Y_{ik} refers to the yield (landings) of species i in year k and TL_i is the Trophic Level of species i.

Changes in this index could provide useful indications of changes in the landings of high trophic level species that are usually the target fishery species (Vivekanandan et al., 2005).

The Marine Trophic Index$^{3.25}$

The use of the MTI as a measure of the impact of fisheries on marine ecosystems was questioned by Caddy et al. (1998), who had in mind ecological processes that may result in a natural increase in the abundance of planktivores, thus lowering mean TL. As a diffuse and general problem eutrophication can modify the ration between predator and prey abundances, which then could be confused with effects of fisheries. To overcome this problem Pauly and Watson (2005) suggested that the MTI should in fact be based on time series that exclude low-TL organisms. This would lead to an indicator labeled as “cutMTI”, with the “cut” referring to the lowest (cut-off) TL value used in the computation. Pauly proposes the cut-off value of 3.25 (MTI$^{3.25}$). With a cut-off value of 3.25, all species (or groups of species) with TL lower than 3.25 are removed from the computation of the MTI, in order to eliminate the herbivores, detrivores and planktivores whose biomass tends to vary widely in response to environmental factors (Pauly and Watson, 2005).

Fishing in Balance Index

The average efficiency of energy transfer is 10% between trophic levels (Pauly and Christensen, 1995). Pauly et al. (2000) predicted that a fall of one trophic level at which a fishery operates would lead to a 10-fold increase in potential catches. To study this effect Pauly et al. (2000) and Christensen (2000) introduced the Fishing-in-Balance (FIB) index as following:

$$FIB_k = \log\left(Y_k \frac{1}{TE}^{MTI_k}\right) - \log\left(Y_0 \frac{1}{TE}^{MTI_0}\right), \tag{2.2}$$

where Y corresponds to landings in year k, TL is the mean TL of the landings in year k, TE is the transfer efficiency (here set at 0.1 following Pauly et al. (2000)), and 0 refers to any year used as a baseline to normalize the index (Pauly et al., 2000; Christensen, 2000; Cury et al., 2005).

The FIB index has the property of increasing if catches increase faster than would be predicted by TL declines, and of decreasing if an increase in catches fails to compensate for a decrease in TL. This is due to the fact that, in the absence of geographic expansion or contraction, and with an ecosystem that has maintained its structural integrity, moving down the food web should result in increased catches (with the converse being true in the event of an increase in TL), with the FIB index remaining constant.

Pelagic/demersal index

Changes in the trophic composition of marine communities can be tested in terms of large trophic groups such as planktivorous, benthivorous, or piscivorous animals (Caddy, 2000). The expected effect of fishing (although not exclusive) is a decrease in the proportion of piscivorous fish. This is an easily understood indicator that can be estimated based on the knowledge of the biology of the species present in the community rather than on extensive diet data. A related index that has been proposed as an indicator for marine environments is the pelagic (P) to demersal (D) fish biomass ratio in fishery landings (de Leiva Moreno et al., 2000). However, the P/D ratio in fisheries catches is not exclusive in that it might be an indicator of eutrophication rather than exploitation (de Leiva Moreno et al., 2000). The pelagic fish are positively influenced by nutrient enrichment when it stimulates the plankton production (Caddy, 1993), while the demersal fish are influenced by the dynamics of benthic community, which generally responds negatively to the conditions of excessive enrichment. It follows that a positive trend over time in the P/D index may depend both on the eutrophication both from the overexploitation of resources (Libralato et al., 2004). In addition, like other catch-based indicators, it will be sensitive to changes in the fishing targets and methods.

2.2 First statistical approach: smoothing techniques and bootstrapping.

A smoothing function is a function that attempts to capture general patterns in stressor-response relationships while reducing the noise and it makes minimal assumptions about the relationships among variables. The result of a smoothing application is a line through the moving central tendency of the stressor-response relationship and is especially useful to visually assess the relationship between two variables for large datasets, where trends can be hard to visualize.

One of the most common types of smoother, which is implemented in many of the most popular statistical and mathematical packages, is the local polynomial regression smoother. Here, each series of indicators was smoothed using locally weighted scatterplot smoothing (lowess), an outlier-resistant method which estimates a polynomial regression curve using local fitting (Cleveland, 1979). The basic idea is to start with a local polynomial least squares fit and then use robust methods to obtain the final fit. Specifically, at each point a low-degree polynomial is fitted to a subset of the data, using explanatory variable values near the point the response of which is being estimated. The polynomial is fitted using weighted least squares, giving more weight to points near the point the response of which is being estimated and less weight to points further away. The value of the regression function for the point is then obtained by evaluating the local polynomial, using the explanatory variable values for that data point. The fit is complete after regression function values have been computed for each of the n data points.

We applied bootstrapping to the data pairs (x,y) in the original plots and then fitted lowess curves to each simulated bootstrap series.

The bootstrap was introduced by Efron (1979). In order to understand the bootstrap tecnique, we suppose it is possible to draw repeated samples (of the same size) from the population of interest, a large number of times. Then, one would get a fairly good idea about the sampling distribution of a particular statistic from the collection of its values arising from these repeated samples. But, that does not make sense as it would be too expensive and defeat the purpose of a sample study. The purpose of a sample study is to gather information cheaply in a timely fashion.

The idea behind bootstrap is to use the data of a sample study at hand as a “surrogate population”, for the purpose of approximating the sampling distribution of a statistic; i.e. to resample (with replacement) from the sample data at hand and create a large number of “phantom samples” known as bootstrap samples. The sample summary is then computed on each of the bootstrap samples (usually a few thousands).

The calculated sampling distribution allows to infer via confidence intervals. Several different techniques were developed to build confidence intervals using the bootstrapped estimate of the sampling distribution. Among them, the percentile method was chosen to deal with the calculations presented in this chapter.

The basic approach can be sum up as follows: a small percentage, say $100(\alpha/2)\%$ (usually $\alpha = 0.05$), is trimmed off from the lower as well as from the upper end of these numbers. The range of remaining $100(1\text{-}\alpha)\%$ values is declared as the confidence limits of the corresponding unknown population summary number of interest, with level of confidence $100(1\text{-}\alpha)\%$.

The resulting bootstrapped series can be used to assess, both pointwise and globally, the variability in the original lowess fit. With this methodology, each ecological indicator would have a 95% confidence interval for the original lowess.

Finally, for the MTI, FIB and $\text{MTI}^{3.25}$ indicators, which depend on TLs (obtained with their corresponding Standard Errors), it is possible to include random normal variation in the bootstrap procedure in order to propagate the SE of the indicators in all the years and final results.

2.3 A case study: the Mediterranean Large Marine Ecosystem.

The Mediterranean Large Marine Ecosystem (LME) area comprises all the marine waters bounded from the Strait of Gibraltar to the Suez Canal and is considered to be one of the 64 LMEs (see www.lme.noaa.gov for details). The Mediterranean Sea is bordered by 21 countries. It has a narrow continental shelf and covers a surface area of about 2.5 million km^2, of which 1.43% is protected, with 0.4% of the world's seamounts. A warm-temperate climate and several district bio-geographical sub-units characterize this LME. The Mediterranean LME is considered a Class III low productivity ecosystem (Robinson, 2000). It presents a composite structure of environmental conditions, with local areas of upwelling, wind-driven currents, high water temperatures at least in some periods of the year, and nutrient inputs from rivers and human activities (Caddy, 1993). The major inflow into the Mediterranean is nutrient-poor, oxygenated Atlantic surface water through the Strait of Gibraltar, resulting in generally well-oxygenated bottom-waters.

The highest levels of productivity occur along the coasts, near major cities and estuaries, while the lowest levels occur in the southeastern Mediterranean (Stergiou, 1997). Generally, the Western and Eastern basins are considered separately, mainly due to climate and atmospheric circulation. Long-term variability of the Western Basin is linked to the North Atlantic Oscillation, whereas Eastern Basin variability is linked to the Indian monsoon.

The dataset

Fishery data were extracted from *FishStat plus*, the FAO Fishery Statistic Database (www.fao.org). The *FishStat plus* system allows users to select and deselect separate fishery data time series easily and directly in the software. This is a complete world-wide database, thus we refer only to the data set *Landings of GFCM Area 37* from 1970 to 2005. This data set also includes the Black Sea, which in fact is another separate Large Marine Ecosystem (as identified by NOOA) and therefore was analyzed separately in the following section.

Data on capture production are reported to FAO by Member Countries as collected by the national institutions in charge of fishery statistics. The data set consists of recorded nominal catches and does not include discards. In many case the identification and separation of landings by species is rather difficult as landings are reported at a higher taxonomic level than species. Of the 200 species present in the original database, 195 species were selected. The species of sponges and echinoderms were eliminated from the study.

The typical Trophic Level (TL) and the respective standard error (SE) of the fish species were obtained from (Stergiou and Karpouzi, 2002) and FishBase (www.fishbase.org); the latter also provides TL estimates from food items for many fish species. For the invertebrates, the TL estimates and their SE have been based on the ISCCAAP Table of FishBase 2000 (Froese and Pauly, 2011), which is in turn based largely on estimates from food web models (Ecopath), and from SeaLifeBase (www.sealifebase.org). For fish family groups or generic groupings (such as Clupeoids nei), TL estimates were obtained using the weighted mean derived from the TLs of species of the class and their relative landings. Demersal and pelagic species, used to calculate the P/D index, have been defined according to the criteria used by FishBase (www.fishbase.org).

Simulation of Discards & Illegal, Unreported and Unregulated fishing (IUU)

In order to explore the contribution of discards and Illegal, Unreported and Unregulated fishing (IUU) to our approach and to avoid over/understimation of the ecosystem indicator trends, we simulate different scenarios including various percentages of discards and IUU (treated together as a single value).

With the aim to set levels of discards and IUU for the simulation we have taken into account the review carried out by Kelleher (2005), which puts the level of discards in the Mediterranean Sea at less than 5%. However, it should be noted that there is a huge variability in discard rates on a fishery by fishery basis. Some particular trawl fisheries can reach up to 50% discards. On the other hand, most of the small-scale fisheries and small pelagic fisheries show discard rates that can be considered as negligible because the schools tend to be monospecific and the fish tend to be of a similar size (Kelleher, 2005). Although a discard level of 5% in the Mediterranean Sea can be considered low we have to be aware that pelagic catches in particular (with a zero or negligible discard rate) comprise more than 58% of the total catch of the Mediterranean LME for the study period. We also have to consider that traditional fishing in the Mediterranean can be considered quite efficient, with a low level of discards for several reasons: catches are wholly destined for human consumption and there are no exports to other areas. This has produced a better utilization of the resource and discards are lower than in other fishery areas (e.g. Atlantic Ocean).

We suggest simulating a minimum of three different scenarios with diverse percentages of additional discards and IUU to landings (e. g. 5%, 10% and 15%). Moreover, since the discards are usually lower TL species in most fisheries, it is advisable use a cut-off (e. g. 3.25, see Pauly and Watson (2005)) to separate trophic groups into two main categories, i.e. low TLs (≤ 3.25) and high TLs (≥ 3.25). The percentage of discards and IUU applied can then be apportioned differently between low TL and high TL species, e.g. 50% low TL and 50% high TL, 75% low TL and 25% high TL, etc. Thus the scenarios used were finally as follows:

- **Scenario 1 (a), (b) and (c)**
  - Landings + 5% discards & IUU= 50% low TL, 50% high TL;
  - Landings + 5% discards & IUU= 75% low TL, 25% high TL;
  - Landings + 5% discards & IUU= 90% low TL, 10% high TL.
- **Scenario 2 (a), (b) and (c)**
  - Landings + 10% discards & IUU= 50% low TL, 50% high TL;
  - Landings + 10% discards & IUU= 75% low TL, 25% high TL;
  - Landings + 10% discards & IUU= 90% low TL, 10% high TL.
- **Scenario 3 (a), (b) and (c)**
  - Landings + 15% discards & IUU= 50% low TL, 50% high TL;
  - Landings + 15% discards & IUU= 75% low TL, 25% high TL;
  - Landings + 15% discards & IUU= 90% low TL, 10% high TL.

Results

Commercial catches in the Mediterranean LME are dominated by small pelagic species, mainly sardines and anchovies. Catches increased from 1950 to the mid-1980s, levelling off at around 900,000 tonnes in the 1990s, with landings over 1 million tonnes recorded in 1994 and 1995. Demersal landings in the Mediterranean are dominated by young fish and the European hake is the most important demersal resource, followed by red-mullet, nephrops, shrimps and whiting. Landings increased in this LME by about 48% since 1970 with high exploitation of both bottom-living (demersal) and big pelagic stocks, e.g. tuna and swordfish.

The Mediterranean MTI index shows a slight increase of 0.02, from 3.18 to 3.20, in the first decade (from 1970). In contrast, from 1980 to 1990 the values of the MTI index decline from 3.20 to 3.17. The MTI increased again in the 1990s, reaching a maximum value of 3.19 in 1995. After this year the trend decreased to a minimum value of 3.12 in 2005 (Figure 2.1).

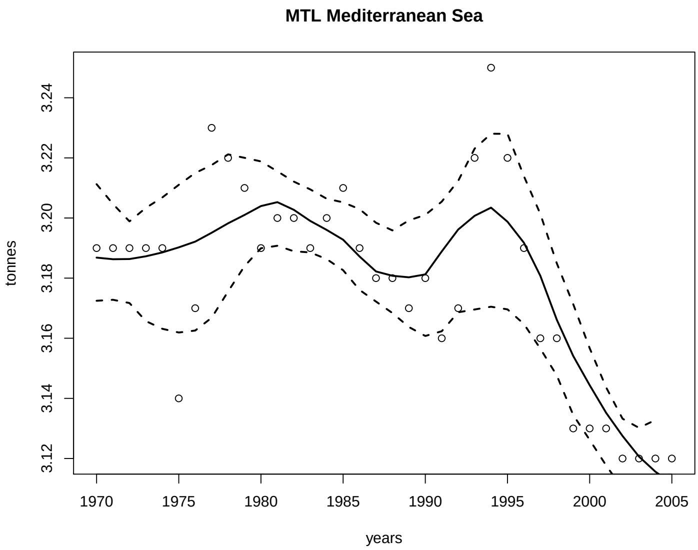


Figure 2.1. The Marine Trophic Index (MTI) for the entire Mediterranean Large Marine Ecosystem from 1970 to 2005. The continuous line represents the Lowess fit and the dotted lines the 95% confidence bands.

The simulations of the MTI with different percentages of discards & IUU show the same trend as the index computed without these (Figure 2.2). In the first decade the index increased and reaches the highest values in all the simulations. From 1980 to 1990 the index decreased, but in the 1992 increased again (Figure 2.2). In the last 10 years the index trend decreased, reaching a value of 3.14 in 2005. The MTI simulations values are greater than 0.02-0.04, compared to the values of the 95% confidence band of the MTI (Figure 2.2). The highest values of MTI are seen in the simulation in which the 10% of additional discards & IUU is split 50-50% between the high and low TL (Figure 2.2). Conversely, the smallest values are seen in the simulation in which the 10% is distributed 25-75% (Figure 2.2).

The FIB index increases to 0.45 from 1970 to 1985. In the second half of the 80s the FIB index remained constant. Since 1990, the values increase from 0.44 to a maximum value of 0.50 in 1995 (Figure 2.3). The last ten years show a continuous decrease. Indeed from 1995 to 2005 the index fell by 0.35, with a minimum value of 0.15 in the last year of the time series examined (Figure 2.3).

In the first decade the P/D index shows a slight increase from values around 2 to 2.1. After 1980, the index decreased steadily, reaching a minimum value of 1.6 in 1993. In the last twelve years, the index value increased continuously to 2.2 in 2005 (Figure 2.4).

Regarding MTI$^{3.25}$ index, in the first five years of the series the index decreased from 3.69 to 3.67. From 1975 to 1985 the index increased to a maximum value of 3.71. From the late 80's until the early 1990's the MTI$^{3.25}$ index decreased to 3.69. From 1990 to 1995 the index values showed an increase of 0.02. Over the past 10 years there was a steady decline to a minimum value of 3.61 in 2005 (Figure 2.5).

Discussion

Results show a declining trend for the MTI, MTI$^{3.25}$ and FIB indeces, while the P/D show an increasing trend. This suggests that the excessive fisheries expansion led to TL changes that are not supported by the ecosystem. In particular, the MTI index show a decline trend, decreasing by 0.02 per decade in the Mediterranean ecosystem in the last 36 years. This result supports the assertion of Pauly et al. (1998) that there have been major changes in the world fish communities, although the decrease observed in Mediterranean data was lower than that estimated by these authors on a global scale (0.1

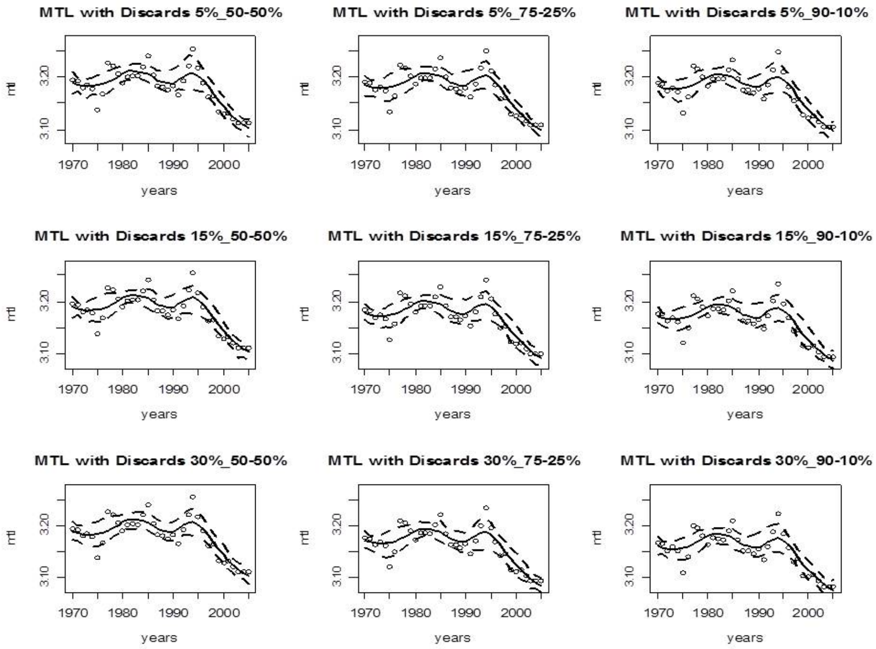


Figure 2.2. Mediterranean Marine Trophic Index with discards & IUU simulations. Three different proportions of discards and IUU are represented: 5%, 10% and 15% of the total landings distributed for low and high Trophic Level species in different ways. In the first column of graphics discards and IUU are represented equally for high and low TL species (50%-50%); in the second 75% are low TL species and 25% high TL species; and in the final column 90% belong to low TL species and 10% to high TL species.

per decade). Comparing the speed at which this decline is occurring in marine ecosystems, we can say that these changes are taking place much more slowly in the Mediterranean Sea. This is probably due to the high biodiversity of this ecosystem, where there is a large functional redundancy of species.

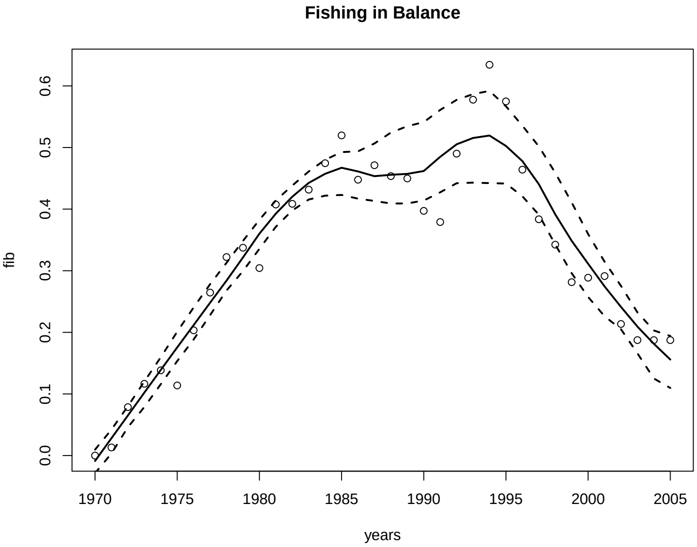


Figure 2.3. Mediterranean Fishing in Balance index between 1970 and 2005.

Pinnegar et al. (2003) found a decrease in average trophic level of landings (about 0.15 TL over 26 years) in the Western Mediterranean, analyzing data on commercial catch and aquaculture, a much higher value than ours (about 0.07 TL over 36 years). It must be emphasized that our study focuses on a longer time series and does not include data coming from aquaculture. Although several different scenarios were tested, with different percentages of discards and different disallocation of theses between the high and low trophic level groups, the trend of the MTI was always the same even if the absolute values differed slightly.

The results show that in the Mediterranean LME fisheries has been supported by high-medium trophic levels (i.e. trophic levels higher than 3.25). The exploitation of these species has evolved over the years much faster than

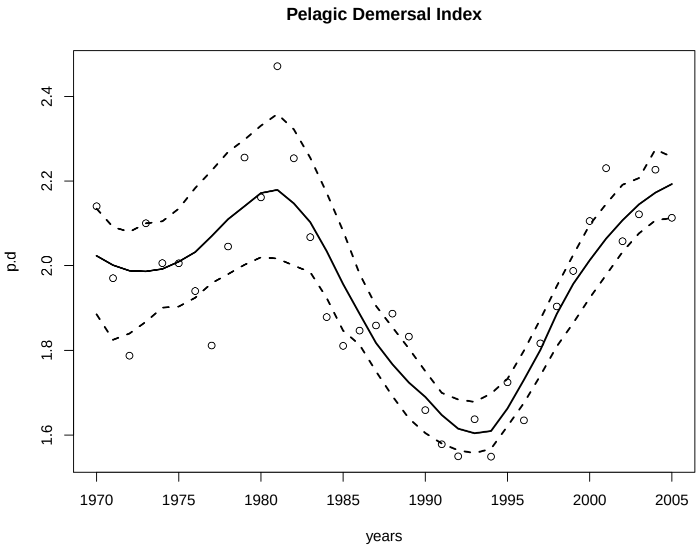


Figure 2.4. Mediterranean ratio of pelagic/demersal landings from 1970 to 2005 (P/D index).

their ability to recover. Indeed, in the last decade Mediterranean landings and MTI have shown downward trends.

The increasing trend of the FIB index in the 1970s probably indicate an expansion of fisheries, both geographically and in terms of the previously unexploited or only lightly exploited stock (Caddy and Garibaldi, 2000). This fishery expansion is probably a consequence of direct government support for the fishing sector as well as the implementation of technological advances. The modernization of small and large-scale fishing fleets (i.e., larger boats, of higher tonnage and engine horsepower, improved fishing gears, use of high-technology equipment) led to the expansion of fishing in areas previously inaccessible to fishing vessels because of strong winds and in deep water areas (Stergiou, 1997). As a result, new resources, for the most part of high trophic levels, have begun

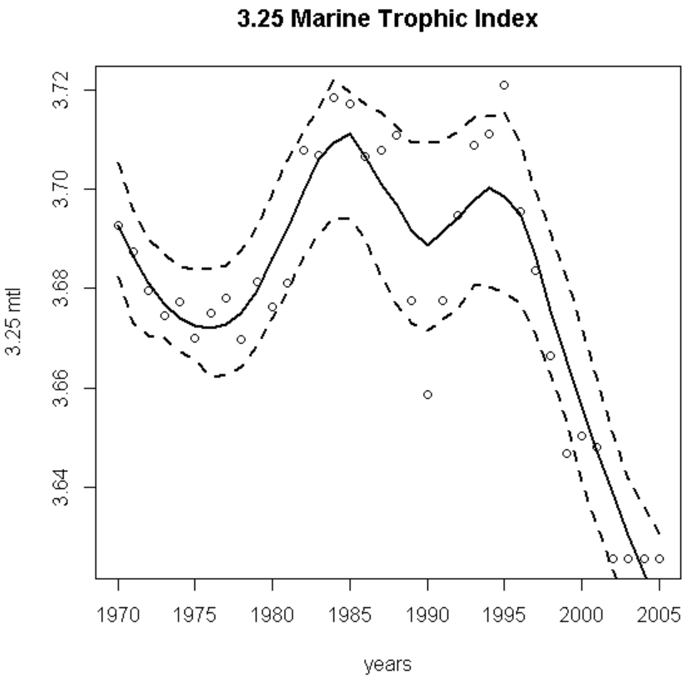


Figure 2.5. Mediterranean Marine Trophic Index$^{3.25}$ from 1970 to 2005.

to be exploited, such as Trachurus spp.

The FIB decrease in recent years could indicate that fisheries have withdrawn an excess of biomass from the Mediterranean ecosystem (Pauly and Watson, 2005). The technological growth and demand for fish in the 1980s has led to the over-exploitation of marine resources, particularly in the case of the highest trophic level demersal resources. This suggests that the changes in the food web could be influenced by the impacts of progress in fishing technologies and changes in market-driven exploitation (Stergiou and Karpouzi, 2002).

The increasing trend of the P/D index over the last decade is probably more the result of reduced demersal abundance than of a real increase in pelagic landings. Pinnegar et al. (2003) suggested that the MTI decline reported by Pauly et al. (1998) in the western Mediterranean may be explained by the increase of landings due to the unwarranted inclusion of farmed bivalves (planktivorous, i.e. low TL) in catch statistics. The decreasing trend in the $\mathrm{MTI}^{3.25}$ index excludes this hypothesis in the Mediterranean LME and confirms a real decrease in the higher TL resources.

The increased importance of small pelagic species in Mediterranean Sea landings may, in part, be a response to the scarcity of more traditional target species, but it might also be related to technical innovation and the development of new fishing gear designs. This, in turn, has greatly increased the vulnerability of small pelagic stocks world-wide (Caddy and Garibaldi, 2000). Almost all the new Mediterranean Sea fisheries of the last 20 years or so (blue whiting, mackerel, horse-mackerel) have targeted species that feed mainly on zooplankton. This is in contrast to the post-war fisheries, which targeted high trophic level species such as hake, haddock and small sharks. Indeed, many of these predatory fishes feed primarily on those species which are now being harvested by the fishery, which implies that modern vessels are operating a full trophic level lower than their post-war counterparts. The composition of historical landings may be affected by phenomena such as natural oscillations in species abundance, changes in fishing technology and economic factors which are likely to have influenced the Mediterranean MTI of landings.

The contents of the Section 2.3 have been submitted in a peer-reviewed journal.

2.4 Assessing the fishing exploitation on the Black Sea Large Marine Ecosystem.

The study area

The Black Sea is the world's most isolated sea. It is connected to the Oceans via the Mediterranean Sea through the Bosphorus, Dardanelle and Gibraltar straits, and linked with the Sea of Azov in the northeast through the Kerch Strait (http://www.icpdr.org). The Black Sea is a highly productive ecosystem with a continental climate. The fluvial discharge, the natural winter production the presence in summer of upwelling and a strong density stratification, making the Black Sea the largest anoxic basin of the global ocean. The deep waters do not mix with the upper layers of water that receive oxygen from the atmosphere. As a result, over 90% of the deeper Black Sea volume is anoxic water. The most peculiar feature of the Black Sea is the absence of marine life at depths beyond 150-200 m., except for a few anaerobic bacteria (www.encyclopediaofukraine.com). In the shallow waters of the continental shelf and river mouths along the northwestern coast are concentrated living organisms. The number of registered alien species at the regional level amount to 217. This number, together the high level of pollution, suggests a serious impact on the Black Sea native biological diversity and negative consequences for human activities.

The data set

In this study we use the fishery landings of the Black Sea Large Marine Ecosystem (see www.lme.noaa.gov for more details) for the years 1970-2005. Fishery data and TLs of the species were extracted from the database in http://www.seaaroundus.org. The data set consists of recorded nominal catches and does not include discarded species. Specific landings were grouped into 11 trophic groups taking into account their trophic level (Table 2.1).

Table 2.1. Trophic groups with the respective trophic levels (TL) and standard error (SE).

Groups	TL	SE
Bivalvia	2.02	0.03
Crustaceans	2.53	0.7
Molluscs	2.82	0.8
Engraulidae	3.11	0.45
Flatfishes	3.13	0.32
Cupleidae	3.15	0.19
Scorpion-fishes	3.56	0.45
Carangidae	3.64	0.32
Percidae	3.66	0.19
Gadidae	4.05	0.68
Sharks and rays	4.15	0.79

Results

Total reported landings in the Black Sea showed critical peaks and troughs, driven primarily by the fluctuation in the landings of European anchovy (*Engraulis encrasicolus*) with a peak landing of 790,000 tones recorded in 1984 (Figure 2.6). The landings have increased following a precipitous decline from 1989 to 1991. However, have not returned to the level achieved in the mid 1980s.

MTI showed an increase of 0.2 in the first two decades. In fact the values of MTI growed from 3.22 to 3.42 from 1970 to 1990 (Figure 2.7). In contrast, from 1990 to 2000 the MTI index showed an abrupt decline, with a decrement of 0.22 (Figure 2.7). Only in the last five years of the time series the MTI index showed a slight increase from a value of 3.20 to 3.25 (Figure 2.7).

The FIB index showed negative values in all 35 years of the series (Figure 2.8). The increase in the FIB index from the 1970s to the mid 1980s is driven by the increased reported landings during this period (Figure 2.8).

In contrast, the decrease in the MTI values since 1990 in not countered by

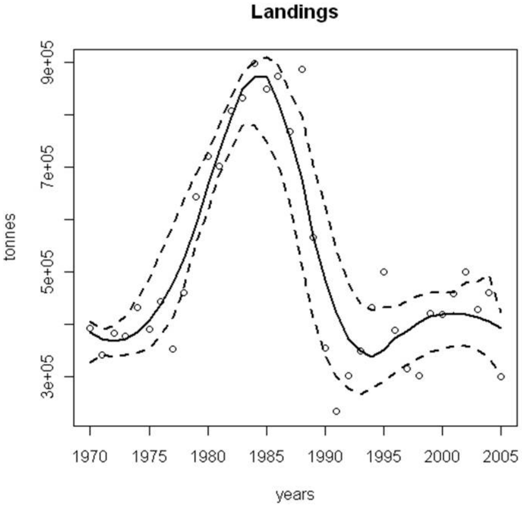


Figure 2.6. Black sea landings from 1970 to 2005. The continue line represented the lowess fit and the dotted lines the 95% confidence bands.

an increase in landings, thus the FIB index has also declined in the early 1990. After reaching the minimum value of -1.8, the FIB index increased by 0.5 from 1995 to 2005 (Figure 2.8).

The MTI$^{3.25}$ showed a drastic decline in the 1984, the same year when the landings of European anchovy peaked (Figure 2.9). In the mid 1990s the index increased from 3.69 to 3.77 in the 1995 (Figure 2.9). In contrast, in the last five years the MTI$^{3.25}$ index showed a strong decrease from 3.77 in 1995 to 3.69 in 2005 (Figure 2.9). The P/D index showed positive values for all the times series (Figure 2.10). This trend showed a decreasing trend since the early 1970s until recorded the minimum values in the 1990. In the following years the index started to increase slowly showed a maximum value of 9.7 in the 2002. In contrast, the last three years of the time series showed a decrease in the index from 9.2 to 4 (Figure 2.10).

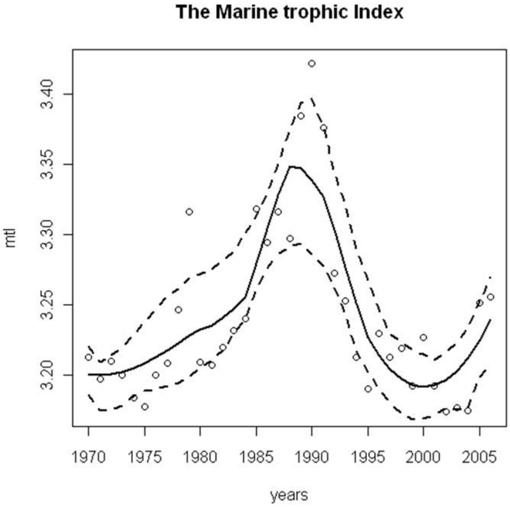


Figure 2.7. Black sea Marine Trophic Index between 1970 and 2005.

Discussion

The indeces show an ecosystem heavily exploited, where the overfishing and the anthropogenic eutrophication are probably the responsible. The decline of the MTI and of the FIB show a *Fishing Down the marine Food Web* situation in this ecosystem. Fishing down has been tested on Mediterranean Large Marine Ecosystems, but the speed with which changes occur in Black Sea fish communities is much higher. Our results support precedent studies that show strong changes fish community of the Black Sea during the last years (Lleonart, 2005).

The fishery eliminated the top predator during the 1970s; that led to reduced predation on planktivores, causing them increasing in the 1980s. Intense and unregulated fishing pressure in these years led to severe overexploitation of most of the major fish stocks (Black Sea Commission, 2002). The MTI and FIB decrease may indicate that fisheries withdraw so much biomass from

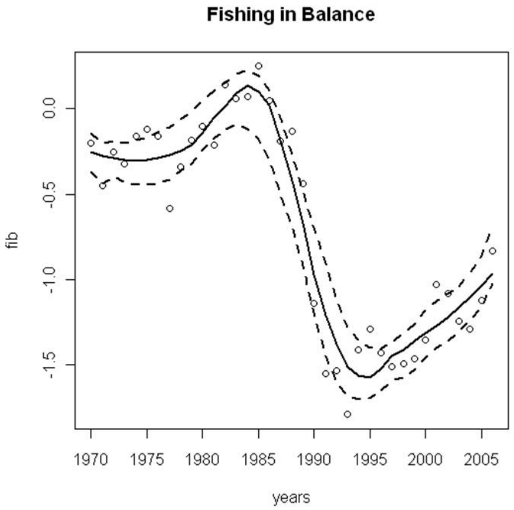


Figure 2.8. Black sea Fishing in Balance Index between 1970 and 2005.

the ecosystem that its functioning could be impaired (Pauly and Watson, 2005). The technological growth and demand of fishery in the years 80s, has developed over-exploitation of marine resources, particularly for highest trophic demersal resources. This suggests that the changes in the food web could be influenced by the impacts of advancement of fishing technologies, and changes in market-driven exploitation (Caddy and Garibaldi, 2000).

Indeed, the top predators as swordfish and tuna, were heavily exploited with the introduction of purse seining and through large scale surface longline and gill-net fisheries in the 1980s in this Large Marine Ecosystem (Caddy, 1993). Some demersal species have pratically disappeared. This has been exacerbated by destructive fishing practices such as catching of under-sized fish. Furthermore, the invasion of the warty comb jelly (*Mnemiopsis leidyi*) contributed to a catastrophic decline of the fish stocks in the mid 1980s. The trend of the MTI$^{3.25}$ and P/D index confirmed this shift in the mid 1980s. The

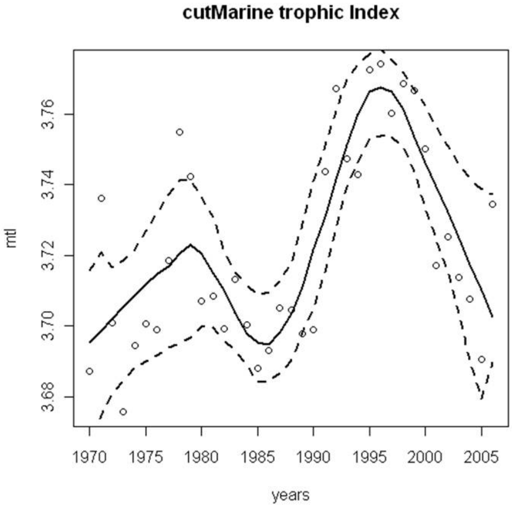


Figure 2.9. Black sea Marine Trophic Index$^{3.25}$ between 1970 and 2005. The cut-off value of 3.25 removed from the computation of the index all species whose biomass tend to vary widely in response to environmental factors.

warty comb jelly eats eggs and larvae of pelagic fish and it caused a dramatic drop in fish populations. The trend of the P/D index showed that the dramatic fall of the black sea fish catch was most pronounced for small pelagic species with a four-fold reduction in the catches between 1988 and 1991, although the landings of these species have partially recovered over the past decade, as showed the trend of the MTI. The P/D index shows this increasing trend in recent years for the small pelagic species, which are r species with a higher turnover, especially in a highly eutrophic ecosystem as the Black Sea. Also, this increase of small planktivorous species might have been a result of the transition of the ecosystem from an oligotrophic to eutrophic stae caused by nutrient enrichment (Caddy, 1993).

