

# Regional and local environmental correlates of native Iberian fish fauna

Ferreira MT, Sousa L, Santos JM, Reino L, Oliveira J, Almeida PR, Cortes RV. Regional and local environmental correlates of native Iberian fish fauna.

Ecology of Freshwater Fish 2007: 16: 504–514. © 2007 The Authors. Journal compilation © 2007 Blackwell Munksgaard

**Abstract** – Knowledge of fishes' habitat requirements is an essential tool for species conservation. We examined the habitat requirements of nine native freshwater fish species using 165 undisturbed or minimally disturbed sites from eight river basins in western Iberia (northern and central Portugal). Six environmental descriptors were considered at the regional level, and six descriptors at a local scale. Statistical associations between local assemblage structure and environmental variables were quantified by canonical correspondence analyses. Logistic regression was performed for each species to identify regional and local factors related to its distribution at the sample sites. The results suggested a reasonable degree of specialisation in the use of local variables and dependence on regional variables. Logistic regression models showed a good adjustment ( $0.75 < \text{AUC} < 0.93$ ). Results of habitat modelling using minimally disturbed sites helped to clarify species preferences for specific abiotic drivers.

**M. T. Ferreira<sup>1</sup>, L. Sousa<sup>1</sup>,  
J. M. Santos<sup>1</sup>, L. Reino<sup>1</sup>,  
J. Oliveira<sup>1</sup>, P. R. Almeida<sup>2</sup>,  
R. V. Cortes<sup>3</sup>**

<sup>1</sup>Departamento de Engenharia Florestal, Instituto Superior de Agronomia, Tapada da Ajuda, Lisbon, Portugal, <sup>2</sup>Departamento de Biologia, Universidade de Évora, Évora, Portugal, <sup>3</sup>Departamento Florestal, Universidade de Trás-os-Montes e Alto Douro, Quinta de Prados, Vila Real, Portugal

**Key words:** habitat requirements; local/regional environment; logistic regression; variation partitioning; Portugal

M. T. Ferreira, Departamento de Engenharia Florestal, Instituto Superior de Agronomia, Tapada da Ajuda, 1349-017 Lisbon, Portugal; e-mail: terferreira@isa.utl.pt

Accepted for publication March 8, 2007

## Introduction

Freshwater ecosystems have frequently been cited as among those that are most altered and threatened by human activities (Mason 1991; Saunders *et al.* 2002; Clavero *et al.* 2004). Although initial approaches to decline and extinction phenomena were mostly centred on terrestrial organisms, more recently increasing attention has been paid to the conservation of freshwater fauna (Harding *et al.* 1998; Ricciardi & Rasmussen 1999). However, to manage aquatic biodiversity we must understand how fish species are distributed and be able to predict their occurrence in a given area (Ormerod 2003). Knowledge of fishes' habitat requirements is an essential tool for species conservation and fishery management (Heggenes *et al.* 1991), but the relationship between fish species and fish assemblages and habitat is an extremely complex one (Hubert & Rahel 1989).

The use of different uni- or multivariate methods to study species distribution and habitat requirements has increased over the last 20 years (Leps & Smilauer

2003). Regression models in particular have been widely used to predict species distribution, abundance and habitat preference (e.g., Beja & Alcazar 2003; Filipe *et al.* 2004; Moreira *et al.* 2005). As predicted by Margalef (1968), presence/absence data permit the detection of macroscale patterns in community ecological studies. However, most studies either consider sites with different degrees of human disturbance, where populations may thus be in a suboptimal habitat, or else include human-mediated variables to describe the habitat. A valid understanding of natural patterns and processes is an essential prerequisite for viable ecosystem management (Stanford *et al.* 1996). We have therefore focused our research efforts on investigating patterns and processes in riverine landscapes that are in a near-natural state.

