What controls the spawning distribution of the Bay of Biscay anchovy: a multi-model approach.

Loots Christophe[1]; Planque Benjamin[2,*]; Vaz Sandrine[1]; Koubbi Philippe [3,4]; Huret Martin[5];

Petitgas Pierre[5]

[1]Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER), Laboratoire

Ressources Halieutiques, 150 quai Gambetta, BP699, F-62321, Boulogne/mer, France

[2]Institute of Marine Research. Postboks 6404. 9294 Tromsø. Norway.

[3]*UPMC Univ Paris 06, UMR 7093, Laboratoire d'Océanographie de Villefranche, 06230*

Villefranche-sur-Mer, France

[4]*CNRS, UMR 7093, LOV, 06230 Villefranche-sur-Mer, France*

[5]*Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER), Département*

Ecologie et Modèles pour l'Halieutique, Rue de l'île d'Yeu, B.P. 21105, 44311 Nantes Cedex

03, Nantes, France.

*Corresponding author: Tel: +47 77 60 97 21. Email: benjamin.planque@imr.no

Abstract. Understanding what controls the spawning distribution of short-lived fish species such as Bay of Biscay anchovy (*Engraulis encrasicolus*) is of crucial importance to understand recruitment process, which drives most of the interannual fluctuations in abundance of the population. Data from the Bay of Biscay PELGAS (PELagiques GAScogne) surveys carried out since 2000, were used to investigate several hypotheses on the control of anchovy spawning distribution in the Bay of Biscay. Six hypotheses were considered. Two extrinsic hypotheses: environmental control and geographical attachment; and four intrinsic hypotheses: spatial dependency, density dependent habitat selection, demographic structure and population persistence. Individual hypotheses, and combinations thereof, were expressed into generalized additive model formulations, and evaluated against data on presence-absence and abundance of anchovy. Models were evaluated, ranked and selected through a multi-model approach. Quantitative ranking of the models was used to identify the main hypotheses explaining distribution of spawning anchovy. We found that presence-absence was mainly determined by physiologically constraining environmental conditions, such as temperature, and geographical attachment at large spatial scale, whereas inside the areas of presence, abundance was primarily determined by trophic conditions, suitable for the successive maturations of the gonads of adults, and demographic structure. Our results also suggest that homing can be an important mechanism to ensure the return of adults to similar spawning areas from year to year.

Keywords. Anchovy; spawning distribution; hypotheses of control; homing; multi-model approach; Bay of Biscay

Introduction

The European anchovy (*E. encrasicolus*) is a short lived, fast growing, early reproductive and highly fecund pelagic fish species (ICES, 2004b). Individuals achieve most of their lifetime somatic growth, and become mature, during their first year of life, and may live up to 4 or 5 years. Spawning takes place in spring/summer, mainly from April to July with a maximal intensity in May. One-year old individuals spawn mainly in June whereas the 2-year old and older individuals spawn earlier, in late April-May (Motos *et al.*, 1996). The spawning stock biomass has strongly declined between 2000 and 2007 (ICES, 2008b). It was estimated to more 100,000 tons in 2000 and 2001, fell under 40,000 tons in 2002 and continued to decline to less than 20,000 tons in 2005. From 2005, the Bay of Biscay anchovy fisheries has been closed, as the stock did not recover to an acceptable abundance level. Because anchovy is a short lived and highly exploited species, the renewal and the maintenance of its stock level is mainly dependent on recruitment process in the most recent years. During 2000-2007, the spawning population was mainly composed of 1-year old individuals. In 2002 and 2005, 2-years old individuals dominated because of a poor recruitment, resulting in the collapse of the stock in these years.

Several environmental factors like North-easterly winds and upwelling intensity have been associated with patterns of recruitment variability (Borja *et al.*, 1998; Borja *et al.*, 2008; Planque and Buffaz, 2008). However, Bellier *et al.* (2007a) suggest that the configuration of the spawning habitat may also affect recruitment. Therefore, identifying factors that control the spawning habitat of anchovy could bring relevant information on what influences recruitment variability during the spawning period. Spawning habitat of anchovy has been mainly described (Motos *et al.*, 1996; Bellier *et al.*, 2007a) and modelled (Planque *et al.*, 2007) through the spatial distribution of eggs. Petitgas (2008) confronted several factors related to both environmental and population size conditions to model the mean spatial pattern

3

of spawning biomass and its associated residual variability. However, while the spatial distribution of eggs relies on this of the spawning adults, little has been done directly on the abundance distribution of the latter.

Most of the current studies related to habitat modelling are only considering the role of abiotic environment while totally ignoring some other relevant factors (Planque *et al.*, accepted). Since this allows for studying the potential habitat, e.g. the set of environmental variables potentially suitable for fish distribution (Planque *et al.*, 2007), this prevents for understanding what controls the observed distribution of fish populations. Recently, Planque *et al.* (in press) proposed a multi-model approach to evaluate and rank a range of ecological hypotheses on the control of the spatial distribution of fish populations. Rather than predicting, this explanatory approach proposed first, to accurately identify the relevant controls among a set of various hypotheses and then, to model the realised habitat, e.g. the areas where fishes are effectively distributed, which is not solely influenced by abiotic conditions (Planque *et al.*, 2007).

Such approach has been tested on benthic and demersal fish species (Loots *et al.*, 2010; Loots *et al.*, in press) but not on small pelagic species. The aim of the present study is to test this approach on the Bay of Biscay anchovy. Prior to analysis, the spatial distribution based on spawning adults' abundance will be described and mapped for the 2000-2007 time period. Then, mechanisms that influence this spatial distribution will be explored and identified using the multi-model approach.

Material and methods

The PELGAS survey

The PELGAS (PELagique GAScogne) survey has taken place each year since 2000 during the spring on the French continental shelf of the Bay of Biscay (Fig. 1, Table 1). It is

mainly designed for stock assessment of small pelagic fish (anchovy and sardine) using hydroacoustic methods (ICES, 2006; ICES, 2008b). During the day, acoustic prospecting is performed along transects that are separated by 12 nautical miles. These transects are oriented perpendicularly to the coast, from 25 m bottom depth to the continental slope (250 m), and cover the whole continental shelf from the Spanish coast in the south to the coast of Brittany in the north. Along these transects, acoustic energy is recorded at each nautical mile using an echo-sounder (SIMRAD EK500 38 kHz) and located as a point called an EDSU (Elementary Distance Sampling Unit). Fish species composition is determined from pelagic trawl hauls. All individuals captured (or a representative sample) are identified, counted and weighted. Otolith sampling, sexing and maturity staging are performed on a representative sample of individuals. Records of each EDSU are then combined following the method outlined in Petitgas *et al.* (2003) to allocate fractions of the total energy recorded (s_A: nautical area scattering coefficient) in term of biomass to the various species captured in the trawl. For each EDSU, this biomass is reallocated in term of number of individuals per size and age according to the mean weight observed in the reference trawl. During the night, vertical profiles using a CTD (Conductivity Temperature Depth) probe (Seabird 19) are performed along the same transects to measure temperature, salinity and density along the water column.

Abundances of spawning adults

During the PELGAS surveys (April-June), all adult anchovies were sexually mature and had started spawning (Motos *et al.*, 1996). Total anchovy abundance calculated at each EDSU were used to depict the spatial distribution of spawning adults between 2000 and 2007.

Distribution mapping

Geostatistics (Webster and Oliver, 2001) were used to map the spatial distribution of spawning adults of anchovy. Abundance of spawning adults were log-transformed ($\log(x+1)$, where x = abundance) to be closer to a normal distribution. For each year, an experimental variogram was computed. A theoretical model, chosen among exponential, circular, spherical and penta-spherical authorised functions was adjusted to this experimental variogram to determine the nugget, the sill and the range. These four models were adjusted using the least-square regression method (Webster and Oliver, 2001) and the one with the best visual and statistical fit to the experimental variogram was retained as the chosen theoretical variogram model. This theoretical variogram was then used to estimate $z(x)$ on the mesh of a regular grid by using the kriging interpolation method. A mesh size of 0.2 decimal degrees was chosen for the interpolation grid (Fig. 1, Table 1). Geostatistics were implemented using Genstat (GenStat Release 7.1., 2004).