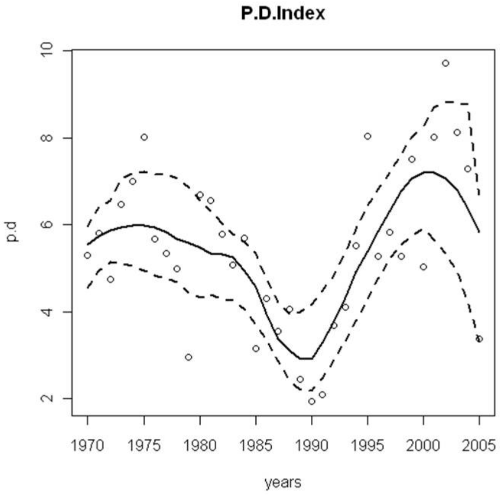


Figure 2.10. Black sea Pelagic/Demersal Index between 1970 and 2005.

This phenomenon is confirmed by the decrease in recent years of the MTI$^{3.25}$ index. This trend indicating that the resources with high TL, which corresponds the demersal fishes, are running out while the low TL species are increasing. These results indicate a lack of sustainability in the ecosystem studied and the need for intervention.

The contents of the Section 2.4 have been published in:
M.G. Pennino, J.M. Bellido, D. Conesa and A. López-Quílez (2011). Trophic indicators to measure the impact of fishing on an exploited ecosystem, Animal Biodiversity and Conservation, **34***(1): 123-131.*

2.4.1 General conclusions

The Large Marine Ecosystem (LME) concept emerges as a possible practical structure upon which the EAFM could be operationalized. In fact, the use of the ecosystem indicators presented have made it possible to identify important trends in two different LMEs, such as the Mediterranean and Black Sea, for a long time-series and in a data-poor situation. The reliability of historical data series of catches in these ecosystems is variable, but data on effort are almost absent. Given the complexity and diversity of the fisheries the available data are probably not sufficient for regular and trustworthy assessments for most species.

We should remember that this approach intends to find a compromise between the different spatial scales and the availability of relevant data to implement an EAFM. This approach is not about understanding the functioning of the ecosystem at a micro-scale level, but about increasing knowledge of the functioning of whole ecosystems.

Indicators rarely give definitive answers but they nearly always suggest the next best question to ask that ultimately will give the answer required. Consequently, indicators very often make people and organisations feel vulnerable, exposed and defensive. This feeling is not likely to change unless more is done to help people understand and accept the strengths, as well as the limitations, of these important measurement tools. One of the biggest problems with developing good indicators is that frequently the best indicators are those for which there is no data. However, there are several advantages to use simple indicators. First, the data are more likely to be readily available and can be used to compare communities. Second, the indicators can help to define problem areas and can be combined with more indicators to overcome each other's weaknesses.

Discussions that include the phrase *but you can't get that data* are not going to lead to better indicators. In fact, if you define a list of indicators and find that the data are readily available for every one of them, you probably have to wait too long to act on the ecosystem improvement. It is better to try to define the best indicators and only settle for less as an interim step while developing data sources for better indicators.

Nevertheless, the interpretation of trophic indicators is still very subjective. References points and limit values have not been established so the interpretaion of the results obtained with these indicators is often unambiguously. The sheer complexity of ecosystems doesn't allow to fix and standardize these values. Each ecosystem trophic structure consists of a specific and unique and it is very difficult to identify critical threshold values that are suitable for all ecosystems.

In order to obtain a comprehensive view of ecosystem dynamics it is essential to assess a number of indicator trends jointly: for instance, trends in the MTI must be accompanied by an examination of the $MTI^{3.25}$, FIB and P/D trends. Jointly these indeces are a great way to compare the dynamics of different ecosystems, since the important aspect for the ecosystem study is the indeces trends over the times and not the values that they assume. Sustainability, however defined, must imply some notion of permanence in at least some of the entities or processes being evaluated. Thus, if there is, in a given ecosystem, a clear trend of the relative abundance of high-TL to low-TL fishes, as indicated by declining MTI values, then this indicates the absence of sustainability and the need for intervention. A multispecies fishery can safely be assumed to be unsustainable if the mean TL of the species it exploits keeps going down (Pauly and Watson, 2005).

In conclusion we recommend the use of these indicators to analyze an ecosystem with a macro-scale approach and get an overview of the ecosystem in this whole. The causes and the drivers that led to the changes highlighted by the analysis must be verified at the micro-scale, when and where you have better data available. We mean that ecosystem indicators are considered promising tools to assess ecosystem conditions because easy to standardize and to estimate with commonly available data. Application of the selected indicators to other marine ecosystems is encouraged so as to evaluate fully their usefulness to a broad selection of LMEs, to evaluate fully their usefulness for an EAFM and to establish international comparability.

2.5 Second statistical approach: nonparametric density estimation, stochastic kernels and transition probability matrices techniques.

In this Section, we present another way for analyzing ecological indicators with the final aim of estimating Mediterranean fisheries exploitation throughout the period 1992-2008 using a meso-scale level. In particular, we have focused on how distributions of fishery ecological indicators evolve over time, through the nonparametric estimation of Gaussian kernel density functions and, according to this, what might be the probable long run distribution. The picture emerging from the results is that the excessive fisheries expansion, both geographically and in terms of exploitation of stocks, led to trophic levels changes that are not supported by the Mediterranean ecosystem. Mediterranean countries persist in their fishery behaviours along the time series, but in the long run they show a tendency to cause similar negative effects on the ecosystem.

Introduction

As previously mentioned, the Mediterranean basin is a complex region that gathers many different ecosystems characterized by a very high level of biodiversity. It is the crossroad between three continents, Asia, Africa and Europe, with very different cultural backgrounds, forms of governments and levels of development (González-Riancho et al., 2009). Mediterranean fisheries are highly diverse and vary geographically, not only because of the existence of different marine environments, but also because of different socio-economic situations. The complex structure of Mediterranean fisheries (e.g. atomised fleet, a huge number of landings points and multispecies catches) makes the data collection and the consequentely fishery management difficult and expensive. Given the complexity and diversity of Mediterranean fisheries the only fairly reliable historical data series available, for all countries, are fish market data.

However, these countries' heterogeneity can be exploited to achieve an efficient fishery management in this LME, through the study of the behaviour of the fishing exploitation in different Mediterranean countries. The overfishing phenomena that is affecting the ecosystems and fish stocks of developed countries, is generating negative effects that could be prevented and smoothed in developing countries.

Here, we address this issue in a complementary way using a meso-scale approach. Specifically, we consider three different ecological indicators that not only assess the fishing exploitation over the time, but also provide an early warning signal of dynamic patterns that might be taking place in the ecosystem. In particular, we use the *Marine Trophic Index*, the *Marine Trophic Index*$^{3.25}$ and the *Fishing in Balance* index. These selected ecological indicators are very useful tools in a data-poor situation, like ours. Moreover, with an only number they can describe a large variety of complex processes occurring within an ecosystem.

In order to accurately capture the fishing exploitation in the Mediterranean basin and its effects on the ecosystem, we propose a methodology that analyzes its evolution over time through the use of ecological indicators, nonparametric density estimation, stochastic kernels and transition probability matrices. These methodologies allows us to ascertain whether there is a tendency for Mediterranean countries to converge in their characteristics, either towards the best or worst practice fisheries, and consequently, on ecosystems.

Our approach is also motivated by the dynamic nature of ecosystems, specially when subjected to an industry, such as fishing, undergoing rapid change. At any particular moment some fisheries will be more, or less, harmful than others, but this ranking will vary over time. Countries will differ in their adjustment to shifts in the fishery industry (new technologies, market prices and national laws), both in their speed of reaction and preparation for such shifts. Looking only at landings statistics (a snapshot of fishing exploitation on a ecosystem) could provide maybe a too restrictive view, as it does not comprehensively capture whether ecosystem are adapting to any potential shocks neither if fisheries could be more sustainable or not over time.

The dataset

Fishery data were extracted from *FishStat plus*, the FAO Fishery Statistic Database (www.fao.org). This is a complete world-wide database. In our case, we have only selected landings from 1992 to 2008 of the 14 Mediterranean countries whose data had continuity over time (Albania, Algeria, Croatia, Egypt, France, Greece, Italy, Lybian, Malta, Montenegro-Serbia, Morocco, Slovenia, Spain, Tunisia).

The database is the same that was used in Section 2.3, except that landings are now divided for the different countries. Here, we use a meso-scale approach, which allows us to reanalyze the patterns previously studied (Section 2.3) and to verify if the use of different spatial scale provides different results in the same ecosystem.

Statistical analysis

Mediterranean fisheries are highly diverse on each country, not only because of the existence of different marine environments, but also because of different socio-economic situations. In order to take into account these differences, and to avoid this affecting the analysis, before applying our techniques, ecological indicators have been normalized using a conditioned approach. In particular, each ecological indicator has been normalized relative to the yearly mean of countries grouped according to their Gross domestic product (GDP). The GDP data have been extracted from official statistics (http://www.kushnirs.org) between 1992 to 2008.

In order to obtain the statistically different groups in terms of the GDP, Cluster Analysis (CA) was performed on a Euclidean similarity matrix with the average method, considering the GDP (1992-2008) of each country to identify similarity between them and to obtain the groups statistically different. In order to evaluate how the hierarchical structure of the CA represented the effective distance between the countries assemblages, the coefficient of cophenetic correlation (Rc, "cophenetic" function, "stats" library of the `R` (R Development Core Team, 2012)) was computed.

Normalizing each ecological indicator has also the benefit of isolating shocks that could bias the analysis, as each indicator could show a tendency either to increase or to decrease. It also permits us to partly offset the distorting effects of outlying observations, to which nonparametric methods are particularly sensitive.

Nonparametric estimation of the univariate density functions

The first step to evaluate how the entire distribution of each ecological indicator evolves over time is to estimate nonparametrically their corresponding density functions in each sample year. The dynamic implications of this analysis are clear: if probability mass tends to be more markedly concentrated around a certain value, convergence is achieved, namely, (normalized) ecological indicator values tend to equalize. If such a value were unity, the outcome would be a convergence process to the average. In other words this would mean that the behaviour of the different fisheries of the Mediterranean countries tend to be gradually more similar. On the contrary, the opposite outcome (divergence) would imply probability mass being increasingly spread across a wider range, implying different behaviours of the Mediterranean fisheries.

Although in many cases the parametric analysis is the most powerful, data might be strongly non-normal, asymmetric or multi-modal. In this sense, one of the most important challenges of data analysis consists of uncovering all complexities that could be hidden and in such attempts, the parametric approach turns out to be clearly unsatisfactory (Tortosa-Ausina, 2002).

However, relying too much on the visual aspect of data has historically been strongly criticized. The first objection the sceptic may argue is logical: this type of analysis is meaningless if graphical representation allows any intrinsic feature in the data to be uncovered. However, in most situations, as the number of observations increases we can see *nothing*. As we already mentioned in Section 2.2, in order to solve such a problem, data must be smoothed, the histogram being the most simple example of smoothing. Indeed, this is the second objection against the nonparametric approach to estimate density functions: why not simply use the histogram to uncover data structure? Although it is not a bad starting point, it has well-known shortcomings (Silverman, 1986) that lead us to choose another way to smooth data.

Specifically, the basis of this approach will be kernel smoothing. It provides a way of uncovering data structure without imposing any parametric model. This allows us to prevent features like a bi-modal structure, which is impossible to uncover through a parametric uni-modal model.

Kernel smoothing consists primarily of estimating the following density function for the ecological indicators computed:

$$\widehat{f}(x) = \frac{1}{Sh}\sum_{s=1}^{S} K\left(\frac{x - EI_s}{h}\right), \tag{2.3}$$

where S is the number of countries being analyzed, EI_s is the specific ecological indicator, K is a kernel function and h is the bandwidth, window width or smoothing parameter.

Multiple options for the kernel selection exist, such as the Epanechnikov, triangular, Gaussian, rectangular, etc. The Gaussian kernel is easily computable, and its expression in the univariate case is the following:

$$K(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}. \tag{2.4}$$

Thus, equation (2.3) becomes:

$$\widehat{f}(x) = \frac{1}{Sh} \sum_{s=1}^{S} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-EI_s}{h})^2}. \tag{2.5}$$

Whereas kernel selection determines the form of the bumps when graphically representing function (2.5), the smoothing parameter h influences it differently, by determining the width of such bumps. However, bandwidth selection is far more important than kernel selection. If h is too small, an excessive number of bumps is generated, thus making it difficult to clearly distinguish data structure. This phenomenon is known as undersmoothing. On the other hand, when h is too large we have oversmoothing, in such a way that some features present in the data (such as multi-modal structures) are hidden. Underlying these concepts lies the traditional trade-off between bias and variance which depends on the smoothing parameter: as h increases, variance decreases and bias increases, and vice versa.

Prior research studies applying the nonparametric estimation of density functions to the analysis of convergence or time evolution of inequalities hardly emphasize the h chosen. In most of them the issue is not even mentioned, while others simply indicate that the smoothing parameter has been chosen *automatically*—referring to the *least squares cross validation* method—. However, as it has been pointed out, choosing different hs significantly influences the results, which forces us to look for a more suitable bandwidth.

Jones et al. (1996) compare different hs, and conclude by stating the importance of this topic. Among them, they state that some first generation methods do not sufficiently smooth data in many circumstances (undersmoothing), while the opposite occurs for others (oversmoothing). Second generation methods offer a reasonable balance between these two extremes or, equivalently, between bias and variance. The higher performance of the second generation methods has been increasingly reported in the literature on kernel smoothing (Cao et al., 1994; Park and Turlach, 1992).

These arguments have led us to choose the bandwidth proposed by Sheather and Jones (1991). It is based on the second generation method solve-the-equation plug-in-approach, and its superior performance relative to the first generation methods has been further verified (Jones et al., 1996). It is known as h_{SJPI}, which has its origins in the authors' names and in the approach followed.

Intra-distribution mobility: stochastic kernels

Nonparametric estimation of the univariate density functions does not inform on some dynamic patterns. In particular, it can be argued that the dynamic evolution of a distribution might not offer a clear pattern either towards convergence or divergence, in the sense described above, but that important intra-distribution movements were taking place. In other words, although the external shape of the density function might not be affected, changes in countries' relative positions could be taking place.

In order to overcome such shortcomings, a law of motion of the cross-section distribution is required. Thus, dynamics can be modelled. Finding out such a law and, therefore, drawing conclusions on the patterns of the variables' cross-section distribution dynamics, requires modelling the stochastic process that takes values that are probability measures (λ_t) associated to the cross-section distribution at time t (F_t), where:

$$\forall\ y \in \mathbb{R}:\ \lambda_t((-\infty, y]) = F_t(y). \tag{2.6}$$

Bearing this in mind, we attempt to build a formal statistical structure which captures the stated phenomena—intra-distribution mobility and, as will be shown later on, long-run behaviour. However, the standard analysis does not provide suitable instruments to model the sequence of the distributions' dynamics. Pursuing such aims, we can consider on Markov Processes Theory and establish a duality to approach the problem.

In the same way that transition probability functions describe the dynamics of a scalar process, *stochastic kernels* describe the dynamics or law of motion of a sequence of distributions. In other words, the stochastic kernel is a mathematical operator that transforms distributions to distributions.

Let λ_t be the probability measure associated to the distribution of each ecological indicator F_t at time t, then the stochastic kernel describing the evolution from λ_t to λ_{t+1} is the mapping M_t to [0,1] of the Cartesian product of a specif ecological indicator and Borel-measurable sets such that (Durlauf and Quah, 1998):

$$\forall \text{ set } A \text{ Borel-measurable}: \;\; \lambda_{t+1}(A) = \int M_t(y, A) d\lambda_t(y). \tag{2.7}$$

Note that the values taken by equation (2.7) are measures or distributions instead of scalars or finite dimensional vectors. Additionally, assuming M_t time-invariant, equation (2.7) could be re-written as:

$$\lambda_{t+1} = M * \lambda_t, \tag{2.8}$$

where M is a representation of the stochastic kernel which encodes information on how starting with a probability measure λ_t associated to the cross-section distribution F_t, we end up with λ_{t+1} (associated to F_{t+1}), i.e., on the different countries' relative positions, which is equivalent to partly knowing the dynamics we attempt to model. Thus, estimation of M from the available data allows empirically quantifying distribution dynamics.

Additionally, considering equation (2.7) and iterating:

$$\lambda_{t+s} = (M * M * \cdots * M) * \lambda_t. \tag{2.9}$$

This expression allows (when $s \rightarrow \infty$) the ergodic distribution to be characterized, thus completely characterizing the dynamics of fishing exploitation. The ergodic distribution should not be considered exactly as a prediction of the future, as future realizations of the variables could be influenced in many ways. This concept should be more properly considered as a characterization of past years' tendencies.

Long-run tendencies: ergodic distribution

Stochastic kernels distributions provide some further insights on the Mediterranean ecosystem dynamic, by identifying fisheries' changes in their relative ecological indicators trends. But it still leaves the long-run behaviour or ergodic distribution unsolved.

In order to computing the ergodic distribution and characterizing long-run behaviour, the ecological indicators space has to be discretized. In such a case, measures λ_t are probability vectors and the stochastic kernel M becomes a transition probability matrix Q. Hence, M and Q both refer to the stochastic kernel, but in the continuous and discrete contexts, respectively:

$$F_{t+1} = Q_{r \times r} \cdot F_t, \tag{2.10}$$

where $Q_{r \times r}$ is a transition probability matrix from one state to another, assuming a countable state space:

$$E = \{e_1, e_2, \ldots, e_r\}, \tag{2.11}$$

for the analyzed variable. The discretization of the observations' space in which the variables may take values in r states e_i, $i = 1, \ldots, r$ allows intra-distribution mobility to be interpreted straightforwardly. In addition, cell p_{ij} in $Q_{r \times r}$ matrix shows the probability that a country initially affiliated to state i transits during the period or periods (l) considered to state j. Cell p_{ij} is defined as:

$$p_{ij} = \frac{1}{T-1} \sum_{t=1}^{T-1} \frac{N_{ij,t}}{N_{i,t}}, \tag{2.12}$$

where T is the number of periods in the sample (17 years), $N_{ij,t}$ is the number of countries transiting during a period from state i to state j and $N_{i,t}$ is the total number of countries starting the period in state i. In addition, each row in the matrix represents a transition probability vector. Such vectors help to better understand the analogy with the continuous case: they are equivalent to the density probability defined for each point in E, when cutting the figure at that point by a plane parallel to $t + l$.

When computing annual transitions (1-year transitions) through transition probability matrices, the available observations for the 17 years are divided into five states $E = \{e_1, e_2, \ldots, e_5\}$. The states' upper limits have been selected in such a way that the initial distribution (1992) is uniform. This strategy gives different limits to the states according to the different definitions of output. If transition probability matrix were the identity matrix, distributions would be invariant and, in addition, no intra-distribution movements would occur. In contrast, if probability tended to be more strongly concentrated off the main diagonal then high intra-distribution mobility would exist (Tortosa-Ausina, 2003).

Results

The similarity dendrogram for the Gross domestic product (GDP) revealed the existence of four different assemblages for the Mediterranean countries (Rc= 0.78). France and Italy form a separate group, as well as Spain and Greece, and Malta and Slovenia, while all other countries are included in a single group. This latter group comprises 8 out of the 14 of Mediterranean countries, including Albania, Algeria, Croatia, Egypt, Lybian, Montenegro-Serbia, Morocco and Tunisia.

As for the analysis of the fisheries exploitation dynamics, we will apply the methodology described in the previous sections to each ecological indicator.

The first stage corresponds to the nonparametric estimation of density functions by means of Gaussian kernel smoothing, for the selected four-year intervals (1992-1995, 1996-1999, 2000-2004, 2005-2008). These periods were chosen, after several attempts, because they allow for the best visual comparison of the shapes of the distributions.

The Marine Trophic Index (MTI)

Figure 2.11(a) shows the time evolution of the distribution of the Marine Trophic Index (MTI) for all sample years with the Gaussian kernel smoothing approach. The MTI displays a clear multi-modality distribution on 1992, which almost vanishes by 2008—even though it seems that the process has undergone a deceleration by 2005 (Figure 2.11(a)). The initial scenario reveals the existence of three different clusters of countries, one below MTI average (the significant peak below 1), another with above average (the "shoulder" at about 1.1), and another lower than the average (the "shoulder" at about 0.85). Densities have become increasingly spread from 1992 to 2005, suggesting that countries are less similar over time. Nevertheless, the last 4 years of the series show a change in the MTI dynamic, and their corresponding densities become more closer to the MTI average. Over time, the three clusters merged, and virtually a main significant peak (mode) emerges.

Stochastic kernels, describing intra-distribution mobility of the MTI, are reported in Figures 2.11(b) and 2.11(c). Specifically, they represent bivariate density functions of the MTI. Each coordinate direction represents a period, and the stochastic kernels attempt to describe the transition from period t to period $t + 1$ (for 1-year transitions, Figure 2.11(b)) and from 1992 to 2008 (for 17-year transitions, Figure 2.11(c)), or how a distribution turns into another one over time.

Through their analysis, it may be inferred that inter-annual mobility is not very high for either output specification, at least compared to 17-year transitions. Such a pattern is suggested by a probability mass concentrated along the positive sloped diagonal in the contour plots, which indicates persistence in the countries positions.

These conclusions are not mirrored when considering Figure 2.11(c), which displays transitions from 1992 to 2008 (17-year transitions). In these cases, intra-distribution mobility is higher. In fact, the probability mass seems to concentrate horizontally at about 1 in 1992, suggesting that initial relative positions are more disperse than the final ones.

In the case of the stochastic kernels, persistence was suggested by probability mass concentrated along the positively sloped diagonal. Equivalently, transition probability matrices suggest persistence if these matrices are the identity matrix. In fact, transition probability matrices are just discretized versions of the stochastic kernels, enabling computation of long run tendencies (ergodic distribution). Thus, Tables 2.2 and 2.3 are the discretized counterparts to the MTI stochastic kernel distributions provided by Figures 2.11(b) and 2.11(c).

Table 2.2. Convergence in Mean Trophic Level (1-year transitions)

	MTI				
	Upper limit				
	0.96	0.98	1.00	1.06	∞
(50)	0.84	0.12	0.04	0.00	0.00
(52)	0.20	0.58	0.16	0.07	0.00
(53)	0.00	0.25	0.53	0.22	0.00
(49)	0.00	0.07	0.18	0.73	0.02
(52)	0.00	0.00	0.00	0.04	0.96
Ergodic distribution	0.27	0.20	0.19	0.22	0.12

Table 2.3. Convergence in Mean Trophic Level (17-year transitions)

	MTI				
	Upper limit				
	0.96	0.98	1.00	1.05	∞
	0.67	0.00	0.00	0.33	0.00
	0.00	0.33	0.00	0.33	0.33
	0.25	0.25	0.25	0.00	0.25
	0.00	0.00	0.50	0.50	0.00
	0.00	0.00	0.00	0.00	1.00
Ergodic distribution	0.00	0.00	0.00	0.00	1.00

In Table 2.3 numbers in parentheses in both 2.2 and 2.3 indicate the total number of observations in each relative MTI state at period t. Thus, as suggested by Table 2.2, there are 53 observations (out of the total number of observations for the 17 years in the sample) initially in state 3 of relative MTI, i.e., ranging between 1.00 and 1.06 times the average. The interpretation of each entry in each matrix is the following: they represent the probability of each countries' to either remain or move to another state of relative MTI. For instance, in Table 2.2, the state 3 shows that in the next period ($t+1$), 53% out of these 53 observations remained in the same state of relative MTI, whereas 22% moved to another state of higher MTI (state 4); the remaining 25% moved to state of less relative MTI (state 2). On the same line, annual transitions show that 84% of countries with initial MTI below 0.96 times the average remained in the same state in the following period, whereas 12%, and 4% moved to states 2 and 3, respectively. The state 5 shows that by the following period $t+1$, 96% out of these 52 observations remained in the same state of relative MTI, whereas 4% moved to the lower MTI state 4. Diagonal entries average to 0.72, show a persistence pattern as the continuous counterpart suggested.

Results differ when only initial and final periods are compared (17-year transitions). Diagonal entries average is much lower (0.55). As Table 2.3 shows, all countries starting in a given state of MTI by 1992 move to other lower or higher states by 2008, except for state 5, that persists to 100%.

The ergodic distributions inform about the probability of a country ending up in a certain state of MTI. Results suggest that probability mass ends up being more concentrated in state 1 of MTI (27%, Table 2.2), including those countries with MTI values ranging between 0.96 and 0.98 times MTI average. The highest MTI state 5 account for almost 12% of probability. On the contrary, the ergodic distribution of the 17-year transitions matrix shows that probability mass ends up being totally concentrated in state 5 of MTI (100%) (Table 2.3).

The Marine Trophic Index$^{3.25}$ (MTI$^{3.25}$)

Gaussian kernel distributions of the Marine Trophic Index$^{3.25}$ (MTI$^{3.25}$) reveals an initial scenario with a clear multi-modality distribution with three different clusters of countries, which persist over time (Figure 2.12(a)). The cluster below

MTI$^{3.25}$ average shows a steady increase with respect to the other clusters, as suggested by probability mass increasingly concentrated at about 1.

The MTI$^{3.25}$ multi-modality distribution is corroborated by the stochastic kernels (Figures 2.12(b) and 2.12(c)). The probability mass concentrated along the positive sloped diagonal in the contour plots indicates persistence in the countries positions.

In Table 2.4 annual transitions show that 95% of countries with initial MTI$^{3.25}$ below 0.96 times the average remained in the same state in the following period, whereas only 5% moved to states 2. Similarly, 93% of countries starting in the highest MTI$^{3.25}$ state, remained in the same state in the following period, whereas only 7% moved to the lower states 4. The diagonal entries average to 0.87 corroborate this persistence pattern.

Table 2.4. Convergence in Mean Trophic Level$^{3.25}$ (1-year transitions)

MTI$^{3.25}$					
	Upper limit				
	0.96	1.02	1.11	1.21	∞
(42)	0.95	0.05	0.00	0.00	0.00
(42)	0.07	0.81	0.12	0.00	0.00
(41)	0.00	0.12	0.80	0.07	0.00
(41)	0.00	0.00	0.10	0.85	0.05
(42)	0.00	0.00	0.00	0.07	0.93
Ergodic distribution	0.32	0.21	0.21	0.16	0.11

Results of ergodic distribution suggest that probability mass ends up being more concentrated in state 1 of MTI (32%, Table 2.4), including those countries with MTI$^{3.25}$ values ranging between 0.96 and 0.98 times MTI$^{3.25}$ average. On the contrary, the ergodic distribution of the 17-year transitions matrix shows that probability mass ends up being more concentrated in state 3 (60%) and 4 (40%) of MTI$^{3.25}$ (Table 2.5).

Table 2.5. Convergence in Mean Trophic Level$^{3.25}$ (17-year transitions)

MTI$^{3.25}$					
	Upper limit				
	0.96	1.02	1.11	1.21	∞
	0.67	0.00	0.00	0.33	0.00
	0.50	0.50	0.00	0.00	0.00
	0.00	0.00	0.33	0.67	0.00
	0.00	0.00	1.00	0.00	0.00
	0.00	0.00	0.00	0.25	0.75
Ergodic distribution	0.00	0.00	0.60	0.40	0.00

The Fishing in Balance Index (FIB)

Gaussian kernel distributions of the FIB show a steady tendency towards divergence in FIB values, as suggested by probability mass increasingly spread across a wider range (Figure 2.13(a)). Density functions show a tighter distribution in the first period of the time series (1992-1995), that has become more wider over the time.

Figure 2.13(a) shows that inter-annual mobility is not very high for either output specification, at least compared to 17-year transitions. Indeed, while Figure 2.13(a) shows that the probability mass is concentrated along the positive sloped diagonal, Figure 2.13(b) shows that the probability is concentrated horizontally at about 0 in 1992, suggesting that initial relative positions are more disperse than final ones.

The intra-distribution mobility patterns are corroborated by the transition probability matrices (Tables 2.6 and 2.7). Table 2.6 shows that, for state 1, by the following period ($t+1$), 84% out of these 43 observations remained in the same state of relative FIB, whereas 14% and 2% moved to higher FIB states (states 2 and 3, respectively). On the same line, annual transitions show that 79% of countries with the higher FIB state remained in the same state in the following period, whereas 14%, 2% and 5% moved to states 4, 3 and 2, respectively.

Table 2.6. Convergence in Fishing in Balance (1-year transitions)

FIB					
	Upper limit				
	-0.20	-0.05	0.03	0.14	∞
(43)	0.84	0.14	0.02	0.00	0.00
(42)	0.17	0.55	0.21	0.04	0.02
(42)	0.04	0.11	0.63	0.20	0.02
(40)	0.00	0.07	0.04	0.62	0.27
(43)	0.00	0.05	0.02	0.14	0.79
Ergodic distribution	0.21	0.16	0.15	0.20	0.28

Table 2.7. Convergence in Fishing in Balance (17-year transitions)

FIB					
	Upper limit				
	-0.20	-0.05	0.03	0.14	∞
	0.84	0.14	0.02	0.00	0.00
	0.19	0.53	0.21	0.04	0.02
	0.04	0.13	0.57	0.23	0.02
	0.02	0.07	0.15	0.50	0.25
	0.00	0.05	0.02	0.16	0.77
Ergodic distribution	0.24	0.17	0.16	0.19	0.25

The ergodic distribution shows that the probability mass ends up being more concentrated in state 5 of FIB (28%, Table 2.6). Nevertheless, the probability is distributed almost uniformly in the other states (about 20%), and more in the the 17-year transitions matrix (Table 2.7).

Conclusions

Over the last decade there has been a strong move towards Ecosystem Approach to Fisheries Management (EAFM) world-wide (Shin and Shannon, 2010). The EAFM is a holistic approach to fisheries management that includes participatory approaches and consideration of a broader view of ecosystem issues including fishing effects on ecosystems (Fluharty, 2005).

In order to achieve these purposes, new methodologies are required since sophisticated ecosystem models are difficult to apply to macro-scale and long time series studies, due to the lack of detailed data. In all these data-poor situations, the most readily available fisheries data are commercial landings, due to their connection to the economy and business, and ecological indicators would appear to be an effective method of highlighting fishing effects on ecosystems, offering ease of interpretation and, ideally, also suggesting the appropriate intervention.

In this Section a new approach and ecological indicators have been considered to estimate Mediterranean fishery exploitation throughout the period 1992-2008. This approach makes possible to answer questions as to how the distribution of the ecological indicators evolve (estimating nonparametrically density functions, via the kernel method), and there are changes in countries relative positions over time and what would be the likely long run (ergodic) distribution of such indicators. This type of analysis permits the uncovering of some features of the distribution that could be hidden if only landings are analyzed, such as a strong multi-modality distribution persistent over time.

Indeed, a clear multi-modality distribution is detected by the analysis of the MTI and in particular of the $\text{MTI}^{3.25}$ index. The two indices show similar trends, however, much more pronounced in the analysis of the $\text{MTI}^{3.25}$ index. The remove of low trophic levels species (trophic levels lower than 3.25) from the calculation of MTI, whose biomass tends to vary widely in response to environmental factors, highlights the persistence of three different clusters in the Mediterranean ecosystem.

One of these clusters shows MTI and $\text{MTI}^{3.25}$ values higher than average of these indices, which implies fisheries supported by high-medium trophic levels (i.e. trophic levels higher than 3.25). Opposite, the cluster with MTI and MTI$^{3.25}$ values lower than the average indicates the existence of fisheries supported by low trophic levels. Both groups show a strong persistence over time. Conversely, intra-distribution mobility is higher in countries in the third cluster, which includes medium MTI and MTI$^{3.25}$ values. This cluster shows a steady tendency towards convergence in MTI and MTI$^{3.25}$ average, reflecting *ecological appropriate* changes in landings.

Nevertheless, ergodic distributions, which inform about the probability of a country ending up in a certain state of MTI and MTI$^{3.25}$, show that most of Mediterranean countries would be in the long run in the lower limit of these indices. This highlights a trend of the increasing prevalence of low trophic levels species over high trophic levels species in the long run, indicating a lack of sustainability in the ecosystem. Currently Mediterranean fisheries are supported by high-medium trophic levels species, but their exploitation could evolve much faster than their ability to recover, generating a fishery shift towards low trophic levels species.

This trend is confirmed by the divergence pattern shown by FIB index. In the first period (1992-1995) all Mediterranean countries showed a similar behaviour, with FIB values close to zero, index that trophic levels changes match *ecological appropriate* changes in landings.

In the following periods the FIB shows a steady tendency towards divergence, resulting in different behaviours of the Mediterranean countries. During the time series the FIB values move away from zero, equivalent to a sustainable fishery behaviour, and assume values greater or lower than zero.

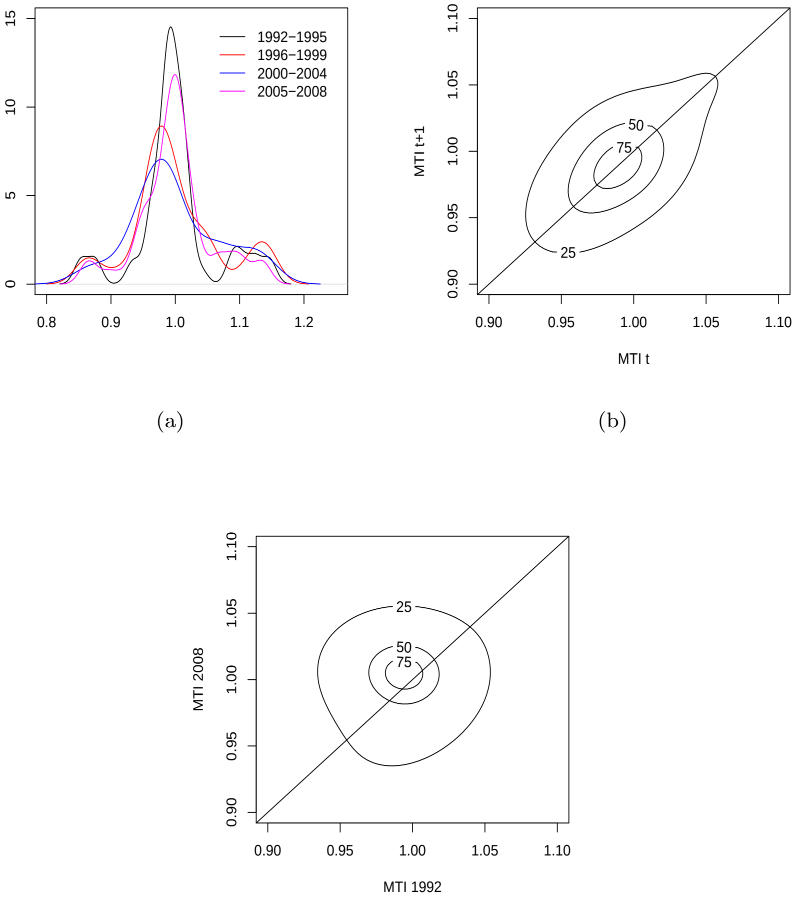
FIB values greater than zero indicate an expansion of fisheries, both geographically and in terms of the previously unexploited or only lightly exploited stock (Pauly et al., 2000). FIB values lower than zero indicate that fisheries, already exploited, have withdrawn an excess of biomass from the Mediterranean ecosystem. The ergodic distribution confirmed these patterns and show that the probability mass ends up being more concentrated in the higher and lower limits of this index. This suggests that changes in trophic levels could be influenced by the impacts of progress in fishing technologies and changes in market-driven exploitation.