A number of studies indicate that fish assemblage organisation can be determined not only by local processes, but also by large-scale environmental factors (Schlosser 1991; Marsh-Matthews & Matthews 2000; Magalhães *et al.* 2002; Hoeinghaus *et al.* 2006). The importance of the scale of environmental variables

that determine the distribution of individual fish species has not received as much attention (Taylor 1996; Wiley *et al.* 1997), because most population studies are focused on local conditions, a particular river or a specific region (e.g., Bovee 1982; Gore & Nestler 1988; Leftwich *et al.* 1997). In some cases, it has been possible to model the occurrence of fish species at the individual drainage basin scale (Oberdorff *et al.* 2001; Filipe *et al.* 2004), and to test the hypothesis of a common response to local and regional environmental variables (Pont *et al.* 2005). In addition, local habitat characteristics, such as water volume, velocity and thermal regime, are often dependent on regional drivers, such as drainage area, hydrological regime and altitude (Poff & Allan 1995; Wiley *et al.* 1997). The relative role and importance of scale factors in shaping and regulating species distribution can also be related to integrity assessment and conservation efforts, because these depend on species-specific perturbation drivers and the spatial scale upon which they act.

In this study we aim: (i) to look at the habitat requirements of nine native Iberian fish species by identifying their environmental correlates at undisturbed or minimally disturbed sites; and (ii) to assess the relative importance of these local and regional environmental variables on the occurrence of each species.

### Study area

The study area is located in western Iberia and includes the hydrological network of central and northern Portugal (Fig. 1). These rivers flow NE–SW towards the West Atlantic coast, and the largest basins (Minho, Lima, Douro and Tagus) are shared with Spain. The inland area is dominated by the pre-Mesozoic Hesperic Massif – a complex geological unit that includes granites, schists and quartzites with a various degrees of metamorphism – whereas most of the western coastal fringe is characterised by Tertiary layers under Quaternary deposits. The southern region below the main course of the River Tagus is dominated by the flat platforms of the Hesperic Massif, and in the lower boundary by the Meso-Cenozoic strata, which is composed of sedimentary deposits (e.g., sandstones, limestones, conglomerates and marls), mainly with a calcareous background. Littoral areas form part of the Cenozoic Sedimentary Basin, with soft clay and marl deposits.

The rainfall pattern exhibits strong seasonal and inter-annual variability, as well as a discharge regime that is typical of Mediterranean rivers. High floods usually occur in autumn or early winter, with a gradual decline in flow that eventually dries out during late spring and summer in the southernmost areas. Asymmetries in the regional climate are well marked, and



Fig. 1. Map of the study area and location of sampling sites in the different river basins.

influenced mainly by elevation and the proximity of the Atlantic Ocean, with longer summer droughts in the continental parts and southern regions. Mean annual temperatures in the latter are typically above 17.5 °C, sometimes with less than 400 mm of mean annual rainfall. By contrast, the northern border of the study area presents a more temperate-like climate: the annual rainfall can reach 2200 mm and mean annual temperature is below 10 °C.

The landscape varies greatly across the studied area, with a patchy pattern of land use in central and western basins, including orchards, vineyards, olive groves and irrigation crops. Northern areas and those inland or at higher altitudes are mostly forested, with *Pinus* and *Eucalyptus* plantations, or Mediterranean woods, whereas the southern areas are more homogenous and characterised by cork-oak and holm-oak woodlands. The largest human settlements and industries are located in littoral areas.

## Material and methods

### Data collection

Fish sampling was carried out in early spring by electrofishing using European Committee for Normalization standards (norm 14011/2002 for wadeable streams) at 165 sites from eight river basins (Minho, Lima, Cávado, Douro, Vouga, Mondego, Lis and Tagus) in western Iberia (northern and central Portugal). Sample sites were considered undisturbed or minimally disturbed and had previously been extracted from a data set of more than 500 available sampling locations. These were identified using existing European guidelines (Refcond 2003) and the EU-FAME

project approach (<http://fame.boku.ac.at>), in which human disturbance is ranked semiquantitatively using all field data and available GIS information. In this case, we considered five disturbance variables: integrity of the riparian galleries and naturalness of the banks, longitudinal connectivity of the river segment, sediment load, input of toxic and acidic effluents, and nutrient input. Each variable was scored from 1 (minimum disturbance) to 5 (maximum disturbance), and only sites scoring 1 or 2 for all five variables were retained for the analysis.

