

---

## Inter-annual variability of North Sea plaice spawning habitat

C. Loots<sup>a,1</sup>, S. Vaz<sup>a,\*</sup>, Corresponding Author Contact Information, E-mail The Corresponding Author,  
P. Koubbi<sup>b,c</sup>, B. Planque<sup>d,2</sup>, F. Coppin<sup>a,1</sup> and Y. Verin<sup>a,1</sup>

<sup>a</sup> Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER), Laboratoire Ressources Halieutiques, 150 quai Gambetta, BP699, 62321, Boulogne/mer, France

<sup>b</sup> UPMC Université Paris 06, UMR 7093, Laboratoire d'Océanographie de Villefranche, 06230 Villefranche-sur-Mer, France

<sup>c</sup> CNRS, UMR 7093, LOV, 06230 Villefranche-sur-Mer, France

<sup>d</sup> Institute of Marine Research — Tromsø, Postboks 6404, 9294 Tromsø, Norway

<sup>1</sup> Tel.: + 33 3 21 99 56 00; fax: + 33 3 2199 56 01.

<sup>2</sup> Tel.: + 47 77 60 97 21.

\*: Corresponding author : S. Vaz, Tel.: + 33 3 21 99 56 32, + 33 3 21 99 56 00; fax: + 33 3 21 99 56 01, email address : [Sandrine.Vaz@ifremer.fr](mailto:Sandrine.Vaz@ifremer.fr)

---

### Abstract:

Potential spawning habitat is defined as the area where environmental conditions are suitable for spawning to occur. Spawning adult data from the first quarter (January–March) of the International Bottom Trawl Survey have been used to study the inter-annual variability of the potential spawning habitat of North Sea plaice from 1980 to 2007. Generalised additive models (GAM) were used to create a model that related five environmental variables (depth, bottom temperature and salinity, seabed stress and sediment type) to presence–absence and abundance of spawning adults. Then, the habitat model was applied each year from 1970 to 2007 to predict inter-annual variability of the potential spawning habitat. Predicted responses obtained by GAM for each year were mapped using kriging. A hierarchical classification associated with a correspondence analysis was performed to cluster spawning suitable areas and to determine how they evolved across years. The potential spawning habitat was consistent with historical spawning ground locations described in the literature from eggs surveys. It was also found that the potential spawning habitat varied across years. Suitable areas were located in the southern part of the North Sea and along the eastern coast of England and Scotland in the eighties; they expanded further north from the nineties. Annual survey distributions did not show such northward expansion and remained located in the southern North Sea. This suggests that this species' actual spatial distribution remains stable against changing environmental conditions, and that the potential spawning habitat is not fully occupied. Changes in environmental conditions appear to remain within plaice environmental ranges, meaning that other factors may control the spatial distribution of plaice spawning habitat.

**Keywords:** Potential Habitat; Generalised Additive Models; Spawning; Inter-annual Variability; North Sea plaice; Spatial Distribution

## Introduction

Knowledge of how environmental factors may impact species distribution lies at the heart of ecology. Understanding the relationship between a species and its surrounding abiotic environment is relevant for the fundamental ecology of a species but also useful in terms of conservation ecology, e.g. through the creation of marine protected areas, for managing marine resources and for evaluating the consequences of climate changes on populations. Hutchinson, in its concept of the ecological niche (Hutchinson, 1957), defines the fundamental niche as an hypervolume of “n” dimensions, each of them being an environmental factor which defines the suitable environmental conditions for a species to survive, grow and reproduce. The concept of habitat is linked to the ecological niche by being its spatial representation: the environmental space, as defined by the environmental relationship between the species and environmental variables, is projected into the geographical space to locate suitable areas (Coyne and Christensen, 1997).

Until now, habitat modelling has mainly been used to predict species distribution (Austin, 2002; Austin, 2007; Guisan and Zimmermann, 2000; Guisan and Thuillier, 2005; Guisan *et al.*, 2006) in order to make inference on the role of tested environmental controls on species distribution according to their ability to explain some degree of variation in the response data (i.e. the presence-absence or abundance). This has led to the development of a new set of methods to better fit the environmental predictor to the biological response. Various response shapes (linear, non-linear, polynomial, bell-shaped form) and data types (Boolean, discrete, continuous) are being used. Generalised additive models (GAM, Hastie and Tibshirani, 1990) are a modelling approach first used in terrestrial ecology (Yee and Mitchell, 1991) and later in marine ecology (Lehmann, 1998). They are often used because they are data driven, rather than model driven, and are based on the addition of their smoothing functions that link biological response and environmental predictors. Moreover, they have been proved as robust as other regression technique such as Generalised Linear Models (GLM, McCullagh and Nelder, 1989) and are no more computationally intensive nor time consuming compared with other methods (Leathwick *et al.*, 2006).

In most studies, the habitat model is applied on a set of environmental values averaged over several years (Vaz *et al.*, 2008), so as to produce a final map which describes the suitable areas using the relationship between species distribution and environmental predictors (Coyne and Christensen, 1997). In some studies, the habitat model is applied on several years and the different maps are combined at the end to locate recurrent, optional and unfavourable suitable

areas (Bellier *et al.*, 2007; Certain *et al.*, 2007). If this can be a practical way to extract general pattern distribution of the species as well as to represent the inter-annual variability, this approach is limited in its possibilities as it does not give any information regarding how the habitat varies from year to year. Loots *et al.* (2007) applied their model on three years of data and mapped the habitat distribution for each year in order to see how the habitat changed across years according to changes in environmental values. However, this approach was purely visual and no real comparison could be made between observed and predicted distribution for each year except with a Spearman correlation coefficient calculated for all the years together.

The common Plaice (*Pleuronectes platessa*) is one of the most studied flatfish species in the North Sea because of its economical importance for several countries (Rijnsdorp and Millner, 1996). Its life cycle has been widely described (Cushing, 1990; Wegner *et al.*, 2003). The spawning period is known for the English Channel and North Sea to occur from late December to April, with a peak of spawning in January-February. Locations of spawning grounds have been studied using eggs surveys (Harding *et al.*, 1978; ICES, 2005). The adults spawn near to the bottom and mainly at night (Simpson, 1971; Nichols, 1989). Because of their positive buoyancy, eggs are pelagic and can be found mainly in the upper layer of the water column (Coombs *et al.*, 1990). Eggs and larvae drift to coastal nursery grounds where juveniles develop during one year (Talbot, 1977). They then progressively migrate offshore where female and male mature at the age of 4-5 and 2-3 years, respectively (Rijnsdorp, 1989). As the capacity of plaice to maintain and replace its population largely depends on its ability to complete its life cycle, pelagic larvae and juveniles have been considered as its main critical life stages (van der Veer *et al.*, 1990; van der Veer *et al.*, 1998; Pastoors *et al.*, 2000). However, little attention has been paid to the adult population, except for some recent studies related to population structure (Hunter *et al.*, 2004; Metcalfe, 2006) and migrations (Hunter *et al.*, 2003; Hunter *et al.*, 2004).