The picture emerging by results is that the excessive fisheries expansion, both geographically and in terms of exploitation of stocks, led to trophic levels changes that are not supported by the Mediterranean ecosystem. Mediterranean countries persist in their fishery behaviours along the time series, but in the long run show a tendency to similar negative effects on the ecosystem. Developing countries have increased their catches and the fishing fleet relatively recently and are developing similar behaviours previously shown by developed countries that are currently in a persistent negative phase.

This complementary nonparametric approach provides a new view to explore the fishing exploitation and on what might be the probable long run ecosystem integrity. Regarding a macro-scale approach (Section 2.3), this study provides more detailed information and enlight some hidden patterns than a broader view does not allows. However, in order to apply this methodology on meso-scale level, we were forced to analyze only half of the time series available. In fact, from 1950 to 1992 data for countries were not available, while in the previous approach (Section 2.3) it was possibile to analyze the entire time series (1950-2005).

In this way a multiscale approach is the best option in order to improve our understanding of the fishery dynamics of fisheries in a particular ecosystem. The use of a multiscale approach, and of complementary methodologies, would provide information benefits by improving the assessment findings, and the applicability of fishery management measures, particulary in data limited fisheries.

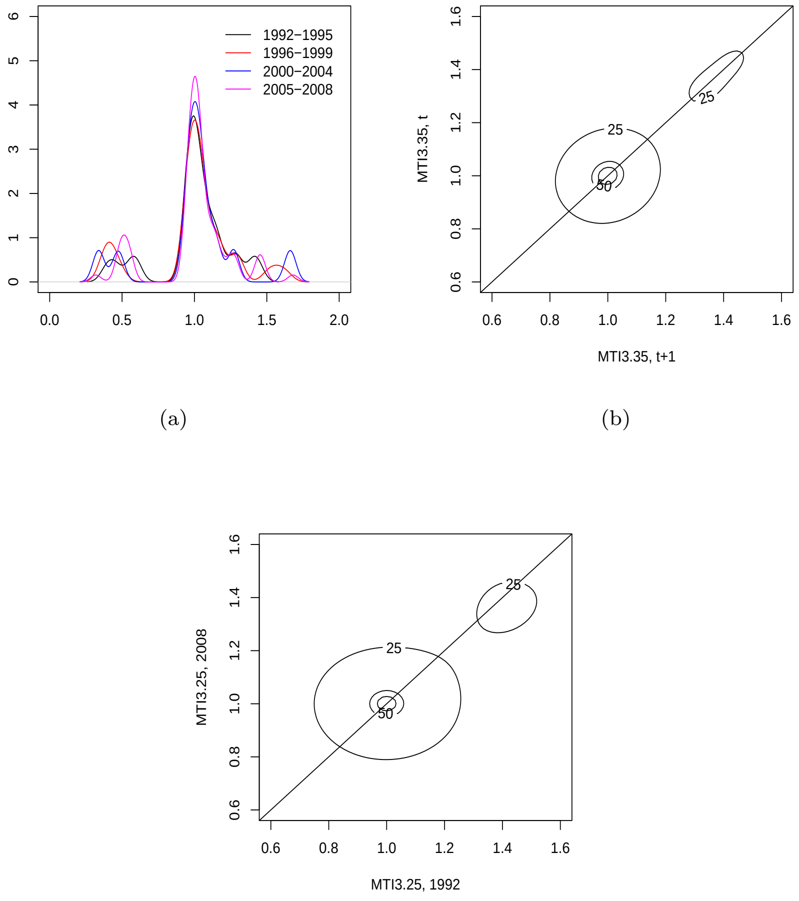
The contents of this Section 2.5 have been submitted in a peer-reviewed journal.



(a) (b)

(c)

Figure 2.11. Gaussian kernel smoothing of the Marine Trophic Index between 1992-2008 (a); Stochastic kernels of the MTI index with 1-year transitions (b) and with 17-year transitions (c).

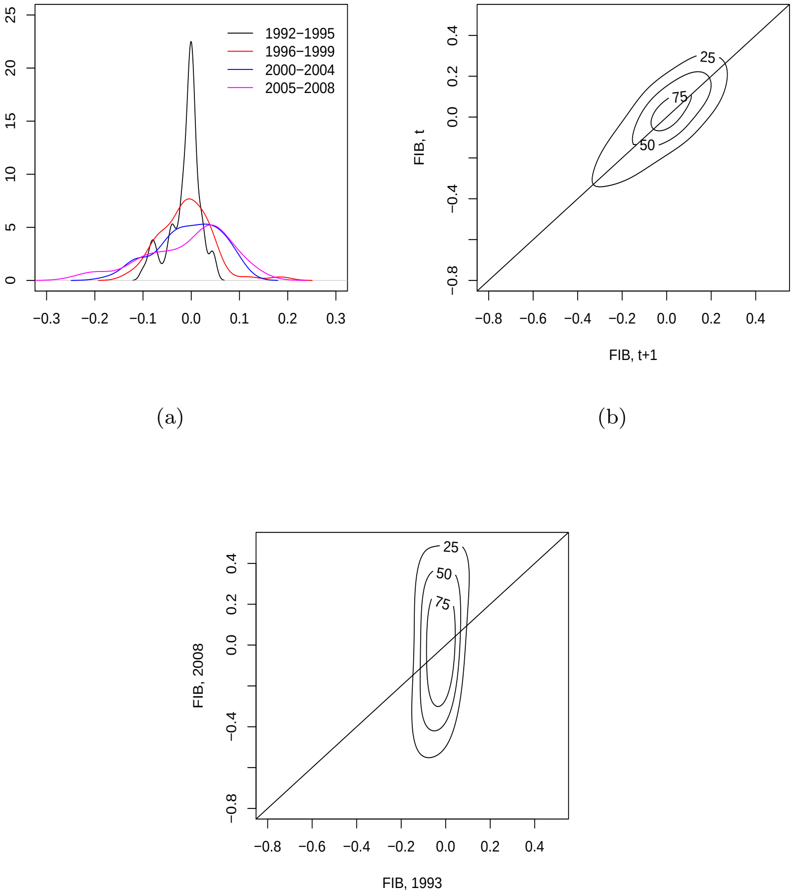


(a)

(b)

(c)

Figure 2.12. Gaussian kernel smoothing of the Marine Trophic Index$^{3.25}$ between 1992-2008 (a); Stochastic kernels of the MTI$^{3.25}$ with 1-year transitions (b) and with 17-year transitions (c).



(a)

(b)

(c)

Figure 2.13. Gaussian kernel smoothing of the Fishing in Balance Index between 1993-2008 (a); Stochastic kernels of the FIB with 1-year transitions (b) and with 17-year transitions (c).

Chapter 3

Modelling the spatial distribution of non target species

In this chapter, we present a Bayesian hierarchical spatial model to estimate and predict the distribution of fishery non target species using fishery-dependent data. In all these cases, the selection of the sampling locations does not depend on the values of the spatial variable and so these are stochastically independent of the field process. As a result, the occurrence of a species is an unbiased indicator of its presence/absence pattern. We present three different types of practical examples. The first two examples use presence/absence data as response variables to estimate and predict the distribution of pelagic and demersal species in determinate studied areas. In contrast, the third example, is based on a Gaussian response variable, in order to address a very topical issue in fisheries: the spatio-temporal distribution of discards.

3.1 Introduction

Modelling patterns of the presence/absence of a species using local environmental factors has been a growing problem in Ecology in the last few years (Chakraborty et al., 2010). This kind of modelling has been extensively used to address several issues, including identifying Essential Fish Habitats (EFHs) in order to classify and manage conservation areas (Pressey et al., 2007), and predicting the response of species to environmental features (Midgley and Thuiller, 2007; Loarie et al., 2008).

Different approaches and methodologies have been proposed for modelling the distribution of species (Guisan and Thuiller, 2005; Hijman and Graham, 2006; Wisz et al., 2008). Generalized linear and additive models (GLM and GAM) (Guisan et al., 2002), species envelope models such as BIOCLIM (Busby, 1991), neural networks (Zhang, 2007; Zhang et al., 2008) and the multivariate adaptive regression splines (MARS) (Leathwick et al., 2005) are some of them.

Most of these applications are only explanatory models that seek to assess the relationship between the presence of species and a suite of one or more explanatory variables (e.g. precipitation, bathymetry, etc.) (Guisan et al., 2002). Moreover, the theory of these methods is based on the fact that the observations are independent, while the fishery data are often inclined to spatial autocorrelation (Kneib et al., 2008). Spatial autocorrelation should be taken into account in the species distribution models, even if the data were collected in a standardized sampling, since the observations are often close and subject to similar environmental features (Underwood, 1981; Hurlbert, 1984). As a consequence, ignoring spatial correlations in this type of analysis could lead to misleading results (Kneib et al., 2008). Note also that extensive spatiotemporal variability, which characterizes dynamic marine ecosystems, presents inherent difficulties for the development of predictive species-habitat models (Valavanis et al., 2008a).

Other complications also arise in the modelling of the occurrence of species due to imperfect survey data such as observer error (Royle et al., 2007; Cressie et al., 2009), gaps in the sampling, missing data, and spatial mobility of the species (Gelfand et al., 2006).

It is also worth mentioning that only a few studies have been developed for predictive models although these models, in addition to offering an estimate of the processes that drive the distribution of species, also provide the probability of the occurrence of species in unsampled areas (Chakraborty et al., 2010).

Our interest here is to propose a hierarchical Bayesian model to predict the occurrence of species by incorporating the environmental and spatial features of each fishing location. The Bayesian approach is appropriate to spatial hierarchical model analysis because it allows both the observed data and model parameters to be random variables (Banerjee et al., 2004), resulting in a more realistic and accurate estimation of uncertainty (see, for instance, Haining et al., 2007, as an example of the advantages over conventional –non-Bayesian– modelling approaches).

Another advantage of the Bayesian approach is the ease with which prior information can be incorporated. Note that prior information can usually be very helpful in discriminating spatial autocorrelative effects from ordinary non-spatial linear effects (Gaudard et al., 1999). Finally, an important feature of our approach is that maps of predicted probabilities of presence in unsampled areas are generated using Bayesian kriging (Handcock and Stein, 1993; Gaudard et al., 1999).

As usual with this kind of hierarchical model, there is no closed expression for the posterior distribution of all the parameters, and so numerical approximations are needed. In our case, we use the integrated nested Laplace approximations (`INLA`) methodology (Rue et al., 2009) and software (`http://www.r-inla.org`) as an alternative to Markov Chain Monte Carlo (MCMC) methods. The main reason for this choice is the speed of calculation: MCMC simulations require much more time to run, and performing prediction has been practically unfeasible. In contrast, `INLA` produces almost immediately accurate approximations to posterior distributions even in complex models. Another advantage of this approach is its generality, which makes it possible to perform

Bayesian analysis in a straightforward way and to compute model comparison criteria and various predictive measures so that models can be compared easily (Rue et al., 2009). `INLA`'s performance has been compared with MCMC and has shown a similar reliability (Held et al., 2010)

In order to provide a realistic view of our methodology, and with the aim to demonstrate that can be used for various species and life stages, in different areas and with a several types of data, we present a variety of practical applications of our approach.

The remainder of this Chapter is organized as follows. After this introduction, in Section 3.2, we present a general Bayesian hierarchical spatial model that accounts for the presence/absence of fish species, allowing both for inference and prediction in unsampled locations. This is commonly known as Bayesian kriging (Banerjee et al., 2004). In Section 3.3, we describe how to implement this model using `INLA`. In Section 3.4, we apply this methodology in a particular setting with fishery-dependent data from purseine fleet in order to describing the distribution of Mediterranean horse mackerel (*Trachurus mediterraneus*) in the bay of Almería. In Section 3.5, we present the application of this model in order to identify the sensitive habitats of the three most caught elasmobranch species (*Galeus melastomus*, *Scyliorhinus canicula*, *Etmopterus spinax*) in the Gulf of Alicante, based on fishery-dependent bottom trawl data. Unlike the previous practical Sections (3.4 and 3.5) in which we demostrate the usefulness of our methodology to estimate and predict the distribution of different species using presence/absence data, in Section 3.6 we apply this approach to a Gaussian response variable to assess the problem of fishery discards in the Gulf of Alicante. Finally, in Section 3.7, we present some concluding remarks and future lines of research.

3.2 Modelling fish presence

This section will describe Bayesian kriging and its application to presence/absence data in fishing. We also discuss the implementation of this kind of model with `INLA` and introduce the SPDE approach to modelling the spatial component.

3.2.1 Bayesian kriging for a binary response

Point-referenced spatial models (Gelfand et al., 2000) are very suitable for situations in which we have observations made at continuous locations occurring within a defined spatial domain. This particular case of spatial models also has the appealing characteristic that the spatial domain is unchanging, even though the precise locations will change over time. In fisheries, this resolves the dimensional control guaranteeing that the inference is realized in relation to the domain instead of the current observed positions, which can change over the years.

In these models, the estimation of the response in unsampled locations can be seen as a statistical prediction problem. When the response is Normal, this is known as *kriging prediction*. Using a Bayesian hierarchical model (Banerjee et al., 2004) such as the one we present in this section allows naturally for non-Gaussian responses, and for taking into account uncertainty in the model parameters. This is known as Bayesian kriging, and the rest of this section discusses its application to fishery data.

Basically, when analyzing fish species distribution, we can encounter two different types of observed data: the amount of catch or just presence/absence data. In the first case it is possible to calculate the absolute abundance of species by standardizing the catch with the fishing effort of the studied fleet, and so it is possible to assess the quantitative spatial distribution of the species within the area of interest. In the second case, presence/absence information can be used as a measure of the relative occurrence of species at each precise observed location, thereby giving a different (but very valid and useful) approximation for the spatial distribution of the species.

For most species, especially for those which are not targeted, information about the absolute abundance of the species is not available. In these situations, the spatial distribution can be obtained by using presence/absence as a response variable of interest instead of absolute abundance. Then, assuming that the probability of catching a species is related to its presence, we model presence/absence by using a point-referenced spatial hierarchical model in line with Diggle et al. (1998).

Specifically, if Z_i represents presence (1) or absence (0) at location i ($i = 1, \dots, n$) and π_i is the probability of presence, then:

$$\begin{aligned} Z_i &\sim \mathrm{Ber}(\pi_i) \\ \mathrm{logit}(\pi_i) &= \mathbf{X}_i\boldsymbol{\beta} + W_i \end{aligned} \tag{3.1}$$

where $\mathbf{X}_i\boldsymbol{\beta}$ represents the linear predictor for observation i; W_i represent the spatially structured random effect; and the relation between π_i and the covariates of interest and random effects is the usual logit link. W_i is assumed to be Gaussian with a given covariance matrix $\sigma^2_W H(\phi)$, depending on the distance between locations, and with hyperparameters σ^2_W and ϕ representing respectively the variance (partial sill in kriging terminology) and the range of the spatial effect:

$$\boldsymbol{W} \sim \mathcal{N}(\mathbf{0}, \sigma^2_W H(\phi))\,. \tag{3.2}$$

This modelling could be augmented by incorporating an additional pure error term (usually Gaussian distributed with variance called nugget effect in kriging terminology) describing the “noise” associated with replication of measurement at each location. Nevertheless, as in this case we are dealing with Bernoulli response, sensitivity to prior assumptions on those random effects precision parameters should be dealt carefully (Roos and Held, 2011).

Once the model is determined, the next step is to estimate its parameters. As we are using the Bayesian paradigm, we have to specify the prior distributions for each parameter involved in the model ($\boldsymbol{\beta}, \sigma^2_W, \phi$). In this context, the usual choice (see, for instance, Banerjee et al., 2004) is to deal with independent priors for the parameters, i.e.

$$p(\boldsymbol{\beta}, \sigma^2_W, \phi) = p(\boldsymbol{\beta})p(\sigma^2_W)p(\phi)\,. \tag{3.3}$$

When there is an aim of expressing initial vague knowledge about the parameters, useful (but not the only) candidates are non-informative Gaussian prior distributions for $\boldsymbol{\beta}$ and inverse gamma distributions for σ_W^2. Specification of $p(\phi)$ will depend on the choice of the correlation function which determines the covariance matrix H. Note that the final choice for the priors will also depend on the type of modelling and parameterization chosen. We will return to this topic later on.

As mentioned above, expressions from (3.1) to (3.3) contain all our knowledge about the spatial occurrence but do not yield closed expressions for the posterior distributions of all the parameters. And so in order to make inference about them, numerical approximations are needed. One possible choice for doing this would be using Markov Chain Monte Carlo (MCMC) methods. This could be done using WinBUGS (Spiegelhalter et al., 1999), flexible software for performing the Bayesian analysis of complex statistical models (see Banerjee et al., 2004) for examples of how to implement spatial hierarchical Bayesian models with WinBUGS). Nevertheless, this option turns out to be very slow when interest is focused on prediction (as in our case), so we have to resort to another approach.

3.2.2 Implementing Bayesian kriging with `INLA`

The key idea underlying what follows is to realize that these hierarchical models can be seen as *Structured Additive Regression (STAR) models* (see, for instance Fahrmeir and Tutz, 2001 for a detailed description of them and Chien and Bangdiwala, 2012 for an applied example of their use). In other words, models in which the mean of the response variable Z_i is linked to a structured predictor that accounts for the effects of various covariates in an additive way. But, more specifically, point referenced spatial hierarchical Bayesian models can also be seen as a particular case of STAR models called *Latent Gaussian models* (Rue et al., 2009), namely those assigning Gaussian priors to all the components of the additive predictor. In this framework, all the latent Gaussian variables can be seen as components of a vector which is the latent Gaussian field.

The great bonus here is that for Latent Gaussian models, we can directly compute very accurate approximations of the posterior marginals using `INLA` (Rue et al., 2009). In spite of its wide acceptance and its good behaviour in many Latent Gaussian models (see for instance, Schrödle and Held, 2011, for a description of how to use `INLA` in spatio-temporal disease mapping), until now it has not been feasible to fit the particular case of continuously indexed Gaussian Fields with `INLA`, as is the case with our spatial component $\boldsymbol{W}$. The underlying reason is that a parametric covariance function needs to be specified and fitted based on the data, which determines the covariance matrix H and enables prediction in unsampled locations. But from the computational perspective, the cost of factorizing these (dense) matrices is cubic in their dimension. Despite computational power today, this problem is still a computational bottleneck in many situations.

Lindgren et al. (2011) have proposed an alternative approach by using an approximate stochastic weak solution to a Stochastic Partial Differential Equation (SPDE) as a Gaussian Markov Random Field (GMRF, Rue and Held, 2005; Rue et al., 2009) approximation to continuous Gaussian Fields with Matérn covariance structure. Specifically, they use the fact that a Gaussian Field $x(\boldsymbol{u})$ with Matérn covariance is a solution to the linear fractional SPDE

$$(\kappa^2 - \Delta)^{\alpha/2} x(\boldsymbol{u}) = \mathcal{W}(\boldsymbol{u}), \qquad \boldsymbol{u} \in \mathbb{R}^d,\ \alpha = \nu + d/2,\ \kappa > 0,\ \nu > 0, \tag{3.4}$$

where $(\kappa^2 - \Delta)^{\alpha/2}$ is a pseudo-differential operator defined in terms of its spectral properties (see Lindgren et al., 2011). They then use a finite-elements method on a triangulation of the region (see Figure 3.1) to construct an approximate GMRF representation of the Matérn Field with parameters κ and $\nu = 1$. They fix ν to 1 for identifiability reasons. An additional parameter τ is used to adjust the scale of the field.

Some important features arise here. Firstly, a GMRF is a discretely indexed Gaussian field $\boldsymbol{x} = (x_1, \ldots, x_n)$, where the full conditionals $\pi(x_i | \boldsymbol{x}_{-i})$, $i = 1, \ldots, n$ depend only on a set of neighbours of each site i. This Markov property makes their precision matrix sparse, enabling the use of efficient (and faster) numerical algorithms.

Secondly, the Matérn covariance function is a really flexible and general family of functions generalizing many of the most-used covariance models in spatial statistics. Its expression, giving the covariance between the values of a random field at locations separated by a distance $d > 0$, can be parameterized as

$$C(d) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)}(\kappa d)^{\nu} K_{\nu}(\kappa d),$$

where K_ν is the modified Bessel function of the second kind and order $\nu > 0$ (Abramowitz and Stegun, 1970, §9.6), $\kappa > 0$ is a scaling parameter and σ^2 is the marginal variance. The parameter ν is a *smoothness* parameter determining the mean-square differentiability of the underlying process, although it is fixed in the SPDE approach since it is poorly identified in typical applications. For more information on the Matérn covariance model see Handcock and Stein (1993); Stein (1999). Finally, GMRFs fit seamlessly with the `INLA` approach, which requires the latent field to be a GMRF.

Under this perspective, for each vertex $i = 1, \dots, n$, the full model can be stated as follows:

$$\begin{aligned} Z_i|\pi_i &\overset{\text{iid}}{\sim} \text{Ber}(\pi_i) \\ \operatorname{logit} \pi_i &= \beta_0 + \mathbf{X}_i \boldsymbol{\beta} + W_i \\ \pi(\beta_0) &\propto 1 \\ \beta_j &\overset{\text{iid}}{\sim} \mathcal{N}(0, 1\text{e-}05) \\ \boldsymbol{W} &\sim \mathcal{N}(\mathbf{0}, \mathbf{Q}(\kappa, \tau)) \\ 2 \log \kappa &\sim \mathcal{N}(m_\kappa, q_\kappa^2) \\ \log \tau &\sim \mathcal{N}(m_\tau, q_\tau^2)\,. \end{aligned} \tag{3.5}$$

In contrast with the previous specification, when using the SPDE approach the correlation function is not modelled directly. Instead, the Gaussian field $\boldsymbol{W}$ is found numerically as a (weak) solution of the SPDE (3.4), depending now on two different parameters κ and τ which determine the range of the effect and the total variance, respectively. More precisely, the range is approximately $\phi \approx \sqrt{8}/\kappa$ while the variance is $\sigma_W^2 = 1/(4\pi\kappa^2\tau^2)$.

Consequently, we have to specify the prior distributions for the parameters involved in this approach $(\beta_0, \boldsymbol{\beta}, \kappa, \tau)$. We set the intercept apart because `INLA` by default specifies a flat improper prior on the intercept, and independent zero-mean Gaussian priors with a fixed vague precision (1e-05) a priori on the fixed effects in $\boldsymbol{\beta}$. The priors for κ and τ are specified over the reparameterizations $\log\tau$ and $2\log\kappa$ as independent Gaussian distributions. We also used the default values for their parameters. Specifically, m_κ is chosen automatically such that the range of the field is about 20% of the diameter of the region, while m_τ is chosen so that the corresponding variance of the field is 1. For instance, in the dataset described in Section 3.4, this gives $m_\kappa = -16.8$ and $m_\tau = 7.16$. Finally, the default a priori precisions for $\log\tau$ and $2\log\kappa$ distributions are $q_\kappa^2 = q_\tau^2 = 0.1$.

The `INLA` program can be used through the `R` (R Development Core Team, 2012) package of the same name. It is worth noting that the SPDE module of `INLA` is still under development and enhancement, but a fully-functional version is readily available by upgrading `INLA` from `R` with the command `inla.upgrade(testing=TRUE)`. As there is still a lack of documentation, there is a downloadable worked-out case study in `http://www.r-inla.org/examples/case-studies` of Lindgren et al. (2011) paper that demonstrates the functionality of the module.

3.3 Estimation and prediction using `INLA`

In what follows we present the basis of how to perform the fitting and prediction in unobserved locations for the Latent Gaussian model in (3.5) using `INLA`'s SPDE module and a brief guide to its syntaxis. It is worth saying that both model fitting and prediction are done simultaneously. Moreover, the fact that `INLA` can be used through `R` provides a familiar interface with the model specification, which is accomplished through the `R`'s *formula* approach.

However, `INLA` provides some additional syntaxis for the definition of random effects, namely the `f()` terms.

Using this syntaxis, the latent field in model (3.5) can be specified as

```
formula = Y ~ 1 + X + f(W, model=spde)
```

where `1` stands for the intercept term, `X` is a fixed linear effect and `W` represents a smooth spatial effect. More terms could be added in the same way if additional covariates were available (for instance, `+ X2 + X3`) or if a noise term were required (`+ f(U, model='iid')`). It is worth noting that while `X` is a variable containing the covariate values at each observation, `W` is only a numeric vector linking each observation with a spatial location.

`INLA` provides different approximation strategies for the posterior marginal distributions. In this study we have used the default ones: the simplified Laplace approximation for the marginalization, and the Central Composite Design for the numerical integration of the hyperparameters. These are the default and recommended settings providing reasonable accuracy with maximum computational efficiency (Held et al., 2010).

The standard output of a run returns the marginal posterior distributions for all the parameters in the model as well as summary statistics, together with several model selection and predictive measures. Specifically, the Deviance Information Criterion (DIC) is a well-known Bayesian model-choice criterion for comparing complex hierarchical models (Spiegelhalter et al., 2002). Additionally, the Conditional Predictive Ordinate (CPO, Geisser, 1993) is defined as the cross-validated predictive density at a given observation, and can be used to compute predictive measures such as the logarithmic score (Gneiting and Raftery, 2007) or the cross-validated mean Brier Score (Schmid and Griffith, 2005). The latter is more adequate for a binary response, measuring the degree to which the fitted probabilities of fish presence at location i coincide with the observed binary outcomes Z_i (Roos and Held, 2011).

As mentioned above, along with the inferential results about the parameters in (3.5), `INLA`'s SPDE module can be used simultaneously to perform prediction in unobserved locations, which constitutes the real interest in this problem. The basic idea is to deal with the species' occurrence at a new location as a random variable with a certain probability of *success* and to calculate a point estimation of this probability, and even its full predictive density.

The SPDE module has a handful of functions to create prediction locations. It allows the construction of a Delaunay triangulation (Hjelle and Dæhlen, 2006) covering the region. As opposed to a regular grid, a triangulation is a partition of the region into triangles, satisfying constraints on their size and shape in order to ensure smooth transitions between large and small triangles. Initially, observations are treated as initial vertices for the triangulation, and extra vertices are added heuristically to minimize the number of triangles needed to cover the region subject to the triangulation constraints. These extra vertices are used as prediction locations. This has at least two advantages over a regular grid. First, the triangulation is denser in regions where there are more observations and consequently there is more information, and more detail is needed. Second, it saves computing time, because prediction locations are typically much lower in number than those in a regular grid. This partition is usually called *mesh* and an example (the one obtained using the data introduced in the section 3.4) can be appreciated in Figure 3.1.

Once the prediction is performed in the selected location, there are additional functions that linearly interpolate the results within each triangle into a finer regular grid. As a result of the process, a faceted surface prediction is obtained which approximates to the true predictive surface.

The prediction in `INLA` is performed simultaneously with the inference, considering the prediction locations as points where the response is missing.

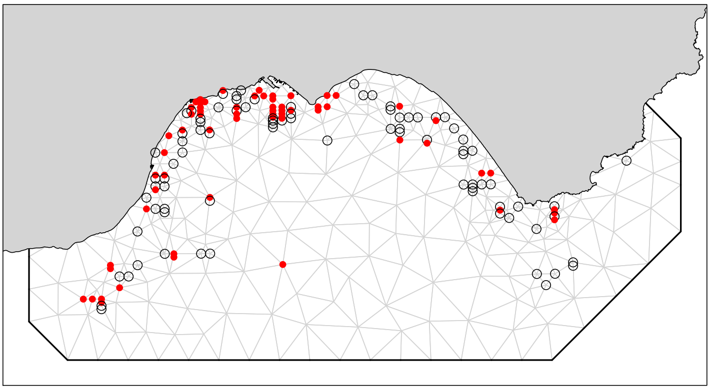


Figure 3.1. Sampling locations for the presence (●) and the absence (○) of the Mediterranean horse Mackerel in the bay of Almería (see the section 3.4).

3.4 Presence of Mediterranean horse mackerel in the bay of Almería

In this Section we present a first application of the model explained in the previous Sections. In particular, we estimate and predict the distribution of a pelagic species, such as Mediterranean horse mackerel (*Trachurus mediterraneus*), using fishery-dependent data from purseine vessels.

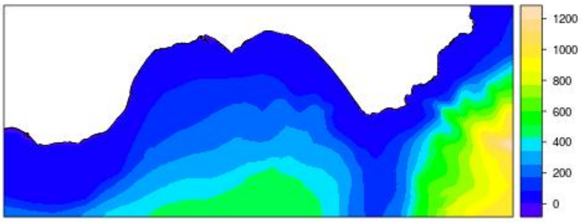
The case study

The study was conducted in the westernmost portion of the Mediterranean Sea, specifically in the bay of Almería, Spain (see Figure 3.1 for a map of the region). The Mediterranean horse mackerel, in spite of its low commercial value, plays an important role in the ecosystem being a food source for other commercially important predators (Froese and Pauly, 2011). But more importantly, this is not a targeted species for commercial fishing, so its occurrence is an unbiased indicator of its presence/absence pattern. Moreover, it also means that the selection of the sampling locations does not depend on the values of the spatial variable and so these are stochastically independent of the field process. This is an important issue as it allows us to predict in all the locations of the bay, including those in which there is no information about the presence/absence of the species.

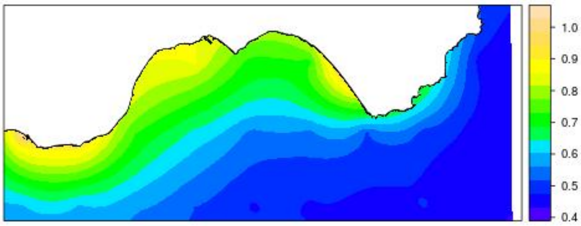
The dataset

The reference fleet for this study was the purseine fleet with landings in the southwestern Spanish ports. This fleet operates in waters on the continental shelf around 200 m. isobaths. The fishing time for each haul lasts around one hour. The data set includes 147 hauls of 15 different purseine vessels and has been provided by the *Instituto Español de Oceanografía* (IEO, Spanish Oceanographic Institute). The IEO provides the national input of the European Plan for collecting fishery data. In particular, they collect samples from the commercial fleet with observers on board. This sampling has been carried out for six years, usually involving about 2-3 observations every month. From this database we have used the geographical location and occurence of the mackerel for each haul.

With respect to the enviromental covariates used in this analysis, we have included those we had information about and those we thought were potentially relevant for a pelagic species like Mediterranean horse mackerel. In particular, the two covariates used were chlorophyll-a (an environmental covariate that usually provides great spatial and temporal coverage Valavanis et al., 2004) and bathymetry (see Figure 3.2 for two maps of both covariates in the region analyzed). The chlorophyll-a data were obtained from satellite data provided by the IEO, while the bathymetry data were obtained from the WFS service of the Spatial Data Infraestruture of the Junta de Andalucía (Andalucian Local Goverment). It is worth noting that if we had had information about other factors such as precipitation, sea surface temperature, etc., they could have



(a) Bathymetry



(b) Chlorophyll-*a*

Figure 3.2. Maps of the covariates in the bay of Almería. The bathymetry map is presented as it was obtained via the Andalucian Goverment, while the Chlorophyll-a is the result of the IEO processing of satellite data.

been included in the analysis via the linear predictor.

Results

All the resulting models obtained from combining those two covariates and the logarithm of the bathymetry were fitted and compared. DIC was used as a measure for goodness-of-fit, while the logarithmic score (LCPO) and the cross-validated mean Brier Score (BS) measured the predictive quality of the models. As shown in Table 3.1, all measures agree on the same model, with a reasonable predictive quality. In particular, the model comparison indicates that (apart from the spatial effect) the logarithm of the bathymetry and the chlorophyll-*a* concentration play a determining role in Mediterranean horse mackerel distribution.

Table 3.1. Model comparison.

	Model	LCPO	BS	DIC
1	1	0.69	0.25	202.77
2	1 + Depth	0.69	0.24	200.87
3	1 + logDepth	0.67	0.24	197.03
4	1 + Chlorophill-a	0.67	0.24	197.60
5	1 + $\boldsymbol{\theta}$	0.66	0.23	195.59
6	1 + Depth + Chlorophill-a	0.67	0.23	196.19
7	1 + Depth + $\boldsymbol{\theta}$	0.67	0.23	195.13
8	1 + logDepth + Chlorophill-a	0.66	0.23	192.18
9	1 + logDepth + $\boldsymbol{\theta}$	0.65	0.23	191.21
10	1 + Chlorophill-a + $\boldsymbol{\theta}$	0.65	0.23	192.18
11	1 + Depth + Chlorophill-a + $\boldsymbol{\theta}$	0.66	0.23	191.48
12	1 + logDepth + Chlorophill-a + $\boldsymbol{\theta}$	0.64	0.22	187.83

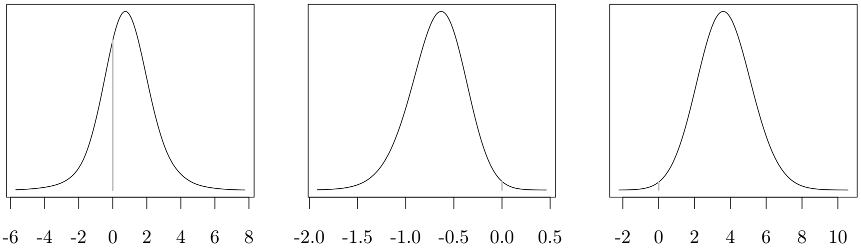
As can be seen in Table 3.2 and Figure 3.3, both covariates have a significant influence on driving the mackerel distribution. Table 3.2 shows a numerical summary of the posterior distribution of the effects, shown in Figure 3.3. In both cases, they show that depth affects the distribution of the species studied negatively, while the chlorophyll-*a* concentration has a positive relationship. Results therefore indicate that the occurrence of Mediterranean horse mackerel is greater in shallow waters (near the coast) where the concentration of the chlorophyll-*a* is higher with to respect to deeper waters. The underlying reason may be that Mediterranean horse mackerel is a pelagic migratory fish occurring at a depth of between 40 and 500 m., usually in surface waters, but at times near the bottom (Ragonese et al., 2003).

In Ecology, chlorophyll-*a* can be used as an indicator of the primary production of an ecosystem. The spatial variability of the primary production

Table 3.2. Numerical summary of the posterior distributions of the fixed effects.

	mean	sd	$Q_{0.025}$	$Q_{0.5}$	$Q_{0.975}$
(Intercept)	0.80	1.56	-2.31	0.78	3.99
log Depth	-0.67	0.29	-1.29	-0.66	-0.14
chlorophyll-a	3.69	1.52	0.79	3.66	6.75

modifies trophic conditions (Katara et al., 2008) of the examined area and thus the distribution of the marine species. Coastal waters are usually zones of high productivity while in surface waters away from coastlines, there is generally plenty of sun but insufficient nutrients. In our case, although captures were scanty in the upper part of the slope (down to 300 m. depth, see Figure 3.1), mackerel was caught on the shelf over practically all the area investigated.