Interpolated abundance of spawning adults were mapped in a geographical information system using ArcMap 9.1 (ESRI, 2005). Extrapolated areas (located outside of the surveyed area) were removed from the maps. The average map, summarising the average spawning area, was computed as the mean of the maps between 2000 an 2007. The variability map, representing the inter-annual variability in spawning, was produced using the standard deviation of maps between 2000 and 2007. Preferential, occasional, rare and unfavourable spawning sites were defined following Bellier *et al.* (2007a) using the average and variability maps. Preferential spawning sites were characterised by a high mean and a low variance, occasional spawning sites by a high mean and a high variance, rare spawning sites by a low mean and a high variance and unfavourable spawning sites by a low mean and a low variance. The thresholds used to split the maps into high and low values of mean and variance were chosen visually according to their distribution histograms. The raster calculator of the Spatial

6

Analyst extension of ArcMap 9.1 was used to combine the average and variability maps into the final map of preferential, occasional, rare and unfavourable spawning sites.

The multi-model approach

The spatial distribution of spawning adults can result from a variety of controls Planque *et al.* (In press). To investigate the role of these controls and how they can possibly interact, six families of models were constructed following Loots *et al.* (2010) and Loots *et al.* (in press): geographical position, environment, spatial dependency, age structure, population size and population memory. These models were then assembled with different level of complexity, from single hypotheses (e.g. pure environmental control) to complex interactions (e.g. density-dependence combined with population memory and environmental controls). The core models are detailed below.

Environment. On the whole, 15 environmental variables were used: bottom depth (m), bottom and surface temperature (˚C) and salinity, difference of temperature and salinity between the surface and the bottom, mixed layer depth (m), potential energy deficit, depth of the pycnocline (m), primary production index ($gC.m^{-2}$), upwelling index ($m.d^{-1}$), frontal index ($kg.m^{-2}.s^{-2}$), eddies index (s^{-2}) and sediment type.

Bottom depth, temperature and salinity were extracted from CTD profiles. Mixed layer depth and potential energy deficit were calculated using the densities extracted from CTD profiles according to the formulae in Planque *et al.* (2006). These variables were used to characterise the vertical stratification of the water column. Potential energy deficit is the amount of energy needed to vertically mix the water column, whilst mixed layer depth is the depth where the gradient of density is the highest. Because of temporal differences in the time periods of the survey (Table 1), anomalies were calculated and used for surface temperature, temperature

difference and potential energy deficit. These nine variables were interpolated using kriging on the same grid as the one used to interpolate spawning adults abundance (see the section on distribution mapping).

Depth of the pycnocline, primary production index, upwelling index, frontal index and eddies index are outputs of the Eco-Mars 3D model that were extracted for the period 2000-2007 (Huret *et al.*, 2009). The Eco-Mars 3D model (Lazure and Dumas, 2008) combines an hydrodynamic model with a biological model and can be used to simulate the main hydrological features for the Bay of Biscay and western part of the English Channel. These indices are available on a 3-days period with a resolution of 4 km. They were averaged on the time period of the PELGAS surveys for each year. Upwelling index, frontal index and eddies index were recoded in term of occurrences (0/1 for absence/presence) following thresholds of, respectively, 0.5 $m.d^{-1}$, 0.002 $kg.m^{-2}.s^{-2}$ and -10 s^{-2}. For each year, the closest value of these indexes was attributed to each cell of the interpolation grid of spawning adults' abundance.

Sediment types originated from the sedimentary structure of the seafloor that was digitalized in 2001 into a GIS on a scale of 1/500,000 from the "map of superficial sediments of the continental shelf of the Bay of Biscay" provided in 1986 by the BRGM (Bureau de recherches géologiques et minières). This map was modified by Chauvet (2001) and sediments were reclassified into four classes: mud, fine sand, coarse sand and gravels. A class of sediment was reallocated to each point of the interpolation grid.

A principal component (PCA) analysis was performed on these 15 environmental variables (columns) available for the 1,106 stations (lines) of the eight years, to study and group the variables according to their collinearity (see appendix 1).

Spatial dependency. Spatial dependency (also termed spatial autocorrelation) describes the spatial structure present in the spatial distribution of a species. This structure may potentially be shared by other controls such as the environment. Spatial dependency in the spatial

distribution of spawning adults was modelled using principal coordinates of neighbour matrices (PCNM) following Loots *et al.* (2010; in press). PCNM are particularly well designed to describe the spatial structure present in the distribution of an organism over a wide range of scales (Borcard and Legendre, 2002; Dray *et al.*, 2006; Bellier *et al.*, 2007b). Extraction of these PCNM was performed following Borcard and Legendre (2002). PCNM that describe the spatial structure present in anchovy distribution were then selected following Borcard and Legendre (2002) and Blanchet *et al.* (2008). Spatial scale of these PCNM was determined following the method outlined in Bellier *et al.* (2007b). PCNM describing equivalent spatial scale were grouped into three sub-models: broad, medium and fine scale sub-models (Bellier *et al.*, 2007b). These sub-models were used as three distinct hypotheses of control.

Age structure. Age structure was defined by annual and spatial age structure. Annual age structure (referred to as annual demography) represents the proportion of each cohort, in each year, for the whole population, whereas spatial age structure (referred to as spatial demography) is the spatial distribution of each cohort at each station for each year. Annual abundance of age classes from 1 to 3 years-old are provided by the ICES working group on anchovy (WGANC, ICES, 2008b). Abundance of each age class at each EDSU (based on the length distribution observed in the respective reference trawls) were interpolated using kriging (see section on distribution mapping) on the same interpolation grid as that used for spawning adults abundance and environmental variables.

Population size. Population size is the size of the population estimated each year. The spawning stock biomass (in tonnes) was used; it is estimated, each year, by the ICES working group on anchovy (WGANC, ICES, 2008b).

Geographical attachment. Geographical attachment can be used to depict the spatial persistence of a population distribution over years, and corresponds to geographical coordinates. Geographical position of each cell of the interpolation grid was used. Geographical attachment models are time-invariant and assume that spatial distribution can be modelled without reference to any specific year.

Population memory. Population memory models depict the degree of persistence of population distributions from one year to the next, and indicate that the current spatial distribution depends on past distributions. Contrary to geographical models (see above), these models are not time-invariant, but instead, they assume some degree of persistence in the spatial distribution over years. Population memory was taken as the observed spawning adult abundance of the preceding year. Each station (grid cell) was allocated the observed abundance of the previous year using the spatial joining tools of ArcMap 9.1. This reduced the total available number of stations from 1,106 to 946 from 2001 to 2007.

Models structure. Generalised additive models (GAM, Hastie and Tibshirani, 1990; Guisan *et al.*, 2002; Wood and Augustin, 2002) have been used for building the models (Loots *et al.*, 2010; in press). A binomial model with a logit link and a Gaussian model with an identity link were built using presence-absence data and non-null log-transformed abundance, respectively. This allows for dealing with non-normality of the data as well as to disentangling between factors that control the presence-absence and those that control the level of abundance.

Calibration and prediction datasets. Models were fitted (i.e. adjusted) on one portion of the dataset (called calibration dataset) and applied for prediction on the remaining portion of the dataset (called prediction dataset) following Loots *et al.* (2010) and Loots *et al.* (in press). For each class of model (either binomial or Gaussian), all the combinations of the hypotheses

were calculated resulting in 2^n-1 candidate models (with n being the number of hypotheses). The data used for the calibration and prediction datasets were selected using the PCA outputs of the analysis performed on the environmental variables.

Model ranking. Models were evaluated, ranked and selected using their ability to predict (when applied to the prediction dataset) the spatial distribution of spawning adults as advised by Planque *et al.* (in press), Loots *et al.* (2010) and Loots *et al.* (in press). For each model, the log-likelihood between observations and predictions (McCullagh and Nelder, 1989; Burnham and Anderson, 2002) was calculated as it allows for comparing non nested models in an evaluation procedure based on prediction. The best model was the one with the highest log-likelihood value. Other models were compared to this model using a likelihood ratio test (LRT, Burnham and Anderson, 2002). Models were ranked according to their LRT from smallest to largest. A selection threshold of 60 for LRT was used in order to keep only a reduced set of selected models. Details on calculation of the log-likelihood and likelihood ratio test can be found in Loots *et al.* (2010).

Variance partitioning. Variance partitioning was used to explore the percentage of variation explained by each of the hypotheses within selected models. Redundancy analysis (RDA, Legendre and Legendre, 1998) was used to relate the retained hypotheses to presence-absence and positive abundance of the prediction dataset. Although RDA is the constrained version of PCA (in that the ordination axes are constrained to become linear combinations of the variables), it has been used here in a single response context, making it equivalent to the multiple regression model described in the PCNM forward selection procedure. Each hypothesis was tested successively and those that were determined as significant ($p<0.05$) were kept for partial analysis. Partial analysis is a way of estimating how much of the variation of the response variable (either presence-absence or positive abundance) can be

attributed exclusively to one hypothesis (variable), once the effect of the other hypotheses (covariables) has been removed (Borcard *et al.*, 1992). For a set of tested hypotheses, each variable was firstly regressed linearly by the covariables to partial out the shared part of variation. Once the shared component had been extracted, the partialled out variable was linearly regressed on the response to measure the amount of variation it purely explained on its own. Hence, the amount of variation attributable or shared among the various hypotheses tested could be measured. The 'rda' and 'varpart' functions of the 'vegan' package (Oksanen *et al.*, 2009) of R were used to perform the Monte Carlo permutation test and partial analysis.

Results

Distribution of spawning adults

Spawning adults' abundance was used to depict the spawning distribution of anchovy in the Bay of Biscay for the period 2000-2007. Its spatial distribution has been described using geostatistical analyses and mapped into a GIS. Eight variograms have been computed and adjusted on the spatial distribution of anchovy between 2000 and 2007. A quadratic trend has been used for most of the years except for 2003. The importance of this spatial trend decreased between 2000 and 2001, followed by a highly variable period between 2001 and 2005 and an increase since 2005 (Fig. 2). It can be observed that the variations of the spawning stock biomass seem to be correlated to those of the spatial trend, except in 2002 where the value of the trend decreased a year before the spawning stock biomass.

Maps of the spatial distribution of spawning adults of anchovy from 2000 to 2007 are shown in Fig. 3. Overall the study period, adults were mainly concentrated south of 46°N, whilst being mostly absent of the Northern part, except in 2001 and 2003 where high concentrations were located along the coast of Brittany. In the southern part, spawning adults were abundant along the coast of les Landes. There was a high degree of inter-annual variability in both the level of abundances and the extent of the spawning areas. The highest

abundances were observed in 2000 and 2001 and the lowest abundances in 2005. The spatial

distribution in 2003 was patchy and spread over most of the shelf whereas it was much

narrower and concentrated to the coast between 2004 and 2006.

Maps of each year between 2000 and 2007 were used to calculate the average and

variability maps. These maps were combined using a threshold of 1 for the mean and 0.8 for

the variance to identify preferential, occasional, rare and unfavourable spawning areas for

anchovy (Fig. 4). Three well-defined preferential spawning sites could be identified. Two

were coastal and located in front of the Adour and Gironde estuaries. The third one was more

off the coast. Occasional spawning sites consisted in a continuous area that surrounded the

preferential sites from the French coast of les Landes to the Gironde estuary. Because

observed abundances were low in the Northern part of the Bay of Biscay during most of the

study period, this area was characterised as an unfavourable site for spawning. Rare spawning

sites were located between occasional and unfavourable areas.

Spatial structures arising from the survey design and relevant to the population

distribution were quantified using PCNM. PCNM were computed and extracted using a

threshold of 0.2 (decimal degrees of latitude) to truncate the distance matrix between stations.

Overall, 105 positive PCNM were extracted from the principal coordinate analysis and 15

were kept at the end of the forward selection (p-value < 0.05 with an adjusted-R^2 of 0.091).

From these 15 PCNM, 3 spatial scales were identified in the spatial distribution of spawning

adults of anchovy (Fig. 5). Two PCNM were used to build the broad scale sub-model (66-114

km), five for the medium scale sub-model (33-66 km) and eight for the fine scale sub-model

(< 33 km). In particular, the first PCNM in the broad scale sub-model depict a clear

distinction between the northern and southern part of the Bay of Biscay. The medium scale

describes smaller structures like the river plume of la Gironde and Adour estuaries. Fine scale

is more difficult to interpret but may be related to minor hydrodynamic structures such as

upwelling areas along les Landes coast, eddies along the shelf break and the 'cold pool', a cold bottom body of water in the northern part of the shelf. These PCNM have been used as spatial sub-model for the multi-model approach.

The multi-model approach

For binomial models, 13 hypotheses of control (Table 2) were compared: geographical attachment, the six environmental hypotheses, the three spatial sub-models, annual demography, population size and population memory. For Gaussian models, spatial demography was also used. On the whole, 8,191 binomial and 16,383 Gaussian models were built, compared and ranked according to their likelihood ratio test. Finally, 191 binomial and 6,882 Gaussian models were selected within a range of LRT between 0 and 60.

Relevance of the hypotheses was determined according to their frequency in the selected models classified within intervals of LRT of 10 (Table 3). For both binomial and Gaussian models, the frequency of each hypothesis little varied between intervals except for hypotheses that were not frequent in first models and that became more frequent in the following ones. Hence, the relevance of the hypotheses could be determined from their frequency in selected models with an LRT restricted to values between 0 and 10. Six hypotheses were frequent (more than 50% of selected models) in binomial models: the first environmental hypothesis (Envt. 1, upwelling, fronts, and difference of salinity), bottom temperature, sediment type, broad and medium scales and population memory. For Gaussian models, eight hypotheses were frequent in selected models with an LRT value between 0 and 10: the first two environmental hypotheses (Envt. 1, upwelling, fronts, and difference of salinity; Envt. 2, pycnocline, eddies and primary production), bottom temperature, broad and medium scales, population size and annual and spatial demography.

The percentage of variation explained by each of the six hypotheses in presence-absence and each of the eight hypotheses in positive abundances was evaluated using variance partitioning (Table 4).

For presence-absence (binomial models), the medium scale was not determined as significant by the redundancy analysis. The five significant hypotheses explained 40% of the variation in presence-absence data. Bottom temperature, broad scale sub-model and sediment type contributed the most to the explained variation whereas Envt. 1 and population memory contributed the least. These hypotheses shared a large amount of explained variation and explain only a small part of variation on their own.

For positive abundances only (Gaussian models), only the first two environmental hypotheses, broad scale, annual and spatial demography and population memory were determined as significant by the redundancy analysis. These hypotheses explained a total of 62% of the variation in the positive abundances. Among these hypotheses, primary production, annual and spatial demography contributed to the larger part of the explained variation.

Discussion

The spawning distribution of anchovy shown spatio-temporal variations between 2000 and 2007. However, a mean pattern of distribution with areas of high abundances could have been extracted overall the different years. This pattern corresponds with what was described in the literature from the distribution of the eggs (Motos *et al.*, 1996; Bellier *et al.*, 2007a). At the time of the survey, in late spring and early summer, spawning adults are preferentially located in the southern part of the Bay of Biscay and high abundances are located in three preferential spawning areas, off the Adour and la Gironde estuaries and along the shelf break. Along the season, this distribution shows a pattern of concentration/expansion (Motos *et al.*,

1996) where reproductive activity begins into the two coastal preferential spawning grounds and disperse afterwards to occupy the whole Bay of Biscay area (Planque *et al.*, 2007).

The multi-model approach applied on the spatial distribution of spawning adults of anchovy enabled to distinguish the main hypotheses that influence the distribution of presence-absence from those that influence the local abundance in the areas of presence of spawning fish. These mechanisms are discussed below.

What controls the presence of spawning adults of anchovy?

The spatial pattern of presence-absence was explained by three main hypotheses: the broad scale spatial dependency, bottom temperature and sediment type.

The larger part of variability explained by broad scale compared to the one explained by bottom temperature and sediments suggests that the spatial pattern of occupation of anchovy in the Bay of Biscay is primarily determined by geographical location rather than by environment. Moreover, the large part of the variation explained by the first PCNM highlighted that the spatial segregation between the northern and the southern parts of the shelf of the Bay of Biscay mainly explains the pattern of presence-absence. In the present study, the northern part was characterised as an unfavourable site for spawning which is in agreement with other studies on eggs distribution (Bellier *et al.*, 2007a).

Bottom temperature and sediment type shared a common effect with broad scale spatial dependency and also explained a significant part of variation in the distribution of presence-absence. Relevance of bottom temperature was also found in other studies. Petitgas (2008) has demonstrated that bottom temperature was an important factor to explain the mean spatial distribution of biomass of anchovy in the Bay of Biscay. Planque *et al.* (2007) modelled the potential spawning habitat of anchovy using eggs distribution and found that bottom temperature was the best predictor. Bay of Biscay anchovy are not closely connected with

16

bottom but bottom temperature appears as a good proxy for the relatively homogeneous hydrographic conditions found below the thermocline, where most of the individuals are observed. Indeed, during the day, anchovies form small schools, distributed between 10 and 25m above the bottom (ICES, 2004b) whereas during the night they migrate up to the surface. At the time of the survey, bottom temperature was warmer south of 46°N and colder on the northern part. The role of bottom temperature and the progressive warming from south to north of the Bay of Biscay might explain the northward extent of the potential spawning habitat noticed by Planque *et al.* (2007).

The percentage of variation explained by geography and bottom temperature showed that the latter has a lower predictive capacity than the former. This may be due to the coverage of the survey which was not sufficient to sample all the spatial variation of bottom temperature and therefore adequately detect the relationship between temperature and the spawning distribution of anchovy. We suggest that extending the survey area toward the Spanish coast may allow for exploring this issue.

Our analysis shows that sediment type appears as an important predictor for spatial occupancy. While individuals were mainly distributed in the southern part of the Bay of Biscay that is characterised mainly by fine sand, they seem to avoid the northern part that is dominated by muddy sediments. As was the case with bottom temperature, anchovy are not closely associated to sediments, but the geographical distribution of different types of sediments is partly the result of processes occurring in the water column above. So sediment type can be interpreted as a proxy for water column characteristics. The geographical match between spring hydrological landscapes (Planque *et al.*, 2004) and sediment types supports this interpretation.

What determines the local abundance of spawning anchovy?

Our results show that, inside the area of presence, the level of abundance of spawning anchovy is mainly determined by trophic conditions and population demography. From the thirteen environmental variables tested in the present study, only few explained a significant part of the observed variation in abundances distribution. We found that including process-related variables derived from a hydrodynamic model was usually more successful that using conventional factors such as temperature and salinity.

The relevance of frontal index and primary production confirmed that food production and concentration are important features which regulate the spatial distribution of high spawning concentrations. In the Bay of Biscay, primary production is mainly driven by river outflows that provide nutrients directly from terrigenous sources and by the coastal upwelling between Adour and Gironde estuaries (Motos *et al.*, 1996). However, the present study reveals that the upwelling index is a poor predictor. Increased productivity resulting from nutrient upwelling, even if necessary, does not appear sufficient to provide suitable spawning conditions for anchovy. Hence the spawning distribution of anchovy relies mainly on the primary production originating from river plumes which also provide areas with an important stability of the water column thus favouring nutrient enrichment and concentration of food particles (Motos *et al.*, 1996). Therefore, associated to highly productive areas, thermohaline fronts support high zooplankton productivity and concentration. The irrelevance of eddies suggests that they are not a so important oceanographic feature for determining the areas of high concentrations of spawning anchovy as they may be for larvae (Bakun, 2006). This is supported by the observed spatial distribution of eggs, larvae and juveniles described by Irigoien *et al.* (2008) who showed a segregation pattern in the distribution of the larvae according to their size, i.e. larger larvae are more distant to the coast than smaller ones. It is likely that anchovy favours more the areas that will provide a sufficient food to ensure feeding of adults during the

spawning period rather than areas that will provide suitable conditions for released eggs and larvae and their potential retention over the French shelf (Irigoien *et al.*, 2007). In a highly variable environment like the Bay of Biscay, the spawning of anchovy relies on the occurrence each year of some perennial pattern of distribution related to food distribution associated with fronts. Because it primarily ensures the survival of the adults rather than that of their eggs, this strategy may not always result in successful recruitment during unfavourable years but may insure the reproductive success of adults over longer time periods.

Spatial demography related to how the individuals belonging to different age classes are distributed in space also appeared to be a key factor. Motos *et al.* (1996) have described the segregation pattern of distribution between young individuals of 1-year old mostly distributed to the coast in front of the Gironde Estuary and old individuals (2-year old and more) that are distributed further offshore, in front of the Adour estuary and along the shelf break. This spatial segregation between young and old individuals is more related to the size of the individuals than strictly to their age, i.e. young but large individuals may be found offshore mixed with older ones. In this context, the demographic structure of anchovy population act through a size dependant habitat selection more than an age dependent habitat selection. Since large individuals benefit from enhanced energy storing capacity linked to their larger body size, old individuals are able to travel larger distance compared to young individuals. This ability offers an advantage to old individuals over young ones; however it does not explain why they systematically spawn in more offshore areas each year. If competition interaction between young and old individuals is probably one of the causal factors, the presence of old individuals on the shelf break may be explained in term of food resources adapted to the planktivorous feeding regime of this species. Irigoien *et al.* (2009) showed a clear spatial pattern between 1998 and 2006 in zooplankton distribution in the Bay

of Biscay. High proportions of small-size mesozooplankton were observed near the coast whereas larger zooplankton predominated near the shelf break. Since larger anchovies might be able to feed on larger preys, the persistent spatial difference in the size structure of anchovy preys may favour and constrain the segregation pattern between old and young individuals.

Although population size was not retained as a significant hypothesis in the present study, the similar temporal trends of spawning biomass and of the spatial trend determined from geostatistics suggest an abundance-area relationship where the extent of the spatial distribution varies according to the size of the population. In the present study, the hypothesis relevant to the population size was frequent only in Gaussian models, meaning that a higher number of individuals in the population will affect the level of abundance of spawning anchovy in a given location but not the probability of its presence in that place. This also supports the result obtained by Planque *et al.* (2007) on the absence of relationship between the size of the population and the probability of presence of eggs. Bellier *et al.* (2007a) found that the spatial distribution of anchovy has expanded northward between two distinct periods 1967-1972 and 2000-2004, which may suggest that on the long term, year to year variation in the spawning biomass may also affect the extent of the occupied spawning area. Population size was not found relevant to predict the presence-absence of anchovy between 2000 and 2007 and was not determined as significant for the level of abundance by the redundancy analysis. This may reflect that variations in population size where not sufficient to cause such important changes in the spawning distribution that these effect could be detected by the models.

Homing as a key mechanism for spawning habitat of anchovy?

The mean spatial pattern detected (Figure 4) as well as the relevance of some variables highly structured in space (like the broad scale) with little variation in time (like bottom temperature, Planque *et al.*, 2006) highlights a temporal persistence in the spawning distribution of the Bay of Biscay anchovy. Brochier *et al.* (in press) suggested that a mechanism of homing is implied and tested for two versions of it, geographical and environmental, as to explain the resulting spatial distribution of eggs and larvae of anchovy in two different upwelling ecosystems. They found that environmental homing is more likely to be the key mechanism involved as it was more able to reproduce the spatio-temporal dynamic of spawning compared to geographical homing.

However, the present study differs from the one by Brochier in both the species studied (*E. encrasicolus* versus *Engraulis ringens* in the Humbolt Current) and the environmental characteristics of the study area. While in the Canary, Humbolt and Benguela Currents areas, primary production relies on upwelling, in the Bay of Biscay, the coastal upwelling intensity is weak and primary production is mostly controlled by river discharges on the shelf, and vertical pumping of nutrients on the shelf break. Doligez *et al.* (2003) showed that the reproductive success according to different strategies of reproduction could vary with the degree of temporal variability in environmental conditions. A strategy of reproduction based on geographical homing would have better success in ecosystems where environment is stable and autocorrelated in time. Hence, in upwelling ecosystems where environmental conditions are highly variable, it is likely that a strategy based on environmental homing would be better suited. In the Bay of Biscay, geographical homing may be favored as a result of stability in the geographical arrangement of environmental conditions, in particular trophic conditions provided by river outflows.

Finally, since we found that factors influencing the spatial distribution of presence-absence of spawning individuals were not the same as those controlling their level of

abundance, we suggest that for pelagic species, the mechanism of homing can only explain the pattern of presence-absence. Indeed, we showed that the level of abundance was explained by other types of controls that were more related to the state of the population. Hence, the relevance of both the broad scale and bottom temperature in explaining the spatial distribution of presence-absence may suggest that bottom temperature controls the spatial distribution of adults but that spawning individuals adopt a strategy based on geographical homing in which they return to the southern par of the Bay of Biscay where they usually find suitable spawning conditions of bottom temperature. This would explain why the whole suitable area of bottom temperature is not occupied and also why the broad scale has a greater power of prediction compared to bottom temperature.

Conclusion

The geographical distribution and abundance of anchovy in the Bay of Biscay greatly varied between 2000 and 2007. The multi-model approach used here revealed that from April to June, the spawning distribution is constrained to the southern part of the Bay of Biscay which is mainly characterised by fine sand sediment and where bottom temperature is sufficiently high to allow for the onset of spawning. Aggregations of high spawner abundances are found mainly in areas suitable for adults feeding and vary also according to the population age-structure. Surprisingly, there is a high degree of year-to-year persistence in anchovy spawning distribution and as a result, spawning adults do not strictly track the distribution of ideal environmental conditions. This may lead to poor recruitment in specific years (e.g. 2002, 2005) but ensure good average recruitment over longer time periods.

References

Bakun, A. (2006) Fronts and eddies as key structures in the habitat of marine fish larvae: opportunity, adaptive response and competitive advantage. *Sci. Mar.* **70S2:** 105-122.

Bellier, E., Planque, B. and Petitgas, P. (2007a) Historical fluctuations in spawning location of anchovy (*Engraulis encrasicolus*) and sardine (*Sardina pilchardus*) in the Bay of Biscay during 1967-73 and 2000-2004. *Fish. Oceanogr.* **16:** 1-15.

Bellier, E., Monestiez, P., Durbec, J. and Candau, J. (2007b) Identifying spatial relationships at multiple scales: principal coordinates of neighbour matrices (PCNM) and geostatistical approaches. *Ecography* **30:** 385-399.

Blanchet, F.G., Legendre, P. and Borcard, D. (2008) Forward selection of explanatory variables. *Ecology* **89:** 2623-2632.

Borcard, D., Legendre, P. and Drapeau, P. (1992) Partialling out the spatial component of ecological variation. *Ecology* **73:** 1045-1055.

Borcard, D. and Legendre, P. (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecol. Model.* **153:** 51-68.

Borja, A., Uriarte, A., Egana, J., Motos, L. and Valencia, V. (1998) Relationships between anchovy (*Engraulis encrasicolus*) recruitment and environment in the Bay of Biscay (1967-1996). *Fish. Oceanogr.* **7:** 375-380.

Borja, A., Fontan, A., Saenz, J. and Valencia, V. (2008) Climate, oceanography, and recruitment: the case of the Bay of Biscay anchovy (*Engraulis encrasicolus*). *Fish. Oceanogr.* **17:** 477-493.

Brochier, T., Colas, F., Lett, C., Echevin, V., Cubillos, L.A., Tam, J., Chlaida, M., Mullon, C. and Fréon, P. (in press) Small pelagic fish reproductive strategies in upwelling systems : a natal homing evolutionary model to study environmental constraints. *Prog. Oceanogr.*

Burnham, K.P. and Anderson, D.R. (2002) *Model selection and multimodel inference. A practical information-theoretic approach*. USA: Springer-Verlag, 488pp.

Chauvet, F. (2001) Développement d'un système d'information géographique permettant d'identifier les zones côtières déterminantes pour le renouvellement des stocks de poissons. Rapport de stage Ifremer, 35 pp.

Doligez, B., Cadet, C., Danchin, E. and Boulinier, T. (2003) When to use public information for breeding habitat selection? The role of environmental predictability and density dependence. *Anim. Behav.* **66:** 973-988.

Dray, S., Legendre, P. and Peres-Neto, P.R. (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecol. Model.* **196:** 483-493.

ESRI (2005) Environmental Systems Research Institute. ArcGis. Version 9.1. Redlands, CA : Environmental Research Institute Inc.

GenStat Release 7.1. (2004) Lawes Agricultural Trust (Rothamsted Experimental Station). R.W Payne, S.A Harding, D.A. Murray, D.M. Soutar, D.B. Baird, S.J. Welham, A.F. Kane,

A.R. Gilmour, R. Thompson, R. Webster, G. Tunnicliffe Wilson. VSN International, Wilkinson House, Jordan Hill Road, Oxford, UK.

Guisan, A., Edwards, T.C. and Hastie, T. (2002) Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecol. Model.* **157:** 89-100.

Hastie, T.J. and Tibshirani, R.J. (1990) *Generalized additive models*. London: Chapman and Hall, 335pp.

Huret, M., Struski, C., Léger, F., Petitgas, P., Lazure, P. and Sourisseau, M. (2009) Modélisation couplée physique-biogéochimie du golfe de Gascogne sur la période 1971-2007, 26 pp.

ICES (2004b) Report of the study group on regional scale ecology of small pelagic fish (SGRESP), 23-26 February 2004, Nantes, France. ICES CM 2004/G:06 Ref. ACFM, ACE, 55 pp.

ICES (2006) Report of the working group on acoustic and egg surveys for sardine and anchovy in ICES areas VIII and IX, 24-28 October 2005, Vigo, Spain. ICES CM 2006/LRC:01 Ref. ACFM, 126 pp.

ICES (2008b) Report of the working group on anchovy, 13-16 June 2008, Copenhagen, Denmark. ICES CM 2008/ACOM:04, 230 pp.

Irigoien, X., Fiksen, O., Cotano, U., Uriarte, A., Alvarez, P., Arrizabalaga, H., Boyra, G., Santos, M., Sagarminaga, Y., Otheguy, P., Etxebeste, E., Zarauz, L., Artetxe, I. and Motos, L. (2007) Could Biscay Bay Anchovy recruit through a spatial loophole?. *Prog. Oceanogr.* **74:** 132-148.

Irigoien, X., Cotano, U., Boyra, G., Santos, M., Alvarez, P., Otheguy, P., Etxebeste, E., Uriarte, A., Ferrer, L. and Ibaibarriaga, L. (2008) From egg to juvenile in the Bay of Biscay: spatial patterns of anchovy (*Engraulis encrasicolus*) recruitment in a non-upwelling region. *Fish. Oceanogr.* **17:** 446-462.

Irigoien, X., Fernandes, J.A., Grosjean, P., Denis, K., Albaina, A. and Santos, M. (2009) Spring zooplankton distribution in the Bay of Biscay from 1998 to 2006 in relation with anchovy recruitment. *J. P. Res.* **1:** 1-17.

Lazure, P. and Dumas, F. (2008) An external-internal mode coupling for a 3D hydrodynamical model for applications at regional scale (MARS). *Adv. Water Resour.* **31:** 233-250.

Legendre, P. and Legendre, L. (1998) *Numerical Ecology*. Amsterdam: Elsevier, 853pp.

Loots, C., Vaz, S., Planque, B. and Koubbi, P. (2010) What controls the spatial distribution of North Sea plaice spawning population ? Confronting ecological hypotheses through a model selection framework. *ICES J. Mar. Sci.* **67:** 244-257.

Loots, C., Vaz, S., Planque, B. and Koubbi, P. (In press) Understanding what controls the spawning distribution of North Sea whiting (*Merlangius merlangus*) using a multimodel approach. *Fish. Oceanogr.*

McCullagh, P. and Nelder, J.A. (1989) *Generalized linear models, Monographs on statistics and applied probability 37*. London: Chapman and Hall, 511pp.

Motos, L., Uriarte, A. and Valencia, V. (1996) The spawning environment of the Bay of Biscay anchovy (*Engraulis encrasicolus* L.). *Sci. Mar.* **60:** 117-140.

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Simpson, G.L., Solymos, P., Stevens, M.H.H. and Wagner, H. (2009) vegan: Community Ecology Package. R package version 1.15-3.

Petitgas, P., Massé, J., Beillois, P., Lebarbier, E. and Le Cann, A. (2003) Sampling variance of species identification in fisheries-acoustic surveys based on automated procedures associating acoustic images and trawl hauls. *ICES J. Mar. Sci.* **60:** 437-445.

Petitgas, P. (2008) Fish habitat mapping with Empirical Orthogonal Functions. ICES CM 2008/M:07.

Planque, B., Bellier, E. and Loots, C. (accepted) Uncertainties in projecting spatial distributions of marine populations. *ICES J. Mar. Sci.*

Planque, B., Lazure, P. and Jégou, A. (2004) Detecting hydrological landscapes over the Bay of Biscay continental shelf in spring. *Clim. Res.* **28:** 41-52.

Planque, B., Lazure, P. and Jegou, A. (2006) Typology of hydrological structures modelled and observed over the Bay of Biscay shelf. *Sci. Mar.* **70S1:** 43-50.

Planque, B., Bellier, E. and Lazure, P. (2007) Modelling potential spawning habitat of sardine (Sardina pilchardus) and anchovy (*Engraulis encrasicolus*) in the Bay of Biscay. *Fish. Oceanogr.* **16:** 16-30.

Planque, B. and Buffaz, L. (2008) Quantile regression models for fish recruitment-environment relationship - four case studies. *Mar. Ecol.-Prog. Ser.* **357:** 213-223.

Planque, B., Loots, C., Petitgas, P., Lindstrøm, U., Vaz, S. (In press) Understanding what controls the spatial distribution of fish populations using a multi-model approach. *Fish. Oceanogr.*

Webster, R. and Oliver, M.A. (2001) *Geostatistics for environmental scientists. Statistics in practice*. Chichester, England: Wiley, 271pp.

Wood, S.N. and Augustin, N.H. (2002) GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. *Ecol. Model.* **157:** 157-177.
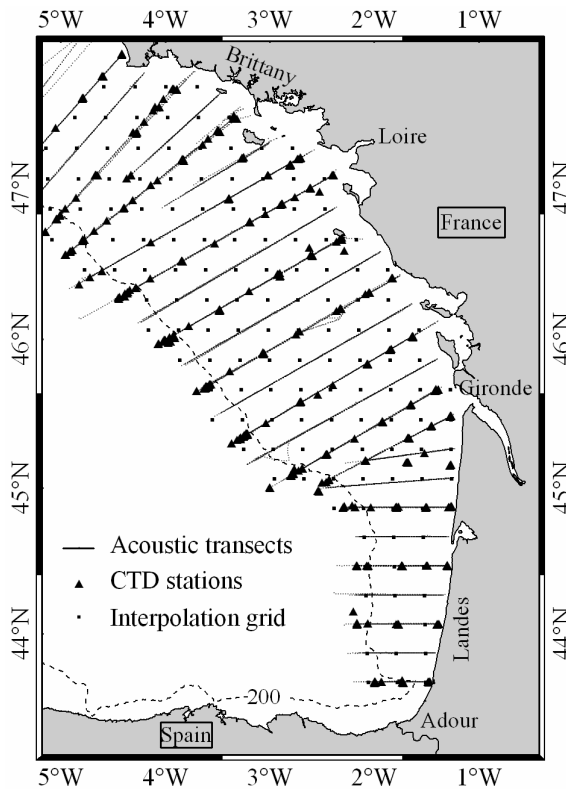
Figures

Figure 1: Map of the Bay of Biscay PELGAS surveys from 2000 to 2007. Acoustic transects, CTD stations and the grid used for the kriging interpolation are indicated. Location of the Adour, la Gironde, la Loire rivers, the coast of les Landes and Brittany and the isobath 200 m delimiting the French continental shelf are also indicated.
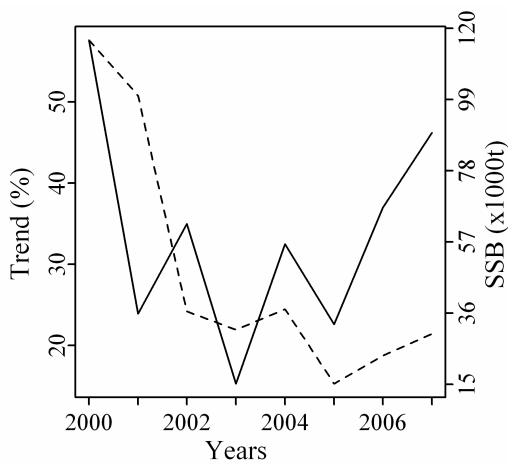

Figure 2: Plots of the percentage of variation explained by the spatial trend obtained within geostatistical analyses (black line) and estimation of the spawning stock biomass (dotted line) for each year between 2000 and 2007.
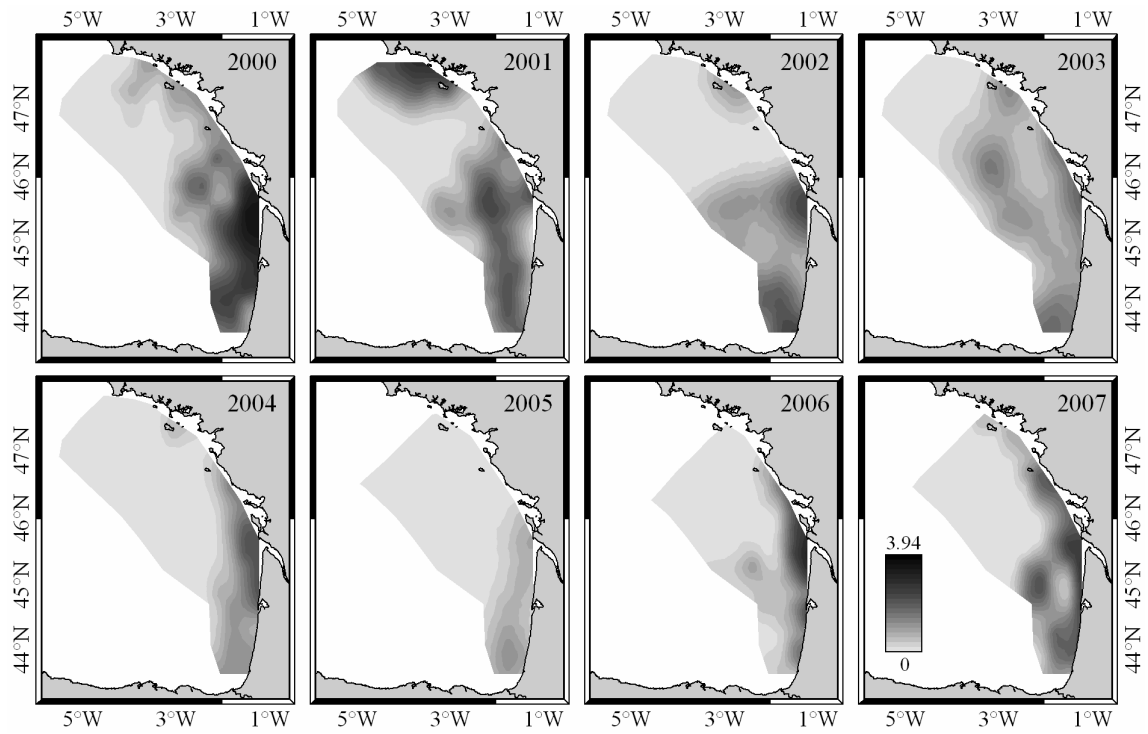
28

Figure 3: Maps of the spatial distribution of spawning anchovy in the Bay of Biscay between 2000 and 2007. A classification of twenty classes of equal interval ranging from 0 to 3.94 (log-transformed) was used.
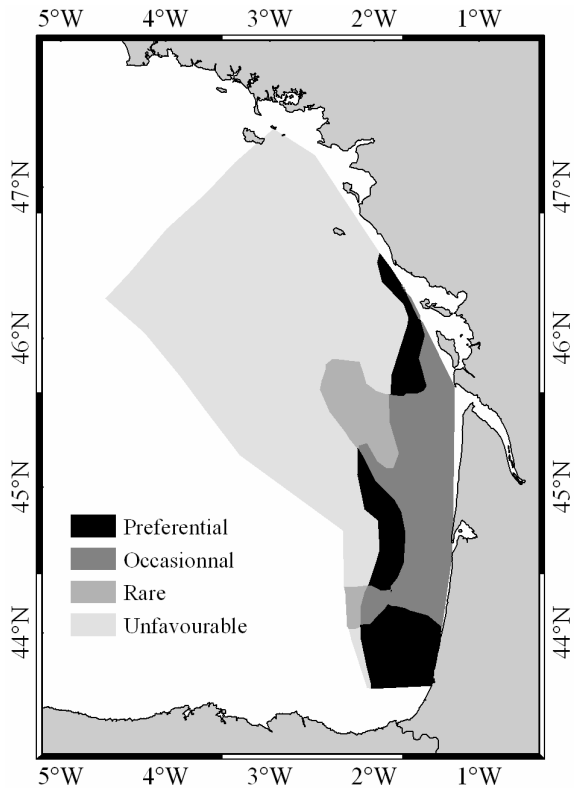
Figure 4: Location of preferential, occasional, rare and unfavourable spawning sites for anchovy in the Bay of Biscay. The map was computed from the average and variability maps that were calculated overall distribution maps between 2000 and 2007 (see Fig. 3).
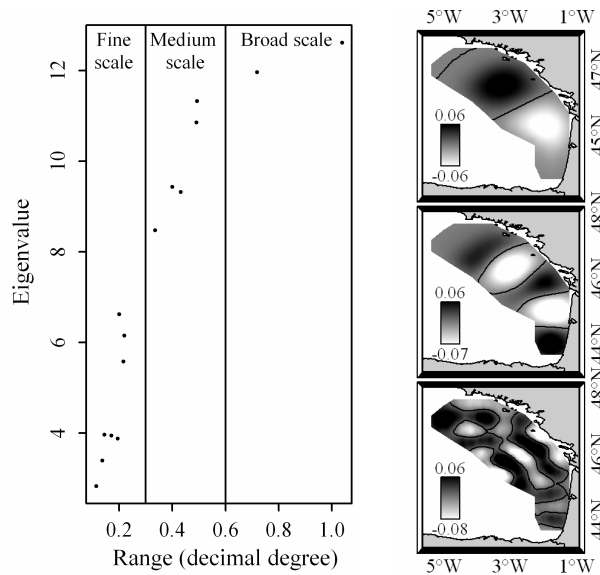

Figure 5: Spatial scale sub-models of spawning anchovy distribution. Left: plot of the range values of the 15 selected PCNM against their eigenvalue. Ranges were obtained by fitting a Gaussian model as a theoretical variogram on each spatial predictor. Three scales were determined. Broad scale: eigenvectors with a range between 66 and 114 km (0.6 to 1.1

decimal degrees of latitude); medium scale: eigenvectors with a range between 33 km and 66 km (0.3 to 0.6 decimal degree of latitude); fine scale: eigenvectors with a range between 0 and 33 km (less than 0.3 decimal degrees of latitude). Right: maps of the three scale sub-models. PCNM 1, 6 and 34 were used to depict each scale. Black line indicates the zero value.

Tables

Table 1. Time period of the PELGAS surveys for each year between 2000 and 2007. Number of CTD stations and EDSU available for kriging interpolation are indicated for each year, along with the final number of stations of the grid used for kriging interpolation.

Year	Time of the survey		Number of stations		
	Start	End	CTD	EDSU	Grid
2000	18-Apr	14-May	93	1500	149
2001	03-May	03-Jun	60	1636	149
2002	10-May	06-Jun	80	1593	149
2003	30-May	24-Jun	75	1625	149
2004	29-Apr	23-May	79	1736	149
2005	05-May	24-May	66	1381	124
2006	02-May	30-May	62	1355	113
2007	27-Apr	19-May	66	1447	124

Table 2. Formulae of the individual hypotheses. An *s* means that a smoothing function has been used and the corresponding degree of smoothing is indicated in parentheses. PCNM are denoted by the term *Ax*. Front Ind.: front index; Upwell. Ind.: upwelling index; Sal. Diff.: difference of salinity; Pycno. Dep.: depth of pycnocline; Edd.: index of eddies; Prim. Prod.: primary production; Pot. Ener. Def.: potential energy deficit; Surf. Temp.: surface temperature; Temp. Diff.: difference of temperature; Mix. Lay. Dep.: mixed layer depth; Surf. Sal.: surface salinity; Bott. Sal.: bottom salinity; Dep.: depth.; ssb: spawning stock biomass; Prev. Yr. Ab.: abundance of the previous year.
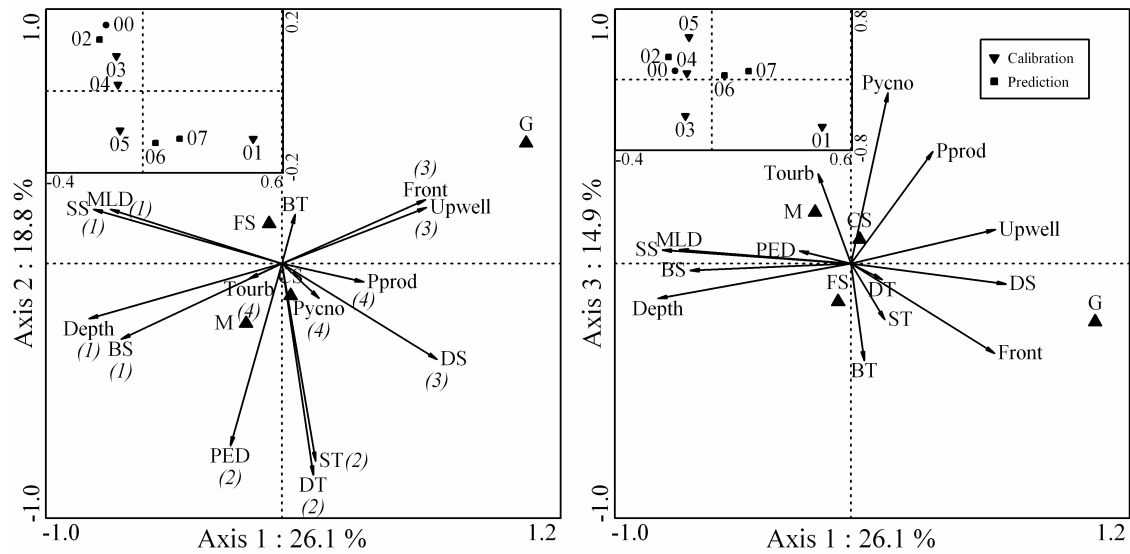
31

Hypotheses	Binomial models	Gaussian models
Geographical Attachment	as.factor(GridKnot)	as.factor(GridKnot)
Environment 1	Front Ind. + Upwel. Ind. + s(Sal. Diff., 4)	s(Front Ind., 4) + s(Upwel. Ind.,2) + s(Sal. Diff., 4)
Environment 2	s(Pycno. Dep., 4) + s(Edd., 3) + s(Prim. Prod., 4)	s(Pycno. Dep., 4) + Edd. + s(Prim. Prod., 4)
Environment 3	s(Pot. Ener. Def., 4) + s(Surf. Temp., 4) + s(Temp. Diff., 4)	s(Pot. Ener. Def., 4) + s(Surf. Temp., 3) + Temp. Diff.
Environment 4	s(Mix. Lay. Dep., 3) + Surf. Sal. + s(Bott. Sal., 2) + Dep.	s(Mix. Lay. Dep., 4) + s(Surf. Sal.,4) + Bott. Sal. + s(Dep.,4)
Bottom Temperature	s(Bott. Temp., 3)	s(Bott. Temp., 4)
Sediments	as.factor(Sediments)	as.factor(Sediments)
Broad Scale	s(Ax1, 1) + s(Ax3, 1)	s(Ax1, 2) + s(Ax3, 3)
Medium Scale	s(Ax6, 4) + Ax7 + s(Ax13, 2) + s(Ax14, 4) + s(Ax18, 2)	s(Ax6, 4) + Ax7 + s(Ax13, 4) + Ax14 + s(Ax18, 1)
Fine Scale	s(Ax30, 1) + s(Ax34, 4) + s(Ax39, 4) + s(Ax53, 3) + Ax54 + Ax56 + Ax62 + s(Ax69, 1)	s(Ax30, 2) + s(Ax34, 3) + s(Ax39, 1) + Ax53 + s(Ax54,2) + Ax56 + Ax62 + Ax69
Population Size	s(ssb, 3)	s(ssb, 2)
Annual Demograhy	Age1a + Age2a + Age3a	Age1a + Age2a + Age3a
Spatial Demography	-	s(Age1s, 4) + Age2s + s(Age3s, 4)
Population Memory	s(Prev. Yr. Ab., 2)	s(Prev.Yr. Ab., 3)

Table 3. Frequencies of the different hypotheses, giving the percentage of the model selected from prediction in which each hypothesis is present, for different threshold of selection (LRT) ranging from 0 to 60. Recoding was according to five levels: 0 (absence), + (<25%), ++ (25-50%), +++ (50-75%), ++++ (75-100%).

		Binomial models						Gaussian models					
	LRT	*[0-10[*	*[0-20[*	*[0-30[*	*[0-40[*	*[0-50[*	*[0-60[*	*[0-10[*	*[0-20[*	*[0-30[*	*[0-40[*	*[0-50[*	*[0-60[*
Geographical Attachment		0	0	0	0	0	0	0	+	++	++	++	++
Environment 1		++++	++++	++++	++++	+++	+++	+++	+++	++	++	++	++
Environment 2		0	0	0	+	+	+	+++	++	++	++	++	++
Environment 3		0	0	0	+	+	+	++	+++	++	++	++	++
Environment 4		0	0	0	0	0	0	0	+	+	+	+	+
Bottom Temperature		++++	++++	+++	+++	+++	+++	++++	+++	+++	+++	++	+++
Sediments		++++	++++	++++	++++	+++	+++	+	++	++	++	++	++
Broad Scale		++++	++++	++++	++++	++++	+++	++++	+++	+++	+++	+++	+++
Medium Scale		+++	+++	+++	+++	+++	++	++++	++++	+++	+++	+++	+++
Fine Scale		0	+	++	++	++	++	++	++	++	+++	+++	+++
Population Size		0	0	0	0	+	+	+++	+++	+++	+++	+++	+++
Annual Demograhy		+	+	+	+	+	+	++++	+++	++	++	++	++
Spatial Demography		-	-	-	-	-	-	++++	++++	++++	++++	++++	+++
Population Memory		+++	++	++	++	++	++	++	++	+++	+++	++	+++

Table 4. Relevance of the hypotheses of control. For each hypothesis retained at the end of the redundancy analysis, the percentage of explained variation is indicated. It is expressed in term of total (the percentage the hypothesis totally explains) and pure (the percentage the hypothesis explains on its own once the shared component with other hypotheses has been removed) explained variation.

Hypothesis	Total (%)	Pure (%)
Presence-absence		
Envt. 1	8	1
Upwel. Ind.	2	0
Front Ind.	3	1
Sal. Diff.	7	0
Bott. Temp.	15	5
Sed.	20	4
Br. Sc.	27	5
Ax1	25	3
Ax3	2	2
Pop. Mem.	9	1
Positive abundances		
Envt. 1	12	3
Upwel. Ind.	0	0
Front Ind.	12	2
Sal. Diff.	2	1
Envt. 2	23	11
Pycno.	7	0
Pprod.	16	11
Edd.	0	0
Br. Sc.	13	5
Ax1	5	5
Ax3	2	0
Ann. Dem.	12	10
Sp. Dem.	20	19
Pop. Mem.	12	0

Appendix 1. Principal component analysis (PCA) performed on the 15 environmental variables of the Bay of Biscay used in the multi-model approach. Left: plot of the first and second axes. Right: plot of the first and third axis. The first axis separated offshore areas on the left with higher depth, from coastal areas on the right that are more productive and characterized mainly by gravels. The second axis identified the middle-north of the continental shelf with muddy bottoms and high values of potential energy deficit. The third axis separated the south of the shelf with high values of bottom temperature from the north with lower values. Six groups of variables were determined from the three first axes of the PCA (Envt. 1, Envt. 2, Envt. 3, Envt. 4, BT, Sediment type).The first group of variables (Envt. 1) gathers frontal index, upwelling index and salinity difference. The second group (Envt. 2) consists of pycnocline depth, eddies index and primary production. The third group (Envt. 3) represents potential energy deficit, surface temperature and temperature difference. The fourth group (Envt. 4) is made of mixed layer depth, surface salinity, bottom salinity and depth. Bottom temperature (BT) did not belong to any of these groups and was considered as a hypothesis on its own. Also, sediment type was also considered as a hypothesis as a

categorical variable, it cannot be grouped with other variables. The different groups were used as the environmental hypotheses of control.

Box inlets: each year of data is represented by the centroid of its stations' scores, so as to illustrate its relative position in the ordinated multidimentional space. The years' centroid distribution highlighted years with similar environmental conditions. Four years (2001, 2003, 2004 and 2005) were used to build the calibration dataset representing 571 and 295 stations used to fit binomial and Gaussian models, respectively. Three years (2002, 2006 and 2007) were used for the prediction dataset, which represents 375 and 195 stations, respectively, used for binomial and Gaussian models.

BS: bottom salinity; BT: bottom temperature; DS: difference of salinity; DT: difference of temperature; Front: frontal index; MLD: mixed layer depth; PED: potential energy deficit; Pprod: primary production; Pycno: pycnocline depth; SS: surface salinity; ssb: spawning stock biomass; ST: surface temperature; Upwell: upwelling index; M: mud; FS: fine sand; CS: coarse sand; G: gravels.