Several terms may be used to spatially describe fish reproduction: spawning grounds, spawning habitat and reproductive habitat (Castro *et al.*, 2005). Spawning grounds define the geographical areas where spawning occurs, i.e. where males and females meet to reproduce. When spawning grounds are studied by the means of pelagic egg survey data, there is no certainty that these eggs have actually been released where they were collected, as they may have drifted from elsewhere due to local hydrodynamic conditions. However, eggs are often used as proxies to spawning grounds as it is unlikely that all adults observed in an area are spawning there and then. To alleviate this problem, direct maturity measurements may be

used to estimate the proportions of spawning adults. In this study, we propose to focus on the adult spawning habitat in order to determine which environmental factors constrain the adult distribution at the time of spawning. This will be referred to as spawning habitat thereafter.

The aims of this study were 1) to build a habitat model for North Sea plaice during its reproductive period based on adult spawning distribution, 2) to map this habitat for each year in order to see how it varied from year to year, and 3) to determine how this habitat model reflected observed distribution using mapping and evaluation method.

## **Material and methods**

### *Data*

Data of the first quarter of the International Bottom Trawl Survey (IBTS) were used to model the potential spawning habitat of North Sea plaice. The first quarter of the IBTS has been carried out each year (since 1980) in the North Sea, from January to March, to collect the necessary data for stock assessment of several important demersal fish (ICES, 2007; ICES, 2008). The sampling strategy is based on statistical rectangles of  $1^\circ$  of longitude by  $0.5^\circ$  of latitude (fig. 1), as defined by the International Council for the Exploration of the Sea (ICES). Each rectangle is visited by two different countries, which each perform one standardised 30 min trawl using a 36/47 GOV (Grande Ouverture Verticale) bottom trawl. Trawling locations are randomly chosen by the country among a pre-defined set of three or four trawling locations inside a given rectangle. This results in a minimum of two trawl hauls per rectangle per survey quarter. Trawling depth (m), temperature ( $^\circ\text{C}$ ) and salinity are recorded at each trawling location. All specimens in each trawl sample are sorted by species and counted. Length measurements, otolith sampling and sexual maturity staging are performed for several key species, on a representative sample of individuals within seven standard “roundfish” areas (ICES, 2004, fig. 1). Individuals are classified into four stages of maturity, with stage 3 corresponding to spawning individuals with fluent gonads (ICES, 2004). From 1980 to 2007, 11,343 bottom trawls were performed, which represents an average of 405 bottom trawl hauls per year. These data are available through the DATRAS database (DAtabase of TRAwl Surveys, <http://datras.ices.dk/Home/Default.aspx>) coordinated by ICES.

Abundance data of spawning adults from the first quarter of the IBTS (January-March) were computed from the available proportion of spawning adults within any given length class. The determination of sexual maturity has only been carried out since 2001, so there was

not enough data to representatively calculate the proportion of stage 3 adults within each length class and for each year and roundfish area. Consequently, data on sexual maturity from 2001 to present were merged to calculate these proportions for northwest areas (areas 2, 3 and 4 pooled together, fig. 2) and southeast areas (areas 5, 6 and 7 pooled together, fig. 2). This choice was based on the time of spawning and spawning activity of plaice in each area (Daan *et al.*, 1990). No data on the sexual maturity of plaice in area 1 was available. Data on males and females were merged as there was not enough data to calculate proportions of spawning adults for each sex separately. This reduced the total number of trawling stations, available for analyses, from 11,343 to 7,317. For each of these stations, total abundance of spawning adults (in individuals.km<sup>-2</sup>) were calculated from the summed product of the total abundance within each length class and the corresponding proportions of spawning adults inside that length class.

Five environmental variables were used: depth, bottom temperature and salinity, seabed stress (N.m<sup>-2</sup>) and sediment type. Depth (m), bottom temperature (°C) and salinity were extracted from the physical part of DATRAS from 1970 to present. Sea-bed stress is a measure of the shear friction of water on the seabed due to the tidal currents. It was estimated using a 2D hydrodynamic model (Aldridge and Davies, 1993) from the Proudman Oceanographic Laboratory (POL), and running on a regular grid of 1/8° of longitude by 1/12° of latitude (WGS 1984 datum). The grid of points was then interpolated using ArcMap's Spatial Analyst extension (ESRI, 2005) to create a continuous raster layer of 1 km<sup>2</sup> resolution. Sediment type originated from the seafloor sediment of the North Sea built during the MARGIS project (Schlüter and Jerosch, 2008). Sediment types were classed into five categories: fine sand, coarse sand, mud, pebbles and gravels. A category of sea-bed stress and sediment type was allocated to each trawl location by resampling the corresponding map at that location using ArcMap's Hawth's Analysis Tools extension (Beyer, 2004). Although these five explanatory variables may share a certain amount of variation, in particular due to similar spatial pattern, it was decided to test all of them to build the habitat model.

### *Habitat modelling*

Generalised additive models (GAM, Hastie and Tibshirani, 1990; Guisan *et al.*, 2002; Wood and Augustin, 2002) were used to relate the spatial distribution of North Sea plaice spawning adults to the five environmental variables. In GAM, the biological response is smoothed according to each environmental predictor using a smoothing spline and a link

function to create a smoothing term. All the terms are then added at the end of the fitting process to do an additive model (1)

$$(1) g(E[Y]) = \beta_0 + f_1X_1 + f_2X_2, \dots, + f_nX_n = \beta_0 + \sum_{i=1}^n f_iX_i$$

where  $g()$  is the link function,  $Y$  the response,  $\beta_0$  the intercept,  $X_{1-n}$  the explanatory variables and  $f_{1-n}$  the smoothing functions.

As it is often the case with ecological data, the abundance of spawning adults displayed zero-inflated distributions (fig. 3). As a result, the modelling procedure was split into two steps. 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 (fig. 3), respectively. A log-transformation was applied to seabed stress, and an exponential transformation to bottom salinity ( $\exp(\text{Salinity})/10^{15}$ ), to make their distribution closer to normality (fig. 3).

Binomial and Gaussian models were adjusted using data from 1980 to 2007. The significance of each predictor was tested in a stepwise selection procedure using the Akaike Information Criterion (AIC, Akaike, 1974). The AIC penalised the log-likelihood ( $\log L$ , Burnham and Anderson, 2002) of the model (i.e. the residuals sum of square between observed and fitted values for a classic linear model), by the number of parameters  $p$  included in the model according to (2).

$$(2) AIC = -2 \times \log L + 2 \times p$$

This allowed selecting the most parsimonious model, i.e. the best adjusted model with the least number of predictors as possible. This selection procedure also accounted for collinearity between variables as variation redundancy tends to be avoided. For each predictor, four degrees of smoothing (1 to 4) were tested, except for sediment types which were integrated as a categorical factor in the modelling procedure. Models were implemented using the R free software (R Development Core Team, 2008). The ‘gam’ package (Hastie, 2006) was used to construct the binomial and Gaussian models. The ‘step.gam’ function of the MASS library (Venables and Ripley, 2002) was used to select the explanatory variables according to the AIC.