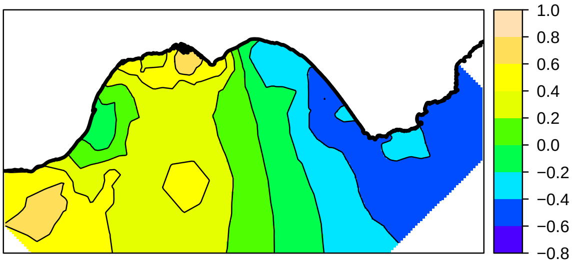
(a) Intercept (b) log(Depth) (c) Chlorophill-a

Figure 3.3. Posterior distributions of the fixed effects.

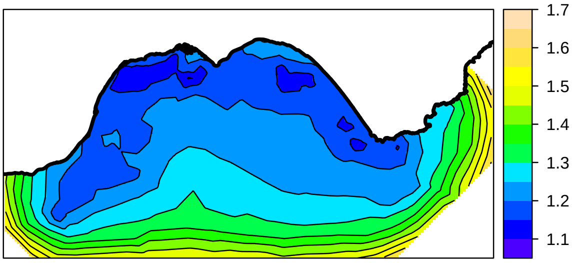
Figure 3.4 displays the posterior mean and standard deviation of the spatial component. This component shows a strong effect with positive values in the western part of the bay of Almería, with values around zero in the middle and with negative values in the eastern part of the area. This results in a clear dependence with respect to longitude in Mediterranean horse mackerel distribution. The western area of the bay of Almería is a protected coastline, the Punta Entinas-Sabinar Natural Park, made up of sand dunes interspersed with a series of freshwater and saline lakes. Its size and development are directly associated with groundwater flows that, jointly with strong hydrochemical variability and an anthropogenic influence due to intensive agriculture, produce a significant concentration of nutrients in the coastal waters. All these factors make this a highly productive area that is the ideal habitat for Mediterranean horse mackerel.

We can also obtain a precise estimation of the complete linear predictor by calculating the corresponding combination of the means of the different effects, as shown in Figure 3.5. The posterior mean of the linear predictor confirms that depth plays a key role in the distribution of Mediterranean horse mackerel, along with the concentration of chlorophyll-a. Along the coast, mean values of the linear predictor show positive values, where the concentration of chlorophyll-a is higher, and as we move away from the coast to the offshore area the mean values become negative.

In order to make the results more understandable, we have also generated maps of predicted probabilities of occurrence using the distribution of the parameter π_i. In this specific case, it is not a linear transformation from the linear predictor, so it is not possible to compute the posterior distribution of the parameter π_i. However, we can obtain any quantile using the corresponding quantiles of the linear predictor. Figure 3.6 shows the median posterior probability of occurrence and the first (b) and third (c) quartiles for this probability. In this way we get not only a point estimate for the probability of occurrence, but also an assessment of the uncertainty of this estimation. Figure 3.6 confirms that the probability of finding this species is greater in areas near the coast at a shallow depth and where the chlorophyll-a concentration is higher. In deeper waters the occurence probability is lower where the nutrient concentration is less. Also, the western part of the bay of Almería shows a higher probability of occurence with respect to the eastern zone due to the presence of the Natural Park and the intensive agriculture that releases a high concentration of organic



(a) Mean



(b) Standard deviation

Figure 3.4. The posterior mean (above) and standard deviation (below) of the spatial effect

material into the sea.

Conclusions

The main goals of this study case have been to predict the occurrence of the Mediterranean horse mackerel in unsampled areas and to estimate its distribution with respect to environmental and geographical factors. The results have

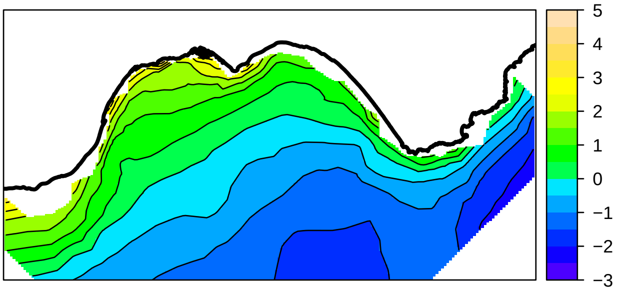
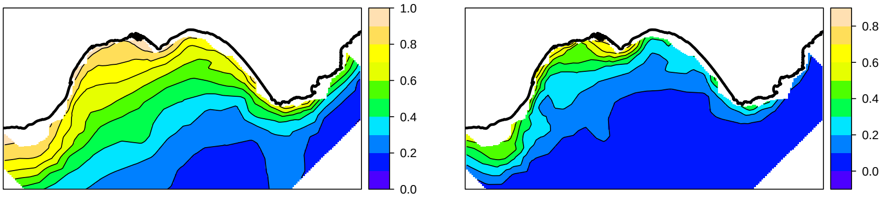


Figure 3.5. Posterior mean of the lineal predictor

shown that the distribution of Mediterranean horse mackerel is influenced by a spatial effect, as well as the depth and the concentration of chlorophyll-*a*. These environmental and geographical factors can play an important role in directing local distribution and variability in the occurrence of this species.

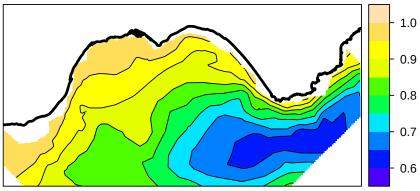
This modelling could be expanded to the spatiotemporal domain by incorporating an extra term for the temporal effect, using parametric or semiparametric constructions to reflect linear, nonlinear, autoregressive or more complex behaviours. A first analysis in this line can be seen in the Chapter 5. Nevertheless, in this case, the information available did not include a reasonable enough number of years for performing any temporal analyses.

The contents of Section 3.2, 3.3 and 3.4 have been published in:
F. Muñoz, M.G. Pennino, D. Conesa, A. López-Quílez and J.M. Bellido (2013). Estimation and prediction of the spatial occurrence of fish species using Bayesian latent Gaussian models, Stochastic Environmental Research and Risk Assessment Journal. In press.



(a) Median

(b) First quartile



(c) Third quartile

Figure 3.6. Median for $\pi_i|\boldsymbol{Z}$ and its corresponding variability.

3.5 Modelling sensitive elasmobranch habitats

In this section we present the application of the model explained in the previous sections in order to identify the sensitive habitats of the three most caught elasmobranch species (*Galeus melastomus*, *Scyliorhinus canicula*, *Etmopterus spinax*) in the Gulf of Alicante, based on fishery-dependent bottom trawl data. This second application shows the flexibility of our approach which in this case is applied to demersal species and in a different area with respect to Section 3.4. Moreover, this second case demonstrates the usefulness that this methodology could provide to fisheries management, identifying sensitive habitats of vulnerable species, such as elasmobranchs, with the aim of improving regional management.

The case study

There is increasing concern worldwide over elasmobranch species because their K-selection life-history traits make them susceptible to population depletion as a result of anthropogenic activity, including unsustainable fisheries, by-catch, and habitat modification (Dell'Apa et al., 2012). Most elasmobranchs are predators at or near the top of marine food chains and thus, play an important role in marine ecosystems, potentially regulating the size and dynamics of their prey populations (Stevens et al., 2000). Their removal could affect the structure and function of marine ecosystems, inducing changes in trophic interactions at the community level due to selective removal of predators or prey species, competitors and species replacement.

In the Mediterranean Sea, this is of particular concern since sharks and rays make up an important percentage of the by-catch (Carbonell and Azevedo, 2003) and their mobile nature makes them potentially accessible to several fisheries at various bathymetric ranges (Ferretti et al., 2008). Bottom trawling is considered responsible for a large proportion of the by-catch of elasmobranch species in the Mediterranean Sea, and throughout the world in general (Maravelias et al., 2012). Evidence of changes in the number of elasmobranchs and the decrease in the abundance of several species (e.g. *Raja clavata* and *Dipturus batis*) over the last decade have been reported for the whole of the Mediterranean Sea and in particular, for the highly exploited area of the Gulf of Lions (Abdulla, 2004). As a result cartilaginous fishes can be used as ecological indicators and their study and monitoring is considered essential for the conservation of the marine ecosystem (Stevens et al., 2000).

In 2009 the European Commission adopted the first Action Plan for the conservation and management of elasmobranchs (EU, 2009) with the aim of rebuilding their stocks under threat, and of setting down guidelines for the sustainable management of the fisheries concerned. Moreover, the implementation of an ecosystem approach to fisheries management (EAFM) and marine spatial planning (MSP) contemplates the protection of priority habitats, a policy of reducing by-catches and the study of current and expected impacts with a view to preparing efficient strategies for the preservation of the marine environment and in particular its living marine resources (Katsanevakis et al., 2009).

In order to achieve these purposes the prerequisites are a solid knowledge of species-environment relationships and the identification of priority areas using robust analysis of existing information and databases (Massutí and Moranta, 2003). Habitat and species mapping is essential for conservation programmes because it provides a clear picture of the distribution and extent of these marine resources, and thus facilitates managing the marine environment (Barberá et al., 2012).

To this end we have analysed a group of georeferenced data of the presence/absence of the most common demersal cartilaginous species collected from fisheries-dependent bottom trawl sampling carried out along the continental shelf and slope of the Western Mediterranean Sea (GFCM Geographical Sub Area 06) during a six-year period of time. In particular, we have modelled the occurrence data of the three most frequently captured species: smallspotted catshark (*Scyliorhinus canicula*, Linnaeus, 1758), blackmouth catshark (*Galeus melastomus*, Rafinesque, 1810) and velvet belly (*Etmopterus spinax*, Linnaeus, 1758), which comprise more than 80% of the total demersal elasmobranch abundance caught during the period 2006-2011. Cluster Analysis (CA) and Multi Dimensional Scaling (MDS) techniques have been applied to observers' data in order to verify whether the three species studied are in fact representative of the whole elasmobranch community of this area.

In this case our methodology allows us both to estimate the processes that drive the distribution of elasmobranchs and also to generate predictive maps of the distribution of species in the study area, especially in non-observed locations.

The establishment of regional marine protected areas for protecting sensitive habitats would benefit from an improved understanding of the spatial distribution of vulnerable species, such as elasmobranchs, and could help towards the more efficient management and control of marine resources.

The study area

The study area was the Gulf of Alicante (Western Mediterranean), between 37º15.6’ and 38º 30.0’ N, and 1º 0.0’ W and 0º 30.0’ E (Figure 3.7). The Gulf of Alicante has a surface area of $3,392\ km^2$ and an average shelf width of approximately $32\ km$.

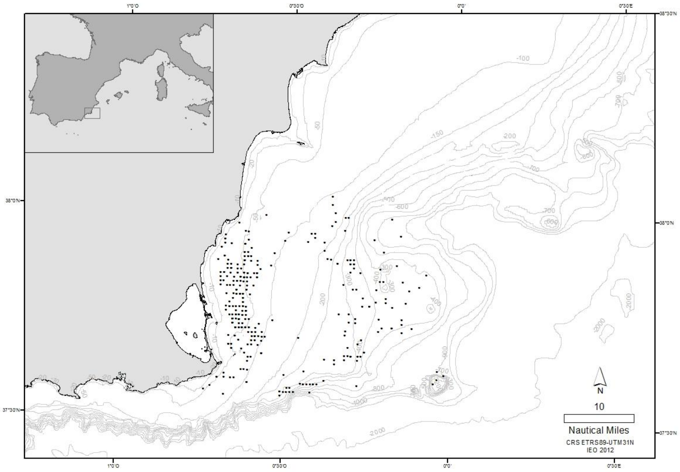


Figure 3.7. Map of the Gulf of Alicante with the sampling locations indicated by black dots.

The largest fleet is the bottom trawl one, with 169 vessels landing an average of 8,000 *t* per year. Seabed trawling usually takes place on the shelf, yielding a multispecific catch with European hake (*Merluccius merluccius*) as the main target species. The elasmobranch species most frequently caught are: *Galeus melastomus*, *Scyliorhinus canicula*, *Etmopterus spinax*, *Raja clavata*, *Raja asterias* and *Squalus acanthias*. Their distribution and abundance vary according to depth.

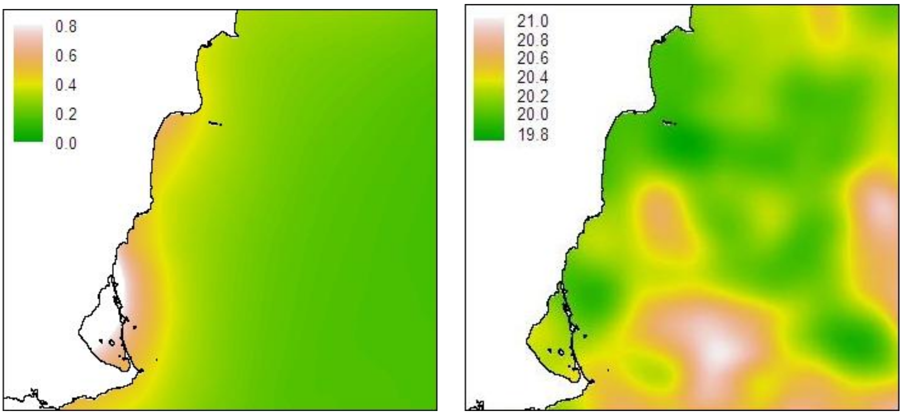
The dataset

The data set includes 400 hauls of 25 different trawler vessels and has been provided by the Spanish Oceanographic Institute (Instituto Español de Oceanografía, IEO). The IEO provides the national input of the European Observers Programme for collecting fishery-dependent data. In particular, they collect samples from the commercial fleet with observers on board. This sampling has been carried out since 2003, usually involving about 2-3 observer samplings every month for the trawler fleet, accounting for an average of 10 hauls monthly. From this database we have used the geographical location and occurrence of the elasmobranch species for each haul. The fisheries were multispecies and none of the elasmobranchs were target species.

Extrinsic factors influencing the spatial distribution of elasmobranch species used were depth, which is often the main gradient along which faunal changes occur when analyzing shelf and upper slope assemblages (Kallianiotis et al., 2000), type of substratum (Demestre et al., 2000), slope of seabed and physical characteristics of the water masses (Maravelias et al., 2007).

For ocean processes, chlorophyll-*a* concentration and Sea Surface Temperature (SST) data can be used to locate thermal and productivity-enhancing fronts and marine productivity hotspots and thus determine the influence of such features on species distribution (Valavanis et al., 2008b). In addition, SST and Chl-*a* are also strong functional links between surface primary productivity and biological activity at the sea floor through the episodic deposition of particulate material (Nodder et al., 2003). Previous studies have shown that the distributions of many demersal fish species are likely to be influenced by overall ecosystem productivity (Matern et al., 2000; Hopkins and Cech, 2003; Martin et al., 2012).

In particular, Chl-*a* concentration can be used as an index of primary production of an ecosystem (de Leiva Moreno et al., 2000). Obviously, primary production depends on a range of factors, including light, light penetration and temperature, which could not be taken into account here due to the absence of data. Nevertheless, the mean value of Chl-*a* concentration can be used as an independent index of primary production in the area of interest, since its variability could modify trophic conditions of the species' habitat from oligotrophic to mesotrophic (Katara et al., 2008).

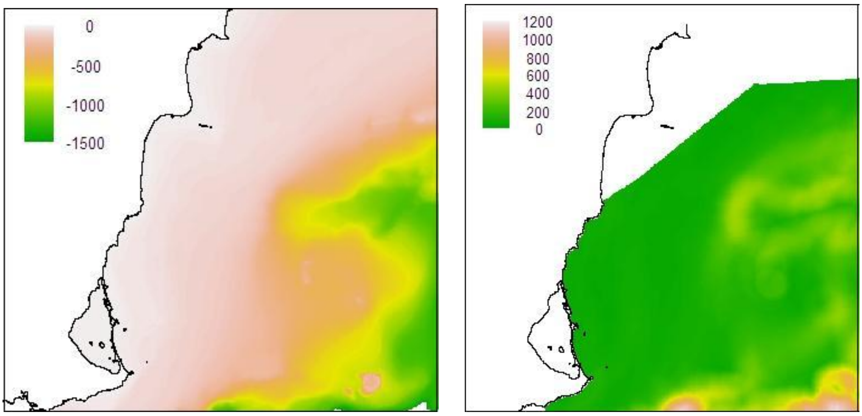


(a) Chlorophyll-a (b) Sea Surface Temperature

Figure 3.8. The spatial patterns of the environmental variables used to map the habitat models, including (a) satellite derived Chlorophyll-*a* mean values; (b) satellite derived sea surface temperature mean values.

Sea Surface Temperature (SST) is strongly related with primary productivity and is thus a possible candidate to explain the distribution of the species (Valavanis et al., 2004). Previous studies on elasmobranchs have implied that SST plays an important role in their distribution (Matern et al., 2000; Martin et al., 2012) from a physiological standpoint. The majority of coastal elasmobranchs are ectothermic and changes in the environmental temperature are rapidly transferred to the body of the animal, thus impacting most physiological processes (Hopkins and Cech, 2003).

The environmental satellite (SST and chlorophyll-a) data has been extracted as a monthly mean from the SeaWiFS (`http://oceancolor.gsfc.nasa.gov`). We have interpolated the raster surface of the SST and chlorophyll-*a* variables, using the *Spline* tool of ArcGIS 10. The *Spline* method is an interpolation method that estimates values using a mathematical function that minimizes overall surface curvature, resulting in a smooth surface that passes exactly through the input points. In particular, we have used the *Tension* method, which controls the stiffness of the surface according to the character of the modelled phenomenon. This method creates a less smooth surface with values more closely constrained by the sample data range (`http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//009z0000006q000000.htm`) (Figure 3.8).



(a) Bathymetry (b) Slope

Figure 3.9. The spatial patterns of the (a) bathymetry and (b) slope of the Gulf of Alicante.

Bathymetry and type of substratum data were obtained from the IEO geoportal, accessible through the website of the Spanish Institute of Oceanography (`http://www.ieo.es`). In order to obtain the value of depth at any precise location of the study area we have interpolated the bathymetric map, using GRASS GIS (http://grass.fbk.eu), first rasterizing contours with a resolution of 500 m. and then using the function *r.surf.contour*, following guidelines given in the website (`http://grass.osgeo.org/wiki/Contourlines_to_DEM`) (Figure 3.9). Log-transformed bathymetry was included in the analysis for smoothing the effect and preserving the linearity of this variable.

The slope map has been derived by the bathymetry map, using the *Slope* tool of the ArcGIS 10. Slope values reflect the maximum rate of change (in degrees) in elevation between neighboring values derived with ArcGIS Spatial Analyst extension (`http://webhelp.esri.com/arcgisdesktop/10/`) (Figure 3.9). The type of substratum polygon shapefile includes a classification of ten categories and a reduced version with four levels: Sand, Mud, Gravel and Rock. In order to reduce the level of variability in the analysis, since we have no observations in all categories, we have used the simplified version (Figure 3.10). Moreover, in the study area there are no areas of gravel, so the categories have been reduced to the remaining three.

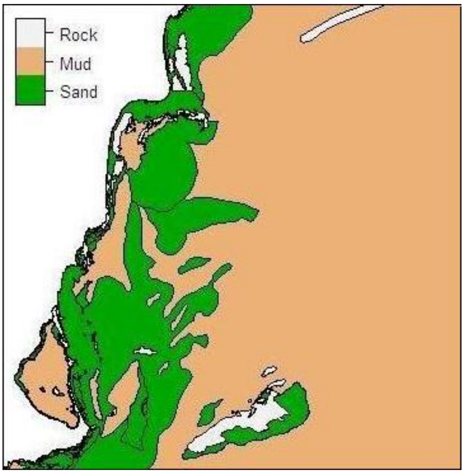


Figure 3.10. Map of the type of substratum in the Gulf of Alicante.

Multivariate analysis

Our data set included 23 different elasmobranch species. We used multivariate analysis techniques in order to verify whether the three most captured species (*Scyliorhinus canicula*, *Galeus melastomus* and *Etmopterus spinax*) were truly representative of the whole elasmobranch community in this area. In particular, we applied Cluster Analysis (CA) and Multidimensional Scaling (MDS) techniques to analyse the 400 bottom trawl hauls. If results show a high degree of similarity in the species assemblage of the different hauls, we can assume that sensitive habitats identified for the three species studied are shared by the remaning elasmobranch species. Consequently, the conservation of these habitats would go a long way to protecting the entire community of elasmobranchs in this area.

Both multivariate analyses were performed on a Euclidean similarity matrix with the average method, considering the occurrence of each species to identify possible differences between the habitats studied.

Model evaluation and calibration

Given that in this application the greater purpose is the spatial prediction of sensitive habitats, we performed a validation predictive procedure to formally evaluate overall model prediction using the area under the receiver-operating characteristic curve (AUC) (Fielding and Bell, 1997), specificity, sensitivity and kappa.

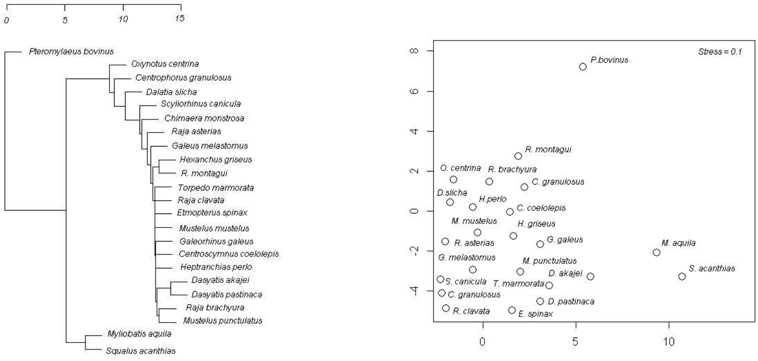
The dataset was randomly split into two main subsets: a training dataset including 70% of the total observations, and a validation dataset containing the remaining 30% of the data. The relationship between occurrence data and the environmental variables was modelled by using the training dataset and the quality of predictions was then assessed by using the validation dataset. We repeated validation 10 times for the best model for each species and results were averaged over the different random subsets.

AUC measures the ability of a model to discriminate between those sites where a species is present and those where it is absent, and has been widely used in the species distribution modelling literature (Elith et al., 2006). AUC ranges from 0 to 1, with values below 0.6 indicating a performance no better than random, values between 0.7-0.9 considered as useful, and values > 0.9 as excellent. AUC is tabulated through the confusion matrix indicating the true positive (TP), false positive (FP), false negative (FN), and true negative (TN) predictions. We can summarize that there are two types of prediction errors: false positive (FP) and false negative (FN). FP leads to an over-prediction while FN or omission error, leads to an under-prediction. From the confusion matrix we calculated the specificity, sensitivity and kappa criteria. Specificity is the proportion of TN correctly predicted and reflects a model's ability to predict an absence given that a species in fact does not occur at a location. Sensitivity is the proportion of TP correctly predicted and reflects a model's ability to predict a presence given that a species in fact occurs at a location. Kappa measures the proportion of correctly classified units after accounting for the probability of chance agreement. It requires a threshold to be applied to the predictions in order to convert them to presence-absence predictions. Kappa provides an index that considers both FP and FN errors. In this study, a maxKappa is used for each model generated.

Results

The similarity dendrogram for the bottom trawl hauls revealed the existence of three different assemblages for the elasmobranchs, which were confirmed by the MDS analysis (Figure 3.11).

Note that picked dogfish (*Squalus acanthias*) and common eagle ray (*Myliobatis Aquila*) form a separate group, bull ray (*Pteromylaeus bovines*) are in a group of their own, and all the other elasmobranch species are included in one single group. This latter group includes 20 out of the 23 of elasmobranch species caught in the study area, including the three most caught species which are the ones used in this study.



(a) CA dendrogram

(b) MDS ordination

Figure 3.11. Analysis Cluster dendrogram (a) and MDS ordination (b) of elasmobranch species caught during bottom trawl commercial hauls carried out in the Gulf of Alicante.

A total of 400 trawl fishery operations were observed during a period of six years. In the case of the smallspotted catshark (*Scyliorhinus canicula*), its presence was recorded in 204 of these hauls, that of the blackmouth catshark (*Galeus melastomus*) in 135 and that of the velvet belly (*Etmopterus spinax*) in 54. The main predictors of elasmobranch habitats in the western Mediterranean Sea were depth, slope and type of substratum.

SST and chlorophyll-*a* concentration show a negative relationship with species occurrence but affect only the distribution of *S. canicula* (Table 3.3). The final models with the best fit for *G. melastomus* and *E. spinax* do not include SST and chlorophyll-*a* concentration as relevant variables (Table 3.3).

No important yearly differences were found in this area for the occurrence of these species. All the models that include the temporal effect, show higher Deviance Information Criterion (DIC) than those without it.

The model selected for its best fit (based on the lowest DIC, LCPO and BS)

Table 3.3. Numerical summary of the posterior distributions of the fixed effects for the best model of the three species studied. This summary contains the mean, the standard deviation, the median and a 95% credible interval, which is a central interval containing 95% of the probability under the posterior distribution.

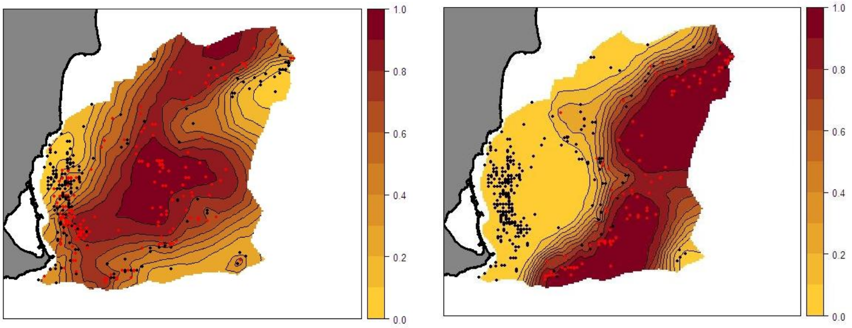
Species	predictor	mean	sd	$Q_{0.025}$	$Q_{0.5}$	$Q_{0.975}$
S. canicula	(Intercept)	0.23	2.14	-4.23	0.73	4.32
	Log Depth	1.06	0.63	-0.32	1.04	2.05
	Seabed(Mud)	-0.32	0.42	-1.16	-0.33	0.51
	Seabed(Rock)	-1.91	0.95	-3.87	-1.88	-0.12
	Slope	0.21	0.11	-0.51	0.22	0.74
	Chlorophyll-a	-13.96	7.83	-29.96	-13.82	1.24
	SST	-0.52	0.27	-0.22	-0.51	1.05
G. melastomus	(Intercept)	-1.72	2.91	7.53	-1.89	4.94
	Log Depth	0.33	-0.27	0.25	0.35	0.85
	Seabed(Mud)	0.41	-0.50	0.40	0.42	1.29
	Seabed(Rock)	-0.72	1.28	-3.36	-0.68	1.69
	Slope	0.19	0.02	-0.06	0.20	0.33
E. spinax	(Intercept)	-2.35	12.76	-32.34	-2.72	28.65
	Log Depth	5.45	3.63	2.70	5.35	10.39
	Seabed(Mud)	0.08	0.85	-1.61	0.09	1.73
	Seabed(Rock)	-0.73	1.28	-3.36	-0.68	1.70
	Slope	0.09	0.03	-0.06	0.08	0.10

with *S. canicula* has log-transformed bathymetry, type of substratum, slope, SST and chlorophyll-a concentration as covariates, and a stochastic spatial component that accounts for the residual spatial autocorrelation. Table 3.3 presents a numerical summary of the posterior distributions of the fixed effects for this final model. This summary contains the mean, the standard deviation, the median and a 95% credible interval, which is a central interval containing 95% of the probability under the posterior distribution.

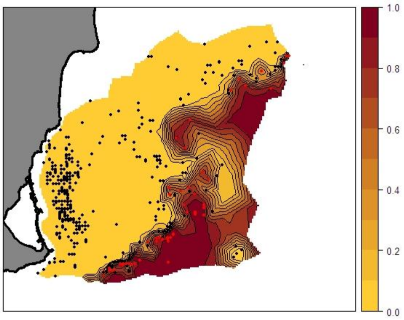
Results showed a positive relationship between bathymetry and the presence of *S. canicula* (posterior mean = 1.06; 95% CI = [-0.32,2.05]). Conversely, chlorophyll-*a* concentration showed a negative relation with respect to the presence of this species (posterior mean = -13.96; 95% CI = [-29.96,1.24]). For low SST values, the occurrence of *S. canicula* is higher (posterior mean = -0.52; 95% CI = [-0.22,1.05]). Rock substratum is the type of seabed that shows the lowest estimated probability of occurrence (posterior mean = -1.91; 95% CI = [-3.87,-0.12]) with respect to the reference level (sand substratum). Muddy substrata also showed a lower estimated coefficient than the reference level (posterior mean = -0.32; 95% CI = [-1.16,0.51]), leaving sandy substrata as the kind of sediment granulometry category with the highest probability of the presence of *S. canicula*. A positive correlation is characterized by a high slope gradient and the probability of occurrence (posterior mean = 0.21; 95% CI = [-0.51,0.74]) of *S. canicula*. As can be appreciated in Figure 3.12 (a), the median posterior probability of the occurrence of *S. canicula* in the Gulf of Alicante, is greater over a high slope gradient, in deeper waters where chlorophyll-*a* and SST values are higher, and where there are sandy seabeds.

Habitats associated with hard substrata and sandy beds, mainly from deeper waters and with a high slope gradient, show a greater probability of the presence of *G. melastomus* than those associated with mud from shallow waters and low slope gradient (Table 3.3 and Figure 3.12 (b)). Table 3.3 showed a positive relation between log-bathymetry and the presence of *G. melastomus* (posterior mean = 0.33; 95% CI = [0.25,0.85]). Rock substratum is the type of seabed that shows the highest estimated probability of occurrence (posterior mean = -0.72; 95% CI = [-3.36,1.69]) with respect to the reference level (sand substratum). Muddy substrata showed a lower estimated coefficient than the reference level (posterior mean = 0.41; 95% CI = [0.40,1.29]). High slope gradient positively influences the probability of occurrence of *G. melastomus* (posterior mean = 0.19; 95% CI = [-0.06,0.33]). Figure 3.12 (b) shows high median posterior probability of occurrence of *G. melastomus* over steeper slopes, in deeper waters and where there are rocky and sandy seabeds.

The results for *E. spinax* are very similar to those for *G. melastomus*. The best model fit for this species included depth, slope of seabed and type of



(a) *S. canicula* (b) *G. melastomus*



(c) *E. spinax*

Figure 3.12. Median of the posterior probability of the presence of the studied elasmobranchs: *S. canicula* (a); *G. melastomus* (b); *E. spinax* (c). Sampling locations for the presence (●) and the absence (●) were plotted.

substratum as relevant covariates. Table 3.3 shows a positive relation between log-bathymetry (posterior mean = 5.45; 95% CI = [2.70,10.39]), slope (posterior mean = 0.09; 95% CI =[-0.06,0.10]) and *E. spinax* occurrence. Habitats associated with hard substrata (posterior mean = -0.73; 95% CI = [-3.36,1.70]) show the highest estimated probability of occurrence with respect to the reference level (sand substratum), the lowest corresponding to muddy beds (posterior mean = 0.08; 95% CI = [-1.61,1.73]).

Figure 3.12 (c) shows a map of the median posterior probability of occurrence of *E. spinax*, with a marked influence of depth. High probabilities appear in deep water, and low probabilities in coastal waters.

Model prediction performance statistics of all models are presented in Table 3.4. All models have achieved AUC values greater than 0.80, which indicates an excellent degree of discrimination between those locations where a species is present and those where it is absent. All maxKappa values are between 0.55 and 0.70, which represents a good degree of similarity between the occurrence of the species and the available real evidence. Specificity and sensitivity also show high values (> 0.8), reflecting a high ability of the model to predict true negative and true positive predictions correctly.

Table 3.4. Model prediction performance statistics for the three species studied. AUC (Area Under the receiver-operated characteristic Curve), maxKappa, sensitivity and specificity.

	AUC	maxKappa	sensitivity	specificity
S. canicula	0.88	0.63	0.87	0.81
G. melastomus	0.84	0.69	0.95	0.90
E. spinax	0.90	0.72	0.84	0.92

Conclusions

The estimated parameters have contributed to quantify habitat use and reveal important relationships of environmental variables with each species' habitat. With the available data, the main predictors of elasmobranch habitats in the western Mediterranean were found to be depth, slope of seabed and type of substrate.

Elasmobranch species from the Gulf of Alicante show different optimum depths, which may indicate certain of fine-tuned bathymetric segregation, although they were in fact found to coexist on shelf and slope bottoms.

However, the direct effect of depth on species occurrence is expected to be relatively small in relation to the indirect effects of bathymetry due to its correlation with many crucial environmental and biological parameters. Due to this, in the absence of such data (e.g. CTD data for oceanographic characteristics in the near bottom), a good knowledge of the bathymetric distribution of species could explain the spatial pattern and it is essential for fisheries management.

From our results, we can summarize that, in our study area, *E. spinax* inhabits the deepest stratum, *G. melastomus* the middle and upper slope and *S. canicula* the shelf and the middle slope. On the upper slope, maps show a habitat overlap between *G. melastomus* and *E. spinax*, and on the middle slope a slight overlap between *G. melastomus* and *S. canicula*. Habitat overlap does not necessarily imply competition, except when resources are in short supply. Otherwise, on rich shelf bottoms, species that are spatially segregated are not driven to differentiate their diets and may easily converge in the use of resources in overlapping areas.

Our spatial results are in accordance with a recent trophic study of these species in the Balearic Islands (Valls et al., 2011). *S. canicula* and *G. melastomus* both prey mainly on euphausiids and share the habitat. However, in our study, *G. melastomus* expressed a wider and deeper distribution trend with respect to *S. canicula*. A possible explanation for this could be a phenomenon of competitive exclusion due to the exploitation of similar resources. Based on the competitive exclusion principle, other species are forced to retreat to the bathymetric and/or geographic range to which they are most highly adapted in relation to the other potential inhabitants. In this case *G. melastomus* retreat to a depth interval of between 400 m. and 700 m., usually home to the biomass peak of decapod crustaceans, which constitute an important part of the diet of this species. *E. spinax* feeds preferentially on cephalopods, while euphausiids are only a small part of the diet. This different trophic pattern with respect to *G. melastomus* would be a mechanism for reducing competition in the deepest stratum.

In general, our results show a negative relationship between all the elasmobranch species studied and depth, and high slope gradient values. In addition, for *S. canicula*, the probability of their presence is higher where SST and chlorophyll-*a* concentration values are lower. This may be explained by the fact that shallow sunlight waters above the continental shelf are usually areas of high productivity and SST mean values, while the deeper waters away from coastlines usually lack sunlight and nutrients, and present low values of chlorophyll-*a* concentration and SST.

This study confirms the importance of the type of substratum in the patterns of elasmobranch spatial distribution, as substrate type was included among the best models for all species. Our analysis shows that elasmobranchs prefer hard and sandy substrates while muddy ones affect their occurrence negatively. This preference has already been documented (Skjæraasen and Bergstad, 2000) and is probably it is partly attributable to the distribution of their preferred prey, which as mentioned earlier, are crustaceans (Holden and Tucker, 1974).

