CHAPTER 6 FISH AND MACRO-CRUSTACEAN RESPONSE SURFACES TO ENVIRONMENTAL GRADIENTS IN THE WESTERSCHELDE ESTUARY

K. Hostens

Marine Biology, Ghent University, Krijgslaan 281-S8, B-9000 Gent, Belgium

Key words: Modelling, Logistic regression, Temporal, Spatial, Juvenile fish, Shrimp, Crab

Submitted for 2nd revision to *Marine Ecology Progress Series*

Abstract. *Juvenile fish and macro-crustaceans of the Westerschelde were intensively sampled with a 3-metre beam trawl in the periods 1988-91 and 1999-2001. Models were developed to predict the occurrence and density of the 15 commonest species in response to a limited set of environmental variables. Single logistic regressions yielded good descriptions of the occurrence of any one species along one environmental gradient. This was related to the maximum likelihood of presence in the field, although several species showed a broad tolerance towards one or more of the 4 environmental variables used. The response curves should only be interpreted as actual distribution patterns of juvenile fish and macro-crustaceans as a function of the four environmental variables. Multiple logistic regressions and normal regressions gave insight into the relative importance of each environmental variable for every single species. All response surfaces, either based on presence/absence or on density data, were highly significant when combining data on temperature, salinity, turbidity, dissolved oxygen concentration and/or their quadratic effects. The addition of other, extrapolated variables (current velocity, mysid prey density, chlorophyll a or suspended particulate matter) did not improve the predictions. For most species the prediction of presence/absence was relatively successful (60-90 % correctly predicted). Sensitivity (% present predicted as present) and specificity (% absent predicted as absent) were equally high in most models, and validation proved the models to be accurate and robust. The models that predicted density patterns could only explain between 20-55 % of the variance. Best 'density' models were built for those species that were present in the estuary during a longer period with only one clear density peak, i.e. Limanda limanda, Pomatoschistus microps, Carcinus maenas, Liocarcinus holsatus, Platichthys flesus, Sprattus sprattus and Pomatoschistus minutus. The least models concerned species belonging to the ecological guild of 'marine juveniles' (e.g. Trisopterus luscus, Merlangius merlangus, Clupea harengus, Solea solea and Pleuronectes platessa). Also, the 'density' models for Crangon crangon, Syngnathus rostellatus and Pomatoschistus lozanoi were less successful. Temperature (and dissolved oxygen concentration) mainly reflected seasonal effects, while salinity and turbidity reflected spatial effects. Still, it is argued that the interaction between several environmental variables (e.g. temperature and salinity) was even more important in predicting species occurrence and density.*

6.1 Introduction

The ecology of estuarine fish and macro-crustaceans has been thoroughly studied. Several studies have shown the importance of estuaries as nursery and wintering sites, migration routes, and areas which naturally support high densities of fish and macro-crustaceans (overview in Elliott & Hemingway 2002). Strong environmental gradients prevent most animals from the adjacent sea or rivers to enter the estuarine ecosystems (McLusky 1989). Following the concept of ecological guilds (Elliott & Dewailly 1995), 24 fish species in the Westerschelde were

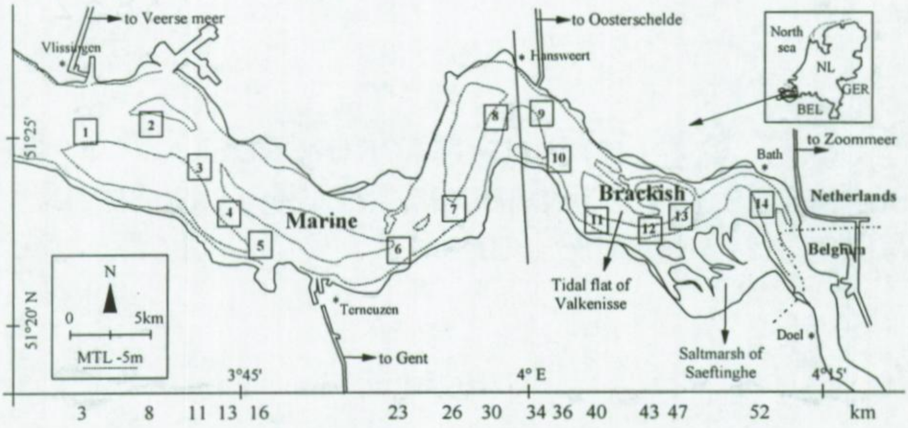


Fig. 6.1 Map of the Westerschelde estuary with 14 sampling locations in the subtidal channel

classified as marine species, 8 as estuarine resident and 5 as diadromous species (Chapter 2).

During the past decades, both seasonal and spatial patterns in fish and macro-crustacean assemblages have been uncovered in different holarctic estuaries, by means of multi-species analyses (*e.g* Elliott & Dewailly 1995, Blaber 1997, Potter *et al.* 1997). Several studies showed that a number of species was present during only a short period, with a clear cyclical succession in time, related to seasonal changes in the environment and to intra-annual and seasonal changes in appearance of several species (Thiel & Potter 2001, Chapter 2). Spatial differences in the distribution pattern of the fish and macro-crustacean assemblages within estuaries, were mainly related to salinity, the degree of exposure and the type of substratum (Henderson 1989, Chapter 2-Add.2).

For management purposes it is necessary to be able to predict the occurrence and distribution of both commercial and non-commercial fish and macro-crustacean species in the estuary. Single-species models can extend the knowledge of the species-environment relationships (Attrill *et al.* 1999). Most probably, single fish and macro-crustacean species within the ecological guilds will respond differently to changes in the surrounding environment. These so-called response curves (to changes in one variable) and surfaces (to combined changes in several variables) can be fitted through mathematical relations (Austin 1987), but the models should be simple to be useful (Bourget *et al.* in press). Logistic regression has been applied in the field of estuarine animal ecology since the mid 1990s (Ysebaert *et al.* 2002).

Several experimental studies generated species tolerance levels to physical factors (overview in Elliott & Hemingway 2002), but few attempts have been made to statistically model trends in abundance for individual crustacean and fish species in the field (Attrill *et al.* 1999, Power *et al.* 2000b). Most biological data sets lack corresponding environmental data, which are necessary to establish conclusions on possible controlling factors. During a period of 4 years and a period of 3 years two 'complete' data sets (with both species and environmental data) have been collected from the Westerschelde (Hostens *et al.* 1996, Hostens unpubl. data).

In this paper, the hypothesis is tested whether a limited set of environmental variables can be used to predict presence/absence and density patterns of the commonest fish and macrocrustacean species in the Westerschelde. It is argued that the used environmental variables adequately represent the seasonal and spatial variation in the estuary and are sufficient to build significant models.

6.2 Material and methods

6.2.1 Sampling methodology

The Westerschelde (55 km, 310 km²) consists of the lower and middle reaches of the Schelde estuary, situated in the southwest part of the Netherlands (Fig. 6.1). It is a well mixed, multiple channel system (average depth between 15 and 20 metres below Mean Tidal Level (MTL)), with large intertidal areas (35% of the area) and subject to large tidal fluctuations (up to 5 metres). Maximum ebb and flood current velocities vary between 2 and 3 knots at average neap and spring tide respectively (recalculated from Anonymous 1992). It is a turbid nutrient-rich and polluted area, under high pressure from dredging activities for shipping purposes. The ecological and environmental properties of the Westerschelde estuary are explained in detail in Meire & Vincx (1993) and Heip & Herman (1995).

Fish and macro-crustaceans were sampled at 14 locations in the main subtidal channel of the Westerschelde estuary (Fig. 6.1). Surveys were repeated monthly between January 1988 and December 1991 and continued on a quarterly basis between

September 1999 and May 2001. August 1988 was sampled twice; May and September 1991 were not sampled. This resulted in 55 surveys and 734 sampling points. A 3-m beam trawl (5x5 millimetres mesh-size in the cod end) was towed over a distance of 1000 m with the R.V. Luctor (34 m, 500 Hp). Sample methodology was fully described in Chapter 2. In summary, fish species were counted and measured, while crustaceans were counted and/or weighed in bulk. Densities were expressed as numbers per 1000 m², assuming a 20 % net-efficiency (see Chapter 3). For the present study, the data of the commonest species, 12 fish and 3 macro-crustacean species, were used (Chapter 2), belonging to the clupeoids (herring *Clupea harengus* and sprat *Sprattus sprattus*), the gadoids (whiting *Merlangius merlangus* and bib *Trisopterus luscus*), Nilsson's pipefish (*Syngnathus rostellatus*), gobiids (common goby *Pomatoschistus microps*, sandgoby *P. minutus* and Lozano's goby *P. lozanoi*), flatfish (plaice *Pleuronectes platessa*, dab *Limanda limanda*, flounder *Platichthys flesus* and sole *Solea solea*), crabs (flying crab *Liocarcinus holsatus* and shore crab *Carcinus maenas*) and brown shrimp (*Crangon crangon*).

Four environmental variables were measured at each sampling point: temperature (°C), dissolved oxygen content ($mg.l^{-1}$), salinity (psu) and turbidity (m^{-1}). The first three were measured near the bottom, whereas turbidity was calculated as the reciprocal of Secchi-depth. Data on other possible determinants were not available, were not measured near the bottom or overlapped only partly with the biotic sampling data. Most samples were taken around spring tide or at an average tidal level. Current velocity was estimated as the average tidal current in the upper water layer in the vicinity of every sampling point, given in hourly intervals in reference to high water near Vlissingen at spring and neap tide in the current atlas (Anonymous 1992), and from the difference between time of sampling and time of high water near Vlissingen. Sampling was only possible below 3 Beaufort, thus the influence of changing weather conditions on the current velocity was considered to be minimal at the moment of sampling. Also, some biological variables were included. For 110 corresponding sampling points, densities of a major prey group, namely the Mysidacea were taken from Mees (1994) and Chavatte (2001). Monthly data on Chlorophyll *a* and suspended particulate matter in the water column near the sampling points were made available through the Centre for Estuarine and Marine Ecology (Yerseke, The Netherlands).

6.2.2 Modelling details

The Generalized Linear Model concept has the advantage that the distribution of the dependent variable does not have to be continuous and the effect of the predictors on the dependent variable can be made linear through link functions (McCullagh & Nelder 1989). All regression equations were established with the statistical package Statistica, which gives an exhaustive explanation of the used techniques (StatSoft 1995).

Second-order polynomials (or quadratic effects) were included in the regressions to achieve gaussian probability functions as response curves and surfaces, which are ecologically more relevant in most cases (Ter Braak & Looman 1986). In the generalized model the response variable *Y* is linearly associated with values on the *X* environmental variables by:

$$Y = g\,(b_0 + b_1X_1 + b_2X_1^2 + b_3X_2 + b_4X_2^2 + \ldots + b_{k-1}X_k + b_kX_k^2) + e \quad (1)$$

where *e* stands for the error variability not accounted for by the predictors (expected to be 0), b_0 is the regression coefficient for the intercept and the b_i values are the regression coefficients (for variables 1 through *k*) computed from the data. The inverse function of *g(...)* is called the link function; so that:

$$f(z_y) = b_0 + b_1X_1 + b_2X_1^2 + b_3X_2 + b_4X_2^2 + \ldots + b_{k-1}X_k + b_kX_k^2 \qquad (2)$$

where z_y stands for the expected value or the probability of *y*, which is bound between 0 and 1.

6.2.3 Presence/Absence models

In a first series of models the density values were reduced to binary data to predict the probability of occurrence of the 15 species as a function of the *a priori* selected environmental variables. Presence/absence data show less variation and are easier to predict than real numbers. For these models the logit link function is used, where $f(z_y) = \log(z_y/(1-z_y))$.

So, equation (2) can be rewritten as:

$$z_y = e^{(b0 + b1X1 + b2X1^2 + b3X2 + b4X2^2 + \ldots + bk\text{-}1Xk + bkXk^2)} / \{1 + e^{(b0 + b1X1 + b2X1^2 + b3X2 + b4X2^2 + \ldots + bk\text{-}1Xk + bkXk^2)}\} \qquad (3)$$

The values of the parameters (b_0 through b_k and the scale parameter) in the generalized linear model were obtained by maximum likelihood estimation. The significance of the models was tested with the –2logL statistic, where the maximized likelihood L approximates a *Chi-square* distribution. The Wald statistic (*Chi-square*, $p<0.05$) was used to test the significance of the regression coefficients.

Single logistic regressions showed the probability of occurrence of the 15 species as a response to each of the 4 environmental variables separately. Linear models were presented when the quadratic effects did not have significant explanatory power.

Next, the presence/absence of the fish and macrocrustacean species was predicted by multiple logistic regressions, where all variables (and their quadratic effects) were entered (or rejected) in the models by means of forward stepwise selection. The resulting presence/absence models (or response surfaces) were validated in several ways. Therefore, the estimated probabilities (on a continuous scale between 0 and 1) were back-transformed to binary data. A threshold p-value was determined for every model as the ranked probability value that corresponded with the observed ratio between presences and absences. Probabilities higher than that p-value were considered to predict presence, while lower values predicted a species to be absent. In this way, the predictive success (percentage correctly predicted), specificity (percentage absent predicted as absent) and sensitivity (percentage present predicted as present) could be calculated.

These diagnostics were calculated for the final models and for five model runs per species with a random selection of 50 % of the data, to evaluate the robustness and classification accuracy of the final models. The 50 % model equations were used in cross-validations with the other half of the respective data sets and with a randomly generated set of 1000 values per environmental variable (within the observed limits of these predictors). For the random environmental sets, linear correlation coefficients between five model runs were calculated.

6.2.4 Density models

In a second approach single-species regression models were based on log-transformed data (calculated as the natural logarithm of density +1 per 1000 m²), to approximate normal distributions. Therefore, an identity link function could be used, where $f(z_y) = z_y$, and equation (2) is appropriate to predict the response surfaces of species abundance by a linear combination of the environmental data. The shape of the response curves (and surfaces) is dependent on the sign of the actual regression coefficients (b) and the ratio between the absolute beta-values of the linear and quadratic effects (i.e. the relative magnitude of both effects) (Fig. 6.2).

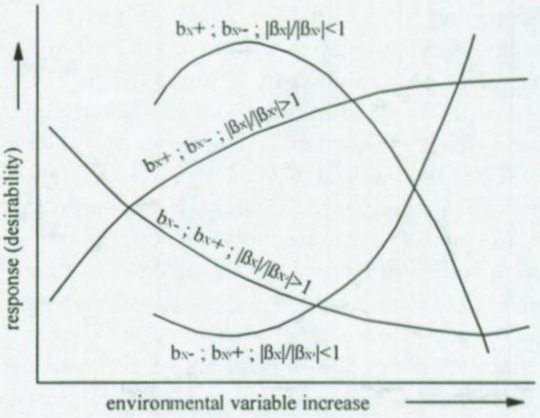


Fig. 6.2 Simplified response curves of a dependent variable, dependent on the signs of the actual regression coefficients (b) and the ratio between the standardized regression coefficients (ß) for the linear and quadratic effects of any independent variable

The model equations were built with the General Stepwise Regression module in Statistica (StatSoft 1995). From an explorative analysis with all environmental variables, the sample points for every single species for which the predicted values deviated more than 60% from the observed values, were considered to be outliers and were eliminated from further analyses. For the final 'density' models, quadratic regression designs and forward stepwise selection of the environmental variables were used. Assumptions on normality and homogeneity of the residuals were graphically inspected. Standardized regression coefficients (ß) indicate the magnitude of the contribution of each withdrawn effect ($p < 0.05$) to the final models. The fit of the regression surface as defined by the predicted values and the estimation of the effects included in the final models were based on the least squares method. Type VI (effective hypothesis) sums of squares can deal with missing cells and provide an unambiguous estimate of the variability of the predicted values attributed to each effect (Hocking 1996).

6.3 Results

6.3.1 Environmental gradients

Both spatial and temporal patterns in the environmental variables are given in Fig. 6.3. Temperature mainly showed a seasonal trend, with summer peaks (averaged 19 °C) and winter troughs (averaged 6 °C), and a cold February 1991 (<1 °C). The upstream part of the Westerschelde was on average 1 degree warmer than the most marine sampling location in a more or less linear trend. An important spatial pattern in the Westerschelde was shown by the linear salinity gradient (averaging between 28 and 13 psu, with minimum and maximum values of 4 and 32 respectively). River runoff clearly showed downstream freshwater intrusions during the winter periods, with low values in 1988. The Westerschelde was well saturated with oxygen (on average between 7 and 11 $mg.l^{-1}$), but oxygen levels linearly dropped with on average 3 units from the marine to the brackish stations with minima between 3 and 5 $mg.l^{-1}$ in the most upstream station 14 (near the Dutch-Belgian border). The seasonal pattern showed high oxygen concentrations during the winter period, with extreme values (maximum 14 $mg.l^{-1}$) in 1991. The seasonal pattern in turbidity resembled the oxygen pattern (averaging between 0.8 and 2.2 m^{-1}), while the spatial pattern showed an opposite gradient with higher values in the upstream part (up to 5 m^{-1}). Both ebb and flood current velocities at the moment of sampling varied between 0 and 3 knots (= 1.5 m.s^{-1}), without a clear seasonal pattern. In most cases the current velocity was higher in the middle reaches of the Westerschelde between stations 9 and 12.

Mysid prey densities were on average 2 m^{-2}, exponentially increasing towards the brackish reaches, with a maximum up to 160 m^{-2} during summer (Mees 1994, Hostens unpubl. data). Chlorophyll *a* concentrations varied between 0.3 and 45 µg.l^{-1} with a clear seasonal peak during spring. SPM values ranged between 4 and 440 mg.l^{-1} (on average 50 mg.l^{-1}), without a clear seasonal or spatial trend. (Chl*a* and SPM data from CEME, the Netherlands).

6.3.2 Single-species response curves

In total, 49 of the 60 single logistic regressions were highly significant. No significant relations were found for brown shrimp *Crangon crangon*.

Five species had a higher probability of occurrence at low temperatures (Fig. 6.4). For *Clupea harengus, Pleuronectes platessa* and *Platichthys flesus* the difference was rather small, but the response curves for *Pomatoschistus microps* and *Limanda limanda* showed a steep decline at temperatures above 12 °C. Four species were predicted to occur at higher temperatures (*Trisopterus luscus*, *Solea solea*, *Carcinus maenas* and *Liocarcinus holsatus*. Three species (*Merlangius merlangus*, *Syngnathus rostellatus* and *Pomatoschistus lozanoi*) showed a bell-shaped curve with an optimum at intermediate temperatures. The opposite holds for *Sprattus sprattus* and *Pomatoschistus minutus*, with increasing probabilities towards the lower and upper end of the temperature range.

Most species showed a linear response and a broad tolerance in relation to salinity. Nine species likely occurred at higher salinities, but only the 2 crab species were clearly limited with a probability of occurrence of more than 50 % at salinities above 15 psu for *C. maenas* and above 27 psu for *L. holsatus*. Four species (*C. harengus*, *P. microps*, *P. flesus* and *S. solea)* were predicted to occur at lower

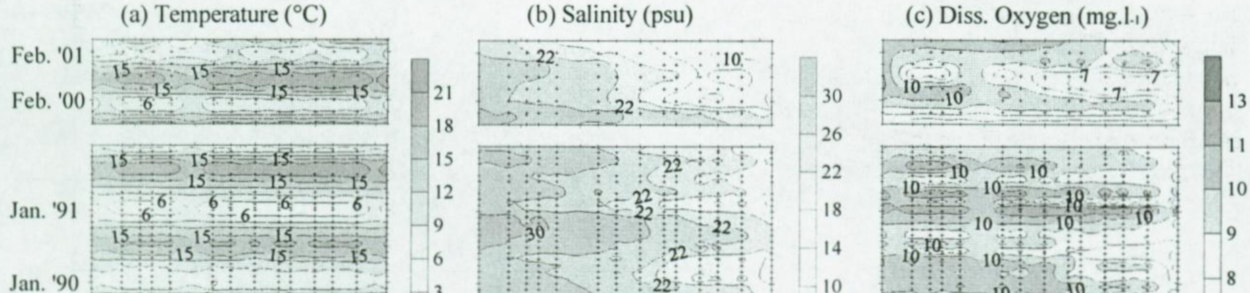


Fig. 6.3 Spatio-temporal contour plots for 5 environmental variables. The abscis represents the spatial variation based on the 14 sampling points; the temporal axis shows 48 monthly surveys between January 1988 and December 1991, and 8 quarterly surveys between September 1999 and May 2001

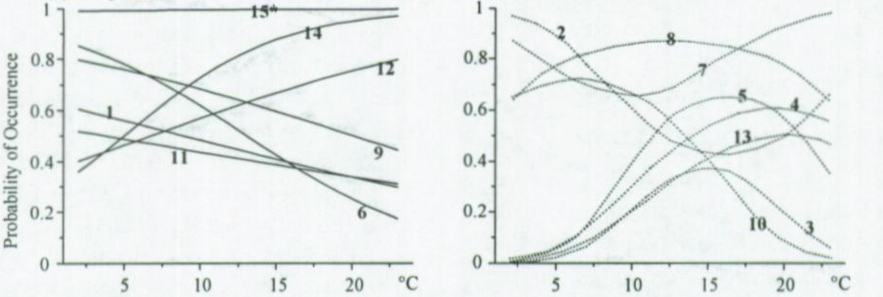
salinities.

Five response curves to dissolved oxygen concentrations were insignificant. Three species (*S. sprattus*, *P. microps* and *P. platessa*) had a higher probability to be found at higher oxygen levels. For *T. Luscus*, *P. minutus* and *S. solea* the opposite was predicted. Four species differed in their tolerance and in the position of their optimum towards the upper (*L. limanda*), the middle (*M. merlangus*), the lower (*S. rostellatus*) or both (*P. flesus*) ends of the oxygen range.

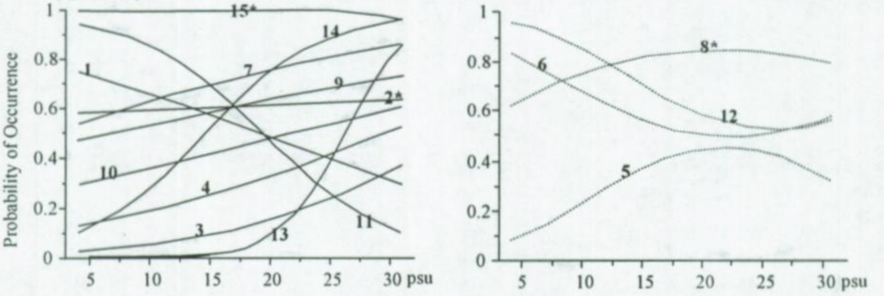
Four response curves showed a higher probability to occur at higher turbidities for four species (*C. harengus*, *S. sprattus*, *P. platessa* and *P. flesus*). Five species were predicted to have a low tolerance towards turbidity, with probabilities of more than 50% at turbidity values <1 m^{-1} for *M. merlangus*, *T. luscus*, *S. rostellatus* and *L. holsatus,* and <2.5 m^{-1} for *C. maenas*. Four other species showed gaussian logit response curves with an optimum towards lower turbidities (*P. minutus* and *P. lozanoi*) or towards higher turbidities (*P. microps* and *L. limanda*).

linear effects only **including quadratic effects**

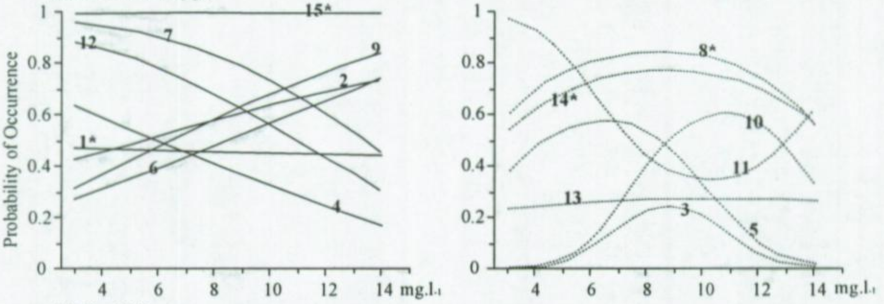
(a) Temperature



(b) Salinity



(c) Dissolved oxygen



(d) Turbidity

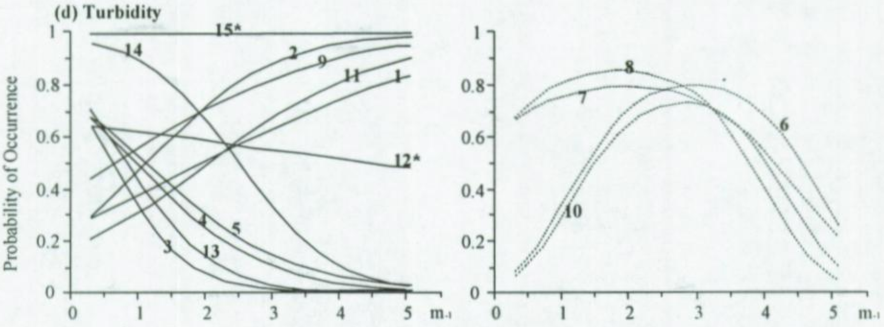


Fig. 6.4 Probability of occurrence fitted with logistic regressions in relation to each of the 4 environmental variables. Linear response curves are given, if the quadratic effects were insignificant. Line numbers correspond with: (1) *C. harengus*, (2) *S. sprattus*, (3) *M. merlangus*, (4) *T. luscus*, (5) *S. rostellatus*, (6) *P. microps*, (7) *P. minutus*, (8) *P. lozanoi*, (9) *P. platessa*, (10) *L. limanda*, (11) *P. flesus*, (12) *S. solea*, (13) *L. holsatus*, (14) *C. maenas*, (15) *C. crangon*. Insignificant regressions were indicated with an asterisk*

Table 6.1 Ecological guild and some diagnostics for the final Presence/Absence models: actual numbers absent and present, -2logL statistic for 'the intercept only' (I) and for 'the intercept with covariates' (I+C), *chi-square*, percentage explained variance, and different environmental variables successively retained in each model

Species	Guild (1)	# Absent	# Present	-2logL		Chi^2	%Var	Selected variables (2)
				I	I+C			
Clupea harengus	MJ	397	334	1008	957	51	5	S T T^2
Sprattus sprattus	MS	281	450	974	823	151	16	T T^2 S Tu O^2 O
Merlangius merlangus	MJ	589	142	720	598	122	17	Tu Tu^2 S
Trisopterus luscus	MJ	469	262	954	775	179	19	T S T^2
Syngnathus rostellatus	ER	441	290	982	791	191	19	T T^2
Pomatoschistus microps	ER	336	395	1009	854	155	15	T Tu Tu^2 S^2
Pomatoschistus minutus	ER	171	560	795	656	139	18	S T^2 T O^2 Tu O Tu^2
Pomatoschistus lozanoi	MS	133	598	693	669	25	4	Tu Tu^2 T T^2
Pleuronectes platessa	MJ	259	472	950	886	65	7	S Tu T
Limanda limanda	MJ	370	361	1013	736	278	27	S^2 Tu T^2 Tu^2 T
Platichthys flesus	ER	425	306	994	820	174	17	S Tu^2
Solea solea	MJ	291	440	983	893	90	9	T S T^2 S^2
Liocarcinus holsatus	MJ	532	199	856	414	442	52	S T O^2
Carcinus maenas	ER	184	547	825	569	256	31	S T
Crangon crangon	ER	3	728	39	32	7	18	S^2

Note: (1) MJ (marine juveniles), MS (marine seasonal/adventitious), ER (estuarine resident) species
(2) T (temperature), S (salinity), O (oxygen), Tu (turbidity) and their corresponding quadratic effects

6.3.3 Presence/Absence modelling diagnostics

As none of the other (extrapolated) environmental variables were withdrawn in the stepwise procedure, the final multiple regressions for the 12 fish and 3 crustacean species were limited to the four main environmental variables (temperature, salinity, oxygen, turbidity, and their quadratic derivations). According to the Wald statistic, temperature and salinity appeared as important variables to predict species occurrences in almost all models, followed by turbidity for 8 models and oxygen for 3 models (Table 6.3). All final logistic models were highly significant ($p << 0.001$) based on the *chi-square* test. The explained variance can be calculated as 1 minus the ratio of the –2logL statistics for the intercept with covariates and the intercept-only model (Table 6.1). Around 50 % of the variance could be explained for *L. holsatus*, around 30 % for *L. limanda* and *C. maenas*, and around 20 % in 8 other models.

Threshold-P values (above which calculated probabilities were converted to presences) were around 0.5 and higher (Table 6.2), which corresponded with the actual ratio of absence and presence

Table 6.2 Threshold-p values, predictive success (% correct), specificity (% absent predicted as absent) and sensitivity (% present predicted as present) for the 15 final presence/absence models. The same diagnostics are given as averages (± std) for each time five fitted models with 50% of the data and five cross-validations of these 50% models with the rest of the data. The last column shows the range of linear correlation coefficients between five fitted models, based on the equations from the 50% models and 1000 random values for each environmental variable

Species	Final P/A model				50% data fitting			50% cross-validation			Correlation
	ThresP	%corr	%AasA	%PasP	%corr	%AasA	%PasP	%corr	%AasA	%PasP	
Clupea harengus	0.46	62	64	59	62 (1.9)	65 (2.0)	59 (3.3)	60 (2.3)	62 (3.3)	58 (2.9)	0.96-0.99
Sprattus sprattus	0.52	70	60	77	70 (1.5)	60 (9.6)	76 (8.1)	70 (1.5)	60 (1.6)	76 (1.6)	0.83-0.95
Merlangius merlangus	0.33	81	88	52	83 (1.1)	89 (1.1)	54 (1.0)	78 (1.2)	86 (1.1)	50 (2.1)	0.89-0.95
Trisopterus luscus	0.52	73	79	62	73 (2.5)	79 (2.1)	62 (3.4)	72 (2.7)	77 (2.0)	62 (4.2)	0.90-0.96
Syngnathus rostellatus	0.58	69	73	62	69 (1.2)	73 (1.8)	62 (5.2)	69 (0.7)	74 (0.8)	62 (1.5)	0.92-0.99
Pomatoschistus microps	0.53	71	69	73	72 (1.9)	70 (2.3)	74 (2.5)	70 (1.6)	66 (2.2)	73 (2.0)	0.84-0.99
Pomatoschistus minutus	0.65	80	53	88	81 (2.7)	57 (5.8)	88 (1.7)	77 (0.7)	51 (3.7)	85 (0.4)	0.84-0.98
Pomatoschistus lozanoi	0.77	74	29	84	75 (1.1)	29 (5.0)	86 (1.4)	74 (0.7)	27 (5.1)	85 (1.0)	0.97-0.99
Pleuronectes platessa	0.58	66	51	73	67 (1.7)	53 (3.6)	75 (1.3)	65 (1.8)	49 (4.2)	74 (1.3)	0.96-0.99
Limanda limanda	0.55	77	78	77	78 (0.9)	78 (1.2)	77 (0.7)	77 (1.3)	77 (1.3)	78 (1.6)	0.84-0.99
Platichthys flesus	0.43	72	76	67	70 (1.7)	74 (1.9)	64 (2.7)	74 (1.9)	77 (1.8)	69 (2.7)	0.93-0.99
Solea solea	0.62	65	56	71	62 (1.3)	52 (3.8)	69 (1.0)	66 (1.7)	55 (1.9)	73 (1.5)	0.92-0.99
Liocarcinus holsatus	0.46	87	91	76	88 (0.8)	92 (0.5)	78 (2.8)	86 (0.5)	90 (0.5)	75 (1.4)	0.85-0.92
Carcinus maenas	0.61	80	60	87	80 (2.2)	60 (2.6)	86 (1.7)	81 (2.1)	60 (2.9)	88 (1.7)	0.98-0.99
Crangon crangon	0.94	99	0	100	99 (0.9)	-	99 (0.6)	92 (11.4)	-	93 (11.7)	0.92-0.95

Table 6.3 Regression coefficients and the Wald statistic between brackets (p <0.05 if Wald >5) for the intercept and the withdrawn environmental variables (including the quadratic effects) for 15 'presence/absence' models (from which the probability of occurrence can be calculated as in formula 3, see text)

Species	Intercept	Temperature	Salinity	Oxygen	Turbidity	Temperature^2	Salinity^2	Oxygen^2	Turbidity^2
Clupea harengus	2.9 (28.7)	-0.26 (8.7)	-0.07 (25.4)	-	-	0.01 (5.7)	-	-	-
Sprattus sprattus	-1.66 (0.6)	-0.75 (33.8)	0.08 (19)	1.26 (8.7)	0.7 (15.1)	0.02 (25.2)	-	-0.08 (12.6)	-
Merlangius merlangus	0.98 (1.8)	-	0.05 (6)	-	-3.6 (47)	-	-	-	0.64 (21.7)
Trisopterus luscus	-6.42 (61.9)	0.56 (20.3)	0.06 (15.5)	-	-	-0.01 (9.3)	-	-	-
Syngnathus rostellatus	-6.77 (74.4)	0.91 (48.3)	-	-	-	-0.03 (31.9)	-	-	-
Pomatoschistus microps	-1.73 (6.9)	-0.1 (30.9)	-	-	2.55 (25.7)	-	0 (3.9)	-	-0.45 (13.2)
Pomatoschistus minutus	-6.15 (5)	-0.74 (28)	0.13 (39.7)	1.77 (10.6)	2.41 (15)	0.03 (34.7)	-	-0.12 (17.7)	-0.46 (9.7)
Pomatoschistus lozanoi	-2.25 (6.5)	0.36 (12.4)	-	-	2.08 (13.7)	-0.01 (11)	-	-	-0.51 (13.3)
Pleuronectes platessa	-1.81 (9.1)	-0.04 (5.8)	0.08 (28.2)	-	0.77 (20.7)	-	-	-	-
Limanda limanda	-7.03 (47.8)	0.46 (19.3)	-	-	3.74 (37.3)	-0.03 (34)	0 (63.9)	-	-0.68 (20.4)
Platichthys flesus	3.16 (59.3)	-	-0.18 (107.4)	-	-	-	-	-	0.1 (7.3)
Solea solea	2.43 (4.5)	0.37 (15.5)	-0.37 (13.1)	-	-	-0.01 (8.5)	0.01 (8.2)	-	-
Liocarcinus holsatus	-19.72 (130)	0.37 (87.9)	0.51 (125.4)	-	-	-	-	0.02 (8.2)	-
Carcinus maenas	-5.16 (109.2)	0.2 (75.5)	0.21 (104.3)	-	-	-	-	-	-
Crangon crangon	11.34 (10.8)	-	-	-	-	-	-0.01 (4.2)	-	-

(Table 6.1). For five species >80 % and for seven other species >70 % of the occurrences could be correctly predicted with the final models. Sensitivity fluctuated around 70 % and specificity was moderate to high, except for *P. lozanoi* and *C. crangon* where <40 % of the absences were predicted to be absent. Model fitting with only 50 % of the data proved to be consistent with the final models. Cross-validation of these 50%-models with the corresponding rest of the respective data achieved similar results. Model testing with random values for the environmental variables showed high correlation coefficients in a pairwise comparison between the five models for every species, ranging between 0.83 and 0.99 (Table 6.2).

6.3.4 Density modelling diagnostics

All 15 multiple stepwise regressions were highly significant (p < 0.0001), and seven models could explain between 58 and 45 % of the variance, i.e. for *L. limanda, P. microps, C. maenas*, *L. holsatus*, *P. flesus, S. sprattus* and *P. minutus* (Table 6.4). The other models explained between 22 and 37 % of the variance. All included regression coefficients were highly significant, except for the intercept in the models for *S. sprattus*, *C. maenas* and *C. crangon.* Temperature and salinity (and/or their quadratic effects) were selected in most models, while oxygen

Table 6.4 Actual (b) and standardized (ß) regression coefficients and t-values (p <0.05 if t >2), for the intercept and the withdrawn environmental variables (including the quadratic effects), next to the F and R^2 statistics for the 15 'density' models (see formula 2 in text)

Species		Intercept	Temperature	Salinity	Oxygen	Turbidity	Temperature^2	Salinity^2	Oxygen^2	Turbidity^2	F	R^2
Clupea harengus	b (ß)	3.52	-0.01 (-0.08)	-0.25 (-1.64)	-	-	-	0.005 (1.24)	-	0.06 (0.17)	74	0.33
	t	11.6	-2.2	-8.5	-	-	-	6.5	-	4.3		
Sprattus sprattus	b (ß)	0.03	-0.54 (-1.98)	0.06 (0.26)	0.91 (0.98)	0.55 (0.25)	0.02 (1.37)	-	-0.06 (-1.2)	-	106	0.50
	t	0.03	-10.1	8.0	4.1	6.9	7.4	-	-5.1	-		
Merlangius merlangus	b (ß)	1.45	-	-	-	-1.29 (-1.7)	-	-	-	0.26 (1.33)	149	0.32
	t	18.0	-	-	-	-14.6	-	-	-	11.4		
Trisopterus luscus	b (ß)	-0.78	0.09 (0.49)	-	-	-	-	0.001 (0.14)	-	-	125	0.28
	t	-8.3	14.5	-	-	-	-	4.2	-	-		
Syngnathus rostellatus	b (ß)	-1.09	0.09 (0.54)	0.08 (0.53)	-	-	-	-0.002 (-0.6)	-	-	95	0.30
	t	-4.1	16.3	2.9	-	-	-	-3.2	-	-		
Pomatoschistus microps	b (ß)	4.23	-0.32 (-1.11)	-	-	0.58 (0.25)	0.01 (0.46)	-	-0.01 (-0.2)	-	188	0.55
	t	8.7	-6.2	-	-	7.6	2.7	-	-5.6	-		
Pomatoschistus minutus	b (ß)	-6.81	-0.30 (-0.86)	0.06 (0.18)	1.85 (1.49)	1.55 (0.53)	0.02 (1.43)	-	-0.12 (-1.73)	-0.28 (-0.37)	73	0.45
	t	-3.9	-4.0	5.3	5.0	4.7	7.0	-	-5.8	-3.5		
Pomatoschistus lozanoi	b (ß)	-2.24	0.33 (1.19)	0.30 (1.24)	-	-	-0.01 (-0.97)	-0.01 (-0.96)	-	-	41	0.22
	t	-3.8	5.9	5.9	-	-	-4.8	-4.6	-	-		
Pleuronectes platessa	b (ß)	-0.60	-	0.04 (0.2)	-	0.84 (0.41)	-0.002 (-0.2)	-	-	-	80	0.27
	t	-2.2	-	5.6	-	10.3	-5.2	-	-	-		
Limanda limanda	b (ß)	-10.26	0.36 (1.16)	0.24 (0.8)	1.22 (1.11)	2.68 (1.05)	-0.02 (-1.6)	-0.004 (-0.55)	-0.07 (-1.26)	-0.43 (-0.64)	104	0.58
	t	-8.3	6.3	5.1	4.9	10.3	-9.2	-3.6	-5.6	-6.7		
Platichthys flesus	b (ß)	5.89	-0.03 (-0.18)	-0.17 (-1.2)	-0.47 (-0.85)	-	-	0.002 (0.6)	0.02 (0.75)	-	132	0.51
	t	10.5	-5.1	-7.2	-4.0	-	-	3.7	3.6	-		
Solea solea	b (ß)	5.17	0.15 (0.66)	-0.44 (-2.1)	-	-	-0.003 (-0.35)	0.01 (1.62)	-	-	97	0.37
	t	12.1	3.7	-11.7	-	-	-2.0	9.1	-	-		
Liocarcinus holsatus	b (ß)	1.59	-	-0.33 (-1.78)	-	-	0.004 (0.42)	0.01 (2.25)	0.004 (0.1)	-	202	0.54
	t	5.9	-	-11.8	-	-	13.6	15.1	3.1	-		
Carcinus maenas	b (ß)	-0.53	0.08 (0.3)	0.12 (0.47)	-	-0.49 (-0.22)	-	-	-	-	259	0.54
	t	-1.9	9.4	16.6	-	-6.5	-	-	-	-		
Crangon crangon	b (ß)	0.89	0.09 (0.38)	-	1.23 (1.53)	-	-	-	-0.07 (-1.61)	-	69	0.22
	t	0.9	10.0	-	6.1	-	-	-	-6.5	-		

and turbidity appeared in half of the models. Most of the actual data points fell within the 95 % predictive limits of the models (Fig. 6.5).

Densities for *C. harengus* and *S. sprattus* were predicted to be high at lower temperature and higher turbidity, in a combination with low salinity for *C. harengus*, and with higher salinity and intermediate oxygen levels for *S. sprattus*. Also, whiting densities will be lower in turbid waters. Densities of *T. luscus* and *S. rostellatus* were predicted to increase with increasing temperature at intermediate to high salinity levels. *P. microps* seemed to prefer lower temperature and oxygen values, while turbidity should be higher. *P. minutus* was predicted to have higher densities at moderate oxygen levels, if turbidity and temperature were high and salinity not too low. For *P. lozanoi* higher salinity and temperature seemed to be the forcing variables. High densities for *P. platessa* were predicted with lower temperature and higher turbidity. For *L. limanda* temperature and oxygen should be intermediately low, while turbidity and salinity should be high. *P. flesus* and *S. solea* were predicted to occur at higher densities at lower salinities (and lower oxygen concentrations for *P. flesus*), with a contrasting behaviour towards temperature. High values of salinity and temperature seemed necessary for both *C. maenas* and *L. holsatus*, in combination with low turbidity for *C. maenas* and somewhat higher oxygen levels for *L. holsatus*. Higher densities of *C. crangon* were predicted at moderate oxygen and higher temperature levels.

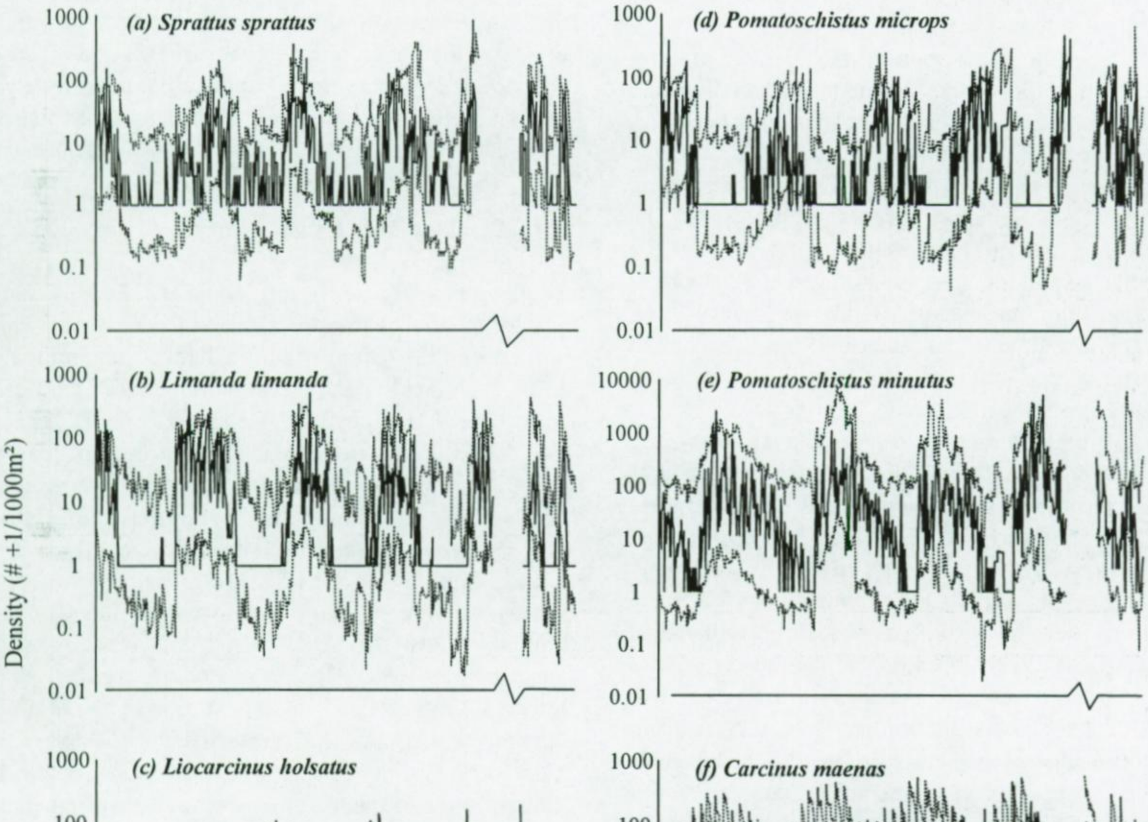


Fig. 6.5 Predictive 95 % confidence limits (broken lines) for six 'density' models and actual density values as individual sampling points (full line), for (a) 1 'marine seasonal', (b-c) 2 'marine juvenile', and (d-f) 3 'estuarine resident' species. Sampling points were ordered from marine to brackish (14 stations) on a monthly basis (48 surveys) between January 1988 and December 1991 and on a quarterly basis (8 surveys) from September 1999 to May 2001

6.4 Discussion

6.4.1 Environmental variables

As the Westerschelde estuary was characterized by a marked gradient in salinity, oxygen and turbidity, these are most likely the factors that will influence the occurrence and density of mobile organisms on a spatial scale. Several studies have shown the importance of a seasonal factor in the distribution of fish and macro-crustaceans, as they only temporarily make use of the estuary as a nursery ground (Hemingway & Elliott 2002). Obviously, temperature is the most suitable variable, although temporal patterns in salinity, oxygen and turbidity were shown as well. Models based on these four (easily monitored) environmental variables are simple, easier to test and cross-validate, and less costly in predicting and controlling the outcome in the future.

The other, extrapolated variables did not contribute significantly to the models. Hydrodynamic variables and sediment characteristics have been recognized as forcing factors on several aspects of the distribution of macrobenthic species (Ysebaert *et al.* 2002). Current velocity was not withdrawn in the models. For fish and macro-crustaceans, currents are indirectly important, mainly during the transport of postlarvae into the estuary (Drake & Arias 1991). The type of sediment was rather uniform (sandy at all 14 stations), and was only determined twice during both survey periods (Chapter 5).

Biotic interactions (*e.g* food availability) are important for the distribution of fish and macro-crustaceans. However, when mysid densities were included, the variance explained only increased slightly (between 0 and 5 %). Moreover, too many sampling points had to be eliminated, which made the models less powerful. Consequently, rather than reflecting causal relationships, the modelled response curves and surfaces should only be seen as descriptions of the actual distribution patterns of the fish and crustacean species in relation to the structuring environmental variables in temperate estuarine environments. Possibly, the combined interaction between the environmental variables is a proxy for prey availability.

Yet, altering the environmental conditions can lead to (drastic) changes in the estuarine populations of fish and macro-crustaceans. Due to major engineering works in the southwest Netherlands, the former Oosterschelde and Grevelingen estuaries were changed into a marine bay and a saline lake, respectively (Nienhuis & Smaal 1994b). For the Oosterschelde, this resulted in a reduced biomass of brown shrimp *Crangon crangon* (Chapter 3). For the Grevelingen, the density of juvenile plaice *Pleuronectes platessa* was reduced with >50% (Doornbos & Twisk 1984), which was only compensated <10% by the increased nursery function of the ebb-tidal delta (Chapter 4-Add.).

In almost all European estuaries, oxygen deficiency has been attributed to excess input of organic matter (Elliott & Hemingway 2002). Oxygen depletion highly restricted the distribution of fish and crustaceans in the upper part of the Schelde estuary till the mid 1990s (Van Damme *et al.* 1994, Mees *et al.* 1993a). Although, water quality improved in this part of the estuary, densities of the commonest species are still very low (Maes *et al.* 1998a). Synergistic interactions between reduced oxygen levels and increased water temperatures, *e.g* through cooling water discharges from power plants, may also occur (Cattrijsse *et al.* 2002).

Physical disturbances, such as dredging and relocation of sediments, may increase turbidity. Although, turbidity is important in lowering predation pressure, extreme values provide a direct negative effect on the health and behavior of fish and macro-crustaceans, not only on recruitment but possibly also on the gill function and respiration of fishes (Marchand 1993, Elliott & Hemingway 2002).

6.4.2 Single-species response curves

Single species response curves were relatively successful to predict the probability of occurrence, and agreed with the real distribution, habitat requirements and ecological tolerances found in literature (Cattrijsse & Hampel 2000, Froese & Pauly 2002).

Estuarine dependent species are considered to be eurytopic, with a high tolerance towards a broad range of environmental conditions. For *Crangon crangon* no significant response curves were found, even by converting low densities into absences or by converting real densities into discontinuous density-classes instead of binary presence/absence data. This is related to the fact that brown shrimp was present almost always and at high numbers throughout the Westerschelde (Chapter 2). Most other species are likely to be found within a limited range of one or more of the 4 environmental variables, although this does not mean the species were physically limited to these ranges.

Within the taxonomic groupings, none of the species showed a comparable preference in all 4 variables. The response curves for the 2 clupeoids (*Clupea harengus* and *Sprattus sprattus*) were only comparable for turbidity; the 2 gadoids (*Trisopterus luscus* and *Merlangius merlangus*) differed in their response to temperature and oxygen; the 3 gobiids (*Pomatoschistus minutus*, *P. lozanoi* and *P. microps*) showed opposite responses in all variables; *Solea solea* showed opposite responses than the other 3 flatfish species to all variables, *Platichthys flesus* to salinity and oxygen, *Pleuronectes platessa* and *Limanda limanda* to turbidity; the 2 crab species (*Car-*

cinus maenas and *Liocarcinus holsatus*) differed in their oxygen preference.

On the other hand, several species responded the same way to the four environmental variables, if compared per ecological guild. The 15 species were classified as 6 estuarine resident species, 2 marine seasonal (or adventitious) species and 7 marine juvenile species (Table 6.1). Four marine juveniles (whiting, bib, sole and flying crab) were more likely to be found at higher temperatures, higher salinities and lower turbidities, while the opposite holds true for 2 estuarine resident species (common goby and flounder). Three estuarine resident species (Nilsson's pipefish, sandgoby and shorecrab) showed similar response curves as the marine juveniles. This is related to the fact that most marine juveniles enter the estuary during springtime and only stay for a relatively short period, which was also the case for the latter 3 (so-called) estuarine residents in the Westerschelde (Chapter 2). Three marine juveniles (herring, plaice and dab) and a marine seasonal species (sprat) rather responded like estuarine residents to temperature and turbidity. In contrast to the other marine juveniles, the latter 4 species arrived at a later moment to winter, and were present during most of the year (except dab) in the Westerschelde (Chapter 2).

6.4.3 Presence/Absence models

The final presence/absence models and their validations achieved similar results, and the standard deviations on the validation diagnostics were very small. This suggested that the regressions were independent of the chosen set of actual observations. For the 8 most common species the sensitivity was higher than the specificity, which means that their presence in the Westerschelde was better predicted than their absence.

Temperature seemed to be an important forcing factor in the presence/absence models as well. This indicated that the prediction of occurrence was mainly dependent on a temporal effect. But the presence/absence models also showed that different environmental variables interact both at temporal and spatial scales. For example, the response curves for *Pleuronectes platessa* and *Pomatoschistus minutus* showed a positive relation with increasing salinities, although it has been shown that flatfish and gobies were more common in the 'brackish water' reaches of the Westerschelde (Chapter 2). Actually, they have a higher probability to be found during the colder months when the saltwedge penetrates deep into the middle reaches of the estuary (Fig. 6.3). The interaction between low temperatures and higher salinities became clear from the negative sign of the regression coefficient for temperature and the positive sign for salinity (Table 6.3).

Little information is found on tolerance for suspended material (or turbidity) by fishes in European estuaries (Cattrijsse 1997). In the Humber estuary (UK) the same correlations for fishes were found with temperature, salinity and dissolved oxygen concentration, but not with turbidity (Marshall & Elliott 1998). Probably, this is one of the first studies to show that several fish and crustaceans have a preference for a limited part of the turbidity range. For *Pomatoschistus lozanoi* this was the only significant response curve in relation to the separate environmental variables. The spatial effect of turbidity was found to significantly contribute to the predictive success of 8 presence/absence models, mainly for species with a higher probability of occurrence at higher turbidities.

Dissolved oxygen concentration was of minor importance in the presence/absence predictions, and the relation with the ecological guilds was less obvious. Fish and crustaceans are highly mobile and can easily avoid low oxygen levels (Elliott & Hemingway 2002). In the lower part of the Schelde estuary, some species were found even at oxygen levels to 2 mg.l^{-1} (Maes *et al.* 1998b). During the study period, dissolved oxygen concentration was never a limiting factor, as it dropped below 5 mg.l^{-1} only once in the Westerschelde.

6.4.4 Density Models

Understanding the nature of possible interactions between variables is necessary to improve the understanding of changes in estuarine fish communities (Power *et al.* 2000b). The density models generally failed to predict extreme peaks and troughs, which is not uncommon to data series that fluctuate through several orders of magnitude (Nisbet & Gurney 1982). Still, most models agreed with the actual distribution of the species, as a vast majority of the sampling points fell within the 95% predictive confidence limits of the models. Only seven models could explain more than 45 % of the variance. No better models were built when only part of the data (*e.g* the monthly data from 1988-1991) was used. The simplicity of the models inevitably incorporated some loss of certainty as probably not all relevant forcing factors have been taken into account.

Fish and macro-invertebrate populations are characterized by high inter-annual variability (see previous chapters). The year-class strength is determined at the early-life history stages through a mechanism that is not yet well understood (see Costa *et al.* 2002). However, water temperature seems to be a key factor affecting the year-class variability. Low winter temperatures are favourable for several fish species (*e.g.* Pihl 1990, see Chapter 4), while mild winter temperatures are preferred by several macro-invertebrates (see Chapter 5).

Habitat complexity and habitat quality have been shown to be important for the distribution of fish and macro-crustaceans in estuaries (Elliott &

Hemingway 2002), but as such, these parameters could not be entered in the models. Food is probably a major factor determining habitat quality (Gibson 1994). However, if food is not a limiting factor (see Chapter 7 and Chapter 8), then temperature is likely to be the most important factor controlling growth (Gibson 1994). Also, the extent of predation might be important to describe habitat quality, but few studies have quantitatively examined predation (in relation to habitat complexity) in different habitats (Costa *et al.* 2002).

The final density models predicted both temporal and spatial patterns, in accordance with the presence/absence models. However, the withdrawn variables per species were not necessarily similar between both types of models. The standardized regression coefficients for temperature and salinity (or their quadratic effects) were highest in 5 and 6 models, respectively. Additionally, higher densities for 8 species were predicted with increasing salinities in the Westerschelde, all or not in combination with lower temperatures. Also in the Humber estuary (UK), salinity was the dominant factor influencing the fish distribution (Marshall & Elliott 1998). This is in contrast with the findings in the inner Severn (UK) and Elbe (Germany) estuaries, where pronounced annual cycles in species composition occurred irrespective of seasonal changes in salinity (Potter *et al.* 1997, Thiel & Potter 2001). Most probably this difference was related to the higher number (and density) of diadromous fish species present in the 2 latter systems.

Turbidity and dissolved oxygen concentration were entered as the first variable in 2 models each. In Australian and South African estuaries, the relation between turbidity and fish distribution was much more pronounced (Blaber 1997). Dissolved oxygen concentration was withdrawn in 7 density models, but only in combination with most other effects. This was probably due to the fact that the seasonal pattern in dissolved oxygen concentration was largely opposite to temperature. However, it seemed to be an important variable in the prediction of *Crangon crangon* densities, comparable with the models built for the shrimp population from the Thames estuary (Attrill *et al.* 1999).

Density estimates were better for most of the estuarine resident species, and least for most of the marine juvenile species. Best models were built for those species that were present in the estuary for a longer period, but with only one clear density peak in either a cold or a warmer season. This was independent of whether (shorecrab, common goby, sprat and flounder) or not (dab, flying crab and sandgoby) they were present in the estuary throughout the year (Hostens *et al.* 1996). The combined interaction between the four environmental variables was the main forcing factor behind the temporal and spatial distribution of these seven species. The lowest R^2 values were allocated to species that either were present in the estuary during only a short period (*e.g* bib, whiting and nilsson's pipefish), or to species with higher densities in both a warmer and a colder season (*e.g* sole, herring, plaice, lozano's goby and brown shrimp) (Chapter 2). Most probably, the temporal and spatial distribution of the latter eight species was more governed by the proper appearance of suitable food resources (Hemingway & Elliott 2002).

In a next step, the robustness of the models will be tested, by applying them to presence/absence and density data from other marine ecosystems, such as the Oosterschelde marine bay (Chapter 4, Chapter 5) and the Humber estuary (Marshall & Elliott 1998). Also, other models like the PISCES model by Henderson & Seaby (1994) or the ones generated for the Thames estuary (Attrill *et al.* 1999), will be applied to the fish and macro-crustacean data from the Westerschelde estuary.

Acknowledgements I would like to thank the people from the R.V. Luctor and the Centre for Estuarine and Marine Ecology (Yerseke, the Netherlands) for the logistics; Dr. Tom Ysebaert (CEME) for the comments on building logistic regressions and the extra environmental data; Dr. Jan Mees (Flanders Marine Institute) and Dr. Ann Vanreusel (Marine Biology Lab) for the invaluable criticism. This study was supported by project no. G.0104.99 of the Flemish Science Foundation and by contract BOF 98-03, 12.0503.98 of the Ghent University.