Local and regional environmental variables were determined for each site according to their scale of definition and spatial variability (Table 1). The local variables were obtained *in situ* and regional variables were obtained from arcview GIS and 1:25,000 topographic maps. Local variables included mean width (m), maximum depth (m), instream cover [defined as any submerged structure in which fish could be hidden from overhead view, undercut banks and surface water turbulence and scored subjectively from 0 (absent) to 3 (abundant)], abundance of macrophytes (%), dominant substrate coarseness (determined visually as 1 = silt, 2 = sand, 3 = gravel/pebble, 4 = cobble, 5 = boulder and 6 = bedrock), mean depth (m), and overhanging tree shading (scored as 0: absent, 1: <30%, 2: 30–60% and 3: >60%). Regional variables included mean air temperature (°C), catchment area (km<sup>2</sup>), altitude (m), gradient slope (%), mean annual rainfall (mm), evapotranspiration (mm), average annual runoff (mm) and distance to source (km). To deal with spatial autocorrelation in the response variable, we included an autocorrelation term (see Augustin *et al.* 1996). This term was generated for each site as the

	Code	Mean/median	Range
Regional variables			
Mean air temperature (°C)	ATEMP	13.00 (±2.10)	7.50–25.0
Altitude (m)	ALT	324.40 (±243.50)	6.7–1034.0
Gradient slope (%)	SLOP	1.9 (±2.4)	0.0–16.7
Evapotranspiration (mm)	EVA	601.4 (±113.3)	425–900
Distance to source (m)	DISTSOURC	27,973.0 (±29,338.2)	1000–136,358
Catchment area (km <sup>2</sup> )	CAREA	264.61 (±475.72)	0.62–2722.50
Mean annual rainfall (mm)	MAR	1166.60 (±534.58)	450.00–3000.00
Average annual runoff (mm)	AAR	8(4)	
Local variables			
Mean depth (m)	MDEP	0.5 (±0.3)	0.1–2.0
Maximum depth (m)	MAXDEP	1.10 (±0.55)	0.20–2.50
Mean width (m)	MWIDTH	9.27 (±9.61)	1.00–57.00
Instream cover (class)	COV	2 (2)	
Abundance of macrophytes (%)	MACRO	7.65 (±9.55)	0.00–51.00
Substrate coarseness (class)	DOMSUBST	3(3)	
Shading (class)	SHAD	2(1)	
Spatial variable			
Auto-correlation	ACORR		

Table 1. Code, mean/median and range of environmental variables measured at 165 sites in western Iberia.

Mean values are given for ATEMP, ALT, SLOP, EVA, DISTSOURC, CAREA, MAR, MAXDEP, MWIDTH and MACRO followed by standard error (in parentheses), while median values (with interquartile distance in parentheses) are given for MAD, COV, DOMSUBST and SHAD.

number of neighbouring sampling locations within 30 km in the river continuum where a given species occurs.

#### Data analysis

A total of 22 species were sampled in western Iberia, but only native fish species with a broad distribution (i.e., occurring at more than 10% of sampling sites) and presenting a reasonable number of individuals along environmental gradients were considered in the analysis. Altogether, nine species were used in the analysis.

Statistical association between assemblage structure and environmental variables was quantified by canonical correspondence analysis (CCA) (Ter Braak & Verdonschot 1995), using CANOCO 4 (Ter Braak & Smilauer 1998) on a 165 sites  $\times$  9 species matrix. Separate CCAs were performed using local and regional environmental data sets. A forward stepwise selection procedure was used to obtain the best predictors for the model, at a significance level of 0.05, using a Monte Carlo permutation test (999 permutations). Fish densities were  $\log(x + 1)$  to minimise the effect of extreme values (Legendre & Legendre 1998).