These patterns were also consistent with those from other studies reporting on habitat utilization by various species (e.g. *R. brachyura*, *R. montagui* and *S. stellaris* in (Ellis et al., 2005); *R. clavata* in (Hunter et al., 2005); *S. canicula* in (Vaz et al., 2008)).

Although the present study was limited to three species, multivariate analysis of elasmobranch assemblage show that 20 out of 23 species of elasmobranchs caught in this area are always fished jointly with one of the species examined. These results indicate that sensitive habitats identified for the three species studied are shared by the other, less frequently caught elasmobranchs. The preservation of these habitats may be useful for protecting the majority of the species of the elasmobranch community in this area.

The contents of the Section 3.5 have been published in:
M.G. Pennino, F. Muñoz, D. Conesa, A. López-Quílez and J.M. Bellido (2013).
Modelling sensitive elasmobranch habitats, Journal of Sea Research. In press.

3.6 Assessing spatio-temporal discard models in a demersal trawl fishery

In previous Sections (Section 3.4 and 3.5) we have illustrated the usefulness of our methodology to estimate and predict the distribution of different species using presence/absence data. Instead, in this section we present a different application for which this method can be used: the problem of fishery discards. We illustrate how the model can be easily adapted to a Gaussian response variable, the abundance of discards, with the aim of improving our understanding of those factors that influence the quantity of discards and to identify their spatial-temporal distribution in the study area.

The case study

Discarding is currently one of the most important issues in fisheries management, from both an economic and environmental point of view (Bellido et al., 2011). Discarding occurs for a range of reasons and is influenced by an even more complex array of factors that still remain poorly understood due, among other things, to incomplete knowledge on the spatio-temporal pattern of discards, which tend to be highly variable in space and time (Feekings et al., 2012).

There are indications that the practice of discarding has altered ecosystem functioning at several levels, causing cascading effects throughout the trophic chains. Various seabird and marine mammals species use discards and offal as feeding resources. This has led to at least two important negative effects. Firstly, to an increase in the numbers of certain species due to the availability of food via discards (Valeiras, 2003; Votier et al., 2004), and secondly to an increase in the bycatch of various species of turtles, sharks, seabirds and marine mammals (Jenkins et al., 2004). But moreover, the remaining discarded biomass ends up on the seabed and its fate is poorly known although some scavenger benthic species, including fish, crabs, shrimps and other invertebrates, may consume them. The decrease in the abundance of the above-mentioned vulnerable species could be followed by increases in that of these other advantaged benthic species.

All these negative trends are the manifestation, expressed by the European Union, that there is a need to quantify discards in order to understand their causes and effects so as to manage them effectively.

However, an assessment of discard trends is not available from many countries owing to the limited spatial and temporal coverage of the fleet for discard sampling. Before 2002, European discard sampling programmes were often based on short-term projects rather than being seen as an integral part of the process of collecting fisheries information. Nevertheless, in 2002, with the establishment of the data collection regulation (EC Regulation 1639/2001), the monitoring of discards became a mandatory part of the European fisheries sampling programme (Viana et al., 2011).

The current literature on discards has mainly been descriptive, with a focus on understanding discard rates of specific species (Welch et al., 2008), estimating the amount or proportion of total catch discarded from particular fisheries (Rochet et al., 2002), as well as global discard assessment (Kelleher, 2005). These studies have contributed to improving our understanding of the discard problem, although discarding behaviours are still mainly unknown (Depestele et al., 2011). Therefore, further knowledge of the factors that influence discard rates as well as of their spatio-temporal pattern is needed.

The use of modelling approaches on discard data provides the opportunity to estimate which different drivers (technical and environmental factors) could influence the discard process while also offering important insights for predicting future catches and discards in terms of both quantity and location.

Furthermore, most of the existing studies have mainly focused on analysing on discards of commercially targeted species (Viana et al., 2011; Feekings et al., 2012). An evaluation of Spanish Mediterranean trawling indicates that discard rates of target species are close to zero, practically negligible. The estimation of total discard rates, however, is considerably higher (Carbonell and Mallol, 2013). The discrepancy between target species and total discard rates lies in the fact that discard is comprised of invertebrate species and fish with a low commercial value. These are benthic and semi-pelagic species caught together with the target species. Consequently, modelling only discards of target species may lead to biased conclusions.

The main goal of this study is to address the discard issue by examining the data collected in the GSA06 (Geographical Sub-Areas) South area, identifying the factors that influence discarding within the Spanish trawling fleet and their spatio-temporal distributions. Specifically, our study focuses on the Gulf of Alicante, which is bounded by the Cape de las Huertas and Santa Pola (see Figure 3.7 in the Section 3.5). It has a surface area of $3,392\ km^2$ and an average shelf width of approximately 32 km.

Two different métiers were analyzed, the bottom otter trawl demersal species métier (OTB-DES), which includes trawlers that operate on the continental shelf with European hake (*Merluccius merluccius*) and the octopus (*Octopus vulgaris*) as target species; and the bottom otter trawl deep-water species métier (OTB-DWS), which includes trawlers that operate in deep waters with red shrimp (*Aristeus Antennatus*) as its target species. On-board sampling data on the fishery is directly related to fishing strategy and is useful for analyzing discard trends (Essington, 2010).

In the first place, we analyzed discards of both métiers in order to understand their quantity and species composition. Secondly, we have focused our analysis on factors influencing discards so as to identify their spatial-temporal patterns in the study area.

A detailed knowledge of spatio-temporal discarding patterns could make it possible to further develop spatial fishery management. Predictive maps could be essential tools for selecting high discard areas and thus facilitate the move to discard-free fisheries as part of the proposed reforms of the Common Fisheries Policy (CFP).

The dataset

Under the European Union Data Collection Framework based on (EC Regulation 199/2008), EU members are obliged to collect biological data including discards. Sampling of discards by the *Instituto Español de Oceanografía* (IEO, Spanish Oceanographic Institute) is based on a métier approach, i.e. a formal segmentation of a fishery by vessel types characterised by the same fishing gear and catch composition.

Discards are sampled at a haul level, by randomly collecting one box of discarded catch from as many hauls as possible during each trip. For each observed haul, an estimate of the total weight discarded is made by the fishermen and the on-board observer, by subtracting the landings weight from the total catch weight. The fraction by weight of the discarded fish in the sample is then multiplied by the total discarded weight of the trip recorded to obtain the total weight of fish discarded per trip (Viana et al., 2011). The discard sample is sorted by the observer into species. Total weights and numbers of each discarded species in the subsample are determined and based on the total approximate discarded weight.

The reference fleet for this study was the trawl fleet which operates in the GSA06 South area (see Section 3.5, Figure 3.10). This trawl fleet has been divided into two different types of métier: the bottom otter trawl demersal species métier (OTB-DES) and the bottom otter trawl deep-water species métier (OTB-DWS).

The OTB-DES includes trawlers that usually operate in continental shelf waters (from 50-200 m. depth) with European hake (*Merluccius merluccius*) and the octopus (*Octopus vulgaris*) as target species. They make short hauls of about 2-4 hours trawling, comprising about 2-3 fishing hauls per trip.

The OTB-DWS involves trawlers that usually operate in deep waters with red shrimp (*Aristeus Antennatus*) as its target species. They generally make a single haul per trip, of about 5-6 hours trawling. The monthly frequency of the sampling usually consists of about 2-3 trips for the OTB-DES métier, and about 1 trip for the OTB-DWS métier.

In our case, 343 OTB-DES hauls and 97 OTB-DWS hauls, sampled from 2009 to 2012, were analyzed. Our study was performed using the logged discards per unit effort (DPUE) as response variable. The logarithm of DPUE was used to downweight extreme values and to ensure that the values fitted were strictly positive. For each métier, DPUE was calculated as discard weight (Kg.) per haul duration (hours).

3.6.1 Modelling discards abundance

Hierarchical Bayesian spatial-temporal models were used to account discards dependence between explanatory variables, as well as to describe the main spatial distribution changes over time. These models can also be considered as an alternative of the models explained in Sections 3.4 and 3.5 that can be used when the response variable is a Gaussian distribution.

In particular, if Y represents the value of DPUE in each haul, we can express its relationship with the spatial, temporal, technical and environmental covariates according to this new general formulation:

$$\begin{aligned} Y_{ij} &\sim \mathcal{N}(\mu, \sigma^2), \quad i = 1, ..., n\,; \quad j = 1, ..., q \\ \mu_{ij} &= X_i\beta + W_i + v_j \\ \boldsymbol{W} &\sim \mathcal{N}(\mathbf{0}, \mathbf{Q}(\kappa, \tau)) \\ 2\log\kappa &\sim \mathcal{N}(m_\kappa, q_\kappa^2) \\ \log\tau &\sim \mathcal{N}(m_\tau, q_\tau^2) \\ v_j &\sim \mathcal{N}(0, 1/\rho_v) \end{aligned} \tag{3.6}$$

where β represents the vector of the regression coefficients, X is the matrix of variables, v is the component of the temporal unstructured random effect at the year j, and W_i represent the spatially structured random effect.

A total of 12 potential variables have been considered for each of the models (for each métier). The exact terms are listed in Table 3.5.

Table 3.5. Summary of variables included in Bayesian Models as potential factors influencing discards.

Variable	Description	Comment
Depth	Mean fishing depth of haul	In meters
Slope	Seabed slope in the fishing location	Derived by the Bathymetry map
Orientation	Seabed aspect in the fishing location	Derived by the Bathymetry map
Type of Seabed	Seabed sediment types in the fishing location	Sand, Mud, Rock, Gravel
Moon	Moon Phase of the trip day	New, Full, Crescent,Waning
Log(CPUE)	Logged catch per unit effort of all species	In kilograms
Light	Daylight when the haul was sampled	Yes/No
Vessel length	Vessel's length overall	In meter
Vessel power	Engine size of vessel	
Vessel GRT	Gross Register Tonnage	
Vessel	Unique code for each vessel	Used as a random effect
Quarter	Quarter in which haul was sampled	1, 2 ,3 ,4

In particular, from the onboard observer data set we have extracted the characteristics of sampled vessels, the spatial location, year, quarter, moon phase, daylight and CPUE of the observed hauls. As previously mentioned in the case of discard abundance we have used a log-transformation of the CPUE variable, computed from the total catch (Kg.) per haul duration (hours).

In order to verify interannual variations in discard abundance we have introduced the quarter variable in the analysis to specify the period when the haul was sampled.

Another potential source of variation in discard quantity is unobserved differences between vessels, which can be the result of either a skipper effect or unobserved gear characteristics. Ignoring such non-independence in the data may lead to invalid statistical inference. In order to differentiate between correlations caused by vessel-specific differences in fishing operation and spatiotemporal patterns in the discards distribution, we have included a vessel effect. Since the exact nature of the vessel effect is of negligible interest, we have included vessel as a random effect.

Bathymetry, slope and type of substratum spatial patterns are the same used in the case study of Section 3.5 (see Figures 3.9 and 3.10).

3.6.2 Bayesian inference

Once the model is determined, the next step is to estimate its parameters. Following Bayesian reasoning, the parameters are treated as random variables, and prior knowledge is incorporated via prior distributions. In particular, for the parameters involved in the fixed effects, we use Gaussian distribution $\beta \sim$ N(0, 100).

For the hyperparameters derived from the spatial effect, κ and τ which represent respectively the range and scale of the spatial effect, we assume prior Gaussian distributions with zero mean and covariance matrix depending of each hyperparameters. For the temporal effect we assume, following Rue and Held (2005), LogGamma prior distribution for the precision ρ_v with hyperparameters

a=1 and b= 5e -05.

As usual in this context, the resulting hierarchical Bayesian model has no closed expression for the posterior distribution of all the parameters, and so numerical approximations are needed. Here, due to speed of calculation, we use the integrated nested Laplace approximation (INLA) methodology and software (http://www.r-inla.org) as an alternative to the Markov chain Monte Carlo (MCMC) methods. INLA provides different approximation strategies for the posterior marginal distributions. Here, we used the Stochastic Partial Differential Equation module (SPDE), which allows us to fit the particular case of continuously indexed Gaussian Fields by INLA, as is the case with our spatial component (Lindgren et al., 2011).

A model selection approach has been used to select relevant variables. Specifically, the Deviance Information Criterion (DIC) is a well-known Bayesian model-choice criterion for comparing complex hierarchical models (Spiegelhalter et al., 2002). The smaller the DIC represent the better compromise between fit and parsimony.

3.6.3 Bayesian kriging

Once the inference is carried out, the next step is to predict the discards abundance in the rest of the area of interest, especially in unsampled locations. Here, we adopted a Bayesian kriging to calculate posterior predictive distributions of the discards abundance for the whole region. Using Bayesian kriging, we incorporated parameter uncertainty into the prediction process by treating the parameters as random variables.

A common method for performing prediction with a Bayesian kriging is to take observations and construct a regular lattice over them. In this study, we have considered a triangulation approach exactly as in the previous Sections (3.4 and 3.5).

Once the prediction is performed in the sampled fishing location, there are additional functions that linearly interpolate the results to the whole area.

As a result of the process, for each point of the area we obtain a predictive posterior distribution of the discards abundance. This means that for each posterior distribution, unlike the mean and confidence interval produced by classical analyses, we are able to explicit the probability statements about the estimation. Thus, the region bounded by the 0.025 and 0.975 quantiles of the posterior distribution has an intuitive interpretation: under the model, the unknown discards estimation is 95% likely to fall within this range of values.

For each métier, maps of the posterior mean from the predictive distribution were plotted to illustrate the predicted discards abundance in this area. In addition, the posterior mean and standard deviation of the spatial component were displayed to detect hidden spatial patterns.

Results

A total of 440 hauls (343 OTB-DES and 97 OTB-DWS) were analysed over the period 2009 to 2012 in the study area. For the OTB-DES the total catch in the whole time series is approximately 81,126 Kg. with a total discard of about 27,406 Kg., representing a discard proportion of 34% of discards. This proportion is about 20% for the OTB-DWS, with a total catch of 15,158 Kg. and about 3,100 Kg. of total discards.

Table 3.6 lists the five most discarded and caught species for the two different métiers, giving the respective quantitties discarded and caught during the time series.

For the OTB-DES métier, the bogue (*B. boops*) represents approximately 23% of total discards between 2009 and 2012, followed by the axillary seabream (*P. acarne*) with 11% and the small-spotted catshark (*S. canicula*) with 8%. In addition to being the most discarded species, the bogue is also the most caught species, representing 8% of the total catch of the OTB-DES. The octopus, which is one of the target species of this métier, is the second most captured species, representing 7% of the total catch and with only 3% of discards. The axillary seabream is the third most captured species, representing about 6% of

Table 3.6. The five most discarded and caught species for the two different métiers sampled, with the respective quantities (Kg.) discarded and caught during the time series.

Métier	Discarded species	Discards	Catch	Caught species	Discards	Catch
OTB-DES						
	B. boops	6,497	6,731	*B. boops*	6,497	6,731
	P. acarne	3,043	4,895	*O.vulgaris*	191	5,583
	S. canicula	2,303	3,244	*P. acarne*	3,043	4,895
	P. erythrinus	1,304	2,592	*M. poutassou*	787	4,438
	T. trachurus	1,165	3,742	*T. trachurus*	1,165	3,742
OTB-DWS						
	L. crocodilus	578	578	*A. antennatus*	8,8	3,893
	G. melastomus	471	1459	*G. melastomus*	471	1,459
	L. caudatus	218	218	*M. poutassou*	18	1,291
	S. canicula	191	422	*P. blennoides*	79	1,279
	E. spinax	164	165	*G.longipes*	25	1,037

the total catch of this métier. European hake, which is one of the main target species of this métier, is only tenth in abundance, accounting for approximately 4% of the total catch.

For the OTB-DWS métier, the jewel lanternfish (*L. crocodilus*) accounts for 19% of total discards, followed by the blackmouth catshark (*G. melastomus*) with 15% and the silver scabbardfish (*L. caudatus*) with 7%. The red shrimp is the target species of this métier and the most caught one, representing 26% of the total catch and with only a 0.3% of discards. Blackmouth catshark is the second most captured species, in addition to being the second most discarded species. It represents 10% of the total catch for this métier and has a discard rate of 32%. Blue whiting (*M. poutassou*) is the third most caught species for this métier, accounting for 9% of the total catch. For both métiers the major proportion of both catch and discards consists of non-target and low commercial value species.

Models showed that the relative importance of each variable was different for each métier, with a certain number of similarities. For both métiers, no important yearly differences were found in this area for discard variability. All the models that include the temporal effect show higher DICs, LCPO and BS than those without it.

In the OTB-DES métier the model selected for its best fit (based on the lowest DIC, LCPO and BS) includes the bathymetry, the log-transformed CPUE, the moon phase and the vessel random effect as covariates, and a stochastic spatial component that accounts for the residual spatial autocorrelation. Table 3.7 presents a numerical summary of the posterior distributions of the fixed effects for this final model.

Among the environmental variables slope, orientation and the type of seabed were found to be irrelevant for the variability of DPUE abundance. No difference was found between day and night trawling. The vessel random effect was relevant for all models, while the characteristics of sampled vessels (GRT, power and length) were of no relevance.

Results showed a negative relationship between bathymetry and the DPUE

Table 3.7. Numerical summary of the posterior distributions of the fixed effects for the OTB-DES métier.

Predictor	Mean	Sd	$Q_{0.025}$	$Q_{0.5}$	$Q_{0.975}$
(Intercept)	-2.54	0.48	-3.48	-2.54	-1.60
Moon(Full)	-0.25	0.16	-0.56	-0.25	0.06
Moon(Crescent)	0.17	0.16	-0.14	0.17	0.48
Moon(Waning)	0.10	0.16	-0.21	0.10	0.41
Depth	-0.22	0.07	-0.35	-0.22	-0.09
Log(CPUE)	1.43	0.05	1.34	1.43	1.52
Quarter 2	1.85	0.55	0.78	1.85	2.92
Quarter 3	1.55	0.57	0.43	1.55	2.67
Quarter 4	0.16	0.55	-0.92	0.16	1.25

(posterior mean = -0.22; 95% CI = [-0.35,-0.09]). Conversely, the log-transformed CPUE showed a positive relation with respect to the amount of DPUE abundance (posterior mean = -1.43; 95% CI = [1.34,1.52]).

The full moon phase shows a lower estimated DPUE abundance (posterior mean = -0.25; 95% CI = [-0.56,0.06]) with respect to the reference level (new moon). Also, the waning moon showed a lower estimated coefficient than the reference level (posterior mean = 0.10; 95% CI = [-1.21,0.41]), leaving the crescent moon as the lunar category with the highest estimated DPUE abundance for the OTB-DES métier (posterior mean = 0.17; 95% CI = [-1.14,0.48]).

All the estimated coefficients of the quarters of the year show higher DPUE abundance than the reference level (first quarter). In particular, the second quarter shows the highest estimated DPUE abundance (posterior mean = 1.85; 95% CI = [0.78,2.92]) with respect to the baseline.

Higher values of DPUE in the OTB-DES métier occur in shallow waters, during the crescent moon and in the second quarter of the year, and when the CPUE is higher.

Figure 3.13 shows the predictive spatial abundance of discards of the OTB-DES métier, influenced by the relevant factors, in the GSA06 South area. These discards show a longitudinal gradient, with the highest values in the central western part of the GSA06 South, along the coastline.

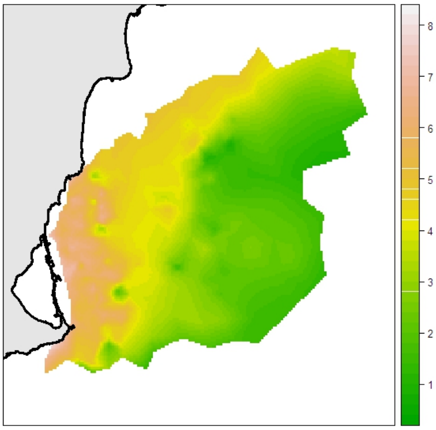
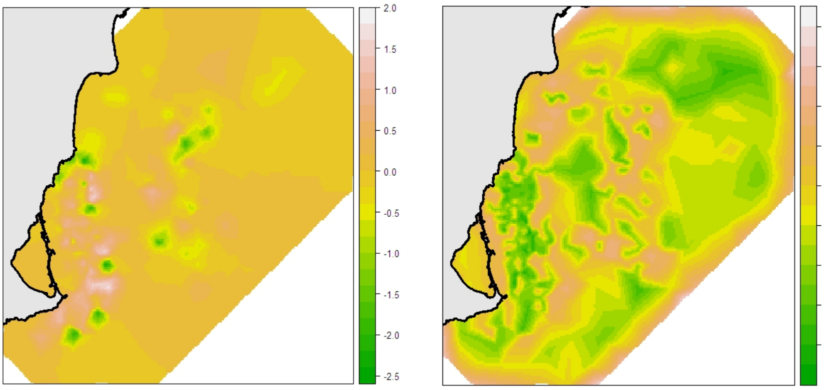


Figure 3.13. Posterior mean of the predictive discard abundance of the OTB-DES métier.

Figure 3.14 displays the posterior mean (a) and standard deviation (b) of the spatial component. The effect of the spatial component was consistent for all models. This component shows different marked hot spots with positive values in the western part, near the coast, and sporadic areas that show negative values.

The best model fit for the OTB-DWS métier includes the log-transformed CPUE, quarter of the year, moon phase, vessel length and type of substratum as relevant covariates together with the vessel and spatial random effects (Table 3.8).

Moon effects change smoothly declining from full moon (posterior mean = -0.07; 95% CI = [-0.50,0.35]) through to the waning moon phase (posterior



(a) Posterior mean (b) Standard deviation

Figure 3.14. The posterior mean (a) and standard deviation (b) of the spatial effect of the OTB-DES métier.

mean = -0.28; 95% CI = [-0.69,0.12]) with respect to the reference level (new moon).

As in the OTB-DES métier, the log-transformed CPUE shows a positive relationship with respect to DPUE abundance (posterior mean = 1.09; 95% CI = [0.93,1.24]).

Regarding seabed, the rock substratum shows the lowest estimated DPUE abundance (posterior mean = -0.07; 95% CI = [-0.75,0.60]) with respect to the reference level (sand substratum). Muddy substrata showed a higher estimated coefficient than the reference level (posterior mean = 0.12; 95% CI = [-0.26,0.49]).

In this case depth is not relevant, and nor is the slope and orientation of the seabed, as in the OTB-DES métier. Neither was the presence or absence of light during the hours of trawling found to be a relevant variable for the DPUE abundance in this métier.

Table 3.8. Numerical summary of the posterior distributions of the fixed effects for the OTB-DWS métier.

Predictor	Mean	Sd	$Q_{0.025}$	$Q_{0.5}$	$Q_{0.975}$
(Intercept)	0.05	0.23	-0.41	0.05	0.51
Moon(Full)	-0.07	0.22	-0.50	-0.07	0.35
Moon(Crescent)	-0.12	0.20	-0.51	-0.12	0.28
Moon(Waning)	-0.28	0.21	-0.69	-0.28	0.12
Log(CPUE)	1.09	0.08	0.93	1.09	1.24
Vessel length	0.06	0.03	0.01	0.06	0.12
Seabed(Mud)	0.12	0.19	-0.26	0.12	0.49
Seabed(Rock)	-0.07	0.35	-0.75	-0.07	0.60
Quarter 2	0.13	0.15	-0.16	0.13	0.42
Quarter 3	0.10	0.16	-0.22	0.10	0.41
Quarter 4	-0.02	0.15	-0.32	-0.02	0.29

The second quarter of the year shows the highest estimated DPUE abundance (posterior mean = 0.13; 95% CI = [-0.16,0.42]) with respect to the reference level (first quarter), while the fourth quarter show the lowest estimated coefficient (posterior mean = -0.02 ; 95% CI = [-0.32,0.29]).

The vessel random effect was relevant in all the models run, while among the technical characteristics of the vessel, only the length is found to be relevant. In particular, longer vessels show higher DPUE values (posterior mean =0.06; 95% CI = [0.01,0.12]).

The higher values of DPUE for the OTB-DWS métier are recorded for longer vessels, on muddy substrata, in the second quarter of the year, when the moon phase is new and the CPUE is higher.

Both the map of the predictive spatial abundance of discards and that of the posterior mean of the spatial effect (Figures 3.15 and 3.16 (a)) show a patchy distribution of the DPUE of the OTB-DWS méteir with three marked hot spots of higher DPUE values.

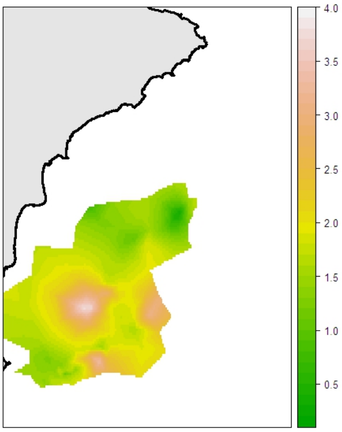
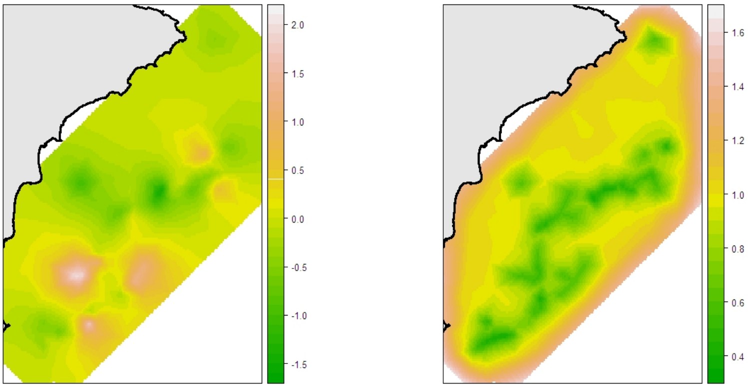


Figure 3.15. Posterior mean of the predictive discard abundance of the OTB-DWS métier.

Conclusions

Estimation of discards and knowledge about the reasons why discarding occurs, have been recognized to be crucial for improving stock assessments and exploring the impacts of fishing on the ecosystem (Tsagarakis et al., 2012). The latter has gained attention during the last decade, ever since an ecosystem approach to fishery management (EAFM) has been established as a priority in fishery science (Bellido et al., 2011). Solving the problem of discards is a complex issue, since discards show high variability across time, space and métiers due to the numerous factors affecting them, including, among others, technical characteristics, environmental conditions and species composition (Rochet and Trenkel, 2005).

In this study we have analyzed about 440 trawl fishing operations, sampled



(a) Posterior mean

(b) Standard deviation

Figure 3.16. The posterior mean (a) and standard deviation (b) of the spatial effect of the OTB-DWS métier.

between 2009 and 2012, to improve our understanding of those factors that influence the quantity of discards and to identify their spatial-temporal distribution in the study area. Two different métiers were analyzed, the bottom otter trawl demersal species métier (OTB-DES), and the bottom otter trawl deep-water species métier (OTB-DWS). Firstly, we analyzed the discards of both métiers in order to understand their quantity and species composition. Secondly, we have focused our analysis on factors influencing discards so as to identify their spatial-temporal patterns in the study area.

It is known that, among different fishing gears, the trawl is responsible for most fisheries discards (Tsagarakis et al., 2008). In the Mediterranean, the discarded fraction of otter-trawl catches ranges from 20 to 70% by weight (Carbonell et al., 1998; Machias et al., 2001; Kelleher, 2005). In our study area, for the whole sampled fleet of trawlers, the discarded fraction accounted for 31% of the total catch, more specifically 34% and a 20% of the total catch for the OTB-DES and OTB-DWS métiers, respectively.

Similar studies on demersal trawls, on a broad scale, reported higher discard ratios, as in the case of the northeastern Mediterranean sea (38-49%; (Machias et al., 2001; Tsagarakis et al., 2008)). However, the discard ratio of the OTB-DES and OTB-DWS métiers in the study area was higher than that of midwater trawls in the Turkish Black Sea (5.1%; (Kelleher, 2005)) and the Adriatic (up to 15%; (Santojanni et al., 2005)), as well as that of the artisanal fishery in Greek waters (10%; (Tzanatos et al., 2007)) and the purse seine fisheries in the western Mediterranean Sea (13-15%; (Kelleher, 2005)).

From the point of view of species composition, the results show that the major proportion of catch and discards is represented by non-target and low commercial value species. In particular, in the OTB-DES métier, bogue and axillary seabream are the most discarded and caught species, representing up to 32% of total discards. In the OTB-DWS métier the most captured species corresponds with the target species (red shrimp) of this métier, unlike the OTB-DES where the main target species is European hake, which is the tenth most captured species. Among the high commercial value species, such as the red shrimp and octopus, discards are negligible for both métiers (0.3% and 3% respectively). Moreover, a significant fraction of the discards of both métiers (10%) consists of elasmobranch species, which are considered vulnerable species due to their biology and K-selection life-history traits. The discarding of non-target species may have negative consequences for both commercial and no-ncommercial species due to its effects on species interactions and cascading effects throughout the trophic web.

Previous studies that have investigated the spatial-temporal variability of, and factors influencing, discards have focused only on target species or species with a high commercial value, as well as global discard estimates. To investigate only the spatio-temporal variability and quantity of target species discards could lead to underestimated and biased conclusions concerning this fleet.

In order to overcome this problem, and to understand which factors influence the variability of the discards, the DPUE of each haul, shared by métier, was modeled with respect to the environmental and technical characteristics of fishing operations for each métier.

Our analysis, perfomed using Bayesian methods, showed that the relative importance of each variable was different for each métier, although with a few similarities were observed. Interestingly, for both métiers, the discarded quantities were not found to be related to factors such as vessel power, GRT, daylight during the haul or enviromental factors such as seabed slope and orientation.

Among the vessel characteristics, only for the OTB-DWS métier does vessel length influence the abundance of discards. A longer vessel implies a greater catch and, implicitly, a higher discard fraction. Indeed, one of the main driving variables that explain the discard variability is the abundance of catch. Our results show a direct and positive relationship between the CPUE and the DPUE, with more catch involving an increase in discard abundance for both métiers.

Surprisingly, for both métiers, the moon phase was relevant for discard variability. As mentioned previously, a considerable proportion of discards consists of elasmobranch species, whose distribution has been related with the lunar phases by several studies (Poisson et al., 2010; Cuevas-Zimbrón et al., 2011).

The type of seabed was only relevant for the OTB-DWS métier. Muddy substrata are those with the highest amount of discards, reflecting the selection of fishing grounds by fishermen and the distribution of the target species of the métier.

Furthermore, discarding is a process decided on board based on the size of the catch and market prices of species, taking into account legal constraints. In our results, the random vessel effect collects this hidden variability that depends only on of the vessel's crew.

Discards fluctuated greatly in each métier, but showed no relevant temporal trends between years. On the contrary, interannual variability was a relevant variable for both métiers. In particular, the second quarter of the year is the period which recorded the highest abundance of discards for both métiers. This seasonal pattern of discards can be attributed to the targeting behaviour of the fishermen and the condition/behaviour of species in the different seasons.

The spatial effect explained much of the variability in discarded quantities for both métiers. The spatial component may reflect the effect of other factors considered and of factors that are not directly taken into account, such as community composition, distance from the coast or productivity gradients, etc., and can contribute to making a good estimate of discards. Maps show a clear spatial longitudinal gradient for the OTB-DES métier, with the highest discards occuring in the central western part of the study area, along the coastline. This trend is confirmed by the relevant negative relationship between the abundance of discards and the fishing depth variable. The abundance of discards is higher in shallow water and along the coastline, and may reflect the selection of fishing grounds in this métier.

The map of the spatial component of the the OTB-DES métier shows several hot spots with high discard values and sporadic areas with lower abundance of discards. This trend reflects resources distribution, and is very useful for identifing sensitive areas that could be avoided by fisheries in order to decrease discards.

Moreover, the spatial predictive discards map and the spatial effect map of the OTB-DWS métier also highlight clear hot spots of discard abundance.

The identification of these areas of high concentration of discards could be an important benefit for the spatial management of the fleet. The interannual/spatial effects could potentially be exploited in an overall strategy of spatial management strategy in order to reduce discard rates, providing the necessary economic incentive for fishermen to adopt selective temporal rotation of fishing grounds. Our findings show that the spatial variability in discard rates can potentially be exploited in a general strategy to reduce discards. A similar approach was proposed for the USA mixed species otter trawl fisheries of the Georges Bank-Southern New England region. By limiting directed fishing to times and places where resources are segregated, the quantity of unintended catch could potentially be reduced (Murawski, 1996). To achieve these purposes, predictive spatial maps, such as the ones generated by our approach could be essential tools for implementing an efficient spatial management and control schemes to reduce discards.

The contents of the Section 3.6 have been submitted in a peer-reviewed journal.

3.7 General conclusions

Improved knowledge of the spatial distribution of commercially fish species and their relationships with the marine environment could form an integral aspect of a precautionary approach (Thrush and Dayton, 2010). There is growing worldwide support for the argument that fisheries should be conducted in a way that is robust to environmental change, and thus that fish stocks should never be exploited to a point where they cease to be resilient to environmental change.

In order to achive this purpose, the European Marine Strategy Framework Directive has specifically requested the incorporation of any existing knowledge about environmental drivers in the assessment of ecosystems and fisheries and in advice (Wiegand et al., 2011).

Implementing plans and control schemes that would specific species assemblages, based on a good knowledge of the species' spatial distribution in relation to the environment, could serve as a sustainable approach to fisheries management and also for other marine activities and uses (Martin et al., 2012). The establishment of marine protected areas for protecting sensitive habitats, in line with the recent trends for a holistic ecosystem-based approach to management and Marine Spatial Planning, would also benefit from an improved understanding of the spatial distribution of vulnerable species. To achieve these objectives, predictive habitat maps, such as those generated by the approach presented in this chapter, could be useful source of information in the selection of areas for improved regional management or technical measures adoption.

Species habitat analysis should be able to identify those areas within the distribution of a species that contribute most to sustain the long-term viability of a population. Although it may be complicated to define the boundaries of sensitive habitats, the definition of these areas, combined with an efficient fishery management that recognizes the importance of such areas, represents the first step towards facilitating an EAFM approach (Valavanis et al., 2008b). Accuracy is not always easy to achieve because there is often a large amount of variability surrounding the measurements of response and environmental variables (Latimer et al., 2006). This variability leads to uncertain predictions, and consequently to uninformed decision making. It is therefore important to develop tools which account for measurements with significant variability. Here, we have shown a Bayesian spatial hierarchical model that makes it possible to identify sensitive habitats together with a full specification of associated uncertainty.

However, in all the presented cases, species, discards and environmental data are sampled during a limited period of time and space, and thus the models fitted can only reflect a snapshot view of the expected relationship. Future studies should compare these temporal and spatial trends from additional sources of data with a wider spatial and temporal coverage, such as fishery-independent scientific survey, which is often considered as being a more reliable abundance index because of its scientifically rigorous design (Cao et al., 2011). However, this first approximation could be serve to promote a greater effort in data collection and to identify the areas on which future research attempts should concentrate.

It should be noted that although the fishery-dependent data that we used were collected with a systematic sampling design, Bayesian spatial models may also aid analyses of data with geographically uneven levels of survey effort because such bias can be incorporated within the spatial random-effect term, thereby reducing its influence on estimates of the effects of environmental variables. By treating spatial effects as a variable of interest, hierarchical Bayesian spatial models can suggest the identity of additional environmental covariates that may improve model fit or the existence of area effects that may limit population viability.

In all the practical cases presented, the main advantage of the Bayesian model formulation is the computational ease in model fit and prediction compared to classical geostatistical methods. Both the stationary and especially the non-stationary models have a large number of parameters. Also, in classical geostatistical applications, the full range of uncertainties that are always associated with species distribution models is not correctly measured, as many parameters that are considered to be “known" are actually estimated through the statistical model (Diggle and Ribeiro, 2007), a potential cause of optimistic assessments of predictive accuracy. Using Bayesian kriging, we have incorporated parameter uncertainty into the prediction process.

The main goal of these studies has been to predict the occurrence of the species and the discard abundance in unsampled areas. To do so, instead of MCMC we have used the novel integrated nested Laplace approximation approach. More precisely, we have applied the work of Lindgren et al. (2011), which provides an explicit link between Gaussian Fields and Gaussian Markov Random Fields through the Stochastic Partial Differential Equation approach. Thanks to the `R INLA` library, the SPDE approach can be easily implemented providing results in reasonable computing time (in contrast to MCMC algorithms). The simplicity of the SPDE parameter specifications provides a new modelling approach that allows an easy construction of non-stationary models that provides a good, computationally very efficient, local interpretation. For these reasons, the SPDE approach has proved to be a powerful strategy for modelling and mapping complex spatial occurence phenomena.

Finally, we would like to mention that the analytical approach we used here to document the spatial patterns of these specific cases can be extended to different species and life stages to improve knowledge of the role of habitat for populations and marine processes.

Chapter 4

Modelling the spatial distribution of target species

In this chapter, we present a Bayesian hierarchical spatial model to estimate the distribution of fishery target species using fishery-dependent data. In these specific cases in which the data are collected following a preferential sampling method, predicting the distribution of the target species using the traditional application of geostatistical methods could lead to biased predictions. To avoid this bias, the basic idea is to interpret the data as a marked point pattern, where the sampling locations form a point pattern and the measurements taken in those locations are the marks. As a practical example, we model the abundance of European hake (*Merluccius merluccius*) in the Gulf of Alicante.

4.1 Introduction

The increasing interest in species distribution models for management and conservation purposes has focused the attention of ecologists on the use of spatial models (Sangermano and Eastman, 2007). Spatial species distribution models can be addressed for theoretical and practical problems, including estimating the relationship between species and their environment, identifying and managing protection areas and predicting their potential response to ecological changes (Latimer et al., 2006). In all these contexts, the main issue is to use and link information on the presence/absence or abundance of a species and the environmental variables associated with predicting where (and how) the species is likely to be present in unsampled locations.

In species distribution models one of the most difficult steps to overcome is obtaining observation data on the species studied (Kery et al., 2009). With the exception of a few pilot methodological studies (Thogmartin et al., 2006), species distributions must frequently be derived from opportunistic observations due to the high cost and time consuming nature of collecting data in the field (Kery et al., 2009). The development of a well designed sampling scheme to estimate directly the distribution of a specific species over the entire study area is quite often unfeasible (Brotons et al., 2007). As an example, studies on sea mammals commonly resort to the affordable technique of sampling from recreational boats (so-called platforms of opportunity), whose bearings are neither random nor systematic (Rodríguez et al., 2007). Most of the data obtained in bird monitoring are also provided by volunteers who concentrate the sampling process in areas where they expect to find the species of interest (Brian, 2011). Moreover, in a fishery ecology context, the most readily available data are fishery-dependent surveys from commercial fleets, usually monitored by observers on board, due to their connection to the economy and business (Vasconcellos and Cochrane, 2005).

Although most of the time they are all that is available, data of this kind raise a practical complication: the sampling process that determines the data locations and the species observations are stochastically dependent (Diggle et al., 2010). Spatial species distribution models usually assume, if only implicitly, that sampling is non-preferential and that the selection of the sampling locations does not depend on the values of the spatial variable. However, data from commercial fishery are a clear example of preferential sampling. Fishermen clearly go to fish in areas where they are likely to find target species.

This problem not only arises in fishery, preferential designs can arise in other contexts, and the underlying reasons could be either because sampling locations are deliberately concentrated in areas where the values of the species of interest are thought likely to be greater (or lesser) than average (Diggle et al., 2010).

As a result, traditional application of geostatistical methods performed solely on the basis of data obtained from these frameworks could lead to biased results if we do not take into account the choice of monitoring locations. The underlying reason is that geostatistical models focus on inferring a continuous spatial process based on data observed at finitely many locations, with the locations typically assumed to be non-informative. These models assume that sampling location and the process being modelled are stochastically independent. But what should we do when this is not the case?

Diggle et al. (2010) showed how ignoring preferential sampling could lead to misleading inferential conclusions. The authors addressed the issue of preferential sampling indicating how this can be approximately evaluated using Log-Gaussian Cox processes (LGCPs) and likelihood-based inference Monte Carlo methods. However, the resulting methods are computationally intensive, as the authors recognize in their discussion, which implies that it is quite complicated to use in practical situations, especially when the objective of the analysis is to perform the prediction in many unobserved locations using environmental covariates. This is an important issue as prediction is often the main utility of the modelling, making then the effect of preferential sampling rely more on the resulting predictive surface than on parameter estimation.

Rue et al. (2010) indicated, in the discussion on Diggle et al.'s paper, that the preferential model is an example of a latent Gaussian model for which

Bayesian inference based on integrated nested Laplace approximation (`INLA`) is both fast and efficient.[1]

In this Chapter we re-examine the methodology for performing preferential sampling models using the approach proposed by Rue et al. (2010). In order to illustrate how preferential sampling affects the performance of standard methods, we apply the preferential approach to assess the spatial distribution of European hake (*Merluccius merluccius, Linnaeus 1758*) in the Gulf of Alicante.

Furthermore, we present the possibility of incorporating an additional spatially structured effect in the model that reflects how wrong the observers were. Indeed, in some cases, the observers will sample in areas where they believe that the variable studied has high values, but they might not be absolutely right. This is because the sampling locations do not in fact directly depend on the "true" availability of the sample but on the observers' idea about what the availability is.

The remainder of this Chapter is organized as follows. After this introduction, in Section 4.2, we review all the different ingredients required to fit a preferential sampling model. Section 4.2.2 explains the modelling approach for the preferential sampling and its improvement. Section 4.3 validates the performance of this methodology through an application within the context of mapping the distribution of the European hake in the Western Mediterranean Sea. We end with a general discussion in Section 4.4.

4.2 Tools for implementing preferential sampling

In this section we review all the necessary tools to fit a preferential sampling model. For a more detailed explanations of the processes presented here see for instance Stoyan et al. (1995), Møller et al. (1998), Diggle (2003), Møller and Waagepetersen (2004), Illian et al. (2008), Ho and Stoyan (2008), Law et al. (2009) and Illian et al. (2012) and references cited therein.

---

[1] Code for solving Diggle et al.'s presented example is available in the (`INLA`) project web page (http://www.inla-project.com).

Formally, a spatial point pattern describes the spatial location of events in a defined study region. Spatial point processes are mathematical models used to describe and analyse these spatial patterns formed by objects that are irregularly or randomly distributed in one or more dimensional space. Examples include locations of plants in a wood, parasites in a culture microbiology and concentrations of contaminants in the air or in the soil.

The most common theoretical model for a spatial point pattern is the Poisson process, usually described in terms of its intensity function $\Lambda(x)$. This function represents the mean number of events per unit area at the sampling location. Intensity may either be constant over space, resulting in a homogeneous or stationary pattern, or vary in space with a spatial trend, resulting in a non-homogeneous pattern. The homogeneous Poisson process represents an idealized form of spatial randomness, where the intensity function is constant and each event occurs independently of the others, as a result of a uniform distribution in space. Nevertheless, the assumption of stationarity is generally unrealistic in most practical applications, making non-homogeneous Poisson processes a better choice to describe reality.

Among all spatial point processes, the class of Cox processes is a flexible family of models, also called "doubly-stochastic" processes, as they arise as non-homogeneous Poisson processes with a random intensity measure. Cox processes provide a statistically tractable class of models for aggregated point patterns in which the spatial trend can reflect underlying heterogeneity due to unknown conditions such as environmental variability.

In Cox processes the variability is assumed to be a stochastic process in itself. In addition, spatial autocorrelation and variation are expressed with a random structure based on an underlying (or latent) random field, which describes the stochastic intensity $\Lambda(s)$, assuming independence from this field. Log-Gaussian Cox processes (LGCPs) are a particular class of Cox processes in which the logarithm of the intensity surface is a Gaussian field. More formally:

$$\log(\Lambda\,(s)) = Z\,(s) \tag{4.1}$$

where $Z(s)$ is a Gaussian random field, $s \in \mathbb{R}^2$ and is conditional on the unknown conditions. In other words, given the random field, the points are independent and hence form a non-homogeneous Poisson process.

The practical fitting of LGCP models to point pattern data remains difficult due to intractable likelihoods as it requires a complicated computation of the integral of the intensity function and its stochastic nature. Methods based on Markov chain Monte Carlo (MCMC) can be very time-consuming both to tune and to run and fitting complex models can easily become computationally expensive.

Nevertheless, LGCPs are a special case of the more general class of latent Gaussian models, which can be described as a subclass of *Structured Additive Regression (STAR) models*. In these models the mean of the response variable is linked to a structured predictor, which can be expressed in terms of linear and non-linear effects of covariates. In a Bayesian framework, by assigning Gaussian priors to all random terms in the predictor, we obtain a latent Gaussian model. As a result we can directly compute LGCP models using the Integrated nested Laplace approximation (`INLA`)(Rue et al., 2009). `INLA` provides a fast and yet reasonably exact approach to fitting latent Gaussian models, and makes the inclusion of covariates and marked point processes mathematically tractable with computationally efficient inference (Simpson et al., 2011).

Marked point processes provide a very useful tool in spatial statistics. They represent a natural approach to the analysis of data where random variables are observed at random positions. Indeed, in many applications the natural set of covariates under consideration will include attributes to event positions. This can be both a qualitative (such as species identification) or quantitative (such as height) measure, usually referred to as marks.

Formally, a marked point process is a random sequence $\Psi = \{[\mathbf{x_n} : \mathbf{m_n}]\}$ with points $\mathbf{x_n} \in \mathbb{R}^2$ and marks $\mathbf{m_n}$ in some mark space. The character Ψ denotes the random measure and is assumed to be stationary and isotropic.

Two different kinds of marked point process models can be considered: processes where marks are a response variable and processes where marks are included as an explanatory variables.

When marks are treated as a response variable and are modeled near to the pattern in a joint model, it is necessary to consider a separate likelihood and latent field, as well as a set of hyperparameters, specific to the marks. In this case, a joint model is particularly appropriate and model comparison methods can be used to assess whether or not the model fit improves when the fields are consider separate or jointly.

To end this Section we would like to remark that very few attempts have been made at fitting models to both the pattern and the marks, in particular to patterns with multiple dependent continuous marks, whilst joint models of covariates and patterns have not been considered.

4.2.1 Fitting models with `INLA`

Unlike MCMC, which makes use of stochastic simulation for fitting, `INLA` use deterministic approximation. It has been designed to work for a large class of models, namely latent Gaussian models, and these are fitted in a Bayesian setting. While `INLA` cannot be used for all models, the class of latent Gaussian models comprises a lot of models, including time series models, generalised additive models and survival models, among others (Rue et al., 2009). Latent Gaussian models consist of three levels: the observations, an underlying latent structure and a vector of hyperparameters:

- The observations (y) encode information about observed data, including design and collection issues, where:

  $$\pi(y|\eta) = \prod_{k \in \mathcal{K}} \pi\left(y_k|\eta_k\right) \tag{4.2}$$

  where $\mathcal{K}$ is an index for grid cells and given a latent field, the observations are considered independent;

- the latent field (η) is a latent (Gaussian) random field, relating to observed and unobserved (spatial) covariates and is included to reduce (spatial) autocorrelation,

$$\pi(\eta|\theta) = N\left(0, \sum(\theta)\right) \tag{4.3}$$

- and hyperparameters (θ) that are parameters for the latent field

$$\pi(\theta) \tag{4.4}$$

In `INLA`, the Gaussian field is approximated by a Gauss Markov random field (GMRF). As a result it is possible to exploit all the characteristics of the latent GMRF model. Firstly, GMRFs are discretely indexed, so the Markov property makes their precision matrix sparse

$$Q(\theta) = \sum^{-1}(\theta), \tag{4.5}$$

is computationally elegant and allows fast numerical integration if the dimension of the vector of hyperparameters θ is chosen to be small. In summary, the structure of the latent Gaussian model allows `INLA` to approximate and speed up the fitting process by exploiting the fact that the latent field is Gaussian such that a Laplace approximation may be used. As a result, the time required for fitting models is substantially reduced.

Moreover, the Matérn covariance function is a highly flexible and general family of functions covering many of the most-used covariance models in spatial statistics (Stein, 1999) (although in this approach it is limited to $\nu = 1$). Finally, GMRFs fit seamlessly with the `INLA` approach, which requires the latent field to be a GMRF.

4.2.2 Implementing the Preferential sampling approach

In this Section we review the Rue et al. (2010) approach proposed in the discussion of Diggle's paper while presenting some improvements such as the possibility of incorporating an additional spatially structured effect in the model which reflects how wrong the observers were.

In preferential sampling, the data structure is a spatial point pattern formed by the sampling locations $\mathbf{x} = (\xi_1, \ldots, \xi_n)$ and associated quantitative marks

$\mathbf{m} = (m_1, \ldots, m_n)$. The spatial point pattern $\mathbf{x_n}$ depends on an observed random field $\{Z(s)\}$ and forms an LGCP with an intensity function $\Lambda(s)$ with the form:

$$\Lambda(s) = \exp\{\alpha + \beta S(s)\}, \tag{4.6}$$

where parameters α and β are real and the marks $\mathbf{m_n}$ are simply the value of the background random field $\{Z(s)\}$.

Moreover, marks $\mathbf{m}$ are assumed to follow an exponential family distribution $F_{1\theta_1}$ with parameter vector $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_q)$ (such as a normal distribution with parameter $\theta_1 = \mu$ and $\theta_2 = \sigma^2$) and to depend on the intensity of the point pattern.

Since `INLA` uses a GMRF approximation to the Gaussian random field the observations have to be discretised. For this purpose, the observation window is discretised into $N = n_{row} \times n_{col}$ grid cells $\{s_{ij}\}$, each with area $|s_{ij}|$, $i = 1, \ldots, n_{row}, j = 1, \ldots, n_{col}$. The points in the pattern can then be described by $\{\xi_{ijk_{ij}}\}$ with $k_{ij} = 1, \ldots, y_{ij}$, where y_{ij} denotes the observed number of points in grid cell s_{ij}. The observed number of points in each grid cell is assumed to be Poisson distributed given the latent field η_{ij}, i.e.

$$y_{ij}|\eta_{ij} \sim Po(|s_{ij}| \exp(\eta_{ij})), \tag{4.7}$$

see (Rue et al., 2009).

We model η_{ij} as

$$\eta_{ij} = \beta_0 + W_f(s_{ij}) + u_{ij}, \tag{4.8}$$

where $W_f(s_{ij})$ refers to a spatially structured effect (i.e. a GMRF) that reflects spatial variation in the pattern and u_{ij} denotes a spatially unstructured i.i.d. error term with $u_{ij} \sim N(0, 1)$.

For the marks, we construct a model where the marks $\mathbf{m}$ depend on the pattern by assuming that they depend on the same spatially structured effect $W_f(s_{ij})$. Specifically, we assume that $m(\xi_{ijk_{ij}})|\kappa_{ijk_{ij}} \sim N(\kappa_{ijk_{ij}}, \sigma^2)$ with

$$\kappa_{ijk_{ij}} = \beta_{02} + W_f(s_{ij}) + v_{ijk_{ij}}, \tag{4.9}$$

where $v_{ijk_{ij}}$ is another error term.

Usually, the spatial effect is modelled using a second-order random walk on a lattice, using vague gamma priors for the hyperparameter and constrained to sum to zero. Here we propose the use of a Matérn correlation function on the regular lattice. As mentioned in the Section 3.2.2, the Matérn covariance function is a really flexible and general family of functions generalizing many of the most-used covariance models in spatial statistics. Its expression, giving the covariance between the values of a random field at locations separated by a distance $d > 0$, can be parameterized as

$$Corr(d) = \frac{1}{2^{\nu-1}\Gamma(\nu)}(\kappa d)^{\nu} K_{\nu}(\kappa d),$$

where K_ν is the modified Bessel function of the second kind and order $\nu > 0$ (Abramowitz and Stegun, 1970, §9.6), $\kappa > 0$ is a scaling parameter and the range is defined to be

$$r = \sqrt{8}\nu/\kappa$$

which about the distance where the covariance function becomes about 0.1.

The hyperparameters of the spatial effect are the precision parameter τ and the range r,

$$W = (\tau, r)$$

The latent field has marginal variance $1/\tau$ and range (as defined above) r. The hyperparameters are represented internally as

$$(log(\tau), log(r))$$

and the prior are assigned to these quantities. For instance, in the dataset described in Section 4.3, we specify minimally informative priors $\mathcal{L}ogGam$(1,0.0005).

Under this perspective, the Point Process model can be stated as follows:

$$\begin{aligned}\boldsymbol{x}|\boldsymbol{\eta} \sim & \mathcal{P}(\exp{(\boldsymbol{\eta})}) \\ \boldsymbol{\eta} = & \mathbf{1}\beta_{0\eta} + \mathbf{X}_{\eta}\boldsymbol{\beta}_{\eta} + W \\ \pi(\beta_0) \propto & 1 \\ \beta_j \overset{\text{iid}}{\sim} & \text{N}(0, \text{1e-05})\end{aligned} \tag{4.10}$$

and the model for the marks as:

$$\begin{aligned}y_i|\kappa_i \overset{\text{iid}}{\sim} & \text{N}(\kappa_i, \tau^2) \\ \kappa_i = & \beta_{0\kappa} + \mathbf{X}_{i\kappa}\boldsymbol{\beta}_{\kappa} + W_i \\ \pi(\beta_0) \propto & 1 \\ \beta_j \overset{\text{iid}}{\sim} & \text{N}(0, \text{1e-05})\end{aligned} \tag{4.11}$$

In order to feed this model into `INLA` we need to use a specific structure in the dataset (see the section *models with more than one type of likelihood* at http://www.r-inla.org/models/tools). Namely, we have to put the observations y_{ij} corresponding to the point pattern and the observations m_{ij} of the marks into a matrix of the form

$$Y = \begin{bmatrix} \boldsymbol{y}_{p\times 1} & \mathbf{NA}_{p\times 1} \\ \mathbf{NA}_{n\times 1} & \boldsymbol{m}_{n\times 1} \end{bmatrix}.$$

Where $p = n_{row} \cdot n_{col}$ is the size of the grid and $\mathbf{NA}_{n\times 1}$ stands for a $n \times 1$ matrix filled with `NA` values.

Correspondingly, the dataframe with the covariates will also be split into two parts with p elements each. Each covariate will be a column with the corresponding values in the first p elements if the covariate affects the point process model and `NA`s in the last p elements, or viceversa, if the covariate affects the marks models.

Since `INLA` by default uses a common intercept for all the data, we need to remove it (using -1 in the formula specification), and use a factor to represent fixed effect with two levels.

During the modelling process, methods for model comparison based on the deviance information criterion (DIC) (Spiegelhalter et al., 2002), will be used to compare these three different models.

Non-Preferential model

A joint model without preferential sampling will consist of carrying out inference separately over the point process and the marks. In this way, we will be treating the marks as sampled in a random process:

$$\begin{aligned} \eta_{ij} &= \beta_0 + W_f(s_{ij}) + u_{ij}\,, \qquad (4.12) \\ \kappa_{ijk_{ij}} &= \beta_{02} + W_g(s_{ij}) + v_{ijk_{ij}}\,, \qquad (4.13) \end{aligned}$$

where $W_g(s_{ij})$ refers to a separate spatially structured effect of the marks model.

Improving the preferential model

In some cases, the observers will sample in areas where they believe that the variable studied has high values, but they not be absolutely right. This is because the sampling locations do not really directly depend on the "true" availability of the sample but on the observers' idea about the availability is. In these cases we assume an additional spatially structured effect $W_h(s_{ij})$ in the model for the marks, and the equation (4.8) changes as follows:

$$\eta_{ij} = \beta_0 + W_f(s_{ij}) + u_{ij}\,, \qquad (4.14)$$

$$\kappa_{ijk_{ij}} = \beta_{02} + W_f(s_{ij}) + W_h(s_{ij}) + v_{ijk_{ij}}\,. \qquad (4.15)$$

This additional random field would reflect how wrong the observers were.

4.3 Distribution of European hake in the Gulf of Alicante

In this section we present a practical application of the preferential approach within the context of mapping the distribution of the of European hake (*Merluccius merluccius, Linnaeus 1758*) in the Western Mediterranean Sea.

The European hake is among the most important commercial species in the Mediterranean, suffering from high fishing pressure and currently overexploited (Lleonart, 2005). It is a demersal species with wide geographical distribution throughout the Atlantic Ocean and the Mediterranean and, as such, has been extensively studied (Goñi et al., 1999; Alvarez et al., 2001; Kacher, 2005). Nevertheless, none of these studies have taken into account the fact that the fishery-dependent data come from preferential sampling.

The study area and trawl survey data

The study was conducted in the eastern Spanish coastal area of the Mediterranean Sea, corresponding to the GFCM fisheries management Geographical Sub-Area 06 (GSA 06) (see Chapter 3, Section 3.5, Figure 3.7). Specifically, our study focuses on the Gulf of Alicante, which is bounded by the Cape de las Huertas and Santa Pola. It has a surface area of 3,392 km^2 and an average shelf width of approximately 32 km. The largest fleet of the Gulf of Alicante is the bottom trawl, with 169 vessels landing an average of $8,000$ t per year. Seabed trawling usually takes place on the shelf, yielding a multispecific catch, with European hake (*Merluccius merluccius, Linnaeus, 1758*) being one of the main target species (García-Rodríguez et al., 2005).

The reference fleet for this study is comprised of the trawler fleet which lands its catch in the Santa Pola port. This fishery is a medium-distance fleet that usually operates in waters from the continental shelf around the 200 m. isobaths. It makes short hauls of about 2-4 hours trawling, comprising about 2-3 fishing hauls per trip. The fleet has been monitored onboard since 2003, when routine yearly sampling was unified through the EU Data Collection Regulation 1639/2001. Data on European hake abundance were collected for 8 years, with some 3-4 observations usually being deployed every month. The data set includes 481 hauls from 15 different trawler vessels and has been provided by the *Instituto Español de Oceanografía* (IEO, Spanish Oceanographic Institute).

Since the catches vary considerably between vessels, we computed the catch for unit effort (CPUE) as the ratio of totally catches and the hours of the haul duration for each fishing operation. In this case the CPUE are the marks and the fishing locations are the patterns. In this first approximation to the issue of preferential sampling, no covariables were included.

Results

All the resulting models obtained were fitted and compared and DIC values were used as a measure for goodness-of-fit of the models.

In particular, the first model, named `Non-Preferential`, consistsof carrying out inference separately over the point process and the marks. In this way, we are treating the hake CPUE (marks) as if they had been sampled in a random process, while at the same time we have a reference model to compare them with. In the second case (`Preferential sampling`) we link the two components of the model through a spatially structured effect, assuming no *fishermen error*. The third and last case is an extension (`Err-Preferential sampling`) that includes an additional spatial effect to account for *fisherman error*.

As mentioned in the Section 4.2.2, originally this approach used a second-order random walk (*rw2d*) to model the structured spatial effects. Here, we have used the Matérn latent model as an alternative to *rw2d*, obtaining getting better results.

It is worth mentioning that in this approach the inference is based on cell-counts and cell-means.

As shown in Table 4.1, DICs values indicate that the best model is the

`Err-Preferential sampling`.

Table 4.1. Model comparision.

Model	Formula for κ	DIC
Non-Preferential	$\beta_{0\kappa} + W_g(s_{ij})$	3398
Preferential sampling	$\beta_{0\kappa} + W_f(s_{ij})$	3295
Err-Preferential sampling	$\beta_{0\kappa} + W_f(s_{ij}) + W_h(s_{ij})$	2994

In complex systems, such as marine systems, considering preferential sampling only might be too simplistic. In fact, the difference between the `Non-Preferential` model and the `Preferential sampling` model is appreciated, but the real improvement is obtained when the error of the fisherman is introduced in the model.

Figure 4.1 (a) displays the posterior mean of the spatial component of the preferential sampling model with fisherman error. This component shows clear hot spots of high-CPUE density, indicated by the positive values, for the studied species, and particularly low-CPUE density areas, indicated by the negative values.

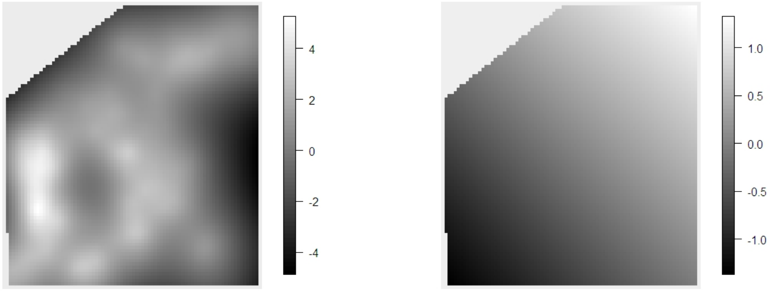
On the contrary, Figure 4.1 (b) shows the spatial component of the non-preferential sampling model. In this case the spatial component surface is smoother and no particular hot spot is marked for this species.

Another interesting result, independent of the species analyzed, is for computational level. Running the three models together took about 15 minutes.

4.4 Conclusions

In this Chapter we presente a modelling approach that could be very useful for the fisheries management and for marine resources in general. In fact, this methodology make it possible to use of opportunistic data, such as those arising from commercial fisheries, to acquire in-depth knowledge on marine species.

In fisheries ecology this approach could suppose a major step forward, given



(a) Err-Preferential sampling (b) Non-Preferential

Figure 4.1. Posterior mean of the spatial effect of the preferential sampling model with fisherman error (a) and (b) the non-preferential sampling model.

that most of the available data are those related to fisheries, although until now they could not be used properly in order to estimate the spatial distribution of the target species. Alternatively, other studies have used data from surveys (when available) to estimate the spatial distribution of the target species. In these cases, the drawback is that, given the high cost that supposes surveys, these are done in a limited time and space. As a result, these studies can only show a restricted temporal distribution over the year of the species studied.

The practical application of the European hake presented here as an example, shows that there is a real improvement in the estimation of the spatial distribution of the target species when a preferential sampling model is used. Moreover, the addition of the error of the fisherman, further enhances this approach.

Another advantage is undoubtedly the use of `INLA` in this context, which principally presents two essential benefits, the first being the speed of computation. As mentioned in the Results Section, running the three models together took about 15 minutes. Secondly, in the fishery context, adjusting spatial models can be problematic due to the boundary effect between sea and land when using the second-order random walk model. In the European hake application we succesfully used the Matérn latent model, given its ease of implementation in `INLA`.

Nevertheless, the next interesting step towards gaining more knowledge on marine species will be to include environmental variables in the model and to provide specific details about the essential habitat of the species studied. In addition, it could be interesting to take into account the differences between years by including a random temporal effect in the model. All these necessary integrations in the model could result in an increase in the computational time.

Moreover, this approach involves two different approximations, the latent Gaussian random field $Z(s)$ and the position of the points. Assessing the position of the data with a high degree of precision requires a great deal of effort because the quality of the likelihood approximation depends on the size of the grid. Consequentely, a much finer grid than the one needed to calculate latent Gaussian random field might be more appropriate. Furthermore, lattice-based methods do not allow the user to refine a specific local approximation to the latent random field. This could greatly affect the posterior inference in all those cases in which there is a large area for which no data is available. To avoid wasting computational resources to generate a high resolution approximation to the latent field over this area, it would be convenient to have the option of changing the resolution of the approximation locally, reducing the resolution in this area without compromising the resolution of sampled areas.

A good solution, could be the use of a computational mesh for representing the latent Gaussian random field using the Stochastic Partial Differential Equation (SPDE) of Lindgren et al. (2011) presented in Chapter 3. The great bonus of this approach could be that, for a finite dimensional, continuously specified Gaussian random field it could be possible to compute the term in the likelihood that depends on the data using the exact positions of the data points for any underlying grid. As a result, these triangulations could allow a more realistic modelling for spatial point processes without being computationally wasteful.

Chapter 5

A Bayesian approach to identifying fish nurseries

In this Chapter we propose a 2-step generalized linear Bayesian hierarchical spatio-temporal model, which consists in applying consecutively a binomial and a Gamma model, to analyze the occurrence and density data on juveniles. In particular, using European hake (*Merluccius merluccius*) as a model species, we estimate the distribution of juvenile occurrence and densities with respect to enviromental and geographical factors, using Bayesian spatial models, Bayesian kriging and fishery-independent data.

5.1 Introduction

The study of the potential areas of distribution of the species is a subject of great interest to fisheries science and marine ecology and questions about the spatial structure of fish populations and its links with environmental features are currently widely debated topics. The principal focus has been on understanding the spatial patterns of juvenile distribution, since a reduction in the fishing mortality of immature fish represents a fundamental prerequisite for sustainable fisheries (Beverton et al., 1957).

Furthermore, reducing any adverse impact, primarily from fishing, on juvenile habitats which are essential to allow the completion of a full life-cycle of a species, is one of the mandates of an Ecosystem Approach to Fishery Management (EAFM) (Garofalo et al., 2011).

Within this framework, a suitable and recommended management tool is the implementation of Marine Spatial Planning (MSP), which protects important ecological areas, such as nursery areas. MSP aims to establish a network of fishery restricted areas (FRA) or marine protected areas (MPA), sited where target species are known to aggregate in critical phases of their life cycle (e.g. juveniles/recruits) (Caddy, 2009). It is well recognised that temporary nursery closures from fishing activity could preserve the reproductive potential of populations and increase recruitment (Berkeley et al., 2004).

Usually, the definition of *nursery* takes into account the spatial persistence of hot spots of recruitment densities over long periods. Any area where juveniles occur in relatively high densities has been considered as a nursery (Dahlgren et al., 2006), but no clear definition has been provided.

Dahlgren et al. (2006) introduced the concept of *Effective Juvenile Habitats*, referring to habitats that make a greater than average overall contribution to the adult population. The spatio-temporal consistency of areas of high-density aggregations of juveniles can be interpreted as an indirect measure of nurseries due to their adaptive response of species life cycles to environmental factors. In fact, the maintenance of a population depends on the successful recruitment of young fish to nursery areas and from nursery areas back to the parent population (Hinckley et al., 2001). Colloca et al. (2009) suggested that this may represent a criterion for identifying a juvenile habitat when the number of juveniles which are successfully recruited to the adult stock cannot be estimated.

Implementing plans and control schemes that would target specific juveniles assemblages, based on a good knowledge of their spatial distribution in relation to the environment, could serve as a sustainable approach to fishery management and also for other marine activities and uses (Martin et al., 2012). The establishment of FRA and MPA for protecting nursery areas, in line with the recent trends for a holistic EAFM and MSP, would also benefit from an improved understanding of the spatio-temporal distribution of species recruits.

To achieve these objectives, predictive juvenile habitat maps could be a useful source of information in the selection of areas for improved regional management or the adoption of technical measures. In this context, the use of modelling approaches can provide the opportunity to estimate which different drivers could influence the recruit processes and to map essential juvenile habitats.

In this study we propose a 2-step generalized linear Bayesian hierarchical spatial-temporal model, which consists of a binomial and a Gamma model, to analyse the occurrence and density data on juveniles. In particular, using European hake (*Merluccius merluccius*) as a model species, we estimate the distribution of juvenile occurrence and densities with respect to enviromental and geographical factors, using Bayesian spatial models, Bayesian kriging (Diggle and Ribeiro, 2007) and fishery-independent data.

The Bayesian inferential paradigm is an attractive alternative to the classical statistical approach because of its philosophical foundation, its unified approach to data analysis, and its ability to formally incorporate random process error both into observed data and model parameters (Banerjee et al., 2004). This is essential in a study like this, where the main goal is to identify juvenile hot spots and to verify their persistence over time, with the least possible error.

We selected European hake as a case study because it is among the most important commercial species in the Mediterranean, suffering from high fishing pressure and currently overexploited (Lleonart, 2005). In fact, in many Mediterranean countries there is still a high marketability of small hake. As a result, this juvenile fraction is particularly exposed, especially to trawl fishery after the bottom settlement stage, when they aggregate over nursery grounds. Reducing fishing effort on these juvenile stages is of the utmost importance and effective management requires a detailed understanding of their spatio-temporal knowledge.

5.2 The data set

Data on juvenile hake were collected during the EU-funded MEDIterranean Trawl Survey (MEDITS) (Bertrand et al., 2002) trawl survey project, carried out from spring to early summer (May to July). The MEDITS project uses a stratified sampling design based on depth (5 bathymetric strata: 10-50, 51-100, 101-200, 201-500 and 501-700 $m.$) and Geographical Sub-Area (GSA). Sampling stations were placed randomly within each stratum at the beginning of the project and were sampled in all subsequent years.

Figure 5.1 shows the sampling locations of the MEDITS survey in the study area, which includes the GSA06 North and South. A total of 90 hauls, divided into the various bathymetric strata, are made every year. Each haul, conducted during daytime hours (06:00 to 18:00 h), lasts from 30 min (hauls up to 200 $m.$ in depth) to 60 min (hauls below 200 $m.$ in depth).

In our analysis, we included all available data from MEDITS collected from the 2000 to 2012. For each haul, we have considered only hake recruits, defined as those individuals less than 15 $cm.$ in total length. This length limit was selected using the *slicing* method. The slicing approach is part of a length-based cohort analysis which assumes that the size of the distributions for each cohort are normal and overlap each other. Thus, it seeks the point where the upper tail of the first cohort and lower tail of the second cohort are balanced in density and therefore it is possible to approximate a size for each age of the species. Once the length of the hake juveniles was established, our next step was to obtain the density of recruits for each haul as the total weight of juveniles ($Kg.$) for each 30 minutes of trawls (min). Our study was performed on the quadratic transformation of the density of recruits to ensure the normality of the response variable.

With respect to the enviromental covariates used in this analysis, we have

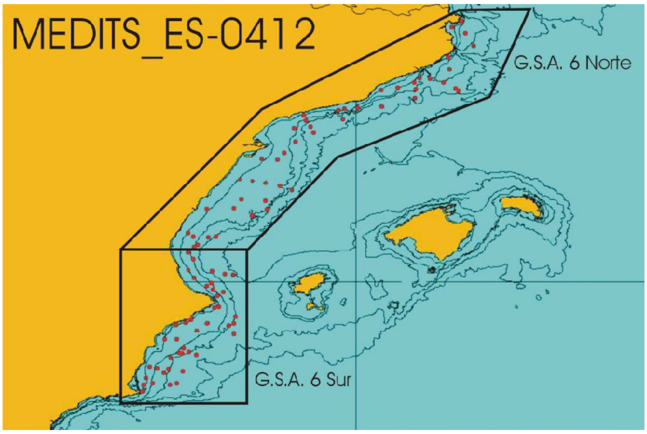


Figure 5.1. Sampling locations (●) of the MEDITS survey in the study area.

included those for which we have information, which also happen to be potentially relevant to decscribe the hake distribution. In particular, the two covariates used were bathymetry and the type of substratum. Both covariates were obtained from the IEO geoportal, accessible through the website of the Spanish Institute of Oceanography (`http:\www.ieo.es`). Quadratic-transformed bathymetry was included in the analysis for smoothing the effect and preserving the linearity of this variable.