Binomial and Gaussian models were reapplied to the environmental values of each year from 1970 to 2007. Probabilities of presence predicted by binomial models and log-transformed abundance predicted by Gaussian models were combined in a delta approach in which one was multiplied by the other (Stefánsson, 1996; Le Pape *et al.*, 2003) so as to predict the spatial distribution. The value of the resulting predicted abundance in a given area depicted the suitability of this area, with high abundance indicating high spawning suitability and low abundance low spawning suitability.

The contributions of each explanatory variable included in the delta model were computed using the coefficient of determination and adjusted coefficient of determination (Legendre & Legendre, pers. com.). The coefficient of determination ( $R^2$ ) measures the proportion of the variation of the response  $Y$  from its mean that is explained by the regression equation. It is calculated as follows (3):

$$(3) R^2 = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2}$$

where  $\hat{Y}_i$  is the  $i^{\text{th}}$  predicted value,  $Y_i$  is the  $i^{\text{th}}$  observed value and  $\bar{Y}$  is the mean of the observed values.

The adjusted coefficient of determination ( $R_a^2$ ) takes into account the respective numbers of degrees of freedom of the numerator and denominator of  $R^2$  :

$$(4) R_a^2 = 1 - (1 - R^2) \times \left( \frac{\text{total } d.f.}{\text{residual } d.f.} \right)$$

where *total d.f.* is the total degrees of freedom and *residual d.f.* is the degrees of freedom of the model. The global and pure effect of each variable was determined following a methodology inspired from Borcard *et al.* (1992) and adapted to the GAM and delta model framework. For each variable, three delta models (combining both binomial and Gaussian models) were built: (i) the selected model, i.e. the model resulting from the stepwise selection, (ii) the selected model excluding the variable of interest (i.e. the variable which contribution needed to be evaluated) and (iii) a simple model containing the variable of interest alone. The adjusted coefficient of determination was calculated using (3) and (4) for these three models.

The global effect of the variable was given by the adjusted coefficient of determination of the model (iii) whereas the pure effect of the variable was obtained by removing the adjusted coefficient of determination of the selected model (i) minus that of model (ii).

Distributions predicted by delta models, originating from the combination of selected binomial and Gaussian models, were graphically compared to observed distributions by the means of a Taylor diagram (Taylor, 2001) and the use of three similarity indices (standard-deviation, root mean squared error RMSE, and Spearman correlation coefficient). A plot of the standard deviation indicates whether the model is able to reproduce the same variability in its prediction as that in the survey data. The RMSE was computed as the root of the mean of the squared differences between each prediction and each survey observation. It incorporates both the variance of the model and its bias. The Spearman correlation coefficient and the RMSE quantify the correspondence between the observed and predicted patterns. Predictions of the model for all the years pooled together and also for each individual year were evaluated separately in order to see how the model accuracy varied across years. The Taylor diagram was computed using the 'plotrix' package in R (Lemon *et al.*, 2008).

### *Geostatistics*

Geostatistics were used to map predictions of the habitat model on a regular grid (Webster and Oliver, 2001). For each year from 1970 to 2007, an experimental variogram was computed according to (5).

$$(5) \gamma(h) = \frac{1}{2 \times n(h)} \times \sum [z(x+h) - z(x)]^2$$

where  $\gamma(h)$  is the experimental variogram,  $n(h)$  is the number of pairs of observations for the distance  $h$ ,  $z(x)$  is the observed abundance at location  $x$  and  $h$  is the distance between two locations. The distance  $h$  was calculated using the latitude and corrected longitude of the trawling location. The longitude correction ( $longitude \times \cos((latitude \times \pi)/180)$ ) transforms decimal degrees of longitude into decimal degrees of latitude which are of constant distance in a Mercator-like projection formula. In case of non stationary of  $z(x)$ , it can be split into two components following (6).

$$(6) z(x) = m(x) + R(x)$$

where  $m(x)$  is the spatial trend and  $R(x)$  are the residuals.

When accounting for more than 20% of the variation of the data, the spatial trend was modelled by fitting a low-order polynomial (linear or quadratic regression) to the spatial coordinates using the least-square regression method. The experimental variogram was then calculated on the residuals. A theoretical model, chosen among exponential, circular, spherical and penta-spherical authorised functions was adjusted to this experimental variogram to determine the nugget, sill and 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 ordinary kriging interpolation method, or the universal kriging in the presence of a spatial drift. A mesh size of 0.2 decimal degrees was chosen for the interpolation grid, as the mean survey resolution of the IBTS was 0.16 decimal degrees. Geostatistics were implemented using Genstat (GenStat Release 7.1., 2004).

#### *Multivariate analysis*

A contingency table of 38 years (columns) and 678 points (lines) was extracted from the interpolation grid of predicted values of the habitat model. A cluster analysis based on a Chi<sup>2</sup> distance and using a flexible link ( $\beta=-0.25$ ) combined to a correspondence analysis was carried out on years 1980-2007 to study temporal variations in the spatial distribution of spawning habitat. Because there was no observed distribution for the seventies, predictions of the habitat model for these years were used as supplementary variables in the correspondence analysis. In this way, they do not interfere in the building of the axes but can be projected on the factorial plane.

## **Results**

### *The habitat model*

Based on the AIC, it was found that all five environmental variables were significant in improving the adjustment of the final binomial (7) and Gaussian (8) models.

$$(7) \ 0/1 = s(\text{Depth}, 4) + s(\text{Temperature}, 4) + s(\text{Salinity}, 4) + s(\text{Bedstress}, 4) + \text{as.factor}(\text{Sediments})$$

(8)  $\log(\text{Ab}>0) = s(\text{Depth}, 4) + s(\text{Temperature}, 3) + s(\text{Salinity}, 4) + s(\text{Bedstress}, 4) + \text{as.factor}(\text{Sediments})$

Plots of fitted values for each smoothing term (fig. 4) clearly showed a non-linear relationship with quadratic smoothing except for bottom temperature for which a concave relationship was modelled using cubic smoothing. For each environmental variable, the response shapes are similar for the binomial or the Gaussian model. There were high presence probability or abundance values at shallow depth (0-50 m), intermediate salinity (around 34.3), weak seabed stress (close to 0) and relatively warm bottom temperature (4-8°C).

The respective contributions of the five variables in the resulting delta model are given in table 1. The delta model explained a total of 25.5 % of the variation in the survey data. Depth accounted for the larger part of this explained variation, followed by salinity whereas temperature, seabed stress and sediment types only explained a small, although significant, part of the variation. Salinity shared a large part of the explained variation with other variables, and especially with depth as shown by the values of the conditional effect of each variable.

#### *Evaluation of the habitat model*

Evaluation of the habitat model by the Taylor diagram is given in fig. 5. Evaluation of the predictions for each year indicated a difference in accuracy of the model depending on the year that is being considered. The correlation between predictions and observations was lower for the nineties than the eighties and lower for the 2000's than the nineties. Predictions for the 2000's had a higher RMSE value than for the nineties and the eighties. This means that there was a general decrease in model accuracy along the three decades, thereby indicating that the model was less able to reproduce observed pattern for the 2000's than for the nineties or the eighties.

#### *Temporal variations of plaice spawning habitat*

The classification separated four groups of samples at the threshold of 20%. These groups were plotted on the first two axes of the correspondence analysis (fig. 6), the first axis explaining 23 % of the variation and the second axis explaining 11.7 %. Groups of sampling stations were mainly distributed along the first axis. The first group, mainly found to the left of the first axis, gathers samples located in the south-eastern part of the North Sea, in the

Flamborough Head area along the English coast and in the Firth of Forth and Moray Firth along Scotland (fig. 6). These sampling stations showed high predicted abundance values, thereby indicating highly suitable habitats. The second group was located close to the first group along the first axis, partially overlapping with it; it showed lower predicted abundance values (i.e. less suitable habitats). The third and fourth groups were located to the right of the first axis and consisted in areas that were located in the central and north western parts of the North Sea, respectively. These areas showed very low or almost null predicted abundance values, indicating non-suitable habitats. Therefore, these four groups along axis 1 were explained by a latitudinal gradient, with high predicted abundance values close to the coast and lower ones in the central part of the North Sea.

Three groups of years were identified using a threshold of 20% and were plotted on the same factorial plan as that used for the groups of sampling stations (fig. 6). These groups of years were also found to be distributed along the first axis. The first group was located to the left of the first axis and was close to the first group of sampling stations. It mainly gathered the eighties and the year 1991. The second group mainly gathered the nineties and the beginning of the 2000's and was located in the centre of the factorial plane, close to the second group of sampling stations. The third group gathered particular years and was located to the right of the first axis, close to the third and fourth groups of sampling stations. The seventies, used as supplementary variables, were distributed to the left of the first axis, close to the eighties, except for years 1973, 1978 and 1979 which were closer to the second group of years.

Habitat maps (fig. 7) for the three years (1987, 1997, 2007) belonging to each group and the supplementary variables (1970) illustrate suitable spawning areas for each group of years that were identified in the correspondence analysis. These maps show that (1) there was a spatial stability of the predicted suitable areas during the three decades in the south-eastern North Sea and along the English and Scottish coasts, and that (2) there also was a spatial expansion of suitable areas between 1970 and 2007. This expansion began with the Moray Firth and Firth of Forth that became more suitable in the eighties compared to the seventies. This was more pronounced in the nineties with a wider suitable area along the English coast that also began to expand in the central part of the North Sea. In the 2000's, this central area was a more suitable for spawning than in previous years. Maps of survey distribution for the same years (fig. 7) indicate that higher survey abundance values were indeed located in areas of higher suitability during the three decades but did not expand further north as suggested by the habitat model.

## Discussion

### *The habitat model*

The variation explained by the final delta model did not exceed 30%, which leaves out a considerable amount of unexplained variation. However, the large number of survey observations and their associated uncertainty due to sampling error generate an important amount of white noise, which cannot be accounted for by the model. Apart from random variation, the remaining unexplained variability may be due to other explanatory variables than those available for this study. All the available environmental predictors were retained in the final adjusted binomial and Gaussian models. This indicates that depth, bottom temperature and salinity, seabed stress and sediment type play an important role in determining the spawning habitat of North Sea plaice. The adjustment of response through the use of GAM smoothing functions has helped identifying complex response curves that were clearly not linear. Moreover, response shapes were quite similar for each response towards a given environmental factor, be it with the binomial or the Gaussian model. For depth, bottom salinity and seabed stress, adjusted response showed a bell-shaped form which indicates an environmental window of preference of North Sea plaice for these three factors during its reproduction. Indeed, spawning adults are known to concentrate in shallow areas of less than 50 m depth, in agreement with what has been described for areas of high concentrations of eggs (Harding *et al.*, 1978). Little has been published regarding the relationship between salinity and spawning except for the southern Bight of the North Sea where high salinity seems to offer optimal conditions for developing eggs (Cushing, 1990). High salinity in the southern Bight is due to inflow of central saline water from the English Channel through the Dover Strait. In this study, we found an optimal response for salinity around 34.3; however, it is not the maximal salinity observed in the whole North Sea area. We argue that the absence of preference for high salinity in our model is due to the lack of regular trawling in areas where salinity is high, such as in the southern Bight. Another explanation would be that IBTS survey takes place too late (February-March) to detect the peak of spawning in the southern Bight, which occurs during the last two weeks of January (Harding and Talbot, 1973; Harding *et al.*, 1978). Relationship with temperature was more linear, with a higher preference for temperatures between 4 to 8°C, which is in the tolerance range of 2-8°C for plaice eggs survival (Harding *et al.*, 1978).

Very few published models used seabed stress as an environmental predictor. Spawning adults seemed to prefer areas with very weak seabed stress, which correspond to areas with low tidal impacts and which cover most of the studied area, the southern Bight excepted. In effect, this area is not adequately covered by the IBTS survey design, due to both timing and scarcity of observations, and this prevents from detecting spawning adults in strong seabed stress conditions. The input of sediments in the model was less clear. There was a high standard deviation for each type of sediment in the adjusted response and a preference for muddy bottom. Sediment distribution in the North Sea is not, however, very contrasted due to the lack of fine resolution map, and is mainly dominated by fine sand: this greatly reduces the usefulness of sediment type as predictor in habitat modelling.

#### *Spawning habitat and spawning grounds*

Highly suitable habitat areas reflected quite well the locations of spawning grounds as described based on egg survey data in the literature (Harding *et al.*, 1978; ICES, 2005). The widest and most suitable area is located in the southern part of the North Sea, in offshore waters along the Dutch, German and Danish coasts and corresponds to the Borkum area (in the German Bight) and the so-called Transition area (south of the Dogger Bank). Other areas are located along the English coast in the Flamborough Head region and along the Scottish coast in the Moray Firth and Firth of Forth. The third main spawning ground is in the southern Bight of the North Sea but this one is not shown on this study's maps because of the lack of regular yearly samplings.

Predictions confirmed the significance of these spawning grounds, already known in terms of egg production, with the German and southern Bights contributing most, and the Flamborough Head being the less productive region. This has remained quite constant during the 20<sup>th</sup> Century, though the German Bight has become more productive than the southern Bight and has expanded since the seventies (Harding *et al.*, 1978).

#### *Expansion of spawning habitat*

A continuous expansion of suitable habitat was observed with time, towards the central and north-western part (south of Orkney Islands) in the 2000s. The study period can be split into at least two periods (1970-1995 and 1995-2007). If environmental conditions were less suitable in the central and north-western part during the seventies, they became more

suitable in the 2000's as shown by suitable habitat expansion. In the habitat models, bottom temperature and salinity were the only predictors that varied from year to year, meaning that temporal change in suitability can only be due to one of these two parameters. However, survey maps indicate that the location of actual spawning areas has remained fairly constant in time, in agreement with spawning grounds remaining spatially stable (ICES, 2005). Spatial stability of actual spawning grounds hence results from environmental conditions varying within the suitable environmental range of the species, i.e. its ecological niche. The potential habitat distribution is hence less variable in time than in space, as expected (Planque *et al.*, 2007). This low variability of suitable areas in the North Sea during the winter period can be an explanation why this species chose this period to reproduce.

#### *Observed distribution versus spawning habitat*

The observed distribution does not seem to have expanded during the last three decades, in contrast with the habitat models. Plaice remained concentrated in areas of highest habitat suitability in the southern part and along the English and Scottish coasts, whereas it remained absent of the central part of the North Sea. This was confirmed by the Taylor diagram that showed that predictions were better at the beginning of the study period than at the end. These results imply that 1) the potential habitat is usually not fully or always occupied, that 2) change in habitat suitability does not automatically lead into a change in actual species distribution, and that 3) potential habitat modelling using environmental predictors cannot fully explain what controls the spatial distribution of a species.

The first point implies that other factors, related to the population itself, affect the way in which a species is distributed and occupies its potential habitat. The actual distribution is usually less spread out than expected, especially during the reproductive period, when individuals tend to aggregate to favour the encounter of gametes of both sexes. The bell-shaped form of response curves found for depth, salinity and seabed stress hence results from the aggregative behaviour (Sutherland, 1983) of the individuals within the preferred environmental range. The other concept which might control potential habitat occupation is related to the density dependent effect (Fretwell and Lucas, 1970; Shepherd and Litvak, 2004) which states that the population will contract in smaller areas at low level population size and expand its distribution range at higher level population size. This refers to the distinction between potential and realised spawning habitat (Planque *et al.*, 2007), realised habitat being the potential habitat which is constrained by characteristics of the spawning population, such

as its size and demography. The realised habitat can be seen as the observed distribution of a population each year.

The second point refers to the persistent character of a population that will play a key role in its spatial distribution. Difference between predicted and actual distribution tend to prove that spawning plaice persists at the same place year after year, whatever the environmental conditions. This faithfulness to spawning areas has also been reported by other studies (Hunter *et al.*, 2003) and in other areas (Solmundsson *et al.*, 2005). This means that more than the environmental conditions, it is the geographical location that is important for plaice reproduction. However, homing mechanisms remain unclear (Harding *et al.*, 1978; Gibson, 1997; Hunter *et al.*, 2003), and adults are not always distributed in the best environmental conditions for themselves but more in order to allow eggs and larvae to encounter favourable conditions thereafter. This should lead to discrepancies between predicted and actual spawning suitable areas. Therefore, climate change impact may not greatly affect spawning spatial distribution, but may have consequences on reproductive success and the survival of eggs and larvae.

Fidelity of the adult plaice population to its spawning grounds can have detrimental effects on population survival under climate change conditions as adults are expected to continue spawning at the same place which may not be environmentally suitable for eggs and larvae to develop. Furthermore in the present study, a model calibrated over thirty years could not predict actual distribution when there was relatively large environmental change (towards the end of the study period). Therefore, it is likely that this model will not be reliable for predicting abundance levels outside of the study period (i.e. simulated climatic conditions) or even in another geographical area. In conclusion, rather than testing additional environmental predictors in order to improve habitat models, it would be wiser to consider and test other factors related to the state of the population such as aggregation behaviour, density-dependent effect and site attachment.

### **Acknowledgments**

The work was part-funded by the EU as part of the RECLAIM project (STREP-FP6, contract n044133). We acknowledge the financial support of the Region Nord-Pas de Calais, ICES for providing the data, and all those involved in the IBTS survey. We thank C. Martin for proof-reading and comments on the manuscript.

## References

- Akaike, H., 1974. A new look at the statistical model identification. *IEEE T. Automat. Contr.* 19, 716-723.
- Aldridge, J.N., Davies, A.M., 1993. A high resolution three-dimensional hydrodynamic tidal model of the Eastern Irish Sea. *J. Phys. Oceanogr.* 23, 207-224.
- Austin, M.P., 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecol. Model.* 157, 101-118.
- Austin, M.P., 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecol. Model.* 200, 1-19.
- Bellier, E., Planque, B., Petitgas, P., 2007. 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.
- Beyer, H.L., 2004. Hawth's Analysis Tools for ArcGIS. Available at <http://www.spatial ecology.com/htools>.
- Borcard, D., Legendre, P., Drapeau, P., 1992. Partialling out the spatial component of ecological variation. *Ecology* 73, 1045-1055.
- Burnham, K.P., Anderson, D.R., 2002. Model selection and multimodel inference. A practical information-theoretic approach, Springer-Verlag, USA.
- Castro, L., Fréon, P., van der Lingen, C.D., Uriarte, A., 2005. Report of the SPACC Meeting on Small Pelagic Fish Spawning Habitat Dynamics and the Daily Egg Production Method (DEPM). Concepcion, Chile GLOBEC report 22. 107 pp.
- Certain, G., Bellier, E., Planque, B., Bretagnolle, V., 2007. Characterising the temporal variability of the spatial distribution of animals : an application to seabirds at sea. *Ecography* 30, 695-708.
- Coombs, S.H., Nichols, J.H., Fosh, C.A., 1990. Plaice eggs (*Pleuronectes platessa* L.) in the southern North Sea: abundance, spawning area, vertical distribution, and buoyancy. *ICES J. Mar. Sci.* 47, 133-139.
- Coyne, M.S., Christensen, J.D., 1997. Biogeography program : habitat suitability index modeling : species habitat suitability index values, technical Guidelines. NOAA , 1-19.
- Cushing, D.H., 1990. Hydrographic containment of a spawning group of plaice in the Southern Bight of the North Sea. *Mar. Ecol.-Prog. Ser.* 58, 287-297.
- Daan, N., Bromley, P.J., Hislop, J.R.G., Nielsen, N.A., 1990. Ecology of North Sea Fish. *Neth. J. Sea Res.* 26, 343-386.

- ESRI, 2005. Environmental Systems Research Institute. ArcGis. Version 9.1. Redlands, CA : Environmental Research Institute Inc.
- Fretwell, S.D., Lucas, H.L., 1970. On territorial behavior and other factors influencing habitat distribution in birds. *Acta Biotheor.* 19, 16-36.
- 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.
- Gibson, R.N., 1997. Behaviour and the distribution of flatfishes. *J. Sea. Res.* 37, 241-256.
- Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. *Ecol. Model.* 135, 147-186.
- Guisan, A., Edwards, T.C., Hastie, T., 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecol. Model.* 157, 89-100.
- Guisan, A., Thuillier, W., 2005. Predicting species distribution: offering more than simple habitats models. *Ecol. Lett.* 8, 993-1009.
- Guisan, A., Lehman, A., Ferrier, S., Austin, M.P., Overton, J.M.C., Aspinall, R., Hastie, T., 2006. Making better biogeographical predictions of species' distributions. *J. Appl. Ecol.* 43, 386-392.
- Harding, D., Nichols, J.H., Tungate, D.S., 1978. The spawning of plaice (*Pleuronectes platessa*) in the southern North Sea and English Channel. *Rapp. P.-v. Réun. Cons. Int. Explor. Mer* 172, 102-113.
- Hastie, T.J., Tibshirani, R.J., 1990. *Generalized additive models*, Chapman and Hall, London.
- Hastie, T., 2006. *gam: generalized additive models*. R package version 0.98.
- Hunter, E., Metcalfe, J.D., Reynolds, J.D., 2003. Migration route and spawning area fidelity by North Sea plaice. *P. Roy. Soc. B-Biol. Sci.* 270, 2097-2103.
- Hunter, E., Metcalfe, J.D., Arnold, G.P., Reynolds, J.D., 2004. Impacts of migratory behaviour on population structure in North Sea plaice. *J. Anim. Ecol.* 73, 377-385.
- Hutchinson, G.E., 1957. Concluding remarks. *Cold Spring Harb. Symp. Quant. Biol.* 22, 415-427.
- ICES, 2004. *Manual for the international bottom trawl surveys. Revision VII. The International Bottom Trawl Survey Working Group. Resource Management Committee. ICES CM 2004.*
- ICES, 2005. *Report of the Planning Group on North Sea Cod and Plaice Egg Surveys in the North Sea (PGEGBS), 10–12 May 2005, Lowestoft, UK. ICES Document CM 2005/G:11.*

- ICES, 2007. Report of the International Bottom Trawl Survey Working Group (IBTSWG), 27-30 March 2007, Sète, France. ICES CM 2007/RMC:05.
- ICES, 2008. Report of the working group on the assessment of demersal stocks in the North Sea and Skagerrak - combined spring and autumn (WGNSSK), 7-13 May 2008, Copenhagen, Denmark. ICES CM 2008/ACOM:09.
- Le Pape, O., Chauvet, F., Mahévas, S., Lazure, P., Guérault, D., Désaunay, Y., 2003. Quantitative description of habitat suitability for the juvenile common sole (*Solea solea*, L.) in the Bay of Biscay (France) and the contribution of different habitats to the adult population. *J. Sea. Res.* 50, 139-149.
- Leathwick, J.R., Elith, J., Hastie, T., 2006. Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions. *Ecol. Model.* 199, 188-196.
- Lehmann, A., 1998. GIS modeling of submerged macrophyte distribution using generalized additive models. *Plant Ecol.* 139, 113-124.
- Lemon, J., Bolker, B., Oom, S., Klein, E., Rowlingson, B., Wickham, H., Tyagi, A., Eterradosi, O., Grothendieck, G., Toews, M., Kane, J., Cheetham, M., Turner, R., Witthoft, C., Stander, J., Petzoldt, T., 2008. plotrix: Various plotting functions. R package version 2.5.
- Loots, C., Koubbi, P., Duhamel, G., 2007. Habitat modelling of *Electrona antarctica* (Myctophidae, Pisces) in Kerguelen by generalized additive models and geographic information systems. *Polar Biol.* 30, 951-959.
- McCullagh, P., Nelder, J.A., 1989. Generalized linear models, Monographs on statistics and applied probability 37, Chapman and Hall, London.
- Metcalf, J.D., 2006. Fish population structuring in the North Sea: understanding processes and mechanisms from studies of the movements of adults. *J. Fish Biol.* 69, 48-65.
- Nichols, J.H., 1989. The diurnal rhythm in spawning of plaice (*Pleuronectes platessa* L.) in the southern North Sea. *J. Cons. int. Explor. Mer* 45, 277-283.
- Pastors, M.A., Rijnsdorp, A.D., Van Beek, F.A., 2000. Effects of a partially closed area in the North Sea ("plaice box") on stock development of Plaice. *ICES J. Mar. Sci.* 57, 1014-1022.
- Planque, B., Bellier, E., 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.
- R Development Core Team, 2008. R : a language and environment for statistical computing, R foundation for statistical computing, Vienna, Austria.

- Rijnsdorp, A.D., 1989. Maturation of male and female North Sea plaice (*Pleuronectes platessa* L.). J. Cons. int. Explor. Mer 46, 35-51.
- Rijnsdorp, A.D., Millner, R.S., 1996. Trends in population dynamics and exploitation of North Sea plaice (*Pleuronectes platessa* L.) since the late 1800s. ICES J. Mar. Sci. 53, 1170-1184.
- Schlüter, M., Jerosch, K., 2008. Digital atlas of the North Sea (DANS). Geo-information regarding geology, geochemistry, oceanography and biology. Available at [http://www.awi.de/en/research/research\\_divisions/geosciences/marine\\_geochemistry/marine\\_gis/digital\\_atlas\\_of\\_the\\_north\\_sea/](http://www.awi.de/en/research/research_divisions/geosciences/marine_geochemistry/marine_gis/digital_atlas_of_the_north_sea/).
- Shepherd, T.D., Litvak, M., 2004. Density-dependent habitat selection and the ideal free distribution in marine fish spatial dynamics: considerations and cautions. Fish Fish. 5, 141-152.
- Simpson, A.C., 1971. Diel spawning behaviour in populations of Plaice, Dab, Sprat and Pilchard. J. Cons. int. Explor. Mer 34, 58-64.
- Solmundsson, J., Palsson, J., Karlsson, H., 2005. Fidelity of mature Icelandic plaice (*Pleuronectes platessa*) to spawning and feeding grounds. ICES J. Mar. Sci. 62, 189-200.
- Stefánsson, G., 1996. Analysis of groundfish survey abundance data: combining the GLM and delta approaches. ICES J. Mar. Sci. 53, 577-588.
- Sutherland, W.J., 1983. Aggregation and the ideal free distribution. J. Anim. Ecol. 52, 821-828.
- Talbot, J.W., 1977. The dispersal of plaice eggs and larvae in the Southern Bight of the North Sea. J. Cons. int. Explor. Mer 37, 221-248.
- Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. 106, 7183-7192.
- Vaz, S., Martin, C.S., Eastwood, P.D., Ernande, B., Carpentier, A., Meaden, G.J., Coppin, F., 2008. Modelling species distributions using regression quantiles. J. Appl. Ecol. 45, 204-217.
- van der Veer, H.K., Pihl, L., Bergman, M.J.N., 1990. Recruitment mechanisms in North Sea plaice *Pleuronectes platessa*. Mar. Ecol.-Prog. Ser. 64, 1-12.
- van der Veer, H.W., Ruardij, P., van den Berg, A.J., Ridderinkhof, H., 1998. Impact of interannual variability in hydrodynamic circulation on egg and larval transport of plaice *Pleuronectes platessa* L. in the southern North Sea. J. Sea. Res. 39, 29-40.
- Venables, W.N., Ripley, D.B., 2002. Modern applied statistics with S, Springer, New York.
- Webster, R., Oliver, M.A., 2001. Geostatistics for environmental scientists. Statistics in practice, Wiley, Chichester.

Wegner, G., Damm, U., Purps, M., 2003. Physical influences on the stock dynamics of plaice and sole in the North Sea. *Sci. Mar.* 67, 219-234.

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

Yee, T.W., Mitchell, N.D., 1991. Generalized additive models in plant ecology. *J. Veg. Sci.* 2, 587-602.

### Figures captions

Fig. 1. The North Sea International Bottom Trawl Surveys (IBTS). Black dots indicate the location of bottom trawls for the first quarter (January-March) of the IBTS, since 1980. Countries involved in the IBTS, ICES statistical rectangles and the seven roundfish areas are also indicated.

Fig. 2. Proportions of adult stage 3 (i.e. spawning adults) North Sea plaice per size class. Left, areas 2–4 pooled. Right, areas 5–7 pooled. No data are available for area 1. The data points represent the proportions observed, and the lines the adjusted proportions using a smoothing spline for missing size classes.

Fig. 3. Histograms of the biological response (log-transformed abundance, presence/absence and non-null log-transformed abundance) and explanatory variables (trawling depth, bottom temperature, bottom salinity and seabed stress) from 1980 to 2007. See text for details on the transformation applied to seabed stress and salinity.

Fig. 4. Individual terms of the generalised additive model. Left, binomial model. Right, Gaussian model. Degrees of smoothing are indicated in brackets. Solid lines denote smoothed values and dotted lines denote the 95% confidence interval.

Fig. 5. Plot of the semi Taylor diagram. Observations and predictions of the habitat model are compared using the standard deviation, the root mean square error (RMSE) and the Spearman correlation. The three indexes were calculated on predictions for each year separately. The same symbol is used for a decade. The observation dataset was symbolised as a point called the reference point and was normalised so that the standard deviation was equal to 1, the correlation equal to 1 (the correlation between the observations and themselves) and the RMSE equal to 0 (the difference between the observations and themselves). Values of the

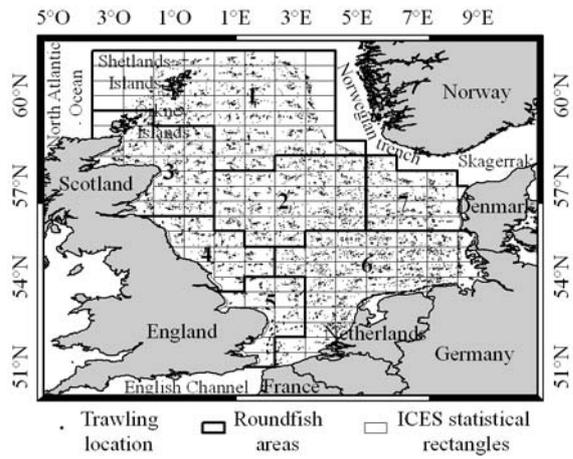
indexes for the predictions were plotted with the standard deviation on the y-axis, the correlation on the radials of the circle and the RMSE on concentric circles around the reference point.

Fig. 6. The correspondence analysis. Left, factorial plan of the first two axes. A classification based on a Chi-square distance was used to compare predictions of the habitat model for years 1980-2007. Predictions for the seventies were used as supplementary variables. Three groups of years and four groups of samples were identified. Right, location of the four groups of samples. Main locations are also indicated. MF : Moray Firth, FF : Firth of Forth, FH : Flamborough Head, SB : Southern Bight, TA : Transition Area, GB : German Bight.

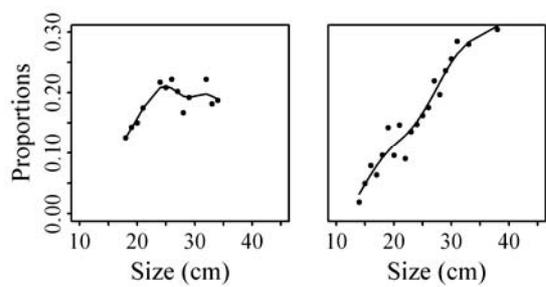
Fig. 7. Maps of predicted (top) and observed (bottom) distribution of North Sea plaice spawning adults for the three groups of years determined by the correspondence analysis. 1987, 1997 and 2007 were used as illustrative years for each of the three groups. Year 1970 was used to illustrate the potential habitat distribution for supplementary variables (seventies).

## Tables

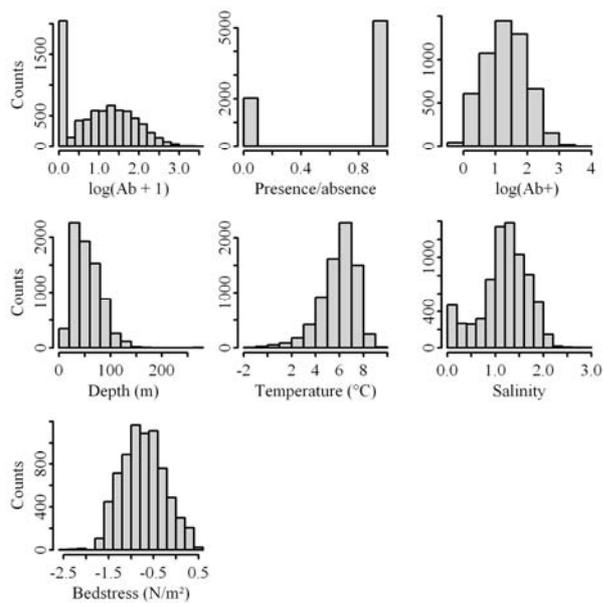
Table 1. Contribution of the explanatory variables. The total and conditional adjusted coefficient of determination are expressed in term of percentage of explained variation. The total effect gives the whole variation explained by the variable whereas the conditional effect gives the amount of variation explained once the effect of the other variables has been removed. The total variation explained by the delta model with all the retained variables is also indicated.



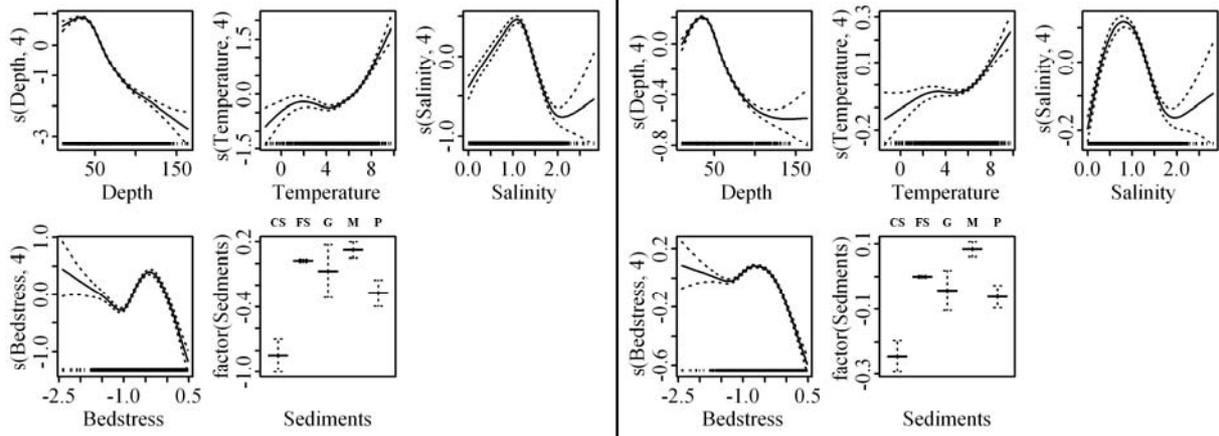
ACCEPTED MANUSCRIPT

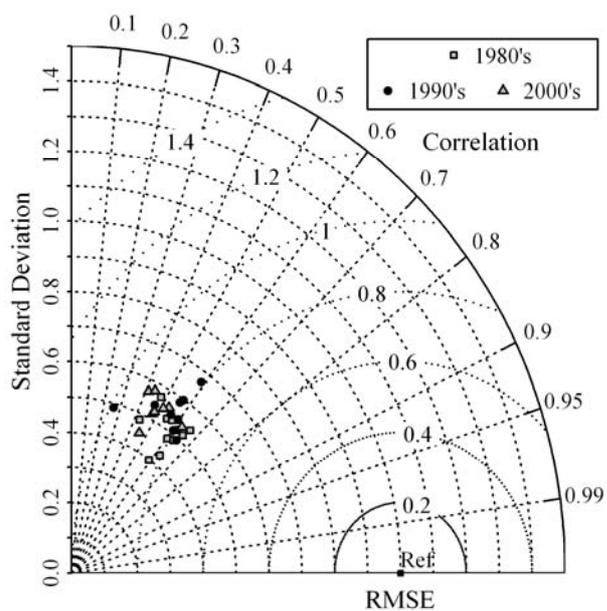


ACCEPTED MANUSCRIPT

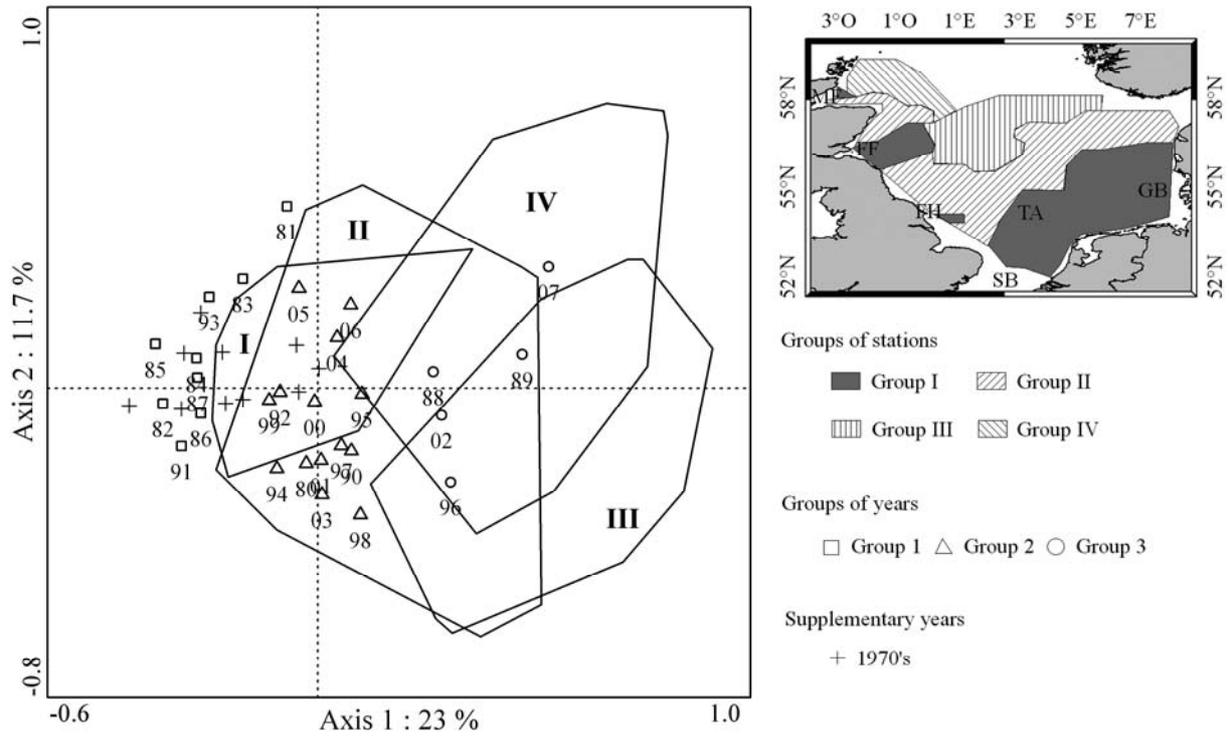


ACCEPTED MANUSCRIPT

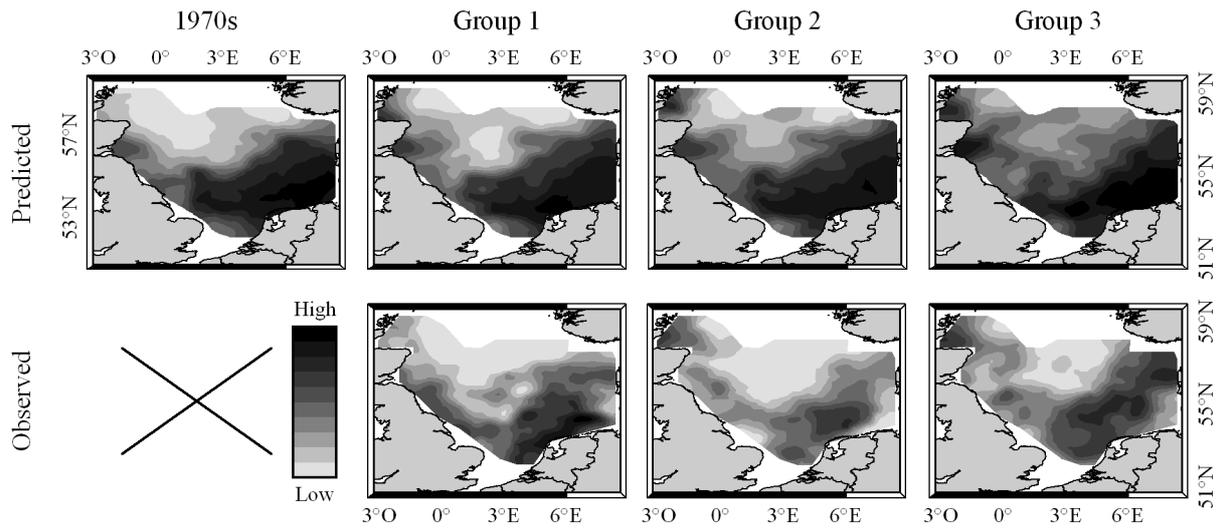




ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT



	Total effect	Conditionnal effect
All variables	25.5	-
Depth	18.5	13.5
Temperature	2.6	0.9
Salinity	11.5	3.3
Bedstress	1.9	1.2
Sediments	1	0.6