Logistic regression analysis was used to model the influence of different variables (local and regional) on the occurrences of nine fish species. In Mediterranean rivers, many native species are restricted to one or a few basins for palaeo-historical and evolutive reasons (Doadrio 2001). To account for this in the models that we developed for each species, we only considered their natural area of occurrence by analysing only basins where a species is potentially present. Firstly, Spearman's analysis of correlations between the environmental variables was used to detect redundant variables, and those for which correlation values were greater than  $|0.8|$  were excluded from the analysis. Twelve variables were considered, plus the variable used to deal with spatial autocorrelation. Secondly, a stepwise selection procedure was applied to each variable with a probability of entry of  $P \leq 0.05$  and removal of  $P > 0.05$  (see Hosmer & Lemeshow 1989). The association between the explanatory variables and interactions with the response variable was tested using the maximum likelihood method (Hosmer & Lemeshow 1989). This procedure was first applied separately for the variables from the two different scales (local and regional), and subsequently using only the significant environmental correlates obtained in these scales to produce the final model.

#### Assessing prediction accuracy

The predictive accuracy of the final models was assessed using the area under the curve (AUC) of the

receiver operating characteristic (ROC) curve (Fielding & Bell 1997; Pearce & Ferrier 2000). We used this graphical method to represent the relationship between the percentage of presences correctly predicted (sensitivity) and one minus the percentage of the absences correctly predicted (specificity). The area comprised between a 45° line and the curve measures the discrimination – i.e., the model's ability to correctly classify a species as present or absent at a given plot (AUC). We used a rough guide to classify the accuracy: 0.50–0.60 = fail; 0.60–0.70 = poor; 0.70–0.80 = fair; 0.80–0.90 = good; 0.90–1 = excellent (Swets 1988).

#### Variation partitioning for GLM models

The variation in the data set was partitioned to assess the relative importance of regional and local components in the selected logistic models. The methodology we followed is an adaptation of that proposed by Borcard *et al.* (1992) and Legendre & Legendre (1998), but extended to logistic regression (e.g., Reino 2005). The Nagelkerke pseudo- $r^2$  (Nagelkerke 1991) was used as a measure of variation explained by each logistic model. The total variation in the expansion model was partitioned into four independent components: (i) a pure local component,  $a$ ; (ii) a shared component – i.e., the local variation that is regionally structured,  $b$ ; (iii) a pure regional component that is probably due to large-scale effects,  $c$ ; and (iv) unexplained variation,  $d$ . The total variation was partitioned by the expression  $Vt = \sum a + b + c + d$ , which was used to extract any single component. This procedure involves calculating the following statistical models: (i) one 'partial' logistic regression using local variables to obtain the significant terms ( $a + b$ ) using a stepwise selection procedure; (ii) another 'partial' logistic regression for regional descriptors ( $b + c$ ), again using a stepwise selection procedure; and (iii) one joint local–regional model with all previously significant variables to obtain the shared effects ( $a + b + c$ ). All data analyses were performed with SPSS v. 12.0.

## Results

#### Fish assemblages

The fish species we examined were well distributed throughout the sample sites and presented a mean occurrence of at least 15% in our data set (Table 2). *Barbus bocagei* Steindachner, *Salmo trutta* Linnaeus and *Squalius carolitertii* Doadrio were the most frequently captured species and occurred at more than 42% of the sampling sites. *Cobitis paludica* de Buen and *Chondrostoma oligolepis* Steindachner were the

Table 2. Mean, standard error (SE), maximum (Max) and minimum (Min) values of relative abundance for the nine species collected at 165 sites in western Iberia.

Species	Mean	SE	Max	Min	FO(%)
<i>Barbus bocagei</i> (Steindachner)	50.7	13.2	1255	0	45.8
<i>Squalius pyrenaicus</i> (Günther)	8.8	2.2	214	0	21.1
<i>Squalius alburnoides</i> (Steindachner)	14.5	3.5	358	0	24.1
<i>Squalius carolitertii</i> (Doadrio)	22.2	4.8	515	0	42.2
<i>Chondrostoma polylepis</i> (Steindachner)	13.9	2.8	321	0	34.9
<i>Chondrostoma duriense</i> (Coelho)	12.8	4.1	486	0	18.7
<i>Chondrostoma oligolepis</i> (Steindachner)	9.0	2.2	217	0	15.7
<i>Cobitis paludica</i> (De Buen)	2.7	0.8	75	0	16.3
<i>Salmo trutta</i> (Linnaeus)	10.0	1.5	100	0	44.0