The type of substratum polygon shapefile includes a classification of eight categories and a reduced version with three levels: sand, mud and rock. In order to reduce the level of variability in the analysis, since we have no observations in all categories, we have used the simplified version (Figure 5.2).

5.3 Modelling nursery areas

We present an extension of the Bayesian hierarchical model explained in Chapter 3.

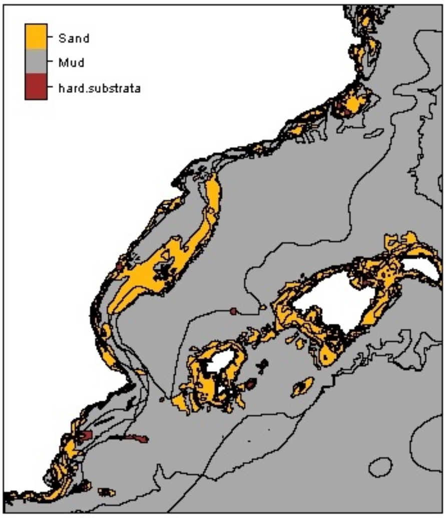


Figure 5.2. Map of the type of substratum in the study area.

Here we develope a 2-step generalized linear Bayesian hierarchical spatio-temporal model with the final aim of analysing both the occurrence and density data on juvenile hake.

Basically, we use two different types of observed data: the density of juveniles and their presence/absence data. In the first case it is possible to approximate the absolute abundance of species and to assess the quantitative spatial distribution of the juveniles within the area of interest. In the second case, presence/absence information can be used as a measure of the relative occurrence of species at each precise observed location, thereby giving a different approximation for the spatial distribution of the species. Note that this 2-step model is a very suitable approach for dealing with high numbers of zero observations, because the model deals with and consequently predicts the juvenile abundance and occurrence of juveniles separately.

Specifically, if Y represents the occurrence of hake recruits for each haul, then we can model it as:

$$\begin{aligned} Y_{ij} &\sim \mathrm{Ber}(\pi_{ij}), \quad i = 1, ..., n\,; \quad j = 1, ..., q \\ \mathrm{logit}(\pi_{ij}) &= X_i\beta + W + v_j \\ \beta_k &\sim \mathcal{N}(\mu_{\beta_k}, \rho_{\beta_k}) \\ W &\sim \mathcal{N}(0, Q(\kappa, \tau)) \\ 2\log\kappa &\sim \mathcal{N}(\mu_\kappa, \rho_\kappa) \\ \log\tau &\sim \mathcal{N}(\mu_\tau, \rho_\tau) \\ v_j &\sim \mathcal{N}(0, \rho_v) \end{aligned} \tag{5.1}$$

where π_i is the probability of hake presence in the location i, $X_i\beta$ represents the lineal predictor, W represents the spatial random effect, v is the temporal unstructured random effect which has been included to verify the persistence of the areas of density aggregation, and the relation between π_i and the covariates of interest and both random effects is the usual logit link.

In the second stage, we model the density of recruits in those places where the fish is present. In fact, we model the quadratic transformation of the density of recruits conditioned to the presence of recruits. If Z represents this new variable, we can express it via the following Gamma model as:

$$\begin{aligned} Z_{ij} &\sim \mathcal{G}a(a, b), \quad i = 1, ..., n\,; \quad j = 1, ..., q \\ \log(\mu_{ij}) &= X_i\gamma + W + v_j \\ \log(\phi_{ij}) &\sim \mathcal{N}(\mu_\phi, \rho_\phi) \\ \gamma_k &\sim \mathcal{N}(\mu_{\gamma_k}, \rho_{\gamma_k}) \\ W &\sim \mathcal{N}(0, Q(\kappa, \tau)) \\ 2\log\kappa &\sim \mathcal{N}(\mu_\kappa, \rho_\kappa) \\ \log\tau &\sim \mathcal{N}(\mu_\tau, \rho_\tau) \\ v_j &\sim \mathcal{N}(0, \rho_v) \end{aligned} \tag{5.2}$$

where each density recruits Z_{ij} has a Gamma distribution with parameters a and b, being respectively $\mu_{ij} = \frac{a}{b}$ and $\varphi_{ij} = \frac{a}{b^2}$, its corresponding mean and variance; $\phi_{ij} = \varphi\mu^2$ is a convenient reparametrization for the precision; $X_i\gamma$ represents the lineal predictor of the covariates of interest; W is the same random spatial effect of model 5.1; v is also the same temporal unstructured random effect previously mentioned model 5.1; and the relation between μ_i and the covariates of interest and random effects is the usual log link. For both models the random spatial effect depends on the parameters κ and τ which determine the range of the effect and the total variance, respectively.

Once the model has been determined, the next step is to estimate its parameters. Following Bayesian reasoning, the parameters are treated as random variables, and prior knowledge is incorporated via prior distributions. In particular, for the parameters involved in the fixed effects of the first model, we use the Gaussian distribution $\beta \sim \mathcal{N}(0, 10)$. In the second model, for the same parameters, we use the Gaussian distribution $\beta \sim$ N(0, 100). Moreover, for the hyperparameters derived from the spatial effect, κ and τ, we assume prior Gaussian distributions with a zero mean and covariance matrix depending on each hyperparameter, in particular, $W \sim \mathcal{N}(0, Q(\kappa, \tau))$, and $2 \log \kappa \sim \mathcal{N}(0, 10)$ and $\log \tau \sim \mathcal{N}(0, 10)$. Finally for the temporal unstructured random effect, we assume a non-informative prior distribution for the precision ρ_v, in particular a very vague LogGamma distribution for the precision ρ_v with hyperparameters a=1 and b= 0.000001.

As usual in this context, the resulting hierarchical Bayesian model has no closed expression for the posterior distribution of all the parameters, and so numerical approximations are needed. Here, as in the previous Sections, we use the integrated nested Laplace approximation (INLA) methodology and software (`http://www.r-inla.org`) as an alternative to the Markov chain Monte Carlo (MCMC) methods. In order to obtain posterior marginal distributions, we used the Stochastic Partial Differential Equation module (SPDE), which allows us to fit the particular case of continuously indexed Gaussian Fields by INLA, this being the case with our spatial component (Lindgren et al., 2011).

A model selection approach has been used to select relevant variables. Specifically, the Deviance Information Criterion (DIC) (Spiegelhalter et al.,

2002) is used as a measure for goodness-of-fit, while the logarithmic score (LCPO) measure the predictive quality of the models. The smaller the DIC and LCPO indicate the better the compromise between fit, parsimony and predictive quality.

5.4 Results

A total of 1080 hauls were analysed over the period 2000 to 2012 in the studied area. All the resulting models obtained from combining the bathymetry, its quadratic transformation, the type of substratum factor, the random spatial component and the temporal effects were fitted and compared. DIC was used as a measure for goodness-of-fit, while the logarithmic score (LCPO) measures the predictive quality of the models. Table 5.1 shows both measures for all the models run with the binomial response variable, while Table 5.2 shows the response variable for the Gamma distribution.

Table 5.1. Model comparison for the binomial response

Model	DIC	LCPO
1 + Depth + Depth^2+ Substratum + year + θ	478.95	-0.21
1 + Depth + Depth^2 +Substratum + θ	479.16	-0.21
1 + Depth + Depth^2 + year+ θ	479.58	-0.21
1 + Depth+ Depth^2 +θ	479.79	-0.21
1+ Depth*Substratum + Depth^2 + year +θ	482.10	-0.22
1 + Depth + Substratum +θ	491.26	-0.22
1 + Depth + year +θ	492.00	-0.22
1 + Depth +θ	492.38	-0.22
1 + Depth*Substratum + year+ θ	495.14	-0.23
1 + Depth*Substratum + θ	495.71	-0.23
1 +Substratum + year +θ	Inf	-0.23
1+ Substratum+θ	Inf	-0.23

In particular, following the Parsimony criterion, we select the final model that includes, the bathymetry, its quadratic transformation and the spatial effect for the binomial response. On the contrary, for the Gamma model the final model involves in addition the type of substratum factor and the temporal effect.

Table 5.2. Model comparison for the Gamma response

Model	DIC	LCPO
1 + Depth*Substratum + Depth^2 + Depth^2:Substratum + year+ θ	906.62	-0.64
1 + Depth*Substratum + Depth^2 + year + θ	907.09	-0.61
1 + Depth + Depth^2 +Substratum + year+ θ	907.87	-0.61
1 + Depth + Substratum + year + θ	909.18	-0.62
1 + Depth + Depth^2 + year+ θ	912.30	-0.61
1 + Depth + year+ θ	915.62	-0.61
1 + Substratum + year+ θ	949.50	-0.66
1 + Depth*Substratum + Depth^2*Substratum+ θ	1026.38	-0.69
1 + Depth + Depth^2 + θ	1027.48	-0.69
1 + Depth*Substratum +θ	1028.36	-0.69
1 + Depth + Substratum + θ	1028.50	-0.69
1 + Depth + θ	1030.35	-0.69
1 + Substratum + θ	1066.44	-0.73

As can be seen in Table 5.3 and Figure 5.3, the bathymetry has a relevant influence on driving hake occurrence. Table 5.3 shows a numerical summary of the posterior distribution of the effects, shown in Figure 5.3.

Table 5.3. Numerical summary of the posterior distributions of the fixed effects for the binomial response variable.

	mean	sd	$Q_{0.025}$	$Q_{0.5}$	$Q_{0.975}$
(Intercept)	2.07	0.78	0.49	2.09	3.57
Depth	1.74e-02	6.79e-03	4.61e-03	1.73e-02	3.13e-02
Depth^2	-6.34e-05	1.37e-05	-9.19e-05	-6.28e-05	-3.79e-05

Specifically, a cuadratic relationship was found between the bathymetry and the occurrence of hake recruits. This relationship shows that the probability

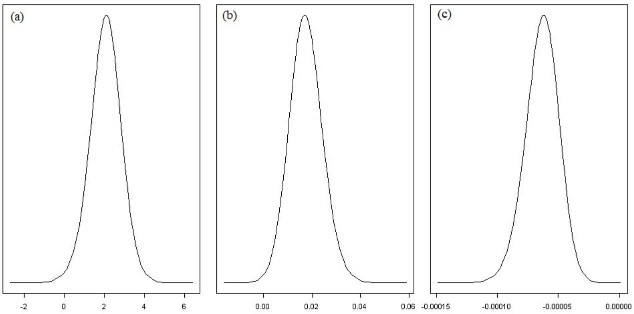
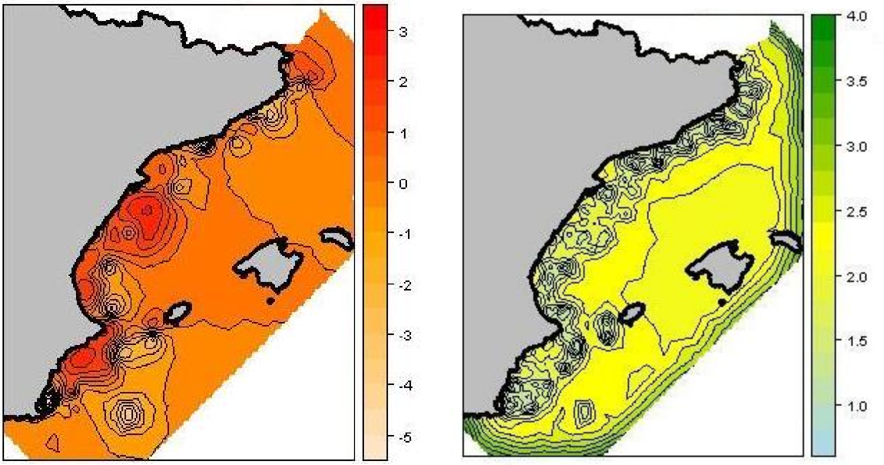


Figure 5.3. Posterior distributions of the fixed effects: (a) Intercept; (b) Depth; (c) Depth^2.

of finding recruits is greater in shallow waters, specially on the continental shelf (from 0 to 300 m.), with to respect to deeper waters (after 300 m.). No relevant difference was found between the type of substratum. Figure 5.4 displays the posterior mean and standard deviation of the spatial component. This component shows a strong effect with positive values along the shelf break and the upper slope of the Spanish Mediterranean coast and with negative values in deeper waters.

Figure 5.5 shows the median posterior probability of occurrence and highlights high-presence aggregation areas, mainly along the shelf break and the upper slope. The probability of finding hake recruits is greater for water between 150-300 m. deep.

With respect the second model, results show again a cuadratic relationship between the bathymetry and the juveniles of hake density (Table 5.4 and Figure 5.6), and again finding hake recruits is highly related to lower waters. Regarding the type of substratum, sand substratum shows the lowest estimated density of hake juveniles (posterior mean = -0.29; 95% CI = [-0.73, 0.13]) with respect to the reference level (rock substratum). Muddy substrata showed a higher estimated coefficient than sandy seabeds.



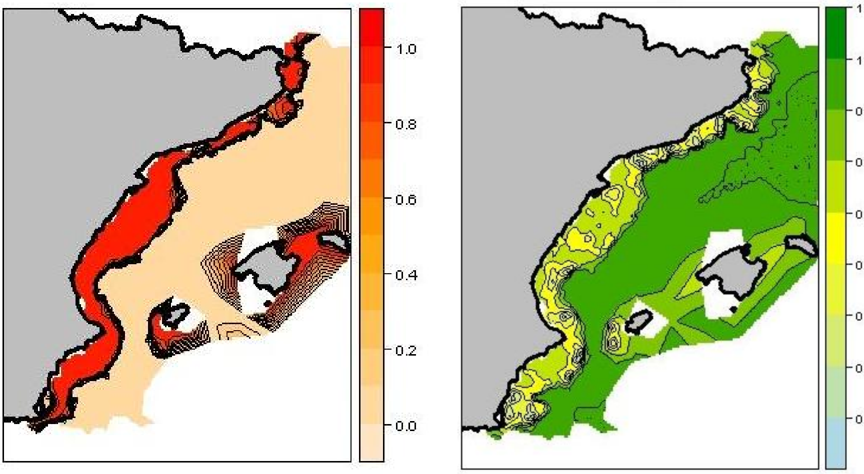
(a) Mean (b) Standard deviation

Figure 5.4. The posterior mean (left) and standard deviation (right) of the spatial effect for the occurrence of the hake juveniles

Table 5.4. Numerical summary of the posterior distributions of the fixed effects for the Gamma response variable.

	mean	sd	$Q_{0.025}$	$Q_{0.5}$	$Q_{0.975}$
(Intercept)	-0.15	0.26	-0.65	-0.15	0.36
Mud	-0.08	0.22	-0.51	-0.07	0.34
Sand	-0.29	0.22	-0.73	-0.29	0.13
Depth	3.10e-03	1.53e-03	5.59e-05	3.11e-03	6.09e-03
Depth^2	-1.61e-05	3.07e-06	-2.20e-05	-1.61e-05	-9.95e-06

Both the map of the predictive spatial density of hake juveniles and that of the posterior mean of the spatial effect (Figures 5.7(a) and 5.8(a)) show marked high-density aggregations (hot spots) of juvenile individuals. The higher-density aggregations are recorded on muddy and rock substrata, in water between 150-



(a) Median

(b) Standard deviation

Figure 5.5. Median (left) and standard deviation (right) for $\pi_i|\boldsymbol{Z}$.

300 m. deep.

5.5 Discussion

We proposed a methodological approach for the identification of fish nurseries on the basis of their spatio-temporal persistence, which can be helpful when quantitative data on the contribution of nursery habitats to the adult population are lacking.

We applied our approach to European hake recruits, using presence/absence and density data collected over 12 years of trawl surveys along the Spanish Mediterranean Sea. Trawl surveys represent one of the most important sources of data for the study of demersal fish, despite the inevitable inherent biases that characterise these datasets.

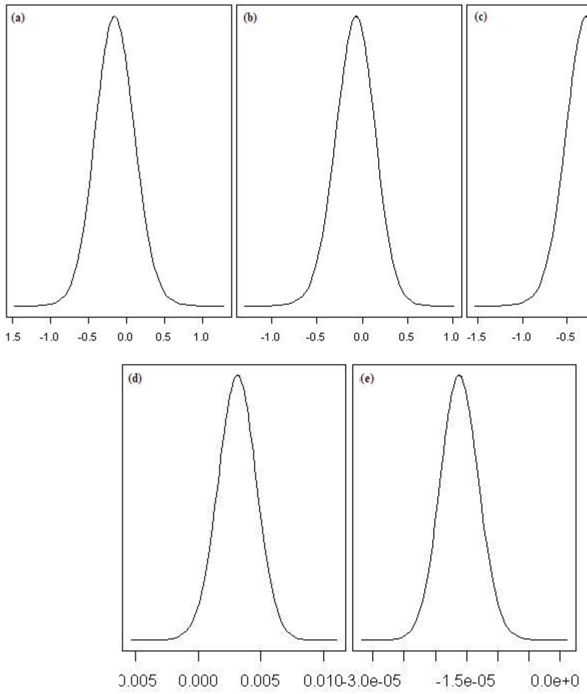
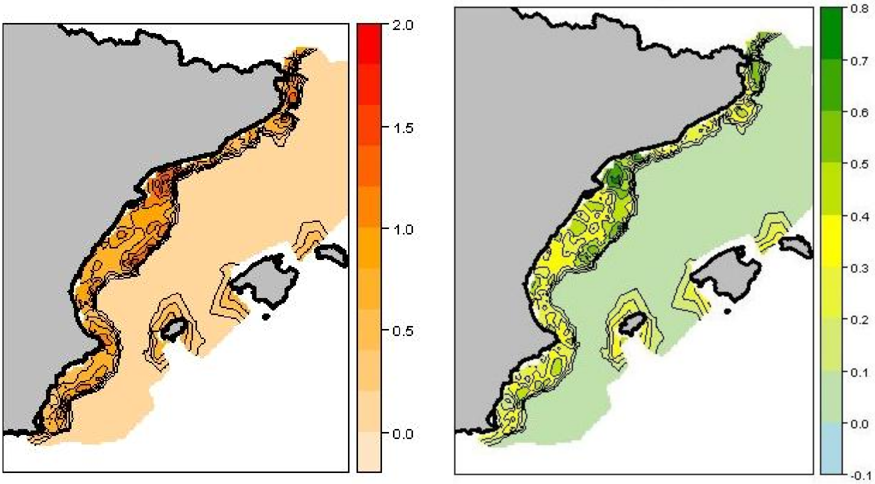


Figure 5.6. Posterior distributions of the fixed effects of the Gamma model: (a) Intercept; (b) Mud Substratum; (c) Sand Substratum; (d) Depth; (e) Depth^2.

However, the proposed Bayesian spatial models can improve analyses of data with biases characterised by treating spatial effects as a variable of interest. These methods make it possible to incorporate the spatial correlation of the variables and the uncertainty of the parameters into the modelling process, resulting in a better quantification of uncertainty.

High-presence and density aggregation areas are recorded mainly along the shelf break and the upper slope, between 150-300 m. deep, and on muddy and rock substrata. This bathymetric distribution agrees with the recruits'



(a) Mean (b) Standard deviation

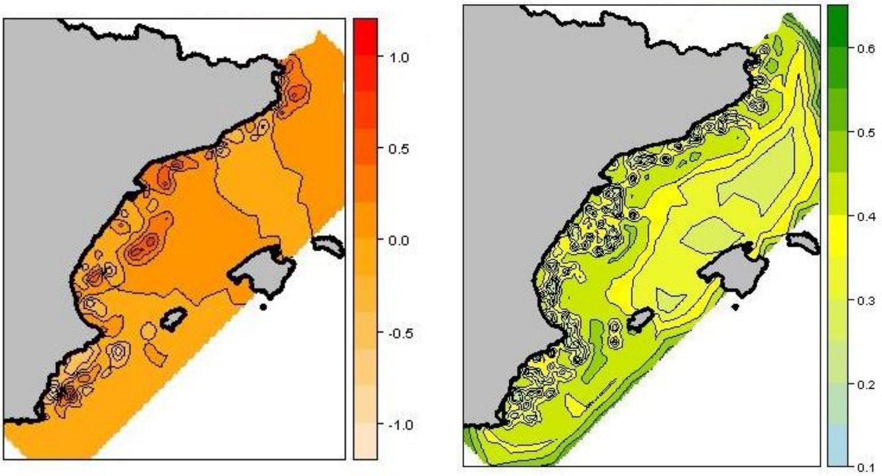
Figure 5.7. The posterior mean (left) and standard deviation (right) of the density of hake juveniles

diet of small planktonic crustaceans (*Euphausiacea*), which inhabit the coastal continental shelf.

Orsi-Relini et al. (2002) found that hake biomass and abundance were affected by depth. They observed the greatest density coinciding with the bathymetric stratum, where most nursery grounds of hake in the Mediterranean Sea are located. Similar results were found by other authors in different areas and with a variety of spatial analytical approaches (Abella et al., 2005).

The type of seabed was only relevant for the density juveniles model. Muddy and rocky substrata are those with the highest amount of density, reflecting the diet selection of this species.

Both the map of the predictive spatial density of hake juveniles and that of the posterior mean of the spatial effect, showed two stable areas characterized



(a) Mean

(b) Standard deviation

Figure 5.8. The posterior mean (left) and standard deviation (right) of the spatial effect of the density of hake juveniles

by high concentrations of juveniles which were identified near C.po Creus and on the continental shelf near Tarragona and Castellón. These areas are high productivity areas due to *downwelling* processes (Revenga, 2012) and groundwater flows.

We considered the stability in space and time of high-density areas to be a useful measure to evaluate their importance as nurseries. Similarly, Colloca et al. (2009) showed that juvenile distribution over time was a good indicator of unit-area productivity of nursery areas of European hake.

Despite many controversial and sceptical arguments regarding the benefits of MPAs for fishery management, it is becoming increasingly acknowledged that the establishment of MPAs is an essential tool for maintaining or rebuilding stocks and protecting their habitat (Roberts et al., 2005).

The importance of the area identified in this study results from a combination of multiple factors: 1) the high persistence per unit area is a proxy of the high recruitment efficiency of these grounds in comparison to that of nearby areas; 2) the high predictability in time and space of nurseries provides more chances of protecting recruits effectively over time; 3) the persistent nurseries are described at a spatial scale relevant to management.

The temporal closures from fishing activity of the high-density areas identified could preserve the reproductive potential of the hake populations, and increase its recruitment.

The analytical approach that we used here to document the persistence of spatial patterns in the distribution of juvenile hake can be extended to different species and life stages to improve knowledge of the role of habitat for populations and communities. This kind of information can be combined with data on the distribution of fishing fleets and with multiple species datasets for an effective EAFM in the Mediterranean.

Chapter 6

Conclusions and further work

The implementation of the Ecosystem Approach to Fisheries Management (EAFM) is the current challenge for a proper management of marine ecosystems and resources. Nevertheless, its practical implementation will take time. A fully-developed EAFM requires scientific knowledge, based on appropriate data collection, that is largely unavailable at present. Its development will be a lengthy, complex and costly process. Consequently, the full implementation of EAFM can only be a long-term objective, to be achieved gradually.

However, nowadays, the development and use of appropriate methodologies to use the data currently available, could be a first step towards the practical implementation of the EAFM, in order to achieve its set of objectives. Indeed, one of the difficulties of the application of the EAFM is that involves a number of different objectives that require several methodologies depending of the kind of data available.

In particular, the EAFM directly aims to:

- Move away from traditional single stock fisheries management to a more broader approach in a manner that addresses the wider interactions between fisheries and the whole ecosystem;

- Develop indicators to measure ecosystem health/stability. These indicators should be integrated in a common tool kit to be applied in different fishery. Once the indicators are developed they can then be applied as a management tool;
- Protect areas/sites, which have particular significance in terms of habitat or species protection, taking due account of fisheries interests. Reinforcement of the protection of special or sensitive habitats;
- Reduce by-catches and discards with, if appropriate, targeted measures to protect vulnerable species;
- Reduce the effects of fishing on juveniles and protect their habitats.
- Use a multiscale approach in order to provides a different approximations of the fishery issues.

In this dissertation, some of those problems have been addressed, depending on the available fisheries source of data, with several methodologies.

In Chapter 1 we have revised the concept and the main aims of the EAFM, introducing the fishing issues that we have discussed, describing the different fishing datasets used, and the types of spatial-temporal scales with which we have addressed these issues.

In Chapter 2 we have proposed specific methodologies that could allow the identification of changes in the dynamics of an ecosystem caused by fishing. The use of ecological indicators associated with statistical techniques such as the bootstrap and smoothing tecniques, have allowed to identify the trends of different ecosystems with different spatial scales. The use of ecological indicators is one of the objectives of the EAFM as they allow the identification of patterns and changes in an ecosystem in a simple manner and also in data-poor situations.

In our particular case, we used them to study the dynamics, from 1950 to today, of two Large Marine Ecosystems such as the Mediterranean and the Black Sea. The proposed methodology is easily reproducible and would be applicable to any ecosystem and any spatial scale, as it uses a data source such as the FAO landings that are available to any country. It is obvious that, despite the macro scales used in this approach, the view provided of the ecosystems is not complete. The environmental, economic and social factors that influence the ecosystems have not been introduced in the study. This is due to the fact that, for a broad spatial and temporal scale as that used, few data are available and therefore usable.

In addition, the ecological indicators should be used only as indeces that identify where there is a problem, which will need to be studied in detail with other methods more specific.

The nonparametric techniques, described in the second approach of Chapter 2, provide a different view of the dynamics of an ecosystem, using the same dataset and ecological indicators, but with a different spatial scale. With respect to a macro-scale approach, the meso-scale one provides more detailed information and hidden patterns than a broader view does not allows. However, in order to apply the second methodology on a meso-scale level, the data available were reduced and the time series analyzed was shortest. This is a common compromise that the use of broader spatial-temporal scales implies.

In this context a multiscale approach results the best option in order to improve our understanding of the fishery dynamics of fisheries in a particular ecosystem. The use of a multiscale approach, and of complementary methodologies, would provide information benefits by improving the assessment findings, and the applicability of fishery management measures, particulary in data limited fisheries.

Other methodologies could be applied using this type of data and to improve our understanding of the dynamics of fishing on marine ecosystems. For example, the use of longitudinal models could provide a more comprehensive view and a different approximation to the dynamics of fishing using FAO landings. Further, bootstrap technique could be improved from a Bayesian point of view and consequently with a greater uncertainty estimation.

In Chapter 3 we have addressed the problem of species distribution models using fishery-dependent data. The approach presented has several advantages. From the point of view of the management of fisheries, the use of fishery-dependent data that otherwise could not be employed and that instead represented a huge source of information that is no used due to the absence of appropiate techniques.

Moreover, using this approach it is possible to estimate the relationship with a particular species and its habitat, to identify its essential habitat and predict what might be its distribution in areas that were not sampled. All these results are supplied with maps that are ideal tools to transmit information from researchers to stakeholders and administrators. From the methodological point of view, the Bayesian approach provides an accurate estimation of the mentioned results, which is essential for effective fisheries management. The data collection has always bias and Bayesian inference addresses these gaps. From the computational point of view, the use of INLA and its module SPDE allows that all these advantages are obtained in a minimum computational time and that the user can manage a familiar programming language as that of R. Until now, the use of other software and the slowness of the computation had slowed the use of Bayesian analysis that instead, as shown, is a useful tool in Ecology.

Moreover, the same methodology can be applied to different cases, in order to achieve the proposed objectives by the EAFM, and to address important issues in fisheries such as discards. The assessment of spatio-temporal distribution of discards could provide the use of additional measures to the management of this problem, such as the temporal of specific hot spots duly identified in which the discards are particularly higher.

Nevertheless, this methodology could be improved with the introduction of splines, if necessary, and a structured temporal component, which could improve the estimation of the temporal variability of data.

In Chapter 4 we review a theme that Diggle et al. (2010) had already addressed and that in Ecology and especially in fishery is of great importance: the use of data collected with a *preferential sampling*. Indeed, until now, the data collected with a preferential sampling were used to assess the distribution of target species of the fishery without taking into account that this could be strongly biased.

In our practical application the lack of covariates implies no particularly interesting results at the biological level for the species studied. Nevertheless, the case study has allowed us to verify that the implementation of traditional application with these data led to biased results.

Moreover, the addition of the error of the fisherman, further enhances this approach.

Finally, the use of INLA to address this modelling, provides the advantage of computational speed, the use of more flexible covariance matrix and an easier programming.

Regarding this type of modelling many future lines remain still open. A real improvement could be obtained by using the module SPDE of INLA. This would allow to address a wider range of more complex models, introducing covariates and the time component, without losing efficiency at a computational level. Moreover the use of a triangulation approach instead of a regular grid one, could improve not only the speed of computation but also the inference, eliminating problems such as the edge effect and improving the accuracy of prediction.

In Chapter 5 we used fishery-independent data to analyze the occurrence and density data on juveniles and thus identify nurseries. The protection of juveniles and specifically of the nursery areas, is an essential objective of the EAFM and methods, like the one presented, could be essential in the future management of fisheries. Indeed, the analysis provided not only an accurate estimation of the relationships between the juveniles and its environment, but also a prediction of its spatial distribution on a meso-scale level. As in Chapter 3, this modeling allows generation of predictive spatial maps that are essential for an easy interpretation of the results. The maps presented have predictive character but also you can generate maps of the spatial effect that, as shown, can highlight hidden trends that would otherwise not be identifiable.

This model could be improved by introducing the model of CPUE as a submodel of the Binomial one. This would provide a detailed spatial distribution of species, quantifying, when it is present, the specific abundance in each location and by relating the two processes with a spatial effect.

In this thesis we presented a multi-scale approach to discuss several spatio-temporal scales and type of data. As mentioned above, in Chapter 1, the choice of the scale implies a compromise in the quality and quantity of data. The use of a global scale and long historical series, implies in most cases, a quality and quantity of data much lower than in a study in micro-scale. Specific approaches in order to understand the dynamics of the fishery on a macro-meso-scales have been proposed. On the contrary, in Chapter 3 and 4, the metodologies presented have been applied on micro-scale studies that have provided very detailed results. It is worth mentioning that all the methodologies presented could be applied in other spatio-temporal scales if the data are available.

Finally, it would be interesting develop a set of methodologies that can be used together to analyze jointly the different sources of fishery data presented in order to take full advantage of existing information and deepen our understanding of marine systems and its resources.

Bibliography

A. Abdulla. Status and conservation of sharks in the Mediterranean Sea. *IUCN Technical Paper*, 144(7), 2004. 98

A. Abella, F. Serena, and M. Ria. Distributional response to variations in abundance over spatial and temporal scales for juveniles of european hake (merluccius merluccius) in the western mediterranean sea. *Fisheries Research*, (71):295–310, 2005. 169

M. Abramowitz and I. A. Stegun. Handbook of mathematical function with formulas, graphs, and mathematical tables. *National Bureau of Standsards, Applied Mathematics Series*, 55, 1970. 85, 146

P. Alvarez, L. Motos, A. Uriarte, and J. Egaña. Spatial and temporal distribution of European hake, Merluccius merluccius (l.), eggs and larvae in relation to hydrographical conditions in the Bay of Biscay. *Fisheries Research*, 50: 111–128, 2001. 149

D. L. Alverson. Discarding practices and unobserved fishing mortality in marine fisheries: an update. 1998. 3

E. A. Babcock, E. K. Pikitch, M. K. McAllister, P. Apostolaki, and C. Santora. A perspective on the use of spatialized indicators for ecosystem-based fishery management through spatial zoning. *ICES Journal of Marine Science: Journal du Conseil*, 62(3):469–476, 2005. 19

S. Banerjee, B. Carlin, and A. Gelfand. *Hierarchical Modeling and Analysis for Spatial Data*. Chapman and Hall/CRC Press, 2004. 22, 79, 80, 81, 82, 83, 157

C. Barberá, J. Moranta, F. Ordines, M. Ramón, A. de Mesa, M. Díaz-Valdés, A. M. Grau, and E. Massutí. Biodiversity and habitat mapping of menorca channel (western mediterranean): implications for conservation. *Biodiversity and Conservation*, 21(3):701–728, 2012. 99

J. M. Bellido, M. B. Santos, M. G. Pennino, J. Valeiras, and G. J. Pierce. Fishery discards and bycatch: solutions for an ecosystem approach to fisheries management? *Hydrobiologia*, 670(1):317–333, 2011. 3, 114, 129

S. A. Berkeley, M. A. Hixon, R. J. Larson, and M. S. Love. Fisheries sustainability via protection of age structure and spatial distribution of fish populations. *Fisheries*, 29(8):23–32, 2004. 156

J. A. Bertrand, L. Gil De Sola, C. Papaconstantinou, G. Relini, and A. Souplet. The general specifications of the medits surveys. *Scientia marina*, 66(S2):9–17, 2002. 158

R. J. H. Beverton, S. J. Holt, et al. On the dynamics of exploited fish populations. *Fishery Investigations Series 2: Sea Fisheries*, 19, 1957. 156

G. Bianchi. Assessment of data-poor fisheries. Proceedings of the Fishery Dependent Information Conference, Galway, Ireland, August 23-26 2010. xvi, xxii

T. A. Branch. The trophic fingerprint of marine fisheries. *Nature*, (468):431–435, 2010. 27

J. A. Brian. *Improving biodiversity monitoring by modeling relative abundance from presence only data*. PhD thesis, Universiteit Hasselt, Diepenbee, Belgium, 2011. 138

L. Brotons, S. Herrando, and M. Pla. Updating bird species distribution at large spatial scales: applications of habitat modelling to data from long term monitoring programs. *Diversity and Distributions*, 13:1–13, 2007. 138

J. R. Busby. BIOCLIM: A bioclimatic analysis and predictive system. In C.R. Margules and M.P. Austin, editors, *Nature conservation: cost effective biological surveys and data analysis*, pages 64–68. CSIRO Australia, 1991. 78

J. F. Caddy. Toward a comparative evaluation of human impacts on fishery ecosystems of enclosed and semi-enclosed seas. *Reviews in Fisheries Science*, 1(1):57–95, 1993. 31, 34, 49, 50

J. F. Caddy. Marine catchment basin effects versus impacts of fisheries on semi-enclosed seas. *ICES Journal of Marine Science: Journal du Conseil*, 57(3): 628–640, 2000. 31

J. F. Caddy. Practical issues in choosing a framework for resource assessment and management of mediterranean and black sea fisheries. *Mediterranean Marine Science*, 10(1):83–120, 2009. 156

J. F. Caddy and L. Garibaldi. Apparent changes in the trophic composition of world marine harvests: the perspective from fao capture database. *Ocean Coastal Management Fisheries Research*, (72):241–252, 2000. 42, 44, 49

J. F. Caddy, J. Csirke, S. M. García, and R. J. R. Grainger. How pervasive is" fishing down marine food webs"? *Science*, 282(5393):1383–1383, 1998. 26, 30

J. Cao, X. Chen, Y. Chen, B. Liu, J. Ma, and S. Li. Generalized linear bayesian models for standardizing cpue: an application to a squid-jigging fishery in the northwest pacific ocean. *Scientia Marina*, 75(4):679–689, 2011. 22, 135

R. Cao, A. Cuevas, and W. Gonzalez Manteiga. A comparative study of several smoothing methods in density estimation. *Computational Statistics & Data Analysis*, 17(2):153–176, 1994. 59

A. Carbonell and M. Azevedo. Application of non-equilibrium production models to the red shrimp (*Aristeus antennatus*, Risso, 1816) fishery in the northwestern Mediterranean. *Fisheries Research*, 65(1):323–334, 2003. 98

A. Carbonell and S. Mallol. Differences between demersal fisheries discards: high and low productivity zones of the northwestern mediterranean sea. *Bolletin Sociedad Historia Natural Balears*, (55):141–155, 2013. 115

A. Carbonell, P. Martín, and S. De Ranieri. Discards of the western mediterranean trawl fleets. *Rapport de Commitee Internationale de Mer Mediteranee*, 35:392–393, 1998. 130

A. Chakraborty, A. E. Gelfand, A. M. Wilson, A. M Latimer, and J. A. Silander. Modeling large scale species abundance with latent spatial processes. *The Annals of Applied Statistics*, 4(3):1403–1429, 2010. 78, 79

L. C. Chien and S. I. Bangdiwala. The implementation of bayesian structural additive regression models in multi-city time series air pollution and human health studies. *Stochastic Environmental Research and Risk Assessment*, 2012. doi: 10.1007/s00477-012-0562-4. In press. 83

V. Christensen. Indicators for marine ecosystems affected by fisheries. *Marine and Freshwater Research*, 51(5):447–450, 2000. 31

R. B. Cleveland. Robust locally-weighted regression and smoothing scatterplots. *Journal of American Statistical Association*, (74):829–836, 1979. 32

F. Colloca, V. Bartolino, G. J. Lasinio, L. Maiorano, P. Sartor, G. Ardizzone, et al. Identifying fish nurseries using density and persistence measures. *Marine Ecology, Progress Series*, 381:287–296, 2009. 156, 170

J. Cotter, P. Petitgas, A. Abella, P. Apostolaki, B. Mesnil, C. Politou, J. Rivoirard, M. J. Rochet, M. T. Spedicato, V. M. Trenkel, et al. Towards an ecosystem approach to fisheries management (eafm) when trawl surveys provide the main source of information. *Aquatic Living Resources*, 22(2):243, 2009. xiv, xx

N. Cressie, C. A. Calder, J. S. Clark, J. M. Van Hoef, and C. K. Wikle. Accounting for uncertainty in ecological analysis: The strengths and limitations of hierarchical statistical modeling. *Ecological Applications*, 19:553–5701, 2009. 78

E. Cuevas-Zimbrón, J. C. Pérez-Jiménez, and I. Méndez-Loeza. Spatial and seasonal variation in a target fishery for spotted eagle ray aetobatus narinari in the southern gulf of mexico. *Fisheries Science*, 77(5):723–730, 2011. 132

P. M. Cury, L. J. Shannon, J. P. Roux, G. M. Daskalov, A. Jarre, C. L. Moloney, and D. Pauly. Trophodynamic indicators for an ecosystem approach to fisheries. *ICES Journal of Marine Science: Journal du Conseil*, 62(3):430–442, 2005. 31

C. P. Dahlgren, G. T. Kellison, A. Adams, B. M. Gillanders, M. S. Kendall, C. A. Layman, J. A. Ley, I. Nagelkerken, and J. E. Serafy. Marine nurseries and effective juvenile habitats: concepts and applications. *Marine Ecology Progress Series*, 312:291–295, 2006. 156

V. H. Dale and S. C. Beyeler. Challenges in the development and use of ecological indicators. *Ecological indicators*, 1(1):3–10, 2001. 29

T. Dalton, R. Thompson, and D. Jin. Mapping human dimensions in marine spatial planning and management: An example from narragansett bay, rhode island. *Marine Policy*, 34(2):309–319, 2010. 18

J. I. de Leiva Moreno, V. N., Agostini, J. F. Caddy, and F. Carocci. Is the pelagic-demersal ratio from fishery landings a useful proxy for nutrient availability? a preliminary data exploration for the semi-enclosed seas around europe. *ICES Journal of Marine Science: Journal du Conseil*, 57(4):1091–1102, 2000. 31, 101

A. Dell'Apa, D. G. Kimmel, and S. Clò. Trends of fish and elasmobranch landings in Italy: associated management implications. *ICES Journal of Marine Science: Journal du Conseil*, 69(6):1045–1052, 2012. 98

M. Demestre, P. Sanchez, and P. Abello. Demersal fish assemblages and habitat characteristics on the continental shelf and upper slope of the north-western Mediterranean. *Journal of the Marine Biological Association of the UK*, 80 (6):981–988, 2000. 101

J. Depestele, S. Vandemaele, W. Vanhee, H. Polet, E. Torreele, H. Leirs, and M. Vincx. Quantifying causes of discard variability: an indispensable assistance to discard estimation and a paramount need for policy measures. *ICES Journal of Marine Science: Journal du Conseil*, 68(8):1719–1725, 2011. 115

P. J. Diggle. *Statistical analysis of spatial point patterns*. Arnold, London UK, 2nd edition, 2003. 140

P. J. Diggle and P. J. Ribeiro. *Model-based Geostatistics*. Springer Series in Statistics. Springer, 2007. 136, 157

P. J. Diggle, J. A. Tawn, and R. A. Moyeed. Model-based geostatistics. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 47(3):299–350, 1998. 82

P. J. Diggle, R. Menezes, and T. Su. Geostatistical inference under preferential sampling. *Journal of the Royal Statistical Society, Series C*, 52:191–232, 2010. 139, 177

T. A. Dinmore, D. E. Duplisea, B. D. Rackham, D. L. Maxwell, and S. Jennings. Impact of a large-scale area closure on patterns of fishing disturbance and the consequences for benthic communities. *ICES Journal of Marine Science: Journal du Conseil*, 60(2):371–380, 2003. 13

F. Douvere. The importance of marine spatial planning in advancing ecosystem-based sea use management. *Marine Policy*, 32(5):762–771, 2008. 16

S. N. Durlauf and D. T. Quah. The new empirics of economic growth. *Discussion Paper*, 384, 1998. 61

B. Efron. Bootstrap methods: Another look at the jacknife. *Annals of Applied Statistics*, (7):1–26, 1979. 27, 33

J. Elith, C. H. Graham, R. P. Anderson, M. Dudık, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, et al. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29(129–151), 2006. 20, 106

J. R. Ellis, A. Cruz-Martinez, B. D. Rackham, and S. I. Rogers. The distribution of chondrichthyan fishes around the British Isles and implications for conservation. *Journal of Northwest Atlantic Fishery Science*, 35:195–213, 2005. 113

T. E. Essington. Ecological indicators display reduced variation in north american catch share fisheries. *Proceedings of the National Academy of Sciences*, 107(2):754–759, 2010. 116

EU. Communication from the commission to the european parliament and the council on a european community action plan for the conservation and management of sharks. Technical report, Com(2009)40 Final, 2009. 98

L. Fahrmeir and G. Tutz. *Multivariate Statistical Modelling Based on Generalized Linear Models*. Springer Series in Statistics. Springer, 2nd edition, 2001. 83

FAO. The state of world fisheries and aquaculture. Fisheries Department, Rome, 2002. 1, 12

J. Feekings, V. Bartolino, N. Madsen, and T. Catchpole. Fishery discards: factors affecting their variability within a demersal trawl fishery. *PLoS One*, 7(4):e36409, 2012. 114, 115

C. Fernández. Stock assessment incorporating discards estimates in some years and implications for prediction of future stock trajectories. *ICES Journal of Marine Science: Journal du Conseil CM 2009/M:10*, 2009. 28

F. Ferretti, R. A. Myers, F. Serena, and H. K. Lotze. Loss of large predatory sharks from the Mediterranean Sea. *Conservation Biology*, 22(4):952–964, 2008. 98

A.H. Fielding and J. F. Bell. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental conservation*, 24(1):38–49, 1997. 105

D. Fluharty. Evolving ecosystem approaches to management of fisheries in the usa. *Marine Ecology Progress Series*, 300:241–296, 2005. 70

R. Froese and D. Pauly. Fishbase. *World Wide Web electronic publication*, 2011. http://www.fishbase.org, version (10/2011). 35, 89

M. García-Rodríguez, A. M. Fernández, and A. Esteban. Characterisation, analysis and catch rates of the small scale fisheries of the Alicante Gulf (SE Spain) over a 10 years time series. *Fisheries Research*, 77:226–238, 2005. 149

S. M. García, D. J. Staples, and J. Chesson. The fao guidelines for the development and use of indicators for sustainable development of marine capture fisheries and an australian example of their application. *Ocean & Coastal Management*, 43(7):537–556, 2000. xiii, xix

G. Garofalo, T. Fortibuoni, M. Gristina, M. Sinopoli, and F. Fiorentino. Persistence and co-occurrence of demersal nurseries in the strait of sicily (central mediterranean): Implications for fishery management. *Journal of Sea Research*, 66(1):29–38, 2011. 156

D. Gascuel, G. Merino, R. Döring, J. N. Druon, L. Goti, S. Guénette, C. Macher, K. Soma, M. Travers-Trolet, and S. Mackinson. Towards the implementation of an integrated ecosystem fleet-based management of european fisheries. *Marine Policy*, 36(5):1022–1032, 2012. xiv, xx

M. Gaudard, M. Karson, E. Linder, and D. Sinha. Bayesian spatial prediction. *Environmental and Ecological Statistics*, 35(6):147–171, 1999. 79

S. Geisser. Predictive inference. monograph on statistics and applied probability, 1993. 87

A. E. Gelfand, N. Ravishanker, and M. D. Ecker. Modeling and inference for point-referenced binary spatial data. In D. Dey, S. Ghosh, and B. Mallick, editors, *Generalized Linear Models: A Bayesian Perspective*, pages 381–394. Marcel Dekker Inc., 2000. 81

A. E. Gelfand, J. A. Silander, S. Jr. Wu, A. M. Latimer, P. Lewis A. G. Rebelo, and M. Holder. Explaining species distribution patterns through hierarchical modeling. *Bayesian Analysis*, 1(1):41–92, 2006. 23, 79

A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian data analysis*. Chapman & Hall/CRC, 2004. 21, 22

T. Gneiting and A. E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007. 87

R. Goñi, F. Alvarez, and S. Adlerstein. Application of generalized linear modeling to catch rate analysis of Western Mediterranean fisheries: the Castellón trawl fleet as a case study. *Fisheries Research*, 42:291–302, 1999. 149

P. González-Riancho, M. Sanò, O. García-Aguilar, and J. Areizaga. A contribution to the implementation of iczm in the mediterranean developing countries. *Ocean & Coastal Management*, 52(11):545–558, 2009. 54

A. Guisan and W. Thuiller. Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, 8:993–1009, 2005. 78

A. Guisan and N. E. Zimmermann. Predictive habitat distribution models in ecology. *Ecological modelling*, 135(2):147–186, 2000. 20

A. Guisan, T. C. Edwards, and T. Hastie. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological modelling*, 157:89–100, 2002. 78

R. Haining, J. Law, R. Maheswaran, T. Pearson, and P. Brindley. Bayesian modelling of environmental risk: a small area ecological study of coronary heart disease mortality in relation to modelled outdoor nitrogen oxide levels. *Stochastic Environmental Research and Risk Assessment*, 21(5):501–509, 2007. 79

M. S. Handcock and M. L. Stein. A Bayesian analysis of kriging. *Technometrics*, 35(4):403–410, 1993. 79, 85

L. Held, B. Schrödle, and H. Rue. *Posterior and Cross-validatory PredictiveChecks: A Comparison of MCMC and INLA*. Springer, 2010. 80, 87

R. Hijman and C. Graham. The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology*, 12(12):2272–2281, 2006. 78

R. Hilborn, C. J. Walters, et al. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. *Reviews in Fish Biology and Fisheries*, 2(2):177–178, 1992. xv, xxi, 15, 21

S. Hinckley, A. J. Hermann, K. L. Mier, and B. A. Megrey. Importance of spawning location and timing to successful transport to nursery areas: a simulation study of gulf of alaska walleye pollock. *ICES Journal of Marine Science: Journal du Conseil*, 58(5):1042–1052, 2001. 156

Ø. Hjelle and M. Dæhlen. *Triangulations and Applications*. Mathematics and Visualization. Springer, 2006. 88

L. P. Ho and D. Stoyan. Modelling marked point patterns by intensity-marked Cox processes. *Statistical Probability Letters*, 78:1194–1199, 2008. 140

M. J. Holden and R. N. Tucker. The food of Raja clavata Linnaeus 1758, Raja montagui Fowler 1910, Raja naevus Müller and Henle 1841 and Raja brachyura Lafont 1873 in British waters. *Journal du Conseil*, 35(2):189–193, 1974. 113

T.E. Hopkins and J.J. Cech. The influence of environmental variables on the distribution and abundance of three elasmobranchs in Tomales Bay, California. *Environmental Biology of Fishes*, 66(3):279–291, 2003. 101, 102

E. Hunter, A. A. Buckley, C. Stewart, and J. D. Metcalfe. Migratory behaviour of the thornback ray, Raja clavata, in the southern North Sea. *Journal of the Marine Biological Association of the United Kingdom*, 85(05):1095–1105, 2005. 113

S. H. Hurlbert. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs*, 54:187–211, 1984. 78

J. B. Illian, A. Penttinen, H. Stoyan, and H. Stoyan. *Statistical Analysis and Modelling of Spatial Point Patterns.* Wiley, Chichester, UK, 2008. 140

J. B. Illian, S. H. Sørbye, and H. Rue. A toolbox for fitting complex spatial point process models using integrated nested laplace approximation (INLA). *The Annals of Applied Statistics*, 6(4):1499–1530, 2012. 140

S. R. Jenkins, C. Mullen, and A. R. Brand. Predator and scavenger aggregation to discarded by-catch from dredge fisheries: importance of damage level. *Journal of Sea Research*, 51(1):69–76, 2004. 114

S. Jennings and N. V. C. Polunin. Effects of fishing effort and catch rate upon the structure and biomass of fijian reef fish communities. *Journal of Applied Ecology*, 33(2):400–412, 1996. xiv, xx, 5

M. C. Jones, J. S. Marron, and S. J. Sheather. A brief survey of bandwidth selection for density estimation. *Journal of the American Statistical Association*, 91(433):401–407, 1996. 59, 60

M. Kacher. Distribution and growth of 0-group European hake in the Bay of Biscay and Celtic Sea: a spatial and inter-annual analyses. *Fisheries Research*, 71:373–378, 2005. 149

A. Kallianiotis, K. Sophronidis, P. Vidoris, and A. Tselepides. Demersal fish and megafaunal assemblages on the Cretan continental shelf and slope (NE Mediterranean): seasonal variation in species density, biomass and diversity. *Progress in oceanography*, 46(2):429–455, 2000. 101

I. Katara, J. Illian, J.P. Graham, B. Scott, and J. Wang. Atmospheric forcing on chlorophyll concentration in the Mediterranean. *Hydrobiologia*, 612:33–48, 2008. 93, 102

S. Katsanevakis, C. D. Maravelias, D. Damalas, A. P. Karageorgis, E. V. Tsitsika, C. Anagnostou, and C. Papaconstantinou. Spatiotemporal distribution and habitat use of commercial demersal species in the eastern Mediterranean Sea. *Fisheries Oceanography*, 18(6):439–457, 2009. 99

K. Kelleher. *Discards in the world's marine fisheries: an update*, volume 470. FAO, 2005. 36, 115, 130, 131

M. Kery, A. Royle, H. Schmid, B. Volet, G. Hafliger, and N. Zbinden. Site-occupancy distribution to correct population trend derived from opportunistic observations. *Conservational Biology*, pages 1388–1397, 2009. 138

T. Kneib, J. Muller, and T. Hothorn. Spatial smoothing techniques for the assessment of habitat suitability. *Enviromental and Ecological Statistics*, 15 (3):343–364, 2008. 78

A. M. Latimer, Shanshan W., A. E. Gelfand, and J. A. Silander. Building statistical modelos to analyze species distributions. *Ecological Applications*, 16(1):33–50, 2006. 20, 135, 138

R. Law, J. B. Illian, D. F. R. P. Burslem, G. Gratzer, C. V. S. Gunatilleke, and I. A. U. N. Gunatilleke. Ecological information from spatial patterns of plants: insights from point process theory. *Journal of Ecology*, 97:616–628, 2009. 140

J. R. Leathwick, D. Rowe, J. Richardson, J. Elith, and T. Hastie. Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish. *Freshwatwer Biology*, 50:2034–2052, 2005. 78

R. S. Leuven and I. Poudevigne. Riverine landscape dynamics and ecological risk assessment. *Freshwater Biology*, 47(4):845–865, 2002. 13

S. Libralato, F. Pranovi, S. Raicevich, F. Da Ponte, O. Giovanardi, R. Pastres, P. Torricelli, and D. Mainardi. Ecological stages of the venice lagoon analysed using landing time series data. *Journal of Marine Systems*, 51(1):331–344, 2004. 32

F. Lindgren, H. Rue, and J. Lindström. An explicit link between Gaussian fields and Gaussian Markov random fields: the SPDE approach (with discussion). *Journal of the Royal Statistical Society, Series B*, 73:423–498, 2011. 84, 86, 121, 136, 153, 162

J. Lleonart. Review of the state of the world fishery resources. *FAO Fish Technical Paper*, 457:49–64, 2005. 48, 149, 157

S. R. Loarie, B. E. Carter, K. Hayhoe, S. McMahon, R. Moe, C. A. Knight, and D. D. Ackerly. Climate change and the future of California's endemic flora. *PLoS ONE*, 3(6):e2502, 2008. 78

C. Lordan, M. Cuaig, N. Graham, and D. Rihan. The ups and downs of working with industry to collect fishery-dependent data: the irish experience. *ICES Journal of Marine Science: Journal du Conseil*, 68(8):1670–1678, 2011. 7

K. E. Lunn and P. Dearden. Monitoring small-scale marine fisheries: An example from thailand's ko chang archipelago. *Fisheries Research*, 77(1):60–71, 2006. xv, xxi

A. Machias, V. Vassilopoulou, D. Vatsos, P. Bekas, A. Kallianiotis, C. Papaconstantinou, and N. Tsimenides. Bottom trawl discards in the northeastern mediterranean sea. *Fisheries Research*, 53(2):181–195, 2001. 130, 131

C. D. Maravelias, E. V. Tsitsika, and C. Papaconstantinou. Environmental influences on the spatial distribution of European hake (Merluccius merluccius) and red mullet (Mullus barbatus) in the Mediterranean. *Ecological Research*, 22(4):678–685, 2007. 101

C.D. Maravelias, G. Tserpes, M. Pantazi, and P. Peristeraki. Habitat Selection and Temporal Abundance Fluctuations of Demersal Cartilaginous Species in the Aegean Sea (Eastern Mediterranean). *PloS one*, 7(4):e35474, 2012. 98

C. S. Martin, S. Vaz, J. R. Ellis, V. Lauria, F. Coppin, and A. Carpentier. Modelled distributions of ten demersal elasmobranchs of the eastern English Channel in relation to the environment. *Journal Of Experimental Marine Biology And Ecology*, 418–419:91–103, 2012. 101, 102, 134, 157

E. Massutí and J. Moranta. Demersal assemblages and depth distribution of elasmobranchs from the continental shelf and slope off the Balearic Islands (western Mediterranean). *ICES Journal of Marine Science: Journal du Conseil*, 60:753–766, 2003. 99

S.A. Matern, J. J. Cech, and T. E. Hopkins. Diel movements of bat rays, Myliobatis californica, in Tomales Bay, California: evidence for behavioral thermoregulation? *Environmental biology of fishes*, 58(2):173–182, 2000. 101, 102

G. F. Midgley and W. Thuiller. Potential vulnerability of Namaqualand plant diversity to anthropogenic climate change. *Journal of Arid Environments*, 70:615–628, 2007. 78

G. G. Moisen and T. S. Frescino. Comparing five modelling techniques for predicting forest characteristics. *Ecological modelling*, 157(2):209–225, 2002. 23

J. Møller and R.P. Waagepetersen. *Statistical inference and simulation for spatial point processes*, volume 100. Chapman and Hall/CRC, 2004. 140

J. Møller, A. R. Syversveen, and R. P. Waagepetersen. Log Gaussian Cox processes. *Scandinavian Journal of Statistics*, 25:451–482, 1998. 140

A. C. Morgan and G. H. Burgess. 11. fishery-dependent sampling: total catch, effort and catch composition. *Management techniques for elasmobranch fisheries*, page 182, 2005. 8

S. A. Murawski. Factors influencing by-catch and discard rates: analyses from multispecies/multifishery sea sampling. *Journal of Northwest Atlantic Fishery Science*, 19:31–40, 1996. 133

NRC National Research Council. *Improving the Collection, Management, and Use of Marine Fisheries Data*. National Academies Press, 2000. xv, xxi

S. Nodder, C. A. Pilditch, P. K. Probert, and J. A. Hall. Variability in benthic biomass and activity beneath the subtropical front, chatham rise, sw pacific ocean. *Deep Sea Research Part I: Oceanographic Research Papers*, 50(8): 959–985, 2003. 101

S. J. Ormerod. A golden age of river restoration science? *Aquatic Conservation: Marine and Freshwater Ecosystems*, 14(6):543–549, 2004. 14

L. Orsi-Relini, C. Papacostantinou, S. Jukic-Peladic, A. Souplet, L. G. De Sola, C. Piccinetti, and S. Kavadas. Distribution of the Mediterranean hake populations (Merluccius merluccius smiridus Rafinesque, 1810) (Osteichthyes: Gadiformes) based on six years monitoring by trawl-surveys: some implications for management. *Scientia Marina*, 26:21–38, 2002. 169

B. Park and B. Turlach. Practical performance of several data driven bandwidih selectors. *Computational Statistics*, 7:251–285, 1992. 59

D. Pauly and V. Christensen. Primary production required to sustain global fisheries. *Nature*, 374(6519):255–257, 1995. 30

D. Pauly and R. Watson. Background and interpretation of the ‘marine trophic index’as a measure of biodiversity. *Philosophical Transactions of the Royal*

Society B: Biological Sciences, 360(1454):415–423, 2005. 26, 30, 36, 43, 49, 53

D. Pauly, V. Christensen, J. Dalsgaard, R. Froese, and Jr F. Torres. Fishing down marine food webs. *Science*, 279(5352):860–863, 1998. 26, 27, 29, 39, 44

D. Pauly, V. Christensen, and C. Walters. Ecopath, ecosim, and ecospace as tools for evaluating ecosystem impact of fisheries. *ICES Journal of Marine Science: Journal du Conseil*, 57(3):697–706, 2000. 30, 31, 71

E. K. Pikitch, C. Santora, E. A. Babcock, A. Bakun, R. Bonfil, D. O. Conover, P. Dayton, P. Doukakis, D. Fluharty, B. Heneman, E. D. Houde, J. Link, P. A. Livingston, M. Mangel, M. K. McAllister, J. Pope, and K. J. Sainsbury. Ecosystem-based fishery management. *Science*, 305(Weekly):346–347, 2004. 2

J. K. Pinnegar, N. V. C. Polunin, and F. Badalamenti. Long-term changes in the trophic level of western mediterranean fishery and aquaculture landings. *Journal of Fisheries Aquatic Science*, (60):222–235, 2003. 40, 44

F. Poisson, J. Gaertner, M. Taquet, J. Durbec, and K. Bigelow. Effects of lunar cycle and fishing operations on longline-caught pelagic fish: fishing performance, capture time, and survival of fish. *Fishery Bulletin*, 108(3): 268–281, 2010. 132

R. L. Pressey, M. Cabeza, M. E., Watts, R. M. Cowling, and K. A. Wilson. Conservation planning in a changing world. *Trends in Ecology and Evolution*, 22:583–592, 2007. 78

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. URL `http://www.R-project.org/`. 57, 86

R. Ragonese, F. Fiorentino, G. Garofalo, M. Gristina, D. Levi, S. Gancitano, G. B. Giusto, P. Rizzo, and G. Sinacori. Distribution, abundance and biological features of picarel (Spicara flexuosa), Mediterranean (Trachurus mediterraneus) and Atlantic (T. trachurus) horse mackerel based on experimental bottom-trawl data (MEDITS, 1994–2002) in the Strait of Sicily. *MedSudMed Technical Documents*, 5:100–114, 2003. 92

S. Revenga. The fisheries importance of four submarine canyons in the spanish Mediterranean sea: Cap de Creus, Palamós, Cape Tiñoso and Alborán platform canyons. *International Union for Conservation of Nature and Natural Resources.*, pages 51–57, 2012. 170

C. M. Roberts, J. P. Hawkins, and F. R. Gell. The role of marine reserves in achieving sustainable fisheries. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360:123–132, 2005. 170

A.R. Robinson. Mediterranean sea circulation. encyclopedia of ocean sciences. *Academic Press*, pages 1689–1706, 2000. 34

M. J. Rochet and V. M. Trenkel. Factors for the variability of discards: assumptions and field evidence. *Canadian Journal of Fisheries and Aquatic Sciences*, 62(1):224–235, 2005. 129

M. J. Rochet, I. Peronnet, and V. M. Trenkel. An analysis of discards from the french trawler fleet in the celtic sea. *ICES Journal of Marine Science: Journal du Conseil*, 59(3):538–552, 2002. 115

J. P. Rodríguez, L. Brotons, J. Bustamante, and J. Seoane. The application of predictive modelling of species distribution to biodiversity conservation. *Diversity and Distribution*, 13(3):243–251, 2007. 138

M. Roos and L. Held. Sensitivity analysis in Bayesian generalized linear mixed models for binary data. *Bayesian Analysis*, 6(2):259–278, 2011. 82, 87

J. A. Royle, M. Kery, R. Gautier, and H. Schmidt. Hierarchical spatial models of abundance and occurrence from imperfect survey data. *Ecological Monographs*, 77:465–481, 2007. 78

H. Rue and L. Held. *Gaussian Markov Random Fields. Theory and Applications.* Chapman and Hall/CRC, 2005. 84, 120

H. Rue, S. Martino, and N. Chopin. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society, Series B*, 71(2):319–392, 2009. 79, 80, 83, 84, 142, 143, 145

H. Rue, S. Martino, D. Mondal, and N. Chopin. Discussion on the paper Geostatistical inference under preferential sampling by Diggle, Menezes and Su. *Journal of the Royal Statistical Society, Series C*, 52:221–223, 2010. 139, 140, 144

K. J. Sainsbury, A. E. Punt, and A. D. M. Smith. Design of operational management strategies for achieving fishery ecosystem objectives. *ICES Journal of Marine Science: Journal du Conseil*, 57(3):731–741, 2000. 3

B. Sainte-Marie, G. J. Piet, and F. J. Quirijns. The importance of scale for fishing impact estimations. *Canadian Journal of Fisheries and Aquatic Sciences*, 66(5):829–835, 2009. 13

F. Sangermano and R. Eastman. *Linking GIS and Ecology: The Use of Mahalanobis Typicalities to Model Species Distribution.* 2007. 138

A. Santojanni, N. Cingolani, E. Arneri, G. Kirkwood, A. Belardinelli, G. Giannetti, S. Colella, F. Donato, and C. Barry. Stock assessment of sardine (sardina pilchardus, walb.) in the adriatic sea with an estimate of discards. *Scientia Marina*, 69(4):603–617, 2005. 131

C. H. Schmid and J. L. Griffith. *Multivariate Classification Rules: Calibration and Discrimination.* John Wiley and Sons, 2nd edition, 2005. 87

B. Schrödle and L. Held. Spatio-temporal disease mapping using INLA. *Environmetrics*, 22:725–734, 2011. 84

S. J. Sheather and M. C. Jones. A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 53(3):683–690, 1991. 59

Y. J. Shin and L. J. Shannon. Using indicators for evaluating, comparing, and communicating the ecological status of exploited marine ecosystems. *ICES Journal of Marine Science: Journal du Conseil*, (67):686–691, 2010. 70

B. W. Silverman. *Density Estimation for Statistics and Data Analysis*, volume 1. Chapman and Hall, London, 1986. 58

D. Simpson, J. Illian, F. Lindgren, S. H. Sørbye, and H. Rue. Going off grid: Computationally efficient inference for log-Gaussian Cox processes. *Journal of the Royal Statistical Society, Series B : Statistical Methodolgy*, pages 1–19, 2011. 142

J. E. Skjæraasen and O. A. Bergstad. Distribution and feeding ecology of Raja radiata in the northeastern North Sea and Skagerrak (Norwegian Deep). *ICES Journal of Marine Science: Journal du Conseil*, 57(4):1249–1260, 2000. 113

D. J. Spiegelhalter, A. Thomas, and N. G. Best. Winbugs version 1.2 user manual. *MRC Biostatistics Unit*, 1999. 83

D. J. Spiegelhalter, N.G. Best, B.P. Carlin, and A. van der Linde. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B*, 64:583–616, 2002. 87, 121, 148, 162

M. L. Stein. *Interpolation of Spatial Data: Some Theory for Kriging*. Springer, 1st edition, 1999. 85, 144

V. Stelzenmüller, J. Lee, E. Garnacho, and S. I. Rogers. Assessment of a bayesian belief network–gis framework as a practical tool to support marine planning. *Marine pollution bulletin*, 60(10):1743–1754, 2010. 18

K. Stergiou and S. K. Karpouzi. Feeding habits and trophic levels of mediterranean fish. *Fish. Biol. Fish*, (11):217–254, 2002. 26, 35, 43

K.I. Stergiou. The hellenic seas: Physics, chemistry, biology and fisheries. *Oceanography Marine Biology Annul Review*, (35):415–538, 1997. 34, 42

J. D. Stevens, R. Bonfil, N. K. Dulvy, and P. A. Walker. The effects of fishing on sharks, rays, and chimaeras (chondrichthyans), and the implications for marine ecosystems. *ICES Journal of Marine Science: Journal du Conseil*, 57 (3):476–494, 2000. 98

D. Stoyan, W. S. Kendall, J. Mecke, and D. G. Kendall. *Stochastic geometry and its applications*. John Wiley and Sons, London, 2nd edition, 1995. 140

W. E. Thogmartin, J. R. Sauer, and M. G. Knutson. Predicting regional abundance of rare grassland birds with a hierarchical spatial count model. *Condor*, 108:25–46, 2006. 138

C. M. Thompson and K. McGarigal. The influence of research scale on bald eagle habitat selection along the lower hudson river, new york (usa). *Landscape Ecology*, 17(6):569–586, 2002. 13

S. F. Thrush and P. K. Dayton. What can ecology contribute to ecosystem-based management? *Annual Review of Marine Science*, 2:419–441, 2010. 134

E. Tortosa-Ausina. Exploring efficiency differences over time in the spanish banking industry. *European Journal of Operational Research*, 139(3):643–664, 2002. 58

E. Tortosa-Ausina. Bank cost efficiency as distribution dynamics: controlling for specialization is important. *Investigaciones económicas*, 27(1):71–96, 2003. 63

V. M. Trenkel, M. J. Rochet, and B. Mesnil. From model-based prescriptive advice to indicator-based interactive advice. *ICES Journal of Marine Science: Journal du Conseil*, 64(4):768–774, 2007. xvi, xxiii, 4, 15

K. Tsagarakis, A. Machias, M. Giannoulaki, S. Somarakis, and I. Karakassis. Seasonal and temporal trends in metrics of fish community for otter-trawl discards in a mediterranean ecosystem. *ICES Journal of Marine Science: Journal du Conseil*, 65(4):539–550, 2008. 130, 131

K. Tsagarakis, V. Vassilopoulou, A. Kallianiotis, and A. Machias. Discards of the purse seine fishery targeting small pelagic fish in the eastern mediterranean sea. *Scientia Marina*, 76(3):561–572, 2012. 129

E. Tzanatos, S. Somarakis, G. Tserpes, and C. Koutsikopoulos. Discarding practices in a mediterranean small-scale fishing fleet (patraikos gulf, greece). *Fisheries Management and Ecology*, 14(4):277–285, 2007. 131

A. J. Underwood. Techniques of analysis of variance in marine Biology and Ecology. *Oceanography and Marine Biology Annual Review*, 19:513–605, 1981. 78

D. V. Valavanis, I. Katara, and A. Palialexis. A GIS-based modelling approach for the mapping of marine productivity hotspots. *Aquatic Science*, 36:234–243, 2004. 90, 102

D. V. Valavanis, G. J. Pierce, A. F. Zuur, A. Palialexis, A. Saveliev, I. Katara, and J. Wang. Modelling of essential fish habitat based on remote sensing, spatial analysis and GIS. *Hydrobiologia*, 612:5–20, 2008a. 78

V. Valavanis, G. Pierce, A. Zuur, A. Palialexis, A. Saveliev, I. Katara, and J. Wang. Modelling of essential fish habitat based on remote sensing, spatial analysis and gis. *Hydrobiologia*, 612(1):5–20, 2008b. 101, 135

J. Valeiras. Attendance of scavenging seabirds at trawler discards off galicia, spain. *Scientia Marina*, 67(S2):77–82, 2003. 114

M. Valls, A. Quetglas, F. Ordines, and J. Moranta. Feeding ecology of demersal elasmobranchs from the shelf and slope off the Balearic Sea (western Mediterranean). *Scientia Marina*, 75(4):633–639, 2011. 112

M. Vasconcellos and K. Cochrane. *Overview of world status of data-limited fisheries: Inferences from landing statistics.* 2005. xiv, xv, xxi, xxii, 138

I. P. Vaughan and S. J. Ormerod. The continuing challenges of testing species distribution models. *Journal of Applied Ecology*, 42(4):720–730, 2005. 13

S. Vaz, C. S. Martin, P. D. Eastwood, B. Ernande, A. Carpentier, G. J.Meaden, and F. Coppin. Modelling species distributions using regression quantiles. *Journal of Applied Ecology*, 45(1):204–217, 2008. 113

M. Viana, N. Graham, J. G. Wilson, and A. L. Jackson. Fishery discards in the irish sea exhibit temporal oscillations and trends reflecting underlying processes at an annual scale. *ICES Journal of Marine Science: Journal du Conseil*, 68(1):221–227, 2011. 22, 115, 117

E. Vivekanandan, M. Srinath, and S. Kuriakose. Fishing the marine food web along the indian coast. *Fisheries research*, 72(2):241–252, 2005. 30

S. C. Votier, R. W. Furness, S. Bearhop andJ. E. Crane, R. W. G. Caldow, P. Catry, K. Ensor, K. C. Hamer, A. V. Hudson, E. Kalmbach, et al. Changes in fisheries discard rates and seabird communities. *Nature*, 427(6976):727–730, 2004. 114

W. Walker, P. Harremoës, J. Rotmans, J. P. van der Sluijs, M. B. A.van Asselt, P. Janssen, and M. P. Krayer von Krauss. Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support. *Integrated assessment*, 4(1):5–17, 2003. 18

D. J. Welch, B. D. Mapstone, and G. A. Begg. Spatial and temporal variation and effects of changes in management in discard rates from the commercial reef line fishery of the great barrier reef, australia. *Fisheries Research*, 90(1):247–260, 2008. 115

J. Wiegand, E. Hunter, and N. K. Dulvy. Are spatial closures better than size limits for halting the decline of the North Sea thornback ray, Raja clavata? *Marine and Freshwater Research*, 62(6):722–733, 2011. 134

M. S. Wisz, R. J. Hijmans, J. Li, A. T. Peterson, C. H. Graham, and A. Guisan. Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, 14:763–773, 2008. 78

W. Zhang. Supervised neural network recognition of habitat zones of rice invertebrates. *Stochastic Environmental Research and Risk Assessment*, 21(6): 729–735, 2007. 78

W. Zhang, X. Zhong, and G. Liu. Recognizing spatial distribution patterns of grassland insects: neural network approaches. *Stochastic Environmental Research and Risk Assessment*, 22(2):207–216, 2008. 78

Z. Zhang and J. Holmes. Use of generalized linear bayesian model to mitigate the impact of spatial contraction in fishing pattern on the estimation of relative abundance. *Fisheries Research*, 106(3):413–419, 2010. 22