ELSEVIER

Are commercial logbook and scientific CPUE data useful for characterizing the spatial and seasonal distribution of exploited populations? The case of the Celtic Sea whiting

Marion Verdoit [a], Dominique Pelletier [a,*], Robert Bellail [b]

[a] Laboratoire de Mathématiques Appliquées à l'Exploitation des Ressources Halieutiques et Aquacoles (MAERHA), IFREMER Centre de Nantes, 2, rue de l'île d'Yeu, B.P. 21105, 44311 Nantes cedex 3, France
[b] Laboratoire Ressources Halieutiques, IFREMER Station de Lorient, 8, rue François Toullec, 56100 Lorient, France

Abstract

A statistical approach is proposed to utilize the information contained both in scientific surveys and commercial fishing logbooks, in order to determine spatial and temporal distributions of recruiting and spawning fish of exploited populations. The approach is based on multivariate descriptive methods including ordination methods and classification techniques. A typology involving a principal component analysis (PCA), followed by a hierarchical ascending classification (HAC), is applied to both scientific and commercial data. PCA that accounts for spatial and temporal contiguities were performed. These analyses allow to compare structures observed at global and local scales, with those observed without focusing on a particular scale. The method is applied to the whiting population in the Celtic Sea. The analyses performed on both scientific and commercial CPUE show that most age groups are found north of 50°N. The analysis of commercial CPUE indicates four periods in the year associated with particular areas. High abundances of adults (age >2 years) are observed during spawning season (January–April) in inshore areas (south of Ireland and off Cornwall). Intermediate abundances of adults are observed in the same area in May and June. Between July and September, all age groups are observed in the Smalls area, where particularly high abundance of recruiting group 1 are found. During the last quarter, the distribution of adults is more widespread, mainly ranging between 51°N and 52°N.
© 2003 Elsevier SAS and Ifremer/IRD/Inra/Cemagref. All rights reserved.

Résumé

Les distributions spatiales et saisonnières des populations exploitées, peuvent-elles être caractérisées par les CPUE commerciales et scientifiques? Le cas du merlan de la mer Celtique. Une approche statistique est proposée pour utiliser les informations issues à la fois des campagnes scientifiques et des carnets de bord de la pêche commerciale, afin de déterminer les distributions spatiales et saisonnières du recrutement et du stock reproducteur de populations exploitées. L'approche est basée sur des méthodes descriptives multivariées qui incluent des méthodes d'ordination et des techniques de classification. On réalise une typologie qui implique une Analyse en Composantes Principales (ACP) suivie d'une Classification Hiérarchique Ascendante (HCA), à la fois sur les données scientifiques et sur les données commerciales. On a effectué des ACP qui prennent en compte la contiguïté spatiale et temporelle. Ces analyses permettent de comparer les structures observées à des échelles globales et locales, par rapport à celles observées, sans privilégier une échelle particulière. Cette approche est appliquée au merlan de la mer Celtique. L'analyse des CPUE scientifiques et commerciales montre que la plupart des groupes d'âge sont trouvés au nord de 50°N. Les analyses des CPUE commerciales indiquent 4 périodes dans l'année associées à des zones particulières. De fortes abondances d'adultes (âge > à 2 ans) sont observées durant la saison de reproduction (de janvier à avril) dans les zones côtières (au Sud de l'Irlande et en Cornouaille). Des abondances intermédiaires d'adultes sont observées dans la même zone en mai et juin. Entre juillet et septembre, tous les groupes d'âge sont observés dans la région des Smalls, où de très fortes abondances du groupe d'âge 1 nouvellement recruté sont observées. Durant le dernier trimestre, la distribution des adultes est plus dispersée, et se répartit surtout entre 51° N et 52° N.
© 2003 Elsevier SAS and Ifremer/IRD/Inra/Cemagref. All rights reserved.

* Corresponding author.
E-mail address: dpellet@ifremer.fr (D. Pelletier).

1. Introduction

Most demersal and benthic fish populations exhibit spatial and seasonal patterns in relation to their annual life cycle, in particular linked to reproduction. After spawning, adults migrate towards feeding areas, and eggs and larvae drift towards nurseries, which correspond in general to habitats that are particularly suitable for larval settlement and juvenile growth. As a consequence, there is often a spatial and/or seasonal segregation between adults and juveniles (Nicholsky, 1968). In spite of this, most population dynamic models, and in particular those used for stock assessment, are not spatially and seasonally structured. Such models are needed for a more accurate evaluation of fishing mortality when fishing effort targets areas of high fish abundance (see e.g. Pelletier and Magal, 1996). Also, models that are structured in space and time are necessary to assess the impact of spatio-seasonal management measures, which are more and more advocated as an additional management tool to reduce overexploitation. Building such models requires a better knowledge of the large-scale spatial and seasonal patterns of distribution for each demographic stage of the population. As for many other species, the main features of whiting, *Merlangius merlangus* (Linnaeus, 1758), life cycle are well identified, but knowledge of stage-specific spatial preferenda, as well as corresponding seasons, is poor. Literature on the subject lacks details, e.g. only mentioning "far from the shore", or "in coastal areas", and may even appear ambiguous or contradictory. Such descriptions are not sufficiently formalized for building a spatial and seasonal model of population dynamics.

Scientific surveys and records from the commercial fishery are the main sources of data available to determine stage-specific spatial and temporal distributions of exploited populations. Scientific survey data have several advantages. First, they are obtained from standardized and controlled sampling protocols, which facilitates their analysis. Second, they are precisely georeferenced and aim at providing a balanced spatial coverage of the area of distribution of the populations studied. A major interest of scientific surveys is to give information on the spatial distribution and abundance level of the youngest individuals (0-group and 1-group), which are not well represented in commercial landings data, because they are either not yet recruited to the fishery or discarded at sea. On the other hand, as surveys are generally conducted once a year, such data cannot provide a sufficient seasonal coverage for describing population dynamics. In contrast, fishery logbooks provide a much larger set of catch per unit effort (CPUE) with a good temporal coverage over the year. However, the location of each fishing operation is only known at the scale of ICES statistical rectangles. Timing is known at the scale of trip duration (i.e. 12 days for French trawlers in the Celtic Sea). Although the spatial and temporal resolutions of commercial fishing data are not as fine as for scientific data, they may be sufficient to depict broad spatial and seasonal features for stages above minimum landing size. CPUE yield relative abundance levels for each characteristic areas (or patches) at specific times of the year (seasons). A major drawback of using commercial CPUE as an index of abundance lies in the lack of information in certain areas not visited by fishermen. Commercial and scientific data exhibit complementary features that could be utilized to determine large-scale spatial and seasonal distributions of exploited populations.

Spatial distributions of marine populations are commonly patchy or display gradient structures, resulting in spatial autocorrelation for abundance (Legendre, 1993). Spatial correlation functions or indices (Moran, 1948; Geary, 1954; Cliff and Ord, 1973) allow to evaluate the similarity of abundances measured at different locations. Some methods of spatial data analysis explicitly handle spatial correlation, e.g. kriging methods (Petitgas, 1993; Pelletier and Parma, 1994), but they require georeferenced data and an appropriate sampling design for estimating correlation structures. In the case of coarse spatial resolution, other methods may be used. For instance, there have been several attempts to introduce spatial information in multivariate descriptive analysis (e.g. Thioulouse et al., 1995), but applications in ecology are scarce (Thioulouse et al., 1995 in terrestrial ecology, Trueblood, 1991; Gaertner, 1997 in marine ecology). Cornillon et al. (1997) developed a method which takes into account both spatial and temporal contiguities in a multitable analysis, but no applications in ecology are published yet. Multivariate analyses accounting for contiguities between individuals are based on a generalization of Geary's index (Lebart, 1969).

In this paper, we aim at delineating (i) patches that identify areas with similar age-specific abundance levels at certain times of the year; and (ii) corresponding seasons as groups of months.

Large-scale spatial and temporal autocorrelation in fish abundance is exemplified by the fact that, each year, specific population stages are confined in particular areas (e.g. nurseries, spawning grounds, wintering grounds) during several months. In order to account for this large-scale autocorrelation of abundance, we propose a method based on multivariate ordination and classification techniques which takes into account temporal and spatial contiguities, unlike typologies based on a simple PCA, which assume that statistical individuals are independent from the standpoint of active variables (age-specific abundances in this case). The approach is illustrated by an application to the whiting population of the Celtic Sea, an important commercial species for the French fishery. We first summarize existing knowledge on the life cycle and spatial and seasonal patterns of the Celtic Sea whiting. In a second step, we describe commercial and scientific CPUE data usually available for exploited populations in the north-east Atlantic, based on the example of the Celtic Sea whiting. In particular, we compare their spatial and temporal resolutions, as well as other features resulting from the sampling design. We then present the method used to delineate seasonal age-specific distribution zones from these data. We finally discuss from a more general standpoint the

questions of data collection aimed at improving the knowledge of spatial and seasonal patterns of exploited populations.

2. Whiting population and exploitation in the Celtic Sea

The Celtic Sea is a continental sea bounded on its eastern and northern sides by France, the UK, and Ireland (Fig. 1). Whiting is one of the main commercial species caught in this area, chiefly by the French fleet, but fishermen from the UK, Ireland, and Belgium also fish this area. The French fleet is made up of medium-sized (20–33 m long) bottom trawlers that operate all year round, making trips of 8–14 days targeting either gadoids, whiting and cod *(Gadus morhua)*, benthic fish, monkfish (*Lophius piscatorius* and *Lophius budegassa*) and megrim *(Lepidorhombus whiffiagonis)*, or Norway lobster *(Nephrops norvegicus)*. Whiting is a target for the gadoid metier (see e.g. Pelletier and Ferraris, 2000 for characterization of metiers in the Celtic Sea fishery), and is incidentally caught by the other two metiers. The minimum legal size of whiting is 27 cm for all areas north of 48°N. The largest share of annual catches is taken south of Ireland and north of Cornwall, during the first quarter of the year where mature individuals concentrate for spawning. In September, a second peak of catches is observed, which may be linked to a feeding concentration on the Smalls (VIIg2) (Charuau and Biseau, 1989).

Whiting is a demersal species living at depths from 10 to 200 m, with a preference for the 30–100 m range (Quéro and Vayne, 1997). The identity of whiting populations in the Celtic Sea and in the Western English Channel is not firmly established (IFREMER and MAFF, 1993). In these areas, spawning takes place from January–February to May–June, with a peak in March–April (Ould el Kettab, 1993). At the beginning of the spawning season, mature whiting migrate towards coastal areas (Ould el Kettab, 1993). After metamorphosis, juveniles of 6–7 cm settle on the bottom in shallow coastal areas (Nagabhushanam, 1965) off Southern Ireland (VIIg3), in St George's Channel (VIIf1) and around Cornwall (VIIe2) (Ould el Kettab, 1993) (Fig. 1). A discard study on the *Nephrops* fishery showed that juveniles are also found during winter on the Smalls (VIIg2) (Charuau and Biseau, 1989). The average length at maturity (L_{50}) is about 32 cm total length, i.e. roughly at age 2 (Charuau and Biseau, 1989).

3. Materials and methods

3.1. Scientific survey data

French Institute for the Exploitation of the Sea (IFREMER) and Center for Environment, Fisheries and Aquaculture Science (CEFAS) have undertaken annual bottom trawl surveys in the Celtic Sea since 1990 and 1989, respectively. These surveys are mainly intended to investigate the biology, abundance and spatial distribution of all species, which can be representatively sampled by bottom trawls. Catch in numbers and weights, and length compositions are routinely recorded for all species caught. Otoliths are sampled, so that age–length keys by sex are available to estimate age distributions from length distributions.

The gear used during IFREMER surveys is a GOV (Large Vertical Opening) 36/47 trawl with a 20 mm mesh codend
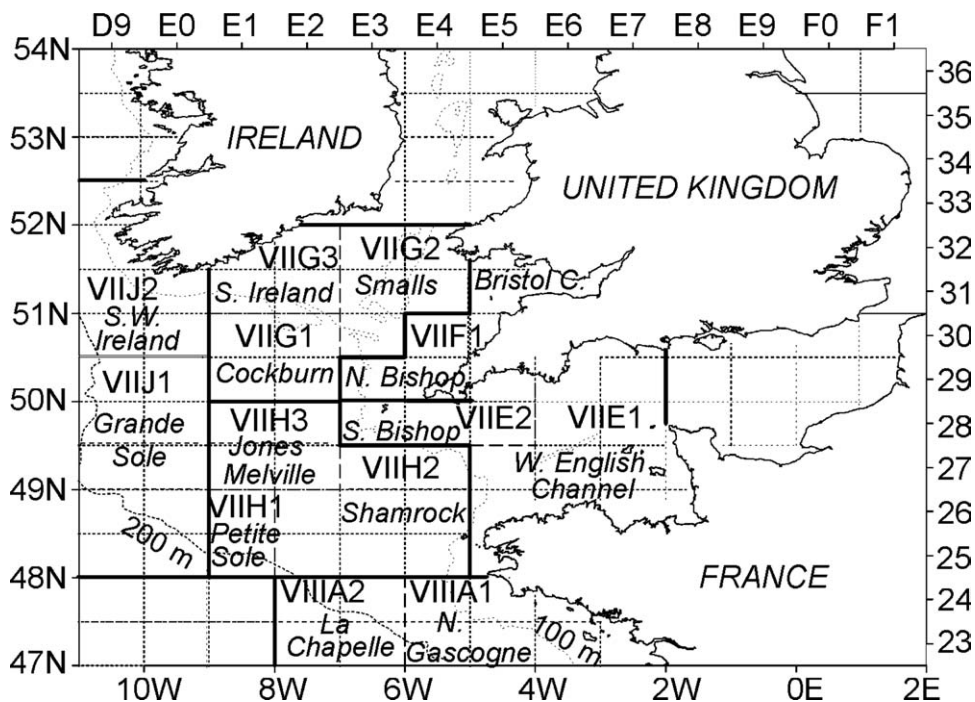

Fig. 1. Map of Celtic Sea with 100 and 200 m isobaths, fishing locations and ICES subdivisions.

liner. The trawl is deployed with plane oval trawl boards of 1300 kg and with 10–20 cm rubber discs on the footrope. Haul duration is 30 min. The survey area is stratified according to latitude and depth with two random hauls in each stratum. Two geographical areas are identified in the Celtic Sea and six depth zones are used (ICES, 1997). The 12 strata are divided into units of 25 square nautical miles. The number of hauls per stratum is proportional to stratum surface. The standard bottom trawl used during CEFAS surveys is a modified Portuguese high-headline trawl (PHHT) which is fitted with 35 cm rubber bobbins on the ground rope, a bunt tickler chain, and a 20 mm mesh codend liner. Haul duration is 60 min. The survey area is stratified per statistical rectangle with two hauls in each stratum. In both surveys, towing speeds are similar (7.41 km h^1) and trawling is carried out at day time only. Technical differences between the UK and French gears pertain to horizontal trawl opening (14.3 vs. 18.4 m), vertical trawl opening (4.4 vs. 4.1 m), door type (4.5/1440 vs. 4.5/1350 m^2 kg^1), ground rope (18.3 vs. 47 m), head line (16 vs. 36 m), warp-length/water depth (3.8 vs. 2–3 m), bobbins, disks and chains (35.6 cm rubber bobbins, rubber disks and 16.5 m bunt tickler chain vs. rubber disks). Possible consequences on relative gear efficiencies are discussed in Section 5.

In this paper, we analyzed data from the IFREMER surveys conducted in 1997, 1998 and 1999. For CEFAS surveys, data only from 1998 and 1999 were used because age–length keys were available only for these two years. IFREMER and CEFAS surveys were, respectively, performed in autumn (October–November) and in spring (March), and will therefore, be referred to as autumn survey and spring survey in the rest of the paper. The sampled area extends from 48°N to 52°N and 3°W to 12°W; it includes ICES Divisions VIIf, g, h, j and the north-western part of VIIe (Fig. 1). For each survey, the number of hauls performed in each sampled ICES statistical rectangle varies from one to five, but is generally one or two. For each age group and each haul, CPUE was calculated as catch weight per towed hour. Summary statistics for these surveys are given in Table 1.

3.2. Commercial logbook data

Fisheries statistics (logbook catch and effort data) are routinely collected by IFREMER and stored in an official database managed by the Regional Center of Statistical Processing (CRTS) of the French administration. In this study, we considered data from every single and paired bottom-trawl vessel working in the Celtic Sea. The mesh size is 90 mm. The areas selected for the analysis are subdivisions VIIe1, VIIe2, VIIf1, and divisions VIIg and VIIh (Fig. 1), thus including the Western English Channel area. Although whiting may be found in VIIj1 and VIIj2 too, these subdivisions were excluded because French catch and effort are negligible there. Effort data consists of nominal fishing effort in trawling hours per ICES statistical rectangle (0.5° latitude × 1° longitude) for each trip recorded in the database. Catch

data (in kg per species and market category) correspond to landings and are provided by the census of sales at market hall. Due to a lack of detailed information about discards, landings per unit effort (LPUE) are assumed to be an abundance index, and will be termed CPUE in the following. The discard problem will be discussed later in the paper.

In this paper, we considered data from 1993 to 1997, based on the expertise of scientists in charge of the fishery. During these years, all vessels use the same gear and vessels have similar lengths. This way, we minimized possible trends in fishing power due to technical characteristics. CPUE were recorded per market category. In order to better match demographic stages, they were converted into CPUE per age group using commercial sampling data collected in French market halls of south-west Brittany where most of Celtic Sea catch is landed. The sampling protocol is a stratified three-stage sampling scheme. For each quarter from 1993 to 1997, a relative age composition is estimated for each market category and applied to catch per month. Summary statistics for commercial data are given in Table 1.

Since fishermen are not mandated to give an estimation of catch by species per rectangle, catch may be unambiguously assigned to a rectangle only if the vessel fished exclusively within a single rectangle during that trip. Such trips are called "single rectangle trips" in the following. They represent on average 37% of all the trips (Table 2). Note that for the purpose of this paper it is not required that these trips be a representative sample of the whole set of commercial trips; rather they should represent a substantial percentage of all trips and a good coverage of the whole Celtic Shelf. This issue is further discussed later.

3.3. Computation of CPUE indices

First we specify the scales at which data are analyzed to investigate recurrent spatial and seasonal patterns of abundance. Regarding spatial scale, commercial CPUE have the coarser resolution, but they represent the bulk of the data. We thus retained the scale of ICES statistical rectangles, and survey CPUE are averaged per rectangle. Recall that there are on average two hauls per rectangle in each survey. Regarding temporal scale, we computed monthly commercial CPUE, which are suitable for studying seasonal features. We are interested in patterns tied to the life cycle, which should be recurrent over years. Evidencing such patterns may be achieved by analyzing either yearly CPUE or CPUE averaged over years. We conducted both types of analyses. However, in the case of commercial data, fishing trips in a given year do not cover all rectangles each month, and CPUE averaged over years provide a better coverage for appraising spatial and seasonal distribution patterns. Furthermore, averaging reduces the variability of the CPUE abundance index. Results of analyses carried out from yearly CPUE and from averaged CPUE were compared to check that analyses of averaged CPUE reflect recurrent patterns and are not too dependent on year-specific patterns. Note that we are only

Table 1

Scientific and commercial data used in the analysis (CV: coefficient of variation). Fishing effort is expressed in trawling hours for scientific data and in number of single rectangle trips for commercial data in the Celtic Sea

Data source	Date		CPUE (kg h^{-1})						Fishing effort
			Age 0	Age 1	Age 2	Age 3	Age 4	Age 5+	
(a) Scientific data									
Autumn									
1997	4–19 November	Maximum	196	251	41	83	159	28	25.5
		Mean	13	16	4	8	14	4	
		CV	7	6	3	4	5	4	
1998	29 October–22 November	Maximum	123	56	85	171	256	56	28
		Mean	6	2	5	9	13	4	
		CV	8	11	8	7	7	5	
1999	13 November–6 December	Maximum	1339	1232	88	276	78	186	30.5
		Mean	63	64	32	9	4	7	
		CV	7	7	10	10	7	9	
Averaged	–	Maximum	223	214	160	54	64	37	–
		Mean	12	12	6	4	5	2	
		CV	3	3	4	3	2	3	
Spring									
1998	1–29 March	Maximum	–	312	394	395	430	296	80
		Mean	–	14	28	29	28	20	
		CV	–	4	3	3	3	3	
1999	4–28 March	Maximum	–	287	212	201	202	151	87
		Mean	–	13	13	12	12	12	
		CV	–	4	4	3	3	3	
Averaged	–	Maximum	–	25	278	269	293	213	–
		Mean	–	14	22	2	22	16	
		CV	–	4	3	3	3	3	
(b) Commercial data									
1993	1993	Maximum	–	15 962	466 939	222 694	48 264	45 131	2 188
		Mean	–	306	7 817	7 235	2 403	1 558	
		CV	–	5	4	3	2	3	
1994	1994	Maximum	–	3 868	107 722	162 937	99 380	49 112	1 695
		Mean	–	71	3 693	10 603	5 774	1 900	
		CV	–	4	3	2	2	3	
1995	1995	Maximum	–	6 950	26 512	123 760	198 355	131 660	1 802
		Mean	–	136	1 451	7 541	10 329	4 530	
		CV	–	5	2	2	2	2	
1996	1996	Maximum	–	514	19 105	58 504	97 333	73 254	1 786
		Mean	–	6	451	5 023	7 181	4 909	
		CV	–	7	4	2	2	2	
1997	1997	Maximum	–	0	28 710	65 814	132 388	97 190	3 647
		Mean	–	0	553	2 911	9 365	6 701	
		CV	–	0	5	3	2	2	
Averaged	1993–1997	Maximum	–	1 260	89 817	162 937	98 577	56 106	2 224
		Mean	–	10	3 464	7 864	7 866	4 411	
		CV	–	9	3	2	2	2	

Table 2

Evolution of the number of single rectangle trips and the number of total commercial trips between 1993 and 1997

	Years				
	1993	1994	1995	1996	1997
Number of total trips	5816	5112	5287	5284	7545
Number of single rectangle trips	2188	1695	1802	1786	3647
Proportion of single rectangle trips	37.6	33.2	34.1	33.8	48.3

interested in possible consequences of year-to-year differences for cluster composition, and not in testing year effects per se. Inferential analyses that test for year effects are not helpful in this purpose. Year effects will inevitably be significant, but this falls beyond the scope of this paper. For commercial data, a mean CPUE was thus calculated per statistical rectangle and per month as the sum of catch divided by the sum of fishing effort, taken over all single rectangle trips between 1993 and 1997. Aggregation over trips is assumed to smooth out as much as possible individual effects due to fishermen behavior (e.g. fishing power, fishing strategy and small-scale allocation of fishing effort). In this analysis, only single rectangle trips are used. They amount from 1700 to 3700 trips per year between 1993 and 1997 (Table 2).

3.4. Characterizing seasonal distributions of population abundance

The objective of the paper is to identify spatial and seasonal patches with similar abundances per age group. To achieve this, we used a multivariate descriptive methodology derived from Pelletier and Ferraris (2000). This approach, termed "typology" in the previous paper, aims at defining clusters of statistical individuals that are similar with respect to variables of interest. Here, in the case of scientific data, statistical individuals are ICES rectangles during survey month. In the case of commercial data, statistical individuals are ICES rectangles during a given month (termed rectangle-months in the paper). Each individual is described by the set of CPUE per age group (in kg h^1).

The typology involves a principal component analysis (PCA), followed by a hierarchical ascending classification (HAC) on the principal axes of the PCA. The advantages of this approach for building clusters of individuals were pointed out by Lebart et al. (1997) and are discussed by Pelletier and Ferraris (2000) in the fisheries science context. Briefly, PCA provides insight into the underlying structure of a large data by synthesizing the relationships between variables. Individuals and variables may be projected on principal axes generated by new variables termed principal components, which are linear combinations (LC) of the initial variables. Principal axes and principal components are ordered according to their decreasing contribution to the variance of the data set. Hence, the first axis is the one along which the set of individuals are the most scattered, i.e. they are best discriminated. In HAC, clusters are built by successive pairwise agglomeration of individuals. HAC is here performed on the principal coordinates of the individuals. Among the possible partitions of individuals illustrated by the hierarchical tree (the dendrogram), we retain partitions with the highest level of the aggregation index, i.e. we cut the tree when the next agglomeration of two clusters leads to a maximum increase of within-cluster variance (Lebart et al., 1997). Clusters are then characterized by computing test-values. Test values are statistically significant when their absolute value is larger than 2 (Lebart, 1969). When a test-

value is significantly positive (respectively, negative) (Lebart et al., 1984), the mean of the quantitative variable (or the frequency of the categorical variable) is significantly higher (respectively, lower) in the cluster than in the population. In the typology, we distinguished active variables used to obtain the factorial axes or the clusters, here the CPUE variables, and illustrative variables that were only used to describe principal axes (by projections on these axes) or to characterize clusters (by computing test-values), e.g. months and rectangles. In the present analysis, HAC groups individuals into clusters according to the abundance of each age group. Hence, each cluster comprises rectangle-months that exhibit similar abundances per age group. HAC results are represented on monthly maps of spatial distribution displaying the cluster to which each rectangle belongs. In order to evaluate how the typology summarized the data, monthly maps of CPUE are also presented for each age group from age 0 to 3+, for both scientific and commercial data.

3.5. Contiguity analyses

A typology based on a simple PCA assumes that individuals are statistically independent from the standpoint of active variables (age-specific abundances in this case). Since the typology of rectangle-months is intended (i) to delineate zones with similar age-specific abundances at a given period of the year; and (ii) to determine seasons during which occupation of space by fish is similar, spatial and temporal contiguities should be considered. In addition to simple PCA, we use the multivariate contiguity analysis described by Thioulouse et al. (1995). This method is described in Appendix A. It yields the same kind of results as simple PCA, but in addition is helpful to check that observed structures are consistent at small and large spatio-temporal scales.

Contiguity analyses followed by HAC on factorial axes were applied to scientific and commercial data. In the case of commercial data, both temporal and spatial contiguities were accounted for defining the neighboring relationship. Regarding temporal scale (the 12 months), the neighbors of a given month are the previous and next months. As for spatial scale, we chose a neighboring relationship of order 8, i.e. two rectangles are neighbors if they have one side or one corner in common. To summarize, the neighbors of a given rectangle-month are: (i) its spatial neighbors in the same month; and (ii) the same rectangle and its spatial neighbors in the next and previous months. In the case of scientific data, only spatial contiguities could be considered since there is 1 month of data per year. In preliminary analyses, the small number of data and their relatively high variability were found to lead to a large number of clusters (six or seven), most of them being composed of only one rectangle. Therefore, only results from a simple PCA will be presented for scientific data.

3.6. Analysis of year effects

In order to ensure ourselves that results were not distorted by year-specific effects, we also carried out typologies from

CPUE that are not averaged over years. These analyses were only conducted for commercial CPUE that form the main piece of information in this study. Several techniques may be used in this purpose: (i) multitable analyses of a series of separate data tables per year, e.g. multiple factorial analyses (MFA) or the STATIS technique (see e.g. Gaertner et al., 1998, 1999 for applications of these techniques to marine ecology); or (ii) conventional PCA of a single data table including all CPUE per rectangle-month for each year. Unfortunately, MFA does not allow for possible missing values and could only be applied to rectangle-months for which a CPUE is available each year from 1993 to 1997, which amounts to only 182 rectangle-months over a total of 502 in the averaged CPUE (i.e. it accounts for only one-third of the information). In addition, many missing values are located in the central Celtic Sea; we thus considered that this analysis would not be interesting for delineating the spatial distribution of interest. The STATIS technique handles missing values, but HAC could not be performed on axes resulting from STATIS in the available softwares. As the analysis of year-to-year variations was not central to our study, we resorted to (ii) and conducted a PCA on the data table including all non-averaged CPUE, followed by a HAC. The importance of year-specific effects was evaluated by considering the year variable as an illustrative variable in the typology, and by analyzing the influence of each particular year on cluster composition and characterization.

Contiguity analyses were performed with ADE-4 (Thiou-louse et al., 1997), freely available at http://pbil.univ-lyon1.fr/ADE-4/ADE-4F.html. Classical multivariate analyses were carried out with SPAD.N (CISIA, 1996).

4. Results

4.1. Scientific CPUE

Simple normalized PCAs were performed on autumn data (46 individuals (ICES statistical rectangles) and six variables (mean CPUE over years for age groups 0 to 5+)) and on spring data (47 rectangles and five variables (mean CPUE over years for age groups 1 to 5+, no age group 0 being observed in spring)). Analyses were also carried out on yearly data tables to check for year effects. Since results were similar to those obtained from pooled data, they were not reported here.

HAC on principal axes leads to a first partition into two well-separated clusters: a set of rectangles where whiting abundance is low for all age groups, and a smaller group of rectangles in the northern Celtic Sea with higher abundances for all age groups. In finer partitions, this latter cluster is further subdivided into age-specific zones comprising few rectangles. For both seasons, retained partitions correspond to four clusters (Table 3) which explain most of the variability of each data set (91% of total variance in autumn and 96% in spring). The first cluster corresponds to rectangles with relatively low abundances for all age groups (see mean CPUE per cluster in Table 3).

In spring, the second cluster is characteristic of very high abundances of ages 2 and 3 and to a lesser extent 4, whereas cluster 3 corresponds to a single rectangle with peak abundances (9–15 times as high as on average over all rectangles) of age groups 2 to 5+. A closer look at data shows that these high values were observed in 1998 and 1999. Cluster 4 is characterized by peak abundances for all ages (same order of magnitude as in cluster 3). Regarding autumn, cluster 2 comprises rectangles with very high abundances of age groups 0 and 1 (on average sevenfold). Rectangles of the third cluster exhibit very high abundances of ages 4 and 3. Cluster 4 is made of one rectangle with peak abundances for all age groups including ages 0 and 1.

The spatial distribution of the rectangles in each cluster (Fig. 2) shows that in both seasons, high abundances of whiting are only found in the northern Celtic Sea and in Cornwall. In the rest of the Celtic Sea, whiting is rare. Differences between spring and autumn illustrate the seasonality of the spatial distribution. In spring, extremely high abundances of age groups 2 to 5+ are found in the northern Celtic Sea (in rectangles 30E4, 32E2 and 32E3), in relation with reproduction. High abundances of young adults are observed in rectangle 32E4 and to a lesser extent in rectangle 29E5. Age group 1 is very abundant in rectangles 32E2 and 32E3. In autumn, whiting appears more widespread distributed, although always in the northern Celtic Sea. Age groups 0 and 1 are concentrated on and around the Smalls and eastwards in the Bristol Channel. All adults are very abundant in rectangle 31E3, and ages 3 and 4 are found in rectangles surrounding the Smalls.

Previous results were compared to simple maps of CPUE. CPUE were represented for age groups 0 to 3+ in autumn (Fig. 3a), and for age groups 1 to 3+ in spring (Fig. 3b). HAC results synthesize well the information contained in Fig. 3.

4.2. Commercial data

We first present results based on CPUE averaged over years, and then we use results based on CPUE per year to evaluate how year-specific effects may influence the patterns revealed by the analysis of averaged CPUE.

4.2.1. Characterizing seasonal distributions of population abundance

This was achieved by a typology of rectangle-months on the basis of their average CPUE over years (502 individuals (the combinations of rectangle and month for which a CPUE is available) described by five active CPUE variables (one per age group 1 to 5+) and two illustrative variables (months and rectangles)). As explained earlier, the typology relies on contiguity analyses followed by an HAC. Contiguity analyses provide local, total and global results.

In all three analyses, the HAC leads to partitions in four clusters (Table 4) which explain 45% of the variance for total

Table 3

Characterization of clusters obtained from HAC of scientific data. Cluster size is the number of rectangles in the cluster. A variable is characteristic if associated test value is significant (in italics). Variables are ranked by decreasing test-value. Rectangle location may be identified in Fig. 2

Cluster (cluster size)	Age (years)	Test value	CPUE (kg h^{-1}) mean (standard error)	Rectangles in the cluster
Autumn				
1 Few whitings (38)	*2*	*–3.5*	0.2 (0.6)	32E1 31D8 31D9 31E0 31E1 30D8 30D9
	5+	*–3.5*	0.9 (3.0)	30E0 30E1 30E2 30E4 29D9 29E0 29E1
	0	*–4.1*	1.1 (3.9)	29E2 29E3 28D8 28D9 28E0 28E1 28E2
	1	*–4.3*	0.6 (3.1)	27D8 27D9 27E0 27E1 27E2 27E3 26D9
	3	*–5.0*	0.3 (1.0)	26E0 26E1 26E2 26E3 26E4 25E0 25E1
	4	*–5.2*	0.7 (1.5)	25E2 25E3 25E4
2 Ages 0 + 1 (4)	*1*	*3.4*	71 (20)	31E2 31E4 32E3 31E5
	0	*3.0*	66 (17)	
	5+	1.1	6 (5)	
	2	1.1	19 (8)	
	4	1.0	11 (3)	
	3	0.8	9 (1.5)	
3 Ages 4 + 3 (3)	*4*	*5.7*	44 (15)	32E4 30E3 32E2
	3	*4.0*	28 (12)	
	5+	0.9	6 (3.5)	
	2	0.2	9 (10)	
	1	–0.5	1.5 (1.2)	
	0	–0.5	0.6 (0.8)	
4 All ages (1)	0	No significant	223 (–)	31E3
	1	values	214 (–)	
	2		160 (–)	
	3		54 (–)	
	4		29 (–)	
	5+		37 (–)	
Spring				
1 Few whitings (42)	*1*	*–5.1*	0.1 (0.4)	32D9 32D8 31D8 31D9 31E0 31E1 31E2
	5+	*–5.5*	1.7 (5.5)	31E3 30D9 30E1 30E0 30E3 29D8 29D9
	4	*–6.1*	2.1 (5.7)	29E0 29E1 29E2 29E3 29E6 28D8 28D9
	2	*–6.4*	0.1 (0.4)	28E0 28E1 28E2 28E3 28E4 28E5 28E6
	3	*–6.4*	1.3 (3.7)	27E0 27D9 27E2 27E1 27E3 27E4 26E0
2 Ages 2 + 3 (2)	*2*	*2.4*	130 (49)	26E1 26E2 26E3 25E0 25E1 25E2 25E3
	3	*2.2*	120 (19)	29E5 32E4
	4	1.6	92 (8)	
	1	1.3	58 (49)	
	5+	0.7	37 (4.4)	
3 Ages 2 to 5+ (1)	1	No significant	25 (–)	30E4
	2	values	191 (–)	
	3		267 (–)	
	4		262 (–)	
	5+		213 (–)	
4 Ages 1 to 5+ (1)	*1*	*6.4*	237 (15)	32E2 32E3
	2	*5.4*	262 (16)	
	4	*5.0*	246 (47)	
	3	*4.8*	236 (33)	
	5+	*4.6*	157 (23)	

and local analyses and 50% for global analysis. It is less than for scientific data because the data set is much larger and more complex, including e.g. temporal variability. Cluster compositions are rather similar in the three analyses, except for April, May and June (Fig. 4). Characterization of clusters by CPUE variables and by variables month and rectangle share common features among the three analyses, but also show slight differences. In the rectangles of the largest cluster (cluster 1), whiting abundance is consistently low for all ages throughout the year. Seasonal patterns of distributions are specific to age groups. In cluster 2, age groups 4, 5+ and to a lesser extent 3 are abundant during the first quarter. In contrast, cluster 3 is characterized by high abundances for age groups 2, 3 and to a lesser extent 4 during the second half
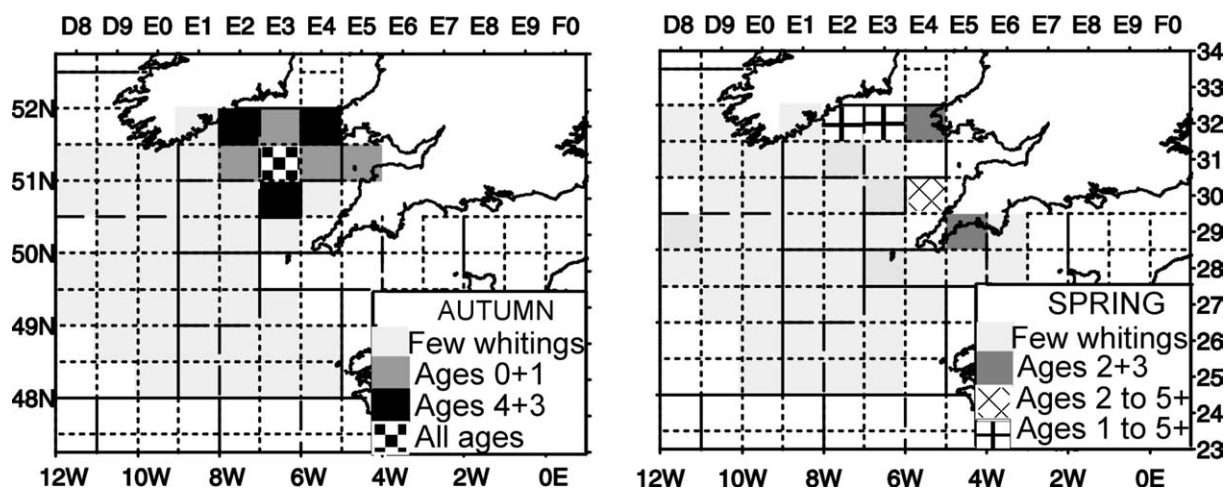
Fig. 2. Spatial distribution of clusters for autumn (left) and spring surveys (right). Shading of rectangle relates to the cluster, not to whiting abundance. Grey levels do not match between left and right, except for cluster 1, which has similar meaning in both cases.

of the year. The smallest cluster (cluster 4) corresponds to concentrations of immature individuals (age 1) from July to September.

The similarities between cluster compositions obtained from total, local and global analyses, show that most spatial and seasonal patterns are consistent at local and global scales. In this respect, we only display maps pertaining to total analysis (Fig. 4a), and maps from April to June for global analysis to illustrate the slight differences found with respect to total analysis (Fig. 4b). Maps indicate four main seasons for whiting distribution. The first, from January to April, where high abundances of adults are found in the northern Celtic Sea and in the North Bishop area in relation to spawning. At this time of the year, adults are mainly found south of Ireland (VIIg3) and in the Bishop area (29E5 and VIIf1). In particular, mature whiting is consistently abundant in rectangles 29E4, 30E4 and 30E5 (Bishop area), and in rectangles 31E2, 32E2, 31E1, 32E3 (south of Ireland). Maps also show a northward shift of distribution between January–February and March–April. Young adults are found in January in the Bristol Channel (rectangle 31E5) due to a peak CPUE (71 440 kg h[1]) of age 2. These individuals may still be immature, but this is impossible to check with available data. Young adults (mostly age 3) are also found in April and May in the Bishop area and south of Ireland, but in total analysis and not in global analysis, indicating that in these areas, abundance of these age groups is quite variable between neighboring rectangle-months (i.e. locally variable). Note also that patterns found in April and May are influenced by data from March and June because of contiguities.

A second season corresponding to May and June is characterized by high abundances of old age groups in the Cornwall area (29E5), associated with young adults in May (see above paragraph). June pattern is intermediate between those of May and July, with increasing abundances of young adults in the northern Celtic Sea. From July to September, high abundances of age 1 are found in the Smalls area (specifically in rectangles 31E4 and 32E3), in relation with recruitment to

the fishery, and young adults (ages 2 and 3) are abundant in the Smalls area. Old matures are mainly observed in rectangle 31E3 (Smalls area) and south of Ireland in August and September.

From October to December, young adults are mostly found in south Ireland and to a lesser extent in the Smalls area. The abundance of old adults remains low, except in December in south Ireland, probably anticipating the forthcoming spawning season (see January pattern on Fig. 4a).

For comparison with typology results, CPUE maps were shown for age groups 1 to 3+ for commercial data (Figs. 5–7). Again, cluster results synthesize well-observed CPUE.

4.2.2. Year-specific patterns

The analysis of year-specific patterns was achieved through the typology of non-averaged commercial CPUE, i.e. 1763 individuals (the combinations of rectangle–month–year for which a CPUE is available) described by five CPUE variables (for each age group 1 to 5+). Month, rectangle, ICES subdivision and year were used as illustrative variables. The first two principal axes explain 82% of the variance of the data. Projections of CPUE variables and months (Fig. 8) are similar to those obtained from the total analysis of averaged data (not reported). On the first axis, areas with high abundances (VIIg3, VIIg2 and VIIf1) are opposed to others. On the second axis, areas with large abundances of adults (VIIf1) in the first quarter are opposed to areas with high abundances of young individuals (ages 1–3) in the second half-year. As for year effects, projection of the year variable indicates that years 1993 and 1994 are opposed to years 1996 and 1997 on the second axis. The contrast between 1993/1994 and 1996/1997 in abundances of ages 1 and 2 vs. 4 and 5+ also shows on summary statistics (Table 1).

Examination of HAC results yields more insight on possible consequences of year effects on the typology of Table 4. The retained partition comprises three clusters and explains 53% of the variance of the data. These results may be compared with those obtained from the total analysis of averaged
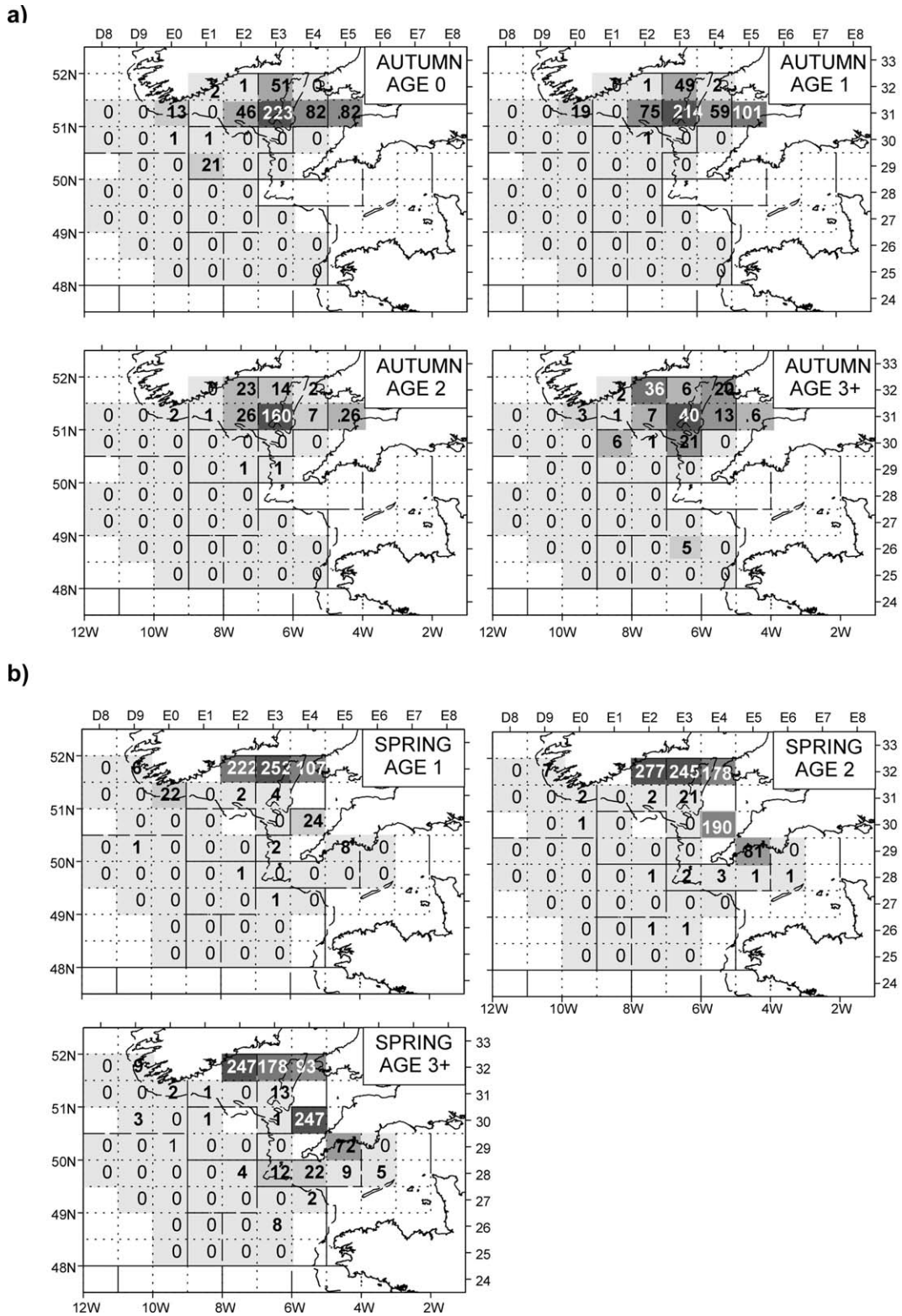
Fig. 3. Spatial distribution of scientific CPUE for age 0 to 3+ in autumn (a) and for ages 1 to 3+ in spring (b). CPUE values in kg h^{-1} are reported for each rectangle.

CPUE data (Table 4). Results obtained from non-averaged CPUE (Table 5) indicate a first cluster (1NA) corresponding to relatively low abundances for all ages. In the second cluster (2NA) are found high abundances of ages 3 to 5+, and to a lesser extent high abundances of age 2. In contrast, the third cluster (3NA) is characterized by very high abundances of age groups 1–3 and by relatively high abundances of ages 4. The first two clusters have similar interpretations in both

Table 4

Characterization of clusters obtained from HAC of averaged commercial CPUE. Cluster size is the number of rectangles in the cluster. A variable is characteristic if associated test value is significant (in italics). Variables are ranked by decreasing test value of variables. The last column indicates how rectangle-months in the cluster are distributed over months

Cluster (cluster size)	Age years	Test value	CPUE, kg h⁻¹ mean (standard error)		Percentages of cluster per month
Total					
T1 Few whitings (450)	1	-6	2	(23)	
	2	-14	1 581	(3 200)	
	5+	-14	2 755	(3 338)	
	3	-16	3 949	(5 321)	
	4	-17	4 355	(4 811)	
T2 Old adults (29)	4	18	51 778	(19 484)	
	5+	17	28 026	(13 955)	
	3	10	37 420	(32 051)	
	2	2	6 644	(6 546)	
	1	-0.04	9	(44)	
T3 Young adults (20)	2	19	40 504	(17 038)	
	3	12	49 788	(30 376)	
	4	4.3	20 749	(17 045)	
	5+	1.6	7 038	(7 171)	
	1	1.1	30	(72)	
T4 Age 1 (3)	1	20	981	(211)	
	3	2.4	29 927	(7 643)	
	4	2.1	24 079	(11 040)	
	2	0.9	8 147	(3 981)	
	5+	0.6	7 144	(3 315)	
Global					
G1 Few whitings (458)	1	-5.6	3	(27)	
	2	-13	1 775	(3 805)	
	5+	-14	2 945	(4 001)	
	3	-16	4 62	(6 223)	
	4	-17	4 666	(5 505)	
G2 Old adults (29)	4	19	54 201	(18 148)	
	5+	16	26 944	(14 094)	
	3	15	51 922	(34 337)	
	2	5.5	12 355	(13 954)	
	1	-0.04	9	(44)	
G3 Young adults (12)	2	16	45 253	(18 605)	
	3	5.5	33 352	(23 114)	
	4	1.6	13 987	(9 955)	
	5+	0.4	5 223	(5 963)	
	1	0.1	12	(29)	
G4 Age 1 (3)	1	20	981	(211)	
	3	2.4	29 927	(7 643)	
	4	2.1	24 079	(11 040)	
	2	0.9	8 147	(3 981)	
	5+	0.6	7 144	(3 315)	

typologies. Hence, clusters 1 and 2 are described by the same ICES subdivisions (Table 5), whereas significant subdivisions for cluster 3NA correspond to those of clusters T3 and T4. Characterization of clusters by rectangles appears rather similar except for clusters 3NA vs. T3 and T4. The influence of rectangle 31E3 on cluster 3NA relates to the peak abundance of age 1 in 1993 (Table 1). As for months, February, March and April characterize clusters 2NA and T2, while clusters 3NA, T3 and T4 correspond to the second half-year. In finer partitions (not reported), the peak recruitment of age 1 in 1993 determines an additional cluster. Although results for age group 1 need be taken cautiously, because of potential variations in discard rates (see zero CPUE in 1997), these patterns may be linked to variations in recruitment rate and/or timing (note that the year effects evidenced here operate within ICES subdivisions VIIg2 and VIIg3). Results
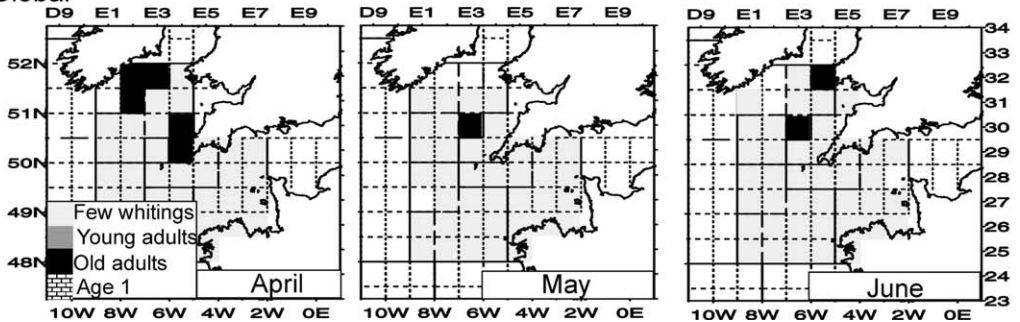
Fig. 4. Spatial distribution of the clusters resulting from the typology of commercial CPUE: total analysis (a), and global analysis (b). Shading of rectangle relates to the cluster, not to whiting abundance. Grey levels have same meaning in (a) and (b) (see legend).

also confirm that there were more rectangles with high abundances in year 1993 (clusters 2NA and 3NA). However, cluster composition is on the whole little affected by the averaging of CPUE over years. We can thus conclude that although several year-specific patterns may be evidenced from the analysis, they do not bias the recurrent spatial and

Fig. 5. Spatial distribution of commercial CPUE for age group 1 for the second half-year. CPUE are always zero in the first half-year. CPUE values in kg h^{-1} are reported for each rectangle.

seasonal patterns displayed by the analysis of averaged CPUE.

5. Discussion

The objective of this paper is to determine age-specific spatial and seasonal patterns, i.e. to delineate zones and periods of the year (seasons) during which abundances per age group are similar. Three kinds of information were available in this purpose: prior knowledge from the literature, scientific survey CPUE and commercial CPUE. We also contemplated using data from a discard study (Cotter et al., 1999), but these data were inappropriate for spatial considerations.

Rather than combining data from different sources, we preferred to analyze them separately to (i) examine whether or not they yield distinct pictures of the spatial distribution, and (ii) stress and contrast their respective information contents, advantages and shortcomings. The same approach was used for scientific and commercial data in order to have comparable results.

5.1. Spatial and seasonal distribution of whiting

Monthly maps obtained from commercial CPUE analyses are relatively consistent with prior knowledge quoted in Section 2. Descriptions from the literature are sometimes contradictory and remain rather coarse. The two main spawning grounds are clearly delineated, as well as the recruitment area, respectively, associated with months of January–April, and July–December. On the continental shelf south of 50°N, whiting abundance is low throughout the year. Since high abundances are less frequent in the northern Celtic Sea dur-

ing the second half-year (Fig. 4), we hypothesize that adults may then disperse over the southern Celtic Sea and the western English Channel. However, this hypothesis remains to be tested since other processes like changes in vulnerability to gear may also potentially explain these patterns. June and December appear as intermediate periods. In June, part of the young adults are not located in the same areas as old adults. This presence on the northern spawning grounds may be related to a first reproduction occurring late in the season. The December pattern may be explained by a progressive arrival of old adults on spawning grounds. However, available data do not allow to check for these hypotheses.

The analysis of scientific CPUE only provides insight on the spatial distribution in October–November and in March, but these data bring information on the abundance of age group 0, these fish being either discarded or not caught by fishermen. Young of the year are not found in spring CPUE because they are not yet vulnerable to the scientific trawl. In autumn, they are located (like age group 1) on the Smalls and in the Bristol Channel.

Concerning age group 1, it is found in commercial landings from July to September in some years and from October to December in other years. This may be explained by lags in recruitment timing, but also by fluctuating discard rates or environmental effects affecting distribution. Whiting recruitment was relatively high in the early 1990s and low in the late 1990s (ICES, 1999). In spring, age 1 is in the northern Celtic Sea like all other age groups. For age group 1, scientific CPUE appear more reliable for delineating spatial patterns, but commercial CPUE are helpful in determining seasonal variations.

With regard to adults, both scientific and commercial maps clearly show that adults are concentrated south of Ireland (and in rectangle 32E4) and in North Bishop from January to April, in relation to reproduction. Consistently
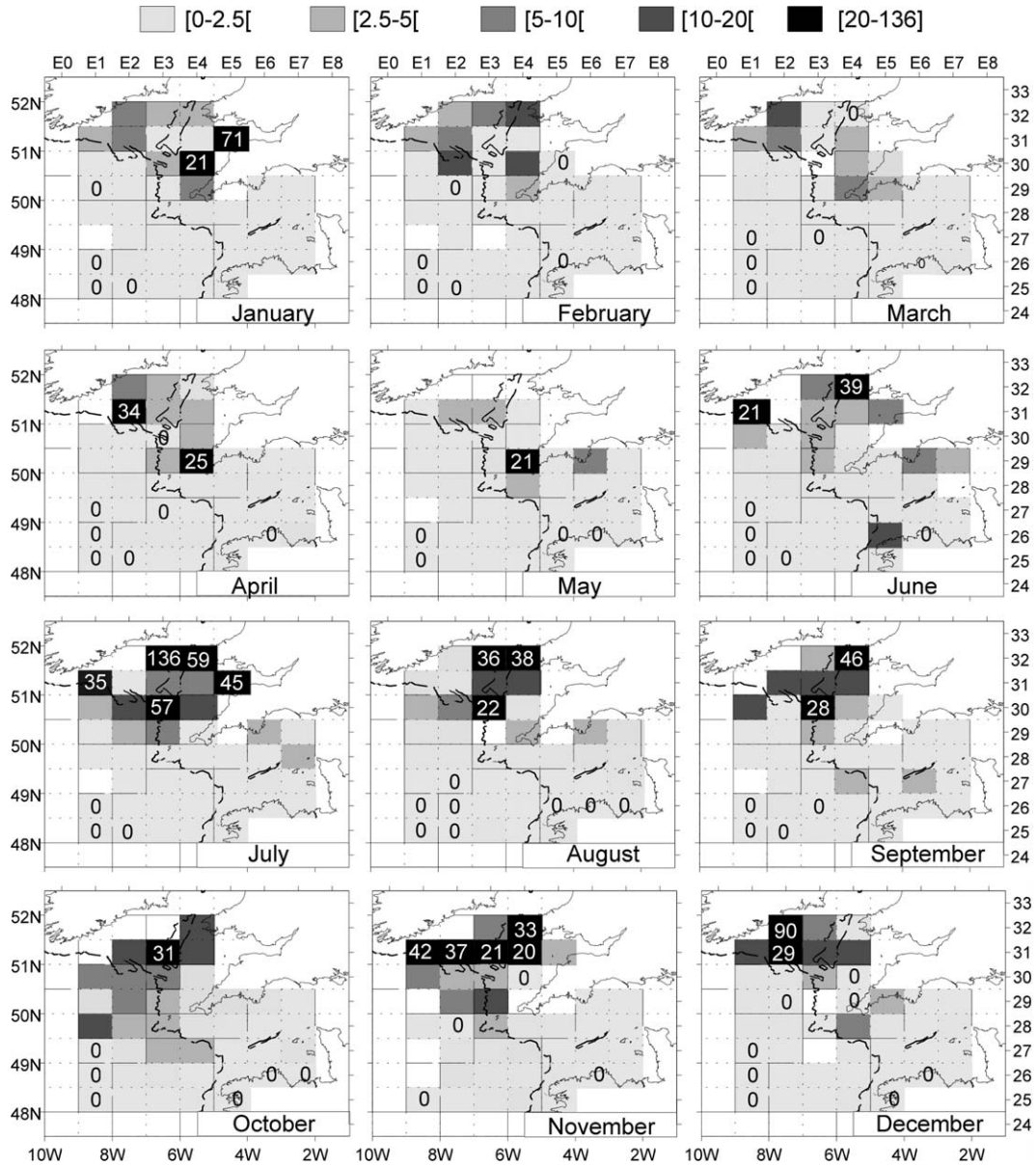
Fig. 6. Spatial distribution of commercial CPUE for each month, for age group 2. CPUE values in kg h^{-1} are reported for each rectangle.

with commercial CPUE, adults are less abundant in autumn than in spring. Spatial patterns for adults are better delineated from commercial CPUE than from scientific CPUE.

Our findings are consistent with prior knowledge, except that, contrary to the literature which only mentions a winter concentration in the northern Celtic Sea, we find a concentration of most age groups in this area in autumn. Present results yield a more precise picture of spatial and seasonal abundance patterns than prior knowledge, since patterns are determined at the scale of ICES rectangles and months. Based on data analysis, these results provide a rationale for determining zones and seasons for whiting distribution. However, the zones obtained are not as compact as we would have expected, in particular in the purpose of constructing a spatial and seasonal model of population dynamics. Rectangles of a given cluster are sometimes scattered in two or

three locations. In addition, 2 months (June and December) apparently correspond to transitional distributions. So far, we have not developed statistical criteria to determine the most appropriate zones (groups of rectangles) and seasons (groups of months) from these monthly maps. These analyses lead to monthly snapshots of spatial distribution. As such, it allows us to build hypotheses on possible migration and dispersion features, but it does not enable us to test them. Tagging experiments are needed to study these questions which fall beyond the scope of this paper. In order to further improve determination of the boundaries of the spatial and seasonal structures, it might be interesting to evidence habitat preferences by studying relationships between abundance and environmental variables. For instance, Zheng et al. (2001) studied the spatial and temporal distribution of the North Sea whiting in relation with depth and temperature.
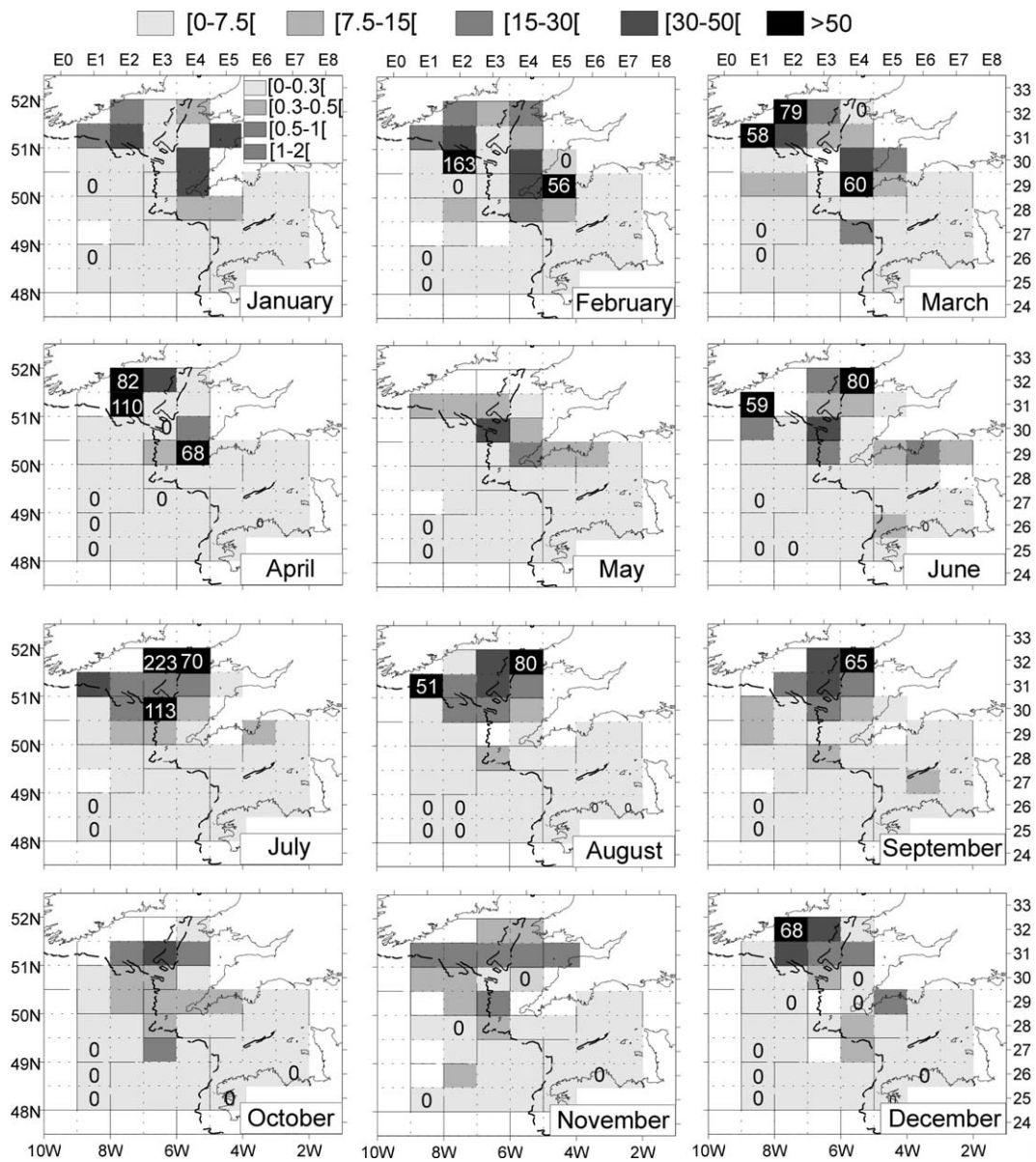
Fig. 7. Spatial distribution of commercial CPUE for each month, for age groups 3+. CPUE values in kg h^{-1} are reported for each rectangle.
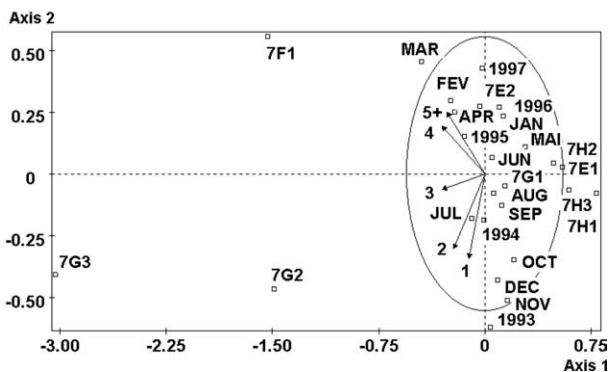
5.2. Methodology

Fig. 8. Projection of CPUE variables (identified by 1 to 5+) on the first two axes of the PCA of commercial CPUE per year. Illustrative variables (month, year, and ICES subdivisions) are also projected.

For commercial data, the typology accounts for spatial contiguities between rectangles and temporal contiguities between months. The contiguity analysis comprises three complementary PCA: local, global and total. Local analysis is useful for identifying local variations, whereas global analysis emphasizes smooth large-scale patterns. Total analysis does not consider any scale in particular, but may be compared to conventional analyses that do not account for contiguities. Contiguity analyses are thus helpful to analyze the fluctuations of variables of interest at several spatial and temporal scales. In the case of whiting abundance, the three analyses yield similar results, indicating that the structures found were consistent across scales. But, a few features emerged from local and total analyses, and not from global analysis. They pertain to sharp variations in abundance be-

Table 5

Characterization of clusters obtained from HAC of non-averaged commercial CPUE per year (partition in three clusters). HAC is performed after a conventional PCA, i.e. not accounting for contiguities. Cluster size is the number of rectangles in the cluster. A variable is characteristic if associated test value is significant (in italics). Variables are ranked by decreasing test value. Only months, rectangles, and ICES subdivision (ICES Sub.) with positive test-values are reported. "Year+" (respectively, "Year–") means positive (respectively, negative) test values for year

Cluster (cluster size)	Age (years)	Test value	CPUE mean (standard error)	Year (–)	Year (+)	Month	ICES Sub.	Rectangles
1NA, few whitings (1655) 93.9%	*1*	*–14*	35 (206)	1997	1996	5	*7E1*	
	2	*–17*	1 519 (4 463)	1995	1993	6	*7H2*	
	5+	*–27*	2 446 (4 069)		1994	10	*7H3*	No significant
	3	*–28*	3 685 (6 724)			12	*7H1*	One
	4	*–32*	3 951 (5 997)			11	7E2, 7G1	Not reported
						7		
						1		
2NA, old adults (90) 5.1%	*4*	*34*	59 622 (29 121)	1997	1996	3	*7G3*	
	5+	*30*	29 957 (19 784)	1995	*1993*	2	*7F1*	*29E4*
	3	*25*	51 218 (41 221)	1994		4	*7G2*	*31E3*
	2	10	12 256 (16 844)					
	1	1	202 (788)					
3NA, young adults and age 1 (18) 1.0%	*1*	*32*	6 014 (4452)	1994	*1993*	11	*7G2*	
	2	*26*	59 953 (40 937)	1997		12	*7G3*	*31E3*
	3	*13*	60 234 (62 328)	1996		9		
	4	4	16 312 (18 205)			10		
	5+	0.6	5 037 (10 140)			7, 8		

tween neighboring rectangle-months for age group 1, and to a lesser extent to the segregation of old and young age groups.

This study rests on several hypotheses that need to be discussed. First, CPUE were averaged over years. This may be justified by the fact that we were not interested in analyzing inter-annual variability of distribution, but rather in characterizing mean patterns, which should be recurrent over the few years studied, since they are tied to life cycle. Distribution shifts in space and time may still occur due to environmental influences. A very strong year class entered the fishery in 1993 and caused a substantial stock increase over the years studied (e.g. in 1995) (ICES, 2000). We checked whether analyses were not too sensitive to particular years, which was not the case except for the abundance of age group 1. In addition, averaging over years was desirable for commercial data that do not cover all rectangles each month in a given year.

Second, we had to consider single rectangle trips in order to be able to assign catch to a rectangle. As mentioned earlier, it is not necessary that these trips are a representative sample of the whole set of commercial trips. They represent a substantial percentage of all trips and provide a good coverage of the whole Celtic Shelf. The same approach was used in Pelletier and Ferraris (2000).

Last, we consider that all trips can be averaged regardless of the metier practiced, the type of vessels and other technical and fishermen-related characteristics of fishing effort. Instead, we could have picked fishing trips targeting whiting. It could be interesting to do this if the objective was to estimate abundance level per patch and season. But, we are rather

aiming at depicting spatio-temporal patterns over the area of distribution, including places where abundance may be low, and we thus need to have the best spatial and seasonal coverage of the area. Regarding individual fishing power and technical characteristics, we assumed these effects were smoothed out by averaging over trips. This approach was also proposed by Hilborn (1985) and Pelletier and Ferraris (2000). We are aware that in places where CPUE are low, they may underestimate whiting abundance relatively to places where they are high, because the former are presumably due to non-targeted fishing effort. Again, this may pose a problem when one wants to estimate average abundances per patch and per season, but not so much to delineate patches and seasons, which is the main objective of this paper. When it comes to abundance estimation, fishing efficiencies per species need to be taken into account. This is also why we do not compare scientific and commercial CPUE values, nor do we combine them in the analysis.

Multivariate descriptive methods were implemented to analyze both commercial and scientific data, because the spatial resolution of commercial data is limited to ICES statistical rectangle. Other methodologies could be considered if information were available at the scale of the fishing operation (haul). For instance, Vignaux (1996) studied the spatial distribution of hoki from precisely georeferenced logbook data. In the case of scientific data, this information is lost by averaging at the rectangle scale. If data were more numerous and from several surveys per year, it would be worthwhile for our purpose to retain geographical coordinates and try to formally combine scientific and commercial information, e.g. Pelletier et al. (1993).

5.3. Data collection

This analysis raises several points related to data collection and the information used for estimating fish abundance. First, we assumed landings were proportional to catch. However, discards of whiting are substantial in this fishery for age groups 0 and 1. The length at which 50% of the fish were discarded (DL_{50}) is 32.6 cm on average, which is larger than the minimum legal landing size. Above 40–45 cm, whiting are generally never discarded. A discard study (Rochet et al., 2002) shows that in the Celtic Sea almost all whiting smaller than 32 cm were discarded in 1997, which explains the zero CPUE of age 1 that year. However, discards are highly variable between years, and CPUE of age 1 is not zero in other years. According to assessment working groups (ICES, 2000), the apparent recent decline in abundance of age 1 might be a consequence of both increased French landings in UK and Ireland, and a higher discarding rate of small fish. Accounting for discards in the analysis would certainly improve the results, but the spatial and temporal coverage of the available discard data were unfortunately not sufficient to take this information into account here. Results from scientific CPUE are here helpful in evaluating the consequences of discards on commercial CPUE results. As for potential misreporting problems for commercial catch, misreporting was deemed unlikely by assessment groups because the TAC of landings is well in excess (ICES, 2000), and commercial data for whiting are considered to be relatively accurate from this standpoint. Third, the resolution of commercial data is not optimal for spatial analyses. Effort is only known at the scale of ICES rectangles, and catch is only known at the trip scale. This shortcoming strongly constrained our methodological choices. Regarding effort, the accuracy of the allocation of trawling time to rectangles is not known. For instance, vessel trajectories during towing likely overlap two rectangles. Starting and ending positions of tows would certainly enable to estimate much more accurate CPUE. The availability of data from vessel monitoring programs would be of great value in this purpose. Regarding catch, the allocation of catch between rectangles is unknown in about 60% of the trips. This information is lacking in data collection systems of many fisheries. It is indeed problematic for studying spatial distributions of exploited populations, but also for characterizing fishing activity (Pelletier and Ferraris, 2000). Still, the sample size and the spatio-temporal coverage of commercial data obtained from compulsory logbook programs are unique, in this case approximately 2000 trips of 12 days per year for commercial data, vs. 50–80 hauls of 30 or 60 min for scientific data. Although there may be considerable reliability problems, these data are very informative about spatial and seasonal patterns.

Concerning scientific data, the fact that data arise from two distinct surveys is likely to result in differences that may affect comparisons of autumn and spring data. First, hauling time is half an hour for the autumn survey and 1 h for the spring survey. It is generally estimated that haul duration may affect size composition of catches in that fish endurance increases with size. The proportion of large individuals in the catch then increases with haul duration (Wardle, 1993). But it seems that the influence of haul duration on catch composition may only be observed at relatively high temperatures (Engås and Godø, 1989). Therefore, we do not believe that such effects may influence the kind of analysis presented here. Second, fishing gears differ slightly in dimensions (see Section 3.1), which may affect selectivity and efficiency. During the analysis, we observed differences in age composition and overall biomass between the CEFAS and in the IFREMER surveys. The autumn survey has lower overall biomass and a younger age composition than the spring survey. We cannot assert that these differences are fully due to seasonal effects, but given the above remarks, we believe that seasonal effects, which are the scope of this paper, are largely predominant in explaining these differences.

6. Conclusion

Our study shows that, to a certain extent, large-scale spatial and temporal patterns of whiting distribution can be delineated from commercial and scientific data, if both data are used in a complementary way. The approach proposed is applicable to any exploited population for which scientific and commercial catch and effort data are available. To ascertain the spatial and temporal distributions obtained, it would be interesting to analyze abundance patterns in relation to environmental variables.

Determining patches and seasons is not a final objective. It may be helpful in several respects. First, it provides a rationale based on statistical data analysis for building hypotheses for a spatio-temporal model of the population dynamics. Such a model is being developed for the Celtic Sea whiting, and will be incorporated in a simulation tool to test management measures which take into account the dynamics of other important species in the fishery (Pelletier et al., 2001). Second, results may be interpreted together with results from tagging experiments in the view of estimating migration coefficients between patches. They may also help designing such experiments. Lastly, patches and seasons defined here may form a basis for defining abundance estimates that are explicitly spatial and seasonal, e.g. from targeted commercial trip data, scientific catch data, or acoustic data.

Table 6
Matrices used in each analysis (simple PCA, total analysis, local analysis and global analysis). $\mathbf{I}_p$ is the $(p \times p)$ identity matrix, $\mathbf{X}_c$ is the $(n \times p)$ standardized data matrix (centered or normalized), $\mathbf{D}$ is the diagonal matrix of neighbor weights (i.e. D_{ii} is the number of neighbors of point i divided by total number of neighbor pairs), $\mathbf{X}_D$ is the D-standardized data matrix. $\mathbf{P}$ (contingence table) is the matrix of the graph divided by the number of neighbor pairs

Simple PCA	Total analysis	Local analysis	Global analysis
$\mathbf{X}_c$	$\mathbf{X}_D$	$\mathbf{X}_D$	$\mathbf{X}_D$
$\mathbf{I}_p$	$\mathbf{I}_p$	$\mathbf{I}_p$	$\mathbf{I}_p$
$1/n\mathbf{I}_n$	$\mathbf{D}$	$\mathbf{D} - \mathbf{P}$	$\mathbf{P}$

Appendix A. Description of contiguity analysis

The multivariate contiguity analysis method requires defining neighboring relationships between individuals. These relationships are derived from both spatial and temporal locations of individuals. The neighboring relationship is represented by a symmetric matrix $\mathbf{A}$ of order n of element m_{ij} such as $m_{ij} = 1$ if two individuals i and j are neighbors, and zero otherwise. A matrix of neighboring weights $\mathbf{P}$ is deduced from $\mathbf{A}$ by $p_{ij} = m_{ij}/(2m)$, where m is the total number of pairs of neighbors, therefore, $\sum_{ij} p_{ij} = 1$.

Based on the neighboring relationship, the total variance of the data table is decomposed into a first component related to neighbor individuals (called local variance) and a second one related to individuals that are not neighbors (called global covariance as it is not always positive). The main structure of the data table is hence described at two spatial scales, local and global (Thioulouse et al., 1995).

Just like simple PCA searches the linear combinations (LC) of variables with maximum variance, local PCA searches the LC of the initial variables which maximizes local variance. This is similar to search the LC of variables that varies most between neighbors. So local analysis will emphasize differences between neighboring points, like sharp variations (Le Foll, 1982). When local variance is close to zero, autocorrelation is strong at the scale of the spatial and temporal neighborhood. In contrast, global analysis searches the LC of initial variables which maximizes global covariance. As such, it mostly detects variation between neighboring individuals, that are consistent at the scale of the whole data set, like a gradient or a smooth variation over space or time (Wartenberg, 1985). In addition to local and global analysis, a total analysis is also performed (Table 6). It is similar to a classical PCA in which individuals are weighted by their number of neighbors. The analyses allow to compare structures observed at global and local scales, with those observed without giving more weight to a particular scale (total analysis).

To further stress differences between these analyses, each of them may be summarized by the data table ($\mathbf{X}$) on which the PCA is performed, the weights of individuals ($\mathbf{C}$) and those of variables ($\mathbf{R}$) (Thioulouse et al., 1995, Table 6). In global analysis, $\mathbf{P}$ is not a positive definite matrix because of possible negative correlations between non-neighboring individuals, and factorial axes may not be described in terms of the fraction of variance they explain. However, associated eigenvalues indicate the relative importance of axes. The spatial correlation coefficient quantifies the strength of the spatial structure associated to a factorial axis.

References

ADE-4, software. Available online at: http://pbil.univ-lyon1.fr/ADE-4/ADE-4F.html.

Charuau, A., Biseau, A., 1989. Study of optimal management of *Nephrops* and demersal fisheries in the Celtic Sea. Final Reports of EEC Contracts DG XIV No. 3133 and 3625.

CISIA, 1996. SPAD.N version 3. CISIA, Saint-Mandé, France.

Cliff, A.D., Ord, J.K., 1973. Spatial Autocorrelation. Pion, London.

Cornillon, P.A., Amenta, P., Sabatier, R., 1997. Three-way data arrays with double neighbourhood relations as a tool to analyse a contiguity structure. Proceedings of the Biannual Meeting of the Classification Group of Società Italiana di Statistica (SIS) Pescara, 3–4 July 1997. In: Maurizo, V., Opiz, O. (Eds.), Classification and Data Analysis. Theory and Application. Springer, pp. 263–270.

Cotter, J., Allen, M., Armstrong, M., Briggs, R., Buckland, S., Clarke, E., Connolly, P., Course, G., Garrod, C., Hermàndez, D., Jolliffe, R., Kay, B., Kilpatrick, D., McCraken, M., Perada, P., Pérez, N., Péronnet, I., Prince, P., Reeves, S., Rolfe, M., Trujillo, V., Wheatley, S., Williams, G., 1999. On-board sampling of fish landed and discarded by commercial vessels. Final Report EC Contract 95/094.

Engås, A., Godø, O.R., 1989. The effect of different sweep lengths on the length composition of trawl catches. J. Cons. Int. Explor. Mer. 45, 263–268.

Gaertner, J.C., 1997. Organisation des assemblages démersaux dans le golfe du Lion: structures spatiales et stabilité temporelle. Spéc. Océanologie biologique Thèse doctorat Marseille.

Gaertner, J.C., Chessel, D., Bertrand, J., 1998. Stability of spatial structures of demersal assemblages: a multitable approach. Aquat. Living Resour. 11, 75–85.

Gaertner, J.C., Mazouni, N., Sabatier, R., Millet, B., 1999. Spatial organisation of demersal assemblages in the Gulf of Lions (north-west of the Mediterranean sea). Mar. Biol. 135, 199–208.

Geary, R.C., 1954. The contiguity ratio and statistical mapping. Incorporated Stat. 5, 115–145.

Hilborn, R., 1985. Fleet dynamics and individual variation: why some people catch more fish than others. Can. J. Fish. Aquat. Sci. 42, 2–13.

ICES, 1997. Report of the International Bottom Trawl Survey Working Group. ICES CM 1997/H:6.

ICES, 1999. Report of the ICES Advisory Committee on Fishery Management, 1999, Part 2. ICES Coop. Res. Rep. No. 236.

ICES, 2000. Report of the Working Group on the Assessment of Southern Shelf Demersal Stocks. ICES CM 2000/ACFM:4.

IFREMER, MAFF, 1993. Biogeographical identification of English Channel fish and shellfish stocks in ICES divisions VIId and VIIe. Final Report of EEC Contract No. 91/1211416/BMF.

Lebart, L., 1969. Analyse statistique de la contiguïté, vol. 28. Publication de l'Institut de Statistiques de l'Université de Paris, pp. 81–112.

Lebart, L., Morineau, A., Warwick, K.M., 1984. Multivariate descriptive statistical analysis. Correspondence Analysis and Related Techniques for Large Matrices. Wiley, New York.

Lebart, L., Morineau, A., Piron, M., 1997. Statistique exploratoire multidimentionnelle. 2ème éd. Dunod, Paris.

Le Foll, Y., 1982. Pondération des distances en analyse factorielle. Statistique et Analyse de données 7, 13–31.

Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm? Ecology 74, 1659–1673.

Moran, P.A.P., 1948. The interpretation of statistical maps. J. R. Stat. Soc. B 10, 243–251.

Nagabhushanam, A.K., 1965. On the biology of the commoner gadoids in Man waters. J. Mar. Biol. Assoc. UK 45, 615–657.

Nicholsky, G.V., 1968. The Ecology of Fishes. Academic Press, London, New York.

Ould el Kettab, M., 1993. La pêcherie de Gadidés de mer Celtique: description, analyse de l'exploitation et évaluation des stocks. Tentative de gestion par un modèle bio-économique. Thèse Doctorat Université de Bretagne occidentale (Brest), France.

Pelletier, D., Magal, P., 1996. Dynamics of a migratory population under different fishing effort allocation schemes in time and space. Can. J. Fish. Aquat. Sci. 53, 1186–1199.

Pelletier, D., Ferraris, J., 2000. A multivariate approach for defining fishing tactics from commercial catch and effort data. Can. J. Fish. Aquat. Sci. 57, 51–65.

Pelletier, D., Parma, A.M., Sullivan, P.J., 1993. Combining different sources of information in estimating abundance maps for exploited fish populations. Session on Recent developments in the quantitative analysis of fisheries data. Invited paper Symposium of the American Fisheries Society, Portland, August 1993.

Pelletier, D., Parma, A.M., 1994. Spatial distribution of pacific halibut: an application of geostatistics to longline survey data. Can. J. Fish. Aquat. Sci. 51, 1506–1518.

Pelletier, D., Mahévas, S., Poussin, B., Bayon, J., André, P., Royer, J.C., 2001. A conceptual model for evaluating the impact of spatial management measures on the dynamics of a mixed fishery. In: Kruse, G.H., Bez, N., Booth, T., Dorn, M., Hills, S., Lipcius, R., Pelletier, D., Roy, C., Smith, S., Witherell, D. (Eds.), Spatial Processes and Management of Marine Populations. University of Alaska Sea Grant, AK-SG-00-04, Fairbanks, pp. 53–66.

Petitgas, P., 1993. Geostatistics for fish stock assessments: a review and an acoustic application. ICES J. Mar. Sci. 50, 285–298.

Quéro, J.C., Vayne, J.J., 1997. Les poissons de mer des pêches françaises. Les encyclopédies du naturaliste. Delachaux et Niestlé, S.A., Lausann.

Rochet, M.J., Péronnet, I., Trenkel, V., 2002. An analysis of discards from the French trawler fleet in the Celtic Sea. ICES J. Mar. Sci. 59, 538–552.

Thioulouse, J., Chessel, D., Champely, S., 1995. Multivariate analysis of spatial pattern: a unified approach to local and global structures. Environ. Ecol. Stat. 2, 1–14.

Thioulouse, J., Chessel, D., Dolédec, S., Olivier, J.M., 1997. ADE-4: a multivariate analysis and graphical display software. Stat. Comput. 7, 75–83.

Trueblood, D.D., 1991. Spatial and temporal patterns of benthic community structure on the Savin Hill Cove mudflat. Diss.-Abst.-Int.-Pt.-B-Sci. and Eng, vol. 51. Ph.D. thesis, Massachusetts University, Boston, USA no. 10.

Vignaux, M., 1996. Analysis of spatial structure in fish distribution using commercial catch and effort data from the New Zealand hoki fishery. Can. J. Fish. Aquat. Sci. 53, 963–973.

Wardle, C.S., 1993. Fish behaviour in fishing gear. second ed. In: Pitcher, T.J. (Ed.), Behaviour of Teleost Fishes. Chapman and Hall, pp. 609–643.

Wartenberg, D., 1985. Multivariate spatial correlation: a method for exploratory geographical analysis. Geograph. Anal. 17, 263–283.

Zheng, X., Pierce, G.J., Reid, D.G., 2001. Spatial patterns of whiting fishery abundance in Scottish waters and relationships with environmental variables. Fish. Res. 50, 259–270.