Mean frequency of occurrence of each species (FO, %) is also shown.

least frequent species, and occurred at 16–18% of the sampling sites – mostly in the Douro (53%) and Lis River basins (100%), respectively. The smaller cyprinids *Squalius pyrenaicus* Günther and *Squalius alburnoides* Steindachner were mainly restricted to the Tagus River basin, whereas *C. paludica* was not present in the northern part of the study area, nor was *Chondrostoma duriense* Coelho found in the southern part of the study area.

Multivariate response of fish communities

The results obtained in CCA based on regional variables (Fig. 2a) revealed a modest species separation along the two canonical axes ( $\lambda_1 = 0.380$ ;  $\lambda_2 = 0.145$ ,  $P < 0.05$ ). The explanatory variables accounted for 16% of the total variation in site-assemblage composition, with the first and second axes displaying 83% of that variation. The first axis revealed a gradient of fish assemblages going from

those composed mostly of *S. trutta*, followed by a dominance of small cyprinids, such as *S. carolitertii* and *S. pyrenaicus*, to assemblages characterised by *B. bocagei* and *C. duriense*. This gradient was associated with altitude ( $t = 4.55$ ,  $P < 0.05$ ), mean annual temperature ( $t = 6.27$ ,  $P < 0.05$ ) and average annual runoff ( $t = 3.66$ ,  $P < 0.05$ ).

The first two axes in the CCA ordination based on local variables (Fig. 2b) were significant ( $P < 0.05$ ), with eigenvalues of 0.259 and 0.187, respectively. The explanatory variables accounted for 16.1% of the total variation in site-assemblage composition, with the first and second axes displaying 85.3% of that variation. As depicted in the biplot, the fish community changed along the first axis from assemblages characterised by the larger sized cyprinids, *B. bocagei* and *C. duriense* and also *S. trutta*, to assemblages characterised by small-sized cyprinids, such as *S. pyrenaicus*, *S. alburnoides* and *C. oligolepis*. This separation was mainly related to instream cover ( $t = 7.43$ ,  $P < 0.05$ ) and abundance of macrophytes ( $t = -4.44$ ,  $P < 0.05$ ). The second axis revealed a weaker gradient associated with substrate coarseness ( $t = -3.89$ ,  $P < 0.05$ ) and shading ( $t = -3.94$ ,  $P < 0.05$ ).

Models for individual species

The variables, evapotranspiration, distance to source and mean depth, were excluded ( $r \geq |0.8|$ ) from the analysis. Twelve variables were therefore retained for subsequent analyses. The autocorrelation variable was not significant for any species univariate models and we did not include this term in further analyses.

The models showed that species occurrences were significantly related to both regional and local

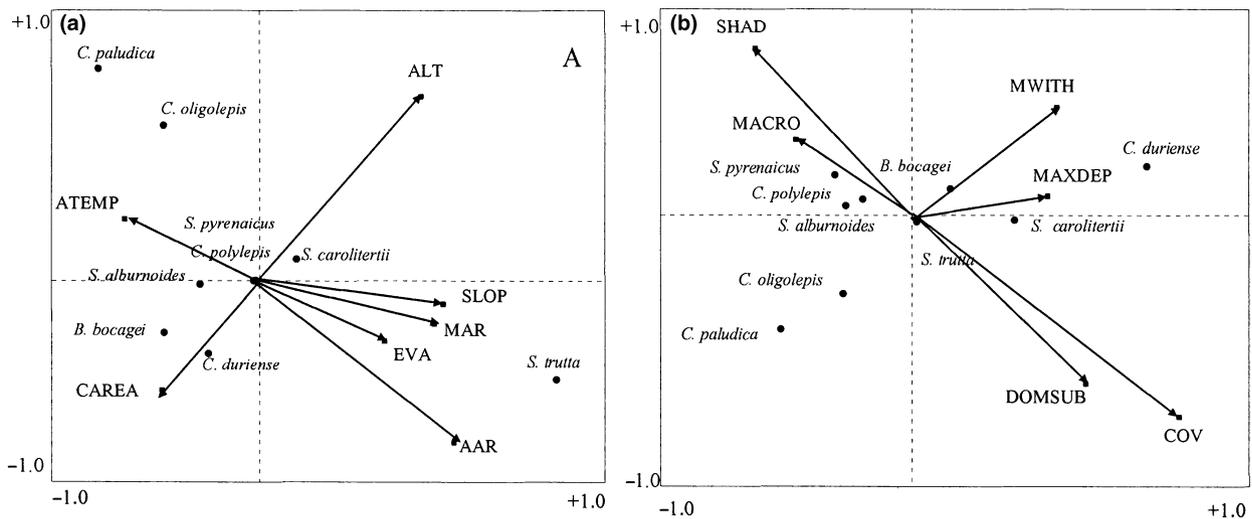


Fig. 2. Axes one and two of the canonical correspondence analysis biplot for the 9 species and (a) regional variables, (b) local variables, assessed in western Iberia.

variables, thereby suggesting an interplay of influences between scales (Table 3). Substrate coarseness and overhanging shading were the most frequently selected local variables. Geomorphological variables (gradient slope and altitude) also played an important role in the occurrence of species – namely *Chondrostoma polylepis* Steindachner *C. oligolepis*, *S. carolitertii*, *S. trutta* and *B. bocagei*. The latter species was mainly associated with larger lowland streams. By contrast, *S. trutta* displayed a positive association with altitude, depth and shading and a negative association with mean annual temperature. Local habitat variables, such as stream width, cover, substrate coarseness and shading, were significant correlates for the occurrence of *S. pyrenaicus*, *S. alburnoides* and *C. duriense*. *Cobitis paludica* was found to be mainly confined to warmer areas presenting finer substrates – i.e., coastal and lower courses of main rivers. In the final model, some species related more to regional (*C. oligolepis*) or local (*C. duriense* and *S. pyrenaicus*) variables.

Accuracy of GLM models

None of the models was classified in the low-accuracy classes (AUC < 0.7) (Table 4). Fit and classification accuracies of logistic regression models were high, indicating a strong predictive power (84%). The model for *S. alburnoides* had the lowest model accuracy (75%). Good accuracy class (0.8 < AUC < 0.9) was assigned to four species belonging to four genera, with *S. trutta* and *C. oligolepis* models displaying the highest levels (93% and 90%, respectively).

Variation partitioning for GLM models

Local, shared and regional components (considering all significant variables obtained in stepwise logistic regression) managed to explain between 70.5% (*S. trutta*) and 23.5% (*S. carolitertii*) of model variation, but the explanation of the distribution patterns varied from species to species (Fig. 3). Some of the studied distributions are highly associated with pure local attributes which are mainly related to habitat (*S. alburnoides*, *S. pyrenaicus* and *C. duriense*), and which ranged from 17.1% to 27.9% of total variation. By contrast, after accounting for local attributes, pure regional covariates are highly represented in three species (*S. trutta*, *C. paludica* and *C. oligolepis*), thus suggesting an underlying large-scale spatial trend in their distribution. Moreover, shared effects also represented a variation in distribution, from 1.9% (*S. alburnoides*) to 35.3% (*B. bocagei*). We should note that species which are highly associated with regional variables also have strong shared effects, thus suggesting that some of the local attributes may be

Table 3. Summary of GLM analysis for the occurrence of fish species in western Iberia including the number of sites (N), the  $\beta$ -estimated coefficients and directions of association (– or +), the levels of significance obtained from likelihood ratio statistics (given in parentheses ( $P < 0.05$ )).

Species	N	Regional variables										Local variables				
		ATEMP	ALT	SLOP	CAREA	MAR	MAD	MAXDEP	MWIDTH	MACRO	COV	DOMSUBST	SHAD			
<i>Barbus bocagei</i>	165		-0.004 (0.000)	-0.629 (0.001)	ns			ns	0.239 (0.000)		ns		1.298 (0.000)	0.037 (0.042)		
<i>Squalius pyrenaicus</i>	53			ns		ns							0.681 (0.012)			
<i>Squalius alburnoides</i>	143			-0.276 (0.002)				1.251 (0.005)								
<i>Squalius carolitertii</i>	112			-0.723 (0.000)				ns								
<i>Chondrostoma polylepis</i>	86		ns		ns			ns	0.064 (0.009)		ns			0.022 (0.003)		
<i>Chondrostoma duriense</i>	79			-1.227 (0.005)				ns			1.231 (0.000)		ns	ns		
<i>Chondrostoma oligolepis</i>	163		-0.011 (0.000)		-0.002 (0.032)			ns								
<i>Cobitis paludica</i>	86	0.455 (0.001)	ns													
<i>Salmo trutta</i>	162	-0.526 (0.001)	0.003 (0.013)					1.236 (0.011)						0.030 (0.002)		

ns represents the variables selected in local or regional models that were not selected in the final global model.

Species	AUC (%)	AUC sig.	Sensitivity	Specificity	Nagelkerke $r^2$
<i>Barbus bocagei</i>	88	$P < 0.001$	0.79	0.75	0.57
<i>Squalius pyrenaicus</i>	87	$P < 0.001$	0.67	0.83	0.49
<i>Squalius alburnoides</i>	75	$P < 0.001$	0.65	0.7	0.28
<i>Squalius carolitertii</i>	78	$P < 0.001$	0.59	0.77	0.24
<i>Chondrostoma polylepis</i>	76	$P < 0.001$	0.75	0.61	0.25
<i>Chondrostoma duriense</i>	88	$P < 0.001$	0.91	0.66	0.49
<i>Chondrostoma oligolepis</i>	90	$P < 0.001$	0.91	0.61	0.48
<i>Cobitis paludica</i>	82	$P < 0.001$	0.66	0.51	0.3
<i>Salmo trutta</i>	93	$P < 0.001$	0.84	0.8	0.7

Table 4. Discrimination ability for optimal GLM models for each species including the percentage of the area under the curve (AUC), level of significance, sensitivity and specificity (see Material and methods for explanation) and Nagelkerke's  $r^2$ .

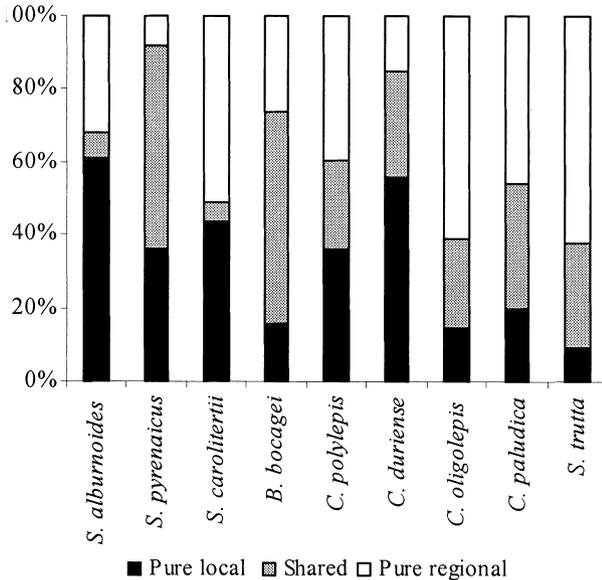


Fig. 3. Variation partitioning applied to the nine GLMs developed for each fish species in western Iberia.

regional or spatially structured. For example, in *S. trutta* this accounts for 20% of total variation.

### Discussion

This study focused on fish distribution in western Iberia and on the relative importance of local and regional factors that shape the fish assemblage structure. The most significant way in which our study differed from most others was that the sampling network consisted of undisturbed or minimally disturbed sites.

Canonical correspondence analysis revealed that generic patterns in the fish assemblage composition in western Iberia were primarily associated with watershