

# Mapping plankton distribution in the Bay of Biscay during three consecutive spring surveys

L. Zarauz<sup>1,\*</sup>, X. Irigoien<sup>1</sup>, A. Urtizbera<sup>2</sup>, M. Gonzalez<sup>1</sup>

<sup>1</sup>Marine Research Division, AZTI Foundation, Herrera Kaia Portualdea, z/g 20110, Pasaia (Guipuzkoa), Spain

<sup>2</sup>Department of Biology, University of Bergen, Høyteknologisenteret, PB 7800, 5020 Bergen, Norway

**ABSTRACT:** Nano-microplankton biomass, mesozooplankton biomass and hydrographic features were measured with mesoscale spatial resolution during 3 surveys in the Bay of Biscay between March and June 2004. Regions of high plankton biomass were associated with mesoscale physical structures. Generalized additive models (GAMs) based upon surface salinity, surface temperature and stratification of the water column explained 83 % of the variability in nano-microplankton biomass distributions, 67 % for small mesozooplankton, and 41 % for large mesozooplankton. The results show that when biological and physical data are collected at the relevant spatial scales, nano-microplankton and small mesozooplankton biomass distributions can be largely explained using simple hydrographic variables and non-parametric regression models.

**KEY WORDS:** Plankton biomass · Hydrography · Generalized additive models · GAM · Image analysis · Size spectra

*Resale or republication not permitted without written consent of the publisher*

## INTRODUCTION

Huntley's temperature dependent growth model (Huntley & Lopez 1992) has provoked an intense discussion within the scientific community about factors controlling zooplankton growth (Kleppel et al. 1996). However, much less attention has been paid to Huntley's observation that the variability in biomass was the dominant factor when estimating overall zooplankton production. Huntley & Lopez (1992) showed that, on scales relevant to the production of zooplankton, variability in biomass may be up to 6 orders of magnitude higher than variability in individual growth rates. Consequently, these investigations identified the necessity for more accurate measurements of biomass for application to production studies. However, obtaining quasi-synoptic measurements of plankton biomass at the relevant spatial resolution is not an easy task. Traditional methods of plankton collection and analysis have always been problematic when trying to measure plankton biomass at the mesoscale (0 to 100 km); a spatial scale at which physical processes are particularly energetic and have strong impacts on biological vari-

ability (Mann & Lazier 1991). The older methods are too expensive and too laborious for mesoscale-solving approaches. Satellite imaging methods provide high-resolution estimations of phytoplankton biomass using chlorophyll measurements (Soulanki et al. 2001), however, they fail to characterise the heterotrophic community (micro-mesozooplankton). Moreover, the non-linear complex relationships between biological and physical variables are difficult to describe with dynamic physical-biological models. Although there is a constant improvement in the ability of models to simulate plankton population dynamics, it is still difficult to reproduce plankton distributions and levels of biomass. Nonetheless, accurate biomass estimates are needed, both for carbon cycle analyses and higher trophic (e.g. fish recruitment) models.

These methodological and modelling problems may explain why the issues raised by Huntley & Lopez (1992) were not considered further until new analytical technologies and statistical tools became available. Recent developments have enabled us to measure plankton biomass at the same spatial scales as hydrographic phenomena. New image analysis systems are

\*Email: lzarauz@pas.azti.es

available to study the distribution of organisms, both autotrophic and heterotrophic, in aquatic ecosystems (Ashjian et al. 2001, Grosjean et al. 2004, Davis et al. 2005, See et al. 2005). By minimizing the time spent in sample analysis, a large amount of data can be obtained in a short period of time, increasing the potential spatial and temporal resolution of such studies. These systems sacrifice taxonomic detail (Hu & Davis 2005), but provide useful information on abundance and size that can be translated into biomass. Using complementary image analysis systems, a broad size range of particles from phytoplankton to mesozooplankton can be studied at the same spatial scales. Furthermore, new statistical methods that model spatial and temporal dependence in ecological data have been recently developed. As an example, generalized additive models (GAMs, Hastie & Tibshirani 1990) provide particularly powerful means for modelling the variation in a given variable as a function of explanatory factors. They accommodate continuous functional forms of almost any shape, and to a large degree, they allow the data to determine the most suitable shape to adopt (Augustin et al. 1998). With the use of spline smoothing functions, they are able to integrate multicollinearity between variables (Yee & Mitchell 1991) and to minimize the effects of extreme observations, which only influence the shape of the curve in the neighbourhood of those points (Wood 2006). The ability to handle the multicollinearity and non-linear relationships existing between variables is the important strength of GAMs, and makes additive modelling a powerful tool in the development of models that improve our understanding of ecological data.

GAMs can be used to predict distributions based upon habitat modelling (Guisan & Zimmermann 2000), and they have been applied successfully in terrestrial ecology (Guisan et al. 2002). In marine systems, statistical modelling has been applied mainly to fish distributions (Planque et al. 2007, Beare & Reid 2002), fish egg distributions (Augustin et al. 1998, Stratoudakis et al. 2003) and plankton species distribution (Beare et al. 2000); however, the potential of GAMs to describe plankton biomass distributions has been scarcely explored. The main objective of the present study was to investigate at relevant spatial scales the relationships between plankton biomass distribution and physical forcing in the Bay of Biscay.

## MATERIALS AND METHODS

**Study area.** The Bay of Biscay is part of the subtemperate eastern North Atlantic Ocean, and it can be defined as an open bay surrounded by the western French and Spanish coastlines. The bay is charac-

terised by strong heterogeneity in environmental forcing and seasonal variations. Topography, hydrographic characteristics, origin and variations of water masses in the Bay of Biscay have been reviewed by Koutsikopoulos & Le Cann (1996).

Three surveys were carried out in the Bay of Biscay during March/April, May, and June 2004. Stations were distributed in transects perpendicular to the coast covering the shelf and shelf-break (Fig. 1). The resolution and geographic limits of each cruise are given in Table 1.

**Sample collection and analysis. Hydrographic data:** The hydrographic characteristics and the fluorescence profiles of the stations were sampled using a conductivity temperature, depth profiler (CTD; model RBR XR420) fitted to the mesozooplankton net. The difference in seawater density between 100 m depth (or 5 m above the bottom) and the surface was used as an index of the water column stratification.

**Nano- and microplankton:** Samples were collected at a depth of 3 m using 1.5 l Niskin bottles. Samples were analyzed onboard using a FlowCAM (Sieracki et al. 1998) to determine the biomass and size structure of the nano- and microplankton community. Fluorescence measurements were not included in the analysis, hence every particle (phytoplankton, zooplankton, detritus, inorganic) was counted and imaged. For each sample, a maximum of either 2000 particles or 10 ml were analyzed. A  $\times 4$  objective was used in the sample analysis. The instrument was calibrated using beads of a known size. Invalid recordings (i.e. bubbles, repeated images) were removed from the image database through visual recognition. The biovolume of each cell was calculated from its equivalent spherical diameter (ESD), and was converted into biomass according to the equation of Montagnes et al. (1994) for marine diatoms and dinoflagellates.

**Mesozooplankton:** Samples were obtained using vertical hauls of a 150  $\mu\text{m}$  Paironet net, except during the March/April cruise when oblique hauls of a 335  $\mu\text{m}$  Bongo net were used. Because of the differences in the capture capacities of the 2 net systems, the zooplankton from the March/April cruise are not included in this study. Nets were lowered to a maximum depth of either 100 m, or 5 m above the bottom at the shallower stations.

Net samples were preserved immediately after collection with 4% borax buffered formalin. The samples were treated for 24 h with 4 ml 1% eosin, which stains cytoplasm and muscle protein effectively; this stain creates sufficient contrast for recognition by image analysis and reduces the counting of detrital material. Net subsamples were scanned in 254 (8 bit) colours at a resolution of 600 dpi using a HP Scanjet 8200 series (Hewlett-Packard) scanner.

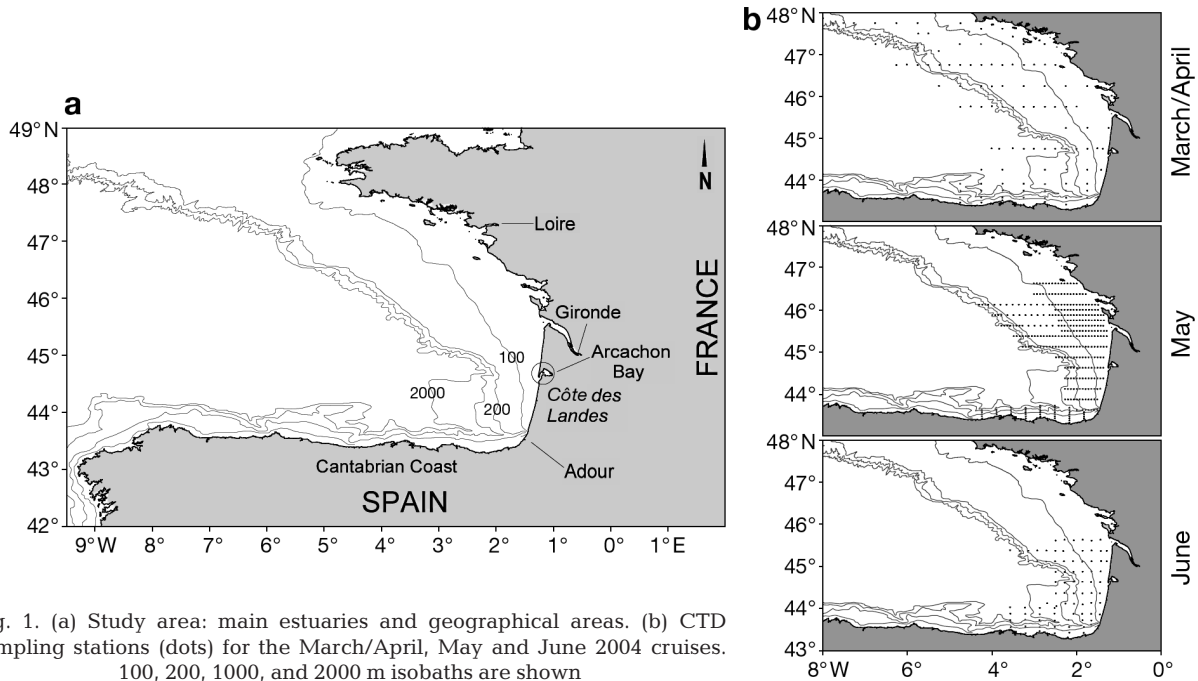


Fig. 1. (a) Study area: main estuaries and geographical areas. (b) CTD sampling stations (dots) for the March/April, May and June 2004 cruises. 100, 200, 1000, and 2000 m isobaths are shown

The resulting jpg images were imported into the Plankton Visual Analyzer (PVA), a custom image analysis software for counting, measuring and classifying objects in digital images and recently integrated into the free software Zooimage available at [www.sciviews.org/zooimage](http://www.sciviews.org/zooimage). The ESD of each particle  $>165 \mu\text{m}$  was detected and measured. The volume of each zooplankton was calculated from the ESD, converted to carbon, and corrected for shrinkage caused by formalin (Alcaraz et al. 2003).

For both nano-microplankton and mesozooplankton, the biomass of every individual was calculated. Very small and very large sizes were considered to be under-represented, because of the relatively low volumes analysed by the flowCAM and PVA, or under-sampling by the collection gear (bottles and nets). Thus, only size ranges represented in  $>90\%$  of the samples were included. Accordingly, nano-microplankton biomass was estimated for the size range  $7$  to  $27 \mu\text{m}$  ESD ( $16$  to  $1024 \text{ pg C particle}^{-1}$ ) and 2 groups were distinguished among mesozooplankton as a function of size, viz. small mesozooplankton, which included organisms from  $165$  to  $560 \mu\text{m}$  ESD ( $0.2$  to  $6.4 \text{ } \mu\text{g C particle}^{-1}$ ), and large mesozooplankton, which covered the size range from  $560$  to  $1100 \mu\text{m}$  ESD ( $6.4$  to  $51.2 \text{ } \mu\text{g C particle}^{-1}$ ).

Distributions of nano-microplankton and mesozooplankton communities, as well as hydrographic features, are described herein for the entire geographical area covered in each of the cruises. However, for comparison of biological and physical conditions between cruises, only the common area sampled during the 3

surveys was considered, in order to avoid changes due to different sampling areas and different variables affecting the community.

**Statistical analysis.** Nano-microplankton and mesozooplankton biomass were modelled as a function of surface salinity, surface temperature and stratification index, using GAMs (Hastie & Tibshirani 1990). The transformed biomass values, i.e.  $\log_{10}(\text{plankton biomass} + 1)$  were used to develop the GAMs. The hydrographic data from the 3 cruises were included in the nano-microplankton models. Due to the lack of data in March/April, only the May and June cruises were used for mesozooplankton. The analyses presented here were performed using the 'mgcv' library in the R statistical software available from the Comprehensive R Archive Network <http://cran.r-project.org>.

GAMs are a non-parametric extension of generalized linear models (GLMs), with the only underlying assumption that the functions are additive and that the components are smooth. The basic concept is replacement of the parametric GLM structure:

$$g(\mu) = \alpha + \sum_{j=1}^p \beta_j x_j \quad (1)$$

with the additive structure:

$$g(\mu) = \alpha + \sum_{j=1}^p s_j(x_j) \quad (2)$$

where the  $s_j$  are spline smooth functions, and  $p$  the number of explanatory variables ( $x_j$ ) (Hastie & Tibshirani 1990). The main strengths of GAMs are their flexibility and their ability to deal with highly non-linear and non-monotonic relationships between the response and the

Table 1. Date, geographical limits, number of stations sampled, plankton biomass mean values and range of plankton data (in brackets), for the cruises in March/April, May and June 2004. Note that biomass data refer to the common area (43.33°–45.64° N 3.57°–1.29° W) sampled in the 3 cruises. Nmi: nano-microplankton; Mzo: mesozooplankton

Cruise	Dates 2004		Limits		Stations (n)			Common area biomass (mg C m <sup>-3</sup> )		
	Start	End	Latitude	Longitude	Nmi	Mzo	CTD	Nmi	Small Mzo	Large Mzo
Mar/Apr	3 Mar	4 Apr	43.41°–47.75° N	7.75°–1.24° W	135	103	107	168 (34–1049)	—	—
May	5 May	18 May	43.37°–46.63° N	4.31°–1.29° W	174	347	314	297 (26–861)	2 (0.26–13)	3 (0.32–12)
June	21 Jun	28 Jun	43.32°–45.64° N	3.57°–1.29° W	78	84	75	337 (81–929)	4 (0.01–22)	2 (0.01–8)

set of explanatory variables. This is mainly because GAMs are data-driven rather than model-driven, and they allow the data to determine the shape of the response curves, rather than being limited by an *a priori* model. Moreover, GAMs allow the possibility of using a higher dimensional smoother to model complex interactions between variables (Yee & Mitchell 1991). In the case of 2 or 3 variables, this would mean fitting:  $g(\mu) = \alpha + s(x_1, x_2)$ , or  $g(\mu) = \alpha + s(x_1, x_2, x_3)$ .

As a first step, GAMs were based upon single explanatory variables to study the influence of individual hydrographic covariates on the plankton biomass distribution. During further analysis, GAMs of increasing complexity were applied, combining multiple explanatory variables. Interrelations between hydrographic covariates, together with their influence on the response variable (plankton biomass), were included within the models using 2- and 3-dimensional smoothing functions.

Comparisons between models were performed in order to select GAM smoothing predictors and smoothing functions. Three criteria were used (Wood 2006): (1) the generalized cross validation score (GCV, the lower the better); (2) the level of deviance explained (0 to 100%, the higher the better); and (3) the confidence region for the smoothing (which should not include zero throughout the range of the predictor). These selection criteria are an adaptation of the method proposed by Wood & Augustin (2002), and have already been applied to GAM model selection when using the 'mgcv' library (Planque et al. 2007).

One of the risks of applying GAMs with few data is 'overfitting' the model by using many parameters; in general, bias decreases and variance (uncertainty) increases as the number of parameters in a model increases (Burnham & Anderson 2002). Accordingly, a Jackknife (JK) procedure was applied to validate the final models, using an independent data set (Lobo & Martín-Piera 2002). For the complete data set (309 stations for nano-microplankton and 354 stations for mesozooplankton), GAMs were recalculated omitting one station at a time. Each of these models based on the  $n-1$  stations was applied to the excluded stations in turn. The p-value and the coefficient of determination

( $r^2$ ) of the least squares linear regression between GAM-predicted and observed biomass data were used to validate the model.

## RESULTS

### Hydrography

In late winter (March/April), the hydrography of the continental shelf was influenced mainly by the presence of cold waters originating from the main estuaries of the area (the Loire, Gironde and Adour). Surface temperature and salinity increased with distance from the shore, and stratification was associated with the plumes from the estuaries (Fig. 2). Surface temperature ranged from 9.7 to 12.7°C (10.9 to 12.6°C in the common area), with minimum values along the French coast, from the Gironde to the Loire estuaries. The water column was mixed, and stratification had a maximum value of 2.5, associated with the Loire River mouth.

In May, there was a general increase in surface temperature, with values ranging from 12.6 to 15.7°C (12.6 to 14.9°C in the common area). The warmest temperatures were recorded over the northern part of the shelf. This area was sampled at the end of the cruise and, by this time, thermal stratification had already commenced. The stratification index was highest (5.6) in the mouth of the Gironde estuary, which had a plume extending over the platform towards the north. Stratification offshore from the Adour estuary also was well-defined during this period.

The Gironde estuary had important seasonal variability in flow, as can be observed in Fig. 3. May is one of the periods with higher outflows, in contrast to March/April and June, when the discharges were relatively low.

In June, high surface temperature water (ranging from 17 to 22°C) covered all the sampled area. Generalized thermal stratification extended over the southern part of the shelf, with a minimum of 1.6 and a maximum of 4 (Fig. 2). A narrow strip of water characterised by lower temperature and stratification oc-

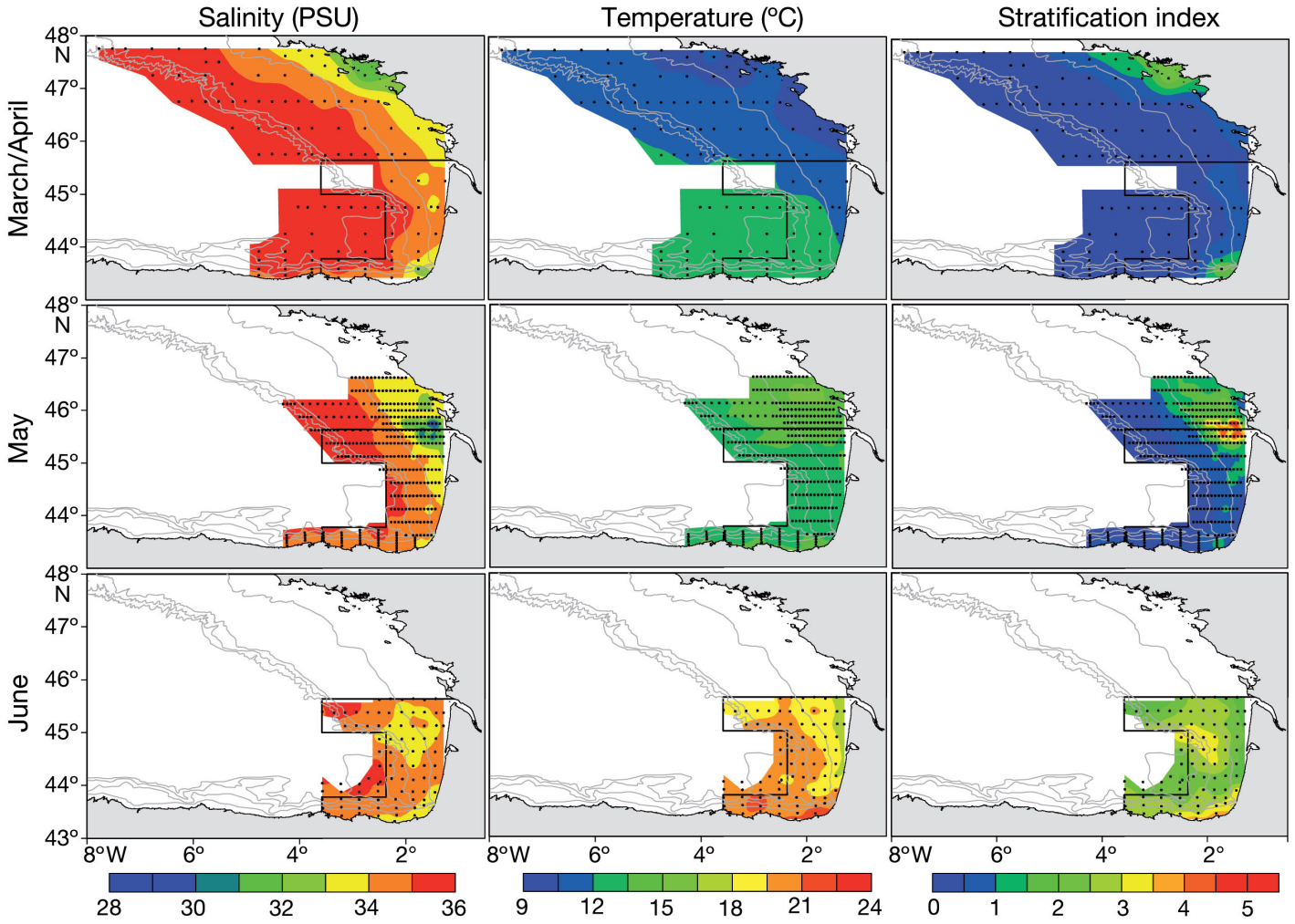


Fig. 2. Spatial distribution of sea surface salinity (PSU), surface temperature (°C) and stratification index, for March/April, May and June cruises. The sampling grid is plotted and the common area sampled in the 3 cruises (used for comparisons) is delimited by a black line. Bathymetry as in Fig. 1

curred along the southern French coast. Upwelling events occur frequently over this area (Pingree 1984, Koutsikopoulos & Le Cann 1996) and may be the cause of this feature.

Average surface temperature and stratification values increased significantly ( $p < 0.0001$ ) from March to June (Fig. 4). Data used in this comparison were selected from the common area sampled during the 3 cruises (Fig. 2).

**Nano-microplankton and mesozooplankton biomass**

Nano-microplankton biomass concentration ranged from 10 to 1174 mg C m<sup>-3</sup>

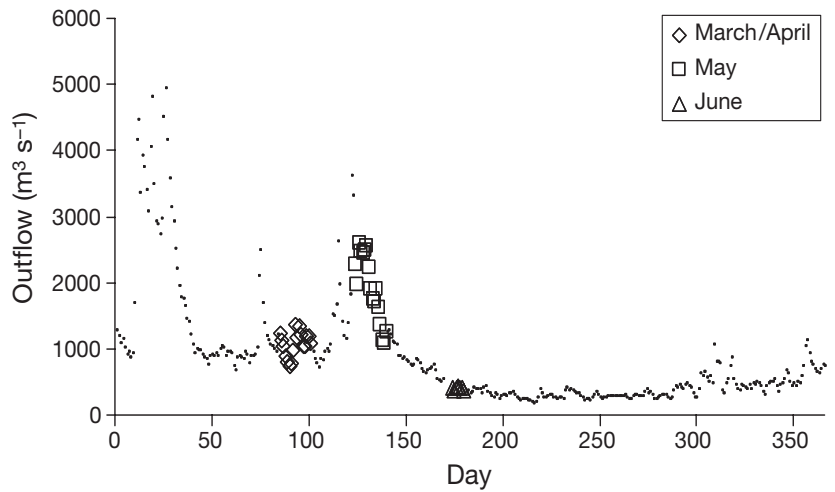


Fig. 3. Daily out-flow of the Gironde estuary during 2004. Different symbols correspond to the dates of cruises

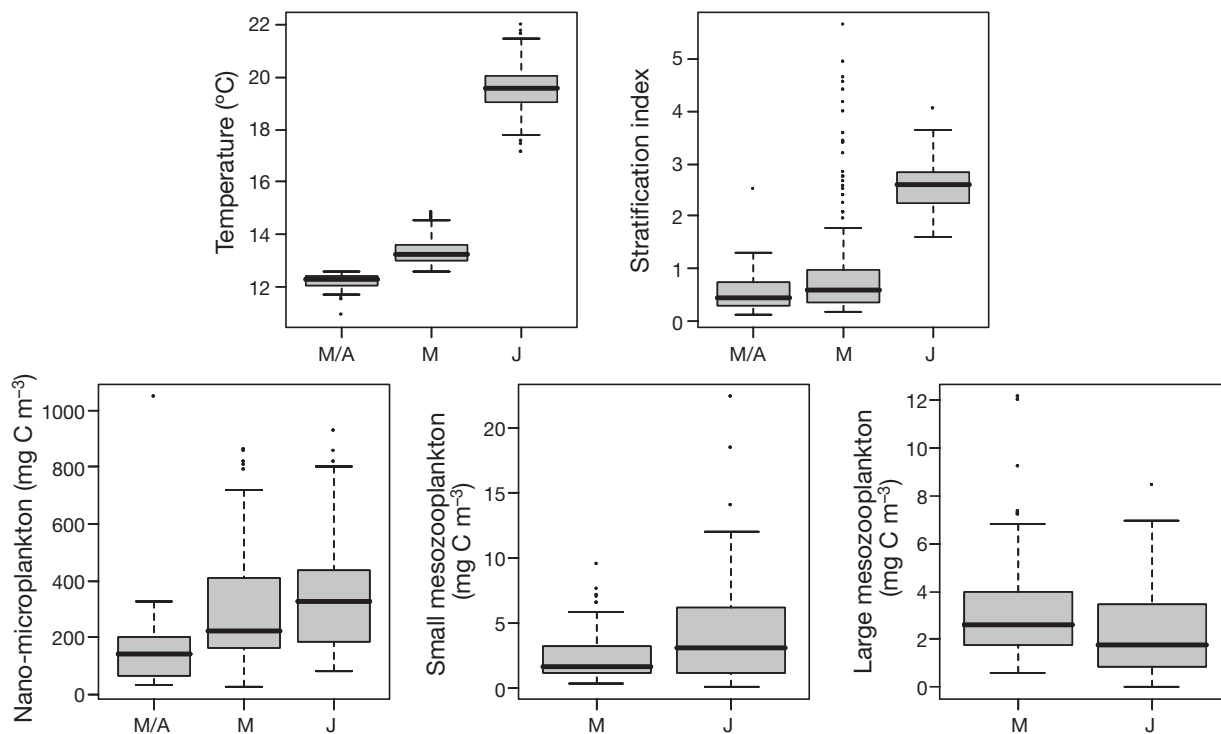


Fig. 4. Boxplot distributions of physical and biological variables during March/April (M/A), May (M) and June (J). Temperature (°C); stratification index; nano-microplankton biomass ( $\text{mg C m}^{-3}$ ); small mesozooplankton biomass ( $\text{mg C m}^{-3}$ ); and large mesozooplankton biomass ( $\text{mg C m}^{-3}$ ). Boxes indicate median (thick black line) and 25 and 75 percentiles. Whiskers expand to 1.5 times the interquartile range. Note: only the common area sampled during the 3 cruises has been used for this comparison

(26 to  $1049 \text{ mg C m}^{-3}$  in the common area) (Table 1). Small and large mesozooplankton biomass concentration ranged from 0.01 to 22 and from 0.01 to  $12 \text{ mg C m}^{-3}$ , respectively (Table 1). The 3 distributions showed seasonal variations, with important differences between the cruises.

In March/April, well-mixed waters of the Bay of Biscay had relatively low biomass of nano- and microplankton, with an intense peak of  $1049 \text{ mg C m}^{-3}$  adjacent to the Gironde river mouth (Fig. 5). Two more areas of higher biomass (around  $400 \text{ mg C m}^{-3}$ ) were identified, one associated with the Loire estuary runoff; and a second located in front of the Cantabrian shelf-break, a region where eddies form at this time of the year (Fig. 6).

In May, with the general warming of the sea surface, the mean value of the nano-microplankton biomass increased significantly. Nano-microplankton biomass extended over the French shelf, with several areas of higher biomass associated with the Gironde river plume, the southeast part of the shelf (Côte des Landes) and the northern shelf-break. Biomass of small mesozooplankton was higher near the Gironde and along the Côte des Landes. Biomass of large mesozooplankton was more uniform over the French shelf, with peaks of biomass at the shelf break and in the Gironde river plume (Fig. 5).

In June, nano-microplankton and small mesozooplankton had similar patterns, with areas of higher biomass in the north of the Bay of Arcachon and in the Côte des Landes, possibly related to upwelling processes. The mouth of the Gironde estuary was characterised by higher levels of biomass (Fig. 5). The main difference between the distributions was the higher biomass of small mesozooplankton over the narrow Cantabrian shelf. By this time, there was a general decrease in the large mesozooplankton biomass. However, some areas of slightly higher biomass occurred, just offshore of the Bay of Arcachon and on the Cantabrian shelf.

It is notable that during all the surveys high biomass areas were associated clearly with waters influenced by river outflows and other physical fronts, such as eddies and upwellings. Moreover, overall community biomass increased gradually, covering the French shelf as surface temperature and stratification increased.

Kruskal-Wallis tests applied to the common area of the surveys showed a significant ( $p < 0.0001$ ) increase of nano-microplankton biomass from March/April ( $168 \text{ mg C m}^{-3}$ ) to June ( $337 \text{ mg C m}^{-3}$ ) (Fig. 4). A significant increase in biomass of small mesozooplankton ( $p < 0.05$ ), together with a significant decrease of large mesozooplankton ( $p < 0.005$ ), occurred from May to June.

### GAM models based on individual explanatory variables: relationship between hydrographic parameters and plankton biomass

Surface temperature was the physical parameter which best explained the nano-microplankton biomass distribution, with the highest deviance explained (43%) and the lowest GCV (0.45) (Table 2). Surface salinity and stratification also explained a high percentage of the deviance (34% each). For the small mesozooplankton biomass distribution, surface salinity was the most important factor, describing 43% of the deviance; surface temperature and stratification explained 18 and 20%, respectively. Large mesozooplankton biomass distribution variability was poorly described by physical variables: surface temperature, salinity and stratification explained 14, 7 and 5% of the deviance, respectively.

Non-linear dependence in the data was handled within the GAMs using spline functions. The smoothed fits indicate the effects of individual hydrographic variables on the plankton biomass distributions (Fig. 7). Approximate confidence interval envelopes were plot-

ted for each function. This is useful because it gives an indication of the parts of the function that are less accurately estimated, often because of fewer data points (Yee & Mitchell 1991). In our calculations, lower accuracy was observed in the estimations for lower values of salinity, higher values of stratification and intermediate values of surface temperature.

**Salinity:** Both nano-microplankton and small mesozooplankton biomass tended to increase in low salinity areas, such as the plumes of the main estuaries (Fig. 7a,d). Slightly higher levels of large mesozooplankton biomass were associated with both low and high salinity values (Fig. 7g), and may be related to the river plume and the shelf break fronts. However, the results for large mesozooplankton must be interpreted with care, because the model explains only 7% of the deviance, and its confidence limits include zero in almost all salinity values.

**Surface temperature:** An increase in nano-microplankton biomass occurred between 14 and 18°C, whereas for small mesozooplankton, the maximum occurred around 16 and 18°C (Fig. 7b,e). Large mesozooplankton experienced a slight increase at temperatures near 14°C (Fig. 7h). The upper and lower limits of

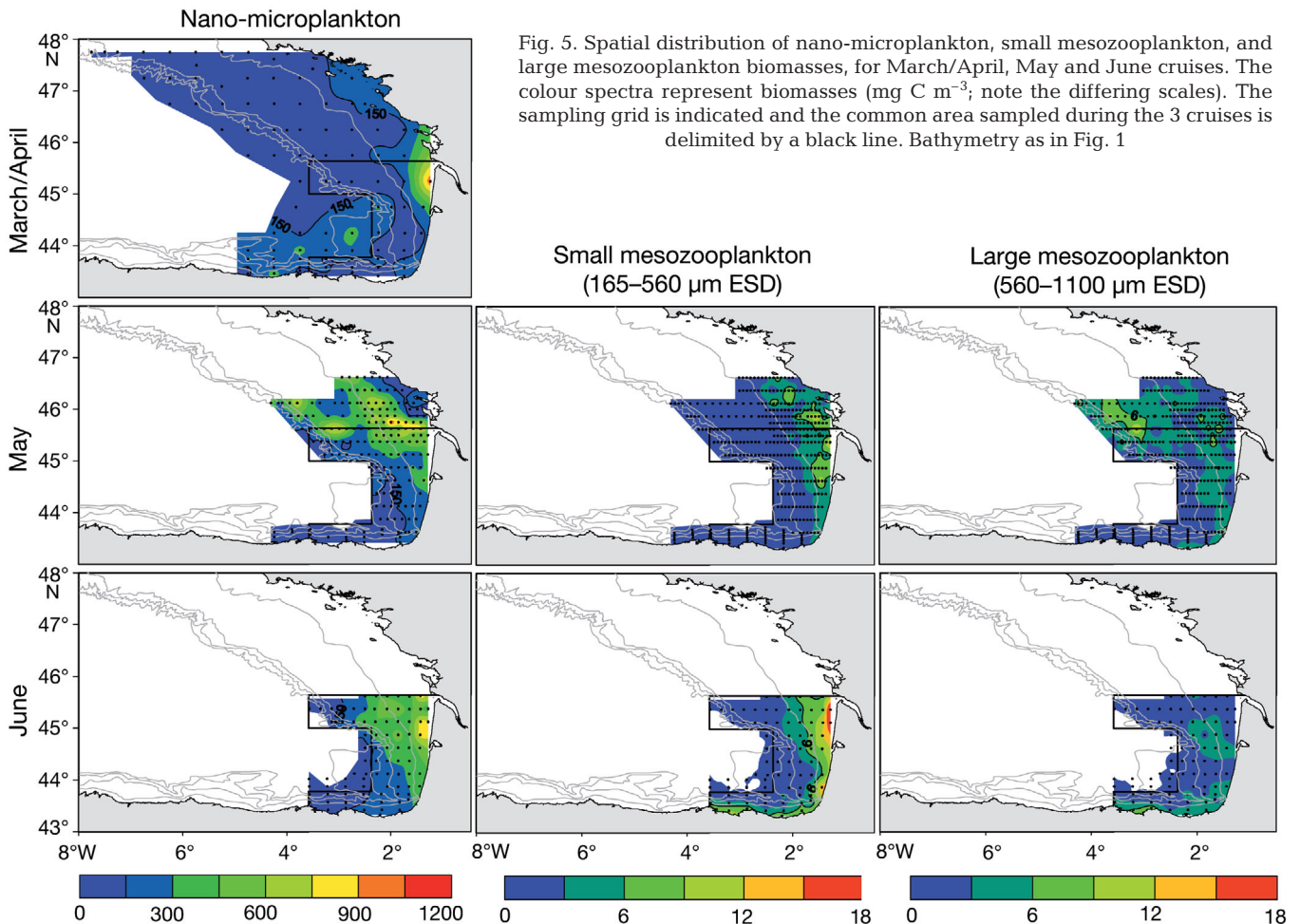


Fig. 5. Spatial distribution of nano-microplankton, small mesozooplankton, and large mesozooplankton biomasses, for March/April, May and June cruises. The colour spectra represent biomasses ( $\text{mg C m}^{-3}$ ; note the differing scales). The sampling grid is indicated and the common area sampled during the 3 cruises is delimited by a black line. Bathymetry as in Fig. 1



Fig. 6. True-colour image of a phytoplankton bloom in the Bay of Biscay. Image taken by the MODIS sensor onboard the 'Terra' satellite, 25 April 2004. Image courtesy of MODIS Rapid Response Project at NASA/GSFC (<http://rapidfire.sci.gsfc.nasa.gov/>)

the 3 distributions cannot be defined clearly, as the confidence interval of these values of temperature include zero.

**Stratification:** Nano-microplankton and small mesozooplankton biomass were related to water column stratification. Fig. 7c,f shows a positive relationship between stratification and biomass, with a first peak around 1.5 and a continuous increase in biomass associated with higher values of stratification. Slightly lower values of biomass of large mesozooplankton occurred at intermediate levels of stratification, between 2 and 3 (Fig. 7i), but this trend is weak and its confidence limits include zero in the majority of the stratification values.

#### **GAM models on multiple explanatory variables: the interaction between hydrographic parameters**

GAMs based upon multiple variables of increasing complexity were established. These GAMs were ranked according to the percentage of deviance they

could explain and their GCV scores (Table 2). The models based on a 3-dimensional smoothing function that included the interactions between the 3 covariates (GAM, highlighted in Table 2) appeared to provide the best fits. They explained 83% of the variability in nano-microplankton, 67% in small mesozooplankton and 41% in large mesozooplankton biomass distributions, and had the lowest GCV in comparison with the rest of the models (0.22, 0.16, and 0.22, respectively). Finally, the results of the model validation showed strong correspondence between observed nano-microplankton and small mesozooplankton biomass values and those predicted by the Jackknife procedure ( $r^2 = 0.70$ ,  $p < 0.001$ ; and  $r^2 = 0.51$ ,  $p < 0.001$ , respectively). Correlation between large mesozooplankton observed and predicted values were lower ( $r^2 = 0.23$ ), but still significant ( $p < 0.001$ ).

The lower overall deviance explained for large mesozooplankton data is consistent with the results obtained from the single variable GAMs. Moreover, stratification processes identified during the surveys were due to changes in density, which were related to salinity and/or temperature gradients in the water column.



Table 2. Single and multiple variable based GAMs, for nano-microplankton, small mesozooplankton, and large mesozooplankton biomass. For each model, the percentage of deviance explained (Dev.), the generalized cross validation score (GCV), and the parameters of the Jackknife procedure,  $r^2$  (JK- $r^2$ ) and p-value (JK-p), are given. The smoothing functions (s) and explanatory variables used in each model appear in brackets: **S**, surface salinity; **T**, surface temperature; **Str**, stratification index. The model analyzed in the discussion is highlighted in grey

Model	Nano-microplankton				Mesozooplankton							
	7–27 $\mu\text{m}$ ESD				165–560 $\mu\text{m}$ ESD				560–1100 $\mu\text{m}$ ESD			
	Dev. %	GCV	JK- $r^2$	JK-p	Dev. %	GCV	JK- $r^2$	JK-p	Dev. %	GCV	JK- $r^2$	JK-p
~s( <b>S</b> )	34.1	0.52	0.32	<0.001	42.7	0.22	0.41	<0.001	7.3	0.27	0.04	<0.001
~s( <b>T</b> )	43	0.45	0.40	<0.001	17.6	0.32	0.13	<0.001	14.3	0.25	0.10	<0.001
~s( <b>Str</b> )	34	0.52	0.31	<0.001	19.6	0.31	0.17	<0.001	5.2	0.28	0.03	0.002
~s( <b>S</b> )+s( <b>T</b> )	53.5	0.38	0.50	<0.001	42.9	0.22	0.50	<0.001	17.2	0.24	0.09	<0.001
~s( <b>T</b> )+s( <b>Str</b> )	54	0.38	0.49	<0.001	33.8	0.26	0.33	<0.001	15.4	0.25	0.10	<0.001
~s( <b>S</b> )+s( <b>Str</b> )	45.9	0.45	0.40	<0.001	43.3	0.22	0.41	<0.001	14.4	0.25	0.10	<0.001
~s( <b>S</b> )+s( <b>T</b> )+s( <b>Str</b> )	59.2	0.36	0.52	<0.001	44.2	0.22	0.50	<0.001	18.1	0.24	0.11	<0.001
~s( <b>S,Str</b> )	64	0.32	0.58	<0.001	53.3	0.20	0.46	<0.001	31.9	0.22	0.21	<0.001
~s( <b>T,Str</b> )	63.9	0.32	0.58	<0.001	47.4	0.22	0.40	<0.001	26.3	0.23	0.16	<0.001
~s( <b>S,T</b> )	70	0.27	0.65	<0.001	58.3	0.17	0.53	<0.001	30.9	0.22	0.20	<0.001
~s( <b>S,Str</b> )+s( <b>T</b> )	66.7	0.3	0.60	<0.001	62	0.17	0.54	<0.001	32.6	0.22	0.20	<0.001
~s( <b>T,Str</b> )+s( <b>S</b> )	67.8	0.3	0.61	<0.001	58.5	0.18	0.51	<0.001	28.6	0.23	0.16	<0.001
~s( <b>S,T</b> )+s( <b>Str</b> )	70.8	0.26	0.65	<0.001	58.7	0.17	0.53	<0.001	32.4	0.22	0.19	<0.001
~s( <b>S,T,Str</b> )	82.7	0.22	0.70	<0.001	67.5	0.16	0.52	<0.001	41.3	0.22	0.23	<0.001

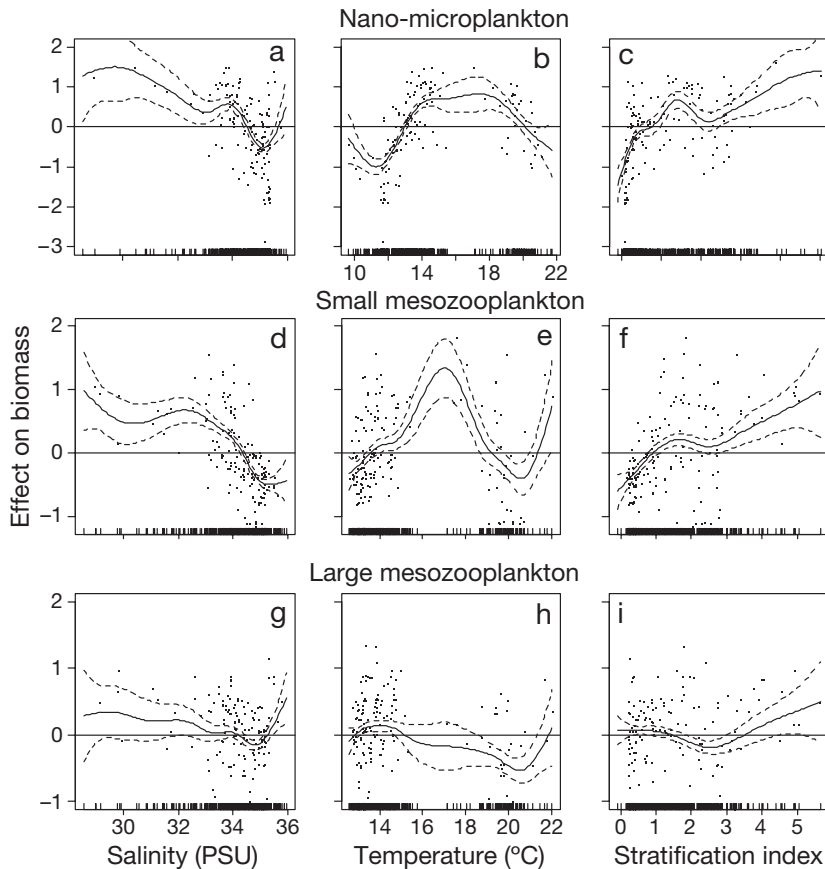


Fig. 7. Outputs of the GAMs based on single variables: (a,d,g) salinity (PSU); (b,e,h) temperature ( $^{\circ}\text{C}$ ); and (c,f,i) stratification index for nano-microplankton, small mesozooplankton and large mesozooplankton biomasses. Broken lines correspond to 2 SE above and below the smoothed estimate plots. Short vertical lines located on the x-axes of each plot indicate the values at which observations were made

We can confirm that the interactions between the 3 variables have the ability to describe plankton distribution to a greater extent than the separate main effects. It is important to note that multi-dimensional splines, though useful for modelling complex interactions between variables, do not allow an examination of the contribution of each variable separately (Yee & Mitchell 1991).

The final models were applied to the original hydrographic data set. The plot made with the estimated plankton biomass for each cruise individually (Fig. 8) agreed well with the measured biomass distribution for the first 2 plankton groups. The models highlighted the same high biomass areas described above for the nano-microplankton and small mesozooplankton communities. The GAM for large mesozooplankton, however, reproduced the general distribution of the community, but failed to describe the magnitude and extension of the areas where higher biomass occurred.

## DISCUSSION

Within the present study, physical and biological variables were measured at a spatial resolution sufficient to identify the main hydrographical structures in the

area, together with the plankton distribution responses to them (Figs. 2 & 5). The average resolution of the surveys was <12 nautical miles, which was considered sufficient to obtain a realistic pattern of the plankton distribution in the Bay of Biscay (Albaina & Irigoien 2004). Using the FlowCAM and the PVA image analysis techniques, a large amount of plankton data was processed, the analysis of which would have been overwhelming with traditional methodology. *In situ* and laboratory-based image analysis methods have been used previously to obtain relevant information on plankton abundance, size and biomass (Ashjian et al. 2001, Grosjean et al. 2004, Davis et al. 2005, See et al. 2005).

Previous studies on the microbial community in the area under investigation have focused generally on either phytoplankton (Varela 1996) or microzooplankton (Quevedo & Anadon 2000). However, the flowCAM analysis represents an overall estimation of biomass, including autotrophic and heterotrophic organisms ranging from 7 to 27  $\mu\text{m}$  ESD. Detrital material of this size was also included in the analysis as it forms part of the energy transfer pathways in the

pelagic food web (Roy et al. 2000). However, technical limitations of the instrument must be considered, particularly for non-organic particle counts. Biomass estimations for waters near the mouth of the estuaries and the coastline are potentially overestimated due to the counting of large (>7  $\mu\text{m}$  ESD) inorganic suspended particles discharged by the rivers.

Our mesozooplankton biomass data fell between the known average ranges for the Bay of Biscay (Poulet et al. 1996, Valdes & Moral 1998). Mesozooplankton-sized marine snow was not included in the analysis, due to the net collection of zooplankton, which destroys aggregates, and eosine used in sample processing, which stains only zooplankton muscle proteins.

The most relevant physical features which could affect the distribution of plankton in the study area are coarse and mesoscale processes (Fig. 2). Such processes relate to recurrent well-documented structures with an important seasonal component, as follows: (1) the plumes of the main estuaries in the Bay (Loire, Gironde and Adour), which are related to river run-off

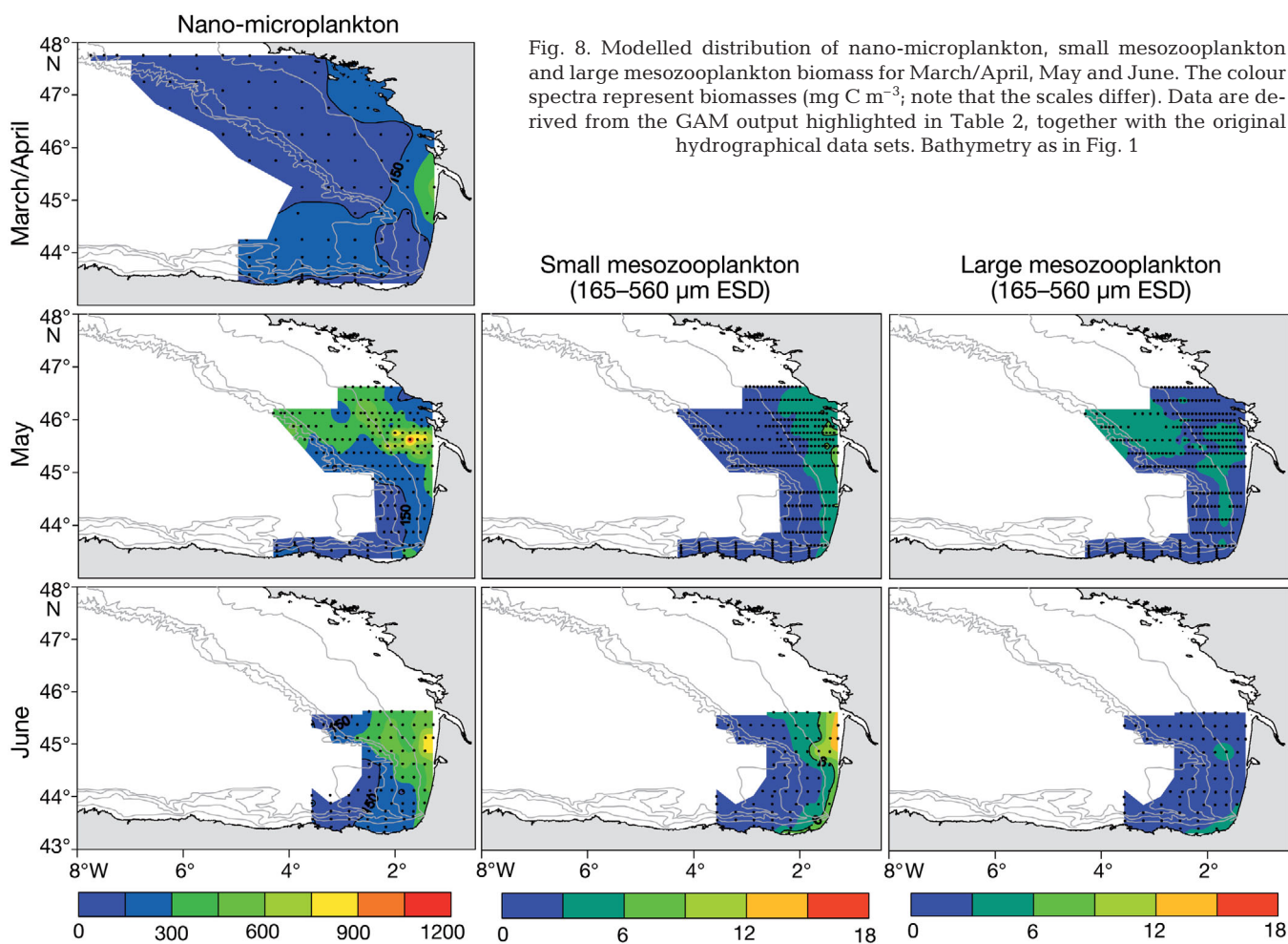


Fig. 8. Modelled distribution of nano-microplankton, small mesozooplankton and large mesozooplankton biomass for March/April, May and June. The colour spectra represent biomasses ( $\text{mg C m}^{-3}$ ; note that the scales differ). Data are derived from the GAM output highlighted in Table 2, together with the original hydrographical data sets. Bathymetry as in Fig. 1

and advection (Fig. 3); (2) the shelf-break front, where internal waves are generated by the interaction between tides and bottom topography (Pingree & Mardell 1981, Pingree et al. 1986, New 1988, Pingree & New 1995); (3) offshore eddies, developing mostly in winter due to instabilities in the poleward slope current (Pingree & Le Cann 1992, van Aken 2002); and (4) occasional summer upwelling events, which depend upon wind regime and are present along the Côte des Landes (Castaing & Lagardere 1983, Jegou & Lazure 1995).

Mesoscale structures are often considered as favourable environments for enhanced productivity (Le Fevre 1986). River outflows provide a continuous input of nutrients through continental drainage. The propagation of internal waves through the thermocline results in an upward transport of cold nutrient-rich waters (Pingree & Mardell 1981, Holligan et al. 1985, Pingree et al. 1986, New 1988). The dynamics of eddies enhances productivity and induces horizontal advection by the pumping of deep waters to the surface (Fernandez et al. 2004). Finally, upwelling events are defined as the processes that bring deeper waters to the surface (Sverdrup et al. 1942, Rochford 1991).

The processes accounting for enhanced production at fronts are diverse. They involve physiological responses of the organisms to the frontal environment (increased nutrient uptake, increased growth rates, depressed photosynthetic response, etc.), as well as the effects of floating, sinking or swimming with convergent and divergent flows at the front (Franks 1992, Lennert-Cody & Franks 2002). In our survey, the main areas of nano-microplankton and mesozooplankton biomass were connected to the supply of nutrients and the hydrodynamics due the mesoscale structures described above (Fig. 5).

GAMs were used in our study to integrate all the information available and to model their interactions. GAMs are statistical models; they are not expected to describe realistic cause-effect relationships, nor to provide explanations about underlying mechanisms for correlations between plankton biomass dynamics and hydrography. However, GAMs highlight significant correlations between the explanatory factors and the response variable (i.e. plankton biomass) in such a way that a significant percentage of the variability in plankton biomass distribution can be described by them. This fact emphasizes the importance of the relationships between physical forcing and biological dynamics, and demonstrates the power of additive modelling as a tool in the development of less parameterized habitat models.

A recent macroecological study (Li et al. 2006) showed that up to 75% of the variance of a community variable such as phytoplankton abundance can be

explained by a single predictor (temperature). This is an example of holistic simplicity arising from underlying complexity. Our research shows that even at the mesoscale, simple hydrographical variables can explain a large percentage of the plankton distribution deviance when data are collected at the appropriate spatial scales and analysed with non linear statistical methods.

Even in regions where single hydrographic variables did not initially reveal specific features, but where higher levels of biomass were found (e.g. eddies in March/April, shelf-break front in May, and upwelling in the northern part of Arcachon Bay in June), the model was able to extract correct information from the salinity, surface temperature and stratification combinations, so as to reproduce the increase of biomass in the regions of dynamic hydrographical features (Fig. 8). In particular, it is interesting to note that the GAM models were able to extract such information and, likewise, to describe correctly increases in nano-microplankton biomass, without having information available on nutrients and light.

The GAM analysis showed that nano-microplankton and small mesozooplankton biomass were influenced by physical variables to a greater degree than large mesozooplankton. This suggests that more complex ecological parameters and relationships may be involved in large mesozooplankton distributions. Several factors might explain this difference. In addition to the physical and chemical properties of the environment, spatial heterogeneity of large mesozooplankton populations is largely determined by the physiological and behavioural properties of the organisms themselves. These encompass interactions between individuals, and the reactions of the zooplankters to their biological environment, including responses to patches of potential food organisms and to predators (Mauchline 1998). Other habitat characteristics, such as depth, may also have a considerable influence on mesozooplankton distributions. As an example, depth preference of large copepods for mid-outer shelf areas is a component of reproductive success (Uye 2000). Size and taxonomy are important factors controlling mesozooplankton distribution over the continental shelf (Albaina & Irigoien 2004), and there are significant differences between day and night abundances of large copepods (Shaw & Robinson 1998). Moreover, there is an intense peak in large mesozooplankton biomass at the shelf break. As the limits of the May and June cruises did not extend beyond the shelf break, our sampling grid is not broad enough to adequately envelop all the variability in the large mesozooplankton distribution patterns, especially in open waters. Broader surveys covering the shelf, shelf-break and oceanic waters will provide relevant information to reproduce the factors limiting distribution. In addition,

further research on image recognition to provide taxonomic information, and the inclusion of more explanatory variables (predator and food abundance, depth, light intensity, etc.) are needed to improve mesozooplankton models.

These simple habitat mapping statistical models we used may be useful for inclusion in the rhomboidal modelling approach proposed by de Young et al. (2004). This approach, conceived to handle the intricate food webs in marine systems, concentrates the biological resolution of the model onto a target species of primary interest; and it decreases the resolution, both up and down the trophic scale with simplified models. De Young et al. (2004) anticipated that different classes of model would be required for the distinct levels of biological resolution. However, both the simplified models for 'food' and 'predators' must provide correct estimations of density or biomass in order to include them in the high-resolution model of the target species. The results of the present study show that statistical habitat mapping techniques have the potential to provide accurate biomass estimations of prey, predators, or other factors affecting target species, at least in the nano-microplankton and small mesozooplankton size range.

*Acknowledgements.* We are grateful to the captains and crew of the RV 'Investigador' and RV 'Vizconde de Eza', and to the onboard scientists and analysts for their support during sampling. Special thanks are due to: E. Martinez and N. Serrano for their help in processing the samples; G. Chust and L. Ibaibarriaga for their help and advice with statistics; Y. Sagarminaga and G. Chust, for providing satellite images of the Bay of Biscay; and V. Valencia and the Harbour Authority of Bordeaux for the Gironde river flow data. This research was supported by several projects funded by the Spanish Ministry of Education and Research, the Department of Agriculture and Fisheries of the Basque Country Government and the European Commission (EIPZI, PRESTIGE, BIOMAN, TRIENNIAL). L.Z.'s work was supported by a doctoral fellowship from the Education, Universities and Research Department of the Basque Country Government. X.I. was supported partially by a Ramon y Cajal Grant from the Spanish Ministry of Education and Research. A.U. was supported by a doctoral fellowship from the Fundación Centros Tecnológicos Iñaki Goenaga. Finally, thanks are extended to M. Collyns (SOES, University of Southampton and AZTI Tecnalia) for critical comments on the manuscript, and to the 3 anonymous reviewers for their useful comments. This paper is contribution no. 375 from AZTI-Tecnalia (Marine Research).

#### LITERATURE CITED

- Albaina A, Irigoien X (2004) Relationships between frontal structures and zooplankton communities along a cross shelf transect in the Bay of Biscay (1995 to 2003). *Mar Ecol Prog Ser* 284:65–75
- Alcaraz M, Saiz E, Calbet A, Trepal I, Broglio E (2003) Estimating zooplankton biomass through image analysis. *Mar Biol* 143:307–315
- Ashjian CJ, Davis CS, Gallager SM, Alatalo P (2001) Distribution of plankton, particles, and hydrographic features across Georges Bank described using the Video Plankton Recorder. *Deep-Sea Res II* 48:245–282
- Augustin NH, Borchers DL, Clarke ED, Buckland ST, Walsh M (1998) Spatiotemporal modelling for the annual egg production method of stock assessment using generalized additive models. *Can J Fish Aquat Sci* 55:2608–2621
- Beare DJ, Reid DG (2002) Investigating spatio-temporal change in spawning activity by Atlantic mackerel between 1977 and 1998 using generalized additive models. *ICES J Mar Sci* 59:711–724
- Beare DJ, Gislason A, Astthorsson OS, McKenzie E (2000) Assessing long-term changes in early summer zooplankton communities around Iceland. *ICES J Mar Sci* 57: 1545–1561
- Burnham KP, Anderson DR (2002) Model selection and multi-model inference: a practical information-theoretic approach. Springer, New York
- Castaing P, Lagardere F (1983) Seasonal water temperature and salinity variations off the Aquitaine continental shelf. *Bull Inst Geol Bass Aquitaine* 33:61–69
- Davis CS, Thwaites FT, Gallager SM, Hu Q (2005) A three axis fast-tow digital Video Plankton Recorder for rapid surveys of plankton taxa and hydrography. *Limnol Oceanogr Meth* 3:59–74
- De Young B, Heath M, Werner F, Chai F, Megrey B, Monfray P (2004) Challenges of modeling ocean basin ecosystems. *Science* 304:1463–1466
- Fernandez E, Alvarez F, Anadón R, Barquero S and 15 others (2004) The spatial distribution of plankton communities in a Slope Water anticyclonic Oceanic eDDY (SWODDY) in the southern Bay of Biscay. *J Mar Biol Ass UK* 84:501–517
- Franks PJS (1992) Sink or swim: accumulation of biomass at fronts. *Mar Ecol Prog Ser* 82:1–12
- Grosjean P, Picheral M, Warembourg C, Gorsky G (2004) Enumeration, measurement, and identification of net zooplankton samples using the ZOOSCAN digital imaging system. *ICES J Mar Sci* 61:518–525
- Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecol Model* 135:147–186
- Guisan A, Edwards TC Jr, Hastie T (2002) Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecol Model* 157:89–100
- Hastie TJ, Tibshirani RJ (1990) Generalized additive models. Chapman & Hall, London
- Holligan PM, Pingree RD, Mardell GT (1985) Oceanic solitons, nutrient pulses and phytoplankton growth. *Nature* 314:348–350
- Hu Q, Davis C (2005) Automatic plankton image recognition with co-occurrence matrices and Support Vector Machine. *Mar Ecol Prog Ser* 295:21–31
- Huntley ME, Lopez MDG (1992) Temperature-dependent production of marine copepods: a global synthesis. *Am Nat* 140:201–242
- Jegou AM, Lazure P (1995) Quelques aspects de la circulation sur le plateau atlantique. In: Cendrero O, Olaso I (eds) *Actas del IV Coloquio Internacional sobre Oceanografía del Golfo de Vizcaya*. Instituto Español de Oceanografía, Santander, p 99–106
- Kleppel GS, Davis CS, Carter K (1996) Temperature and copepod growth in the sea: a comment on the temperature-dependent model of Huntley and Lopez. *Am Nat* 148: 397–406
- Koutsikopoulos C, Le Cann B (1996) Physical processes and hydrographical structures related to the Bay of Biscay anchovy. *Sci Mar* 60:9–19

- Le Fevre J (1986) Aspects of the biology of frontal systems. *Adv Mar Biol* 23:163–299
- Lennert-Cody CE, Franks PJS (2002) Fluorescence patches in high-frequency internal waves. *Mar Ecol Prog Ser* 235: 29–42
- Li WKW, Harrison WG, Head EJH (2006) Coherent assembly of phytoplankton communities in diverse temperate ocean ecosystems. *Proc R Soc B* 273:1953–1960
- Lobo JM, Martin-Piera F (2002) Searching for a predictive model for species richness of Iberian dung beetle based on spatial and environmental variables. *Conserv Biol* 16: 158–173
- Mann KH, Lazier JRN (1991) Dynamics of marine ecosystems. Blackwell Science, Boston, MA
- Mauchline J (1998) The biology of calanoid copepods. *Adv Mar Biol* 33:1–710
- Montagnes DJS, Berges JA, Harrison PJ, Taylor FJR (1994) Estimating carbon, nitrogen, protein and chlorophyll a from volume in marine phytoplankton. *Limnol Oceanogr* 39:1044–1060
- New AL (1988) Internal tidal mixing in the Bay of Biscay. *Deep-Sea Res I* 35:691–709
- Pingree RD (1984) Some applications of remote sensing to studies in the Bay of Biscay, Celtic Sea and English Channel. In: Nihoul JCJ (ed) Remote sensing of shelf seas hydrodynamics. *Proc 15th Int Liege Colloq. Elsevier Oceanography Series* 38, Amsterdam, p 287–315
- Pingree RD, Le Cann B (1992) Anticyclonic Eddy X91 in the southern Bay of Biscay, May 1991 to February 1992. *J Geophys Res C* 97:14353–14367
- Pingree RD, Mardell GT (1981) Slope turbulence, internal waves and phytoplankton growth at the Celtic Sea shelf-break. *Phil Trans R Soc Lond A* 302:663–682
- Pingree RD, New AL (1995) Structure, seasonal development and sunglint spatial coherence of the internal tide on the Celtic and Armorican shelves and in the Bay of Biscay. *Deep-Sea Res I* 42:245–284
- Pingree RD, Mardell GT, New AL (1986) Propagation of internal tides from the upper slopes of the Bay of Biscay. *Nature* 321:154–158
- 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
- Poulet SA, Laabir M, Chaudron Y (1996) Characteristic features of zooplankton in the Bay of Biscay. *Sci Mar* 60: 79–95
- Quevedo M, Anadon R (2000) Spring microzooplankton composition, biomass and potential grazing in the central Cantabrian coast (southern Bay of Biscay). *Oceanol Acta* 23:297–309
- Rochford DJ (1991) 'Upwelling': Does it need a stricter definition? *Aust J Mar Freshw Res* 42:45–46
- Roy S, Silverberg N, Romero N, Deibel D and 6 others (2000) Importance of mesozooplankton feeding for the downward flux of biogenic carbon in the Gulf of St. Lawrence (Canada). *Deep-Sea Res II* 47:519–544
- See JH, Campbell L, Richardson TL, Pinckney JL, Shen R, Guinasso NL (2005) Combining new technologies for determination of phytoplankton community structure in the northern Gulf of Mexico. *J Phycol* 41:305–310
- Shaw W, Robinson CLK (1998) Night versus day abundance estimates of zooplankton at two coastal stations in British Columbia, Canada. *Mar Ecol Prog Ser* 175:143–153
- Sieracki CK, Sieracki ME, Yentsch CS (1998) An imaging-inflow system for automated analysis of marine microplankton. *Mar Ecol Prog Ser* 168:285–296
- Soulanki HU, Dwivedi RM, Nayak SR (2001) Synergistic analysis of SeaWiFS chlorophyll concentration and NOAA-AVHRR SST features for exploring marine living resources. *Int J Remote Sens* 22:3877–3882
- Stratoudakis Y, Bernal M, Borchers DL, Borges MF (2003) Changes in the distribution of sardine eggs and larvae off Portugal, 1985–2000. *Fish Oceanogr* 12:49–60
- Sverdrup HU, Johnson MW, Fleming RH (1942) The oceans: their physics, chemistry and general biology. Prentice-Hall, New York
- Uye S (2000) Why does *Calanus sinicus* prosper in the shelf ecosystem of the Northwest Pacific Ocean? *ICES J Mar Sci* 57:1850–1855
- Valdes L, Moral M (1998) Time-series analysis of copepod diversity and species richness in the southern Bay of Biscay off Santander, Spain, in relation to environmental conditions. *ICES J Mar Sci* 55:783–792
- van Aken HM (2002) Surface currents in the Bay of Biscay as observed with drifters between 1995 and 1999. *Deep-Sea Res I* 49:1071–1086
- Varela M (1996) Phytoplankton ecology in the Bay of Biscay. *Sci Mar* 60:45–53
- Wood SN (2006) Generalized additive models: an introduction with R. Chapman & Hall/CRC, Boca Raton
- Wood SN, Augustin NH (2002) GAMs with integrated model selection using penalized regression splines and applications to environmental modeling. *Ecol Model* 157:157–177
- Yee TW, Mitchell ND (1991) Generalized additive models in plant ecology. *J Veg Sci* 2:587–602

Editorial responsibility: Otto Kinne (Editor-in-Chief), Oldendorf/Luhe, Germany

Submitted: October 19, 2006; Accepted: March 15, 2007  
Proofs received from author(s): August 13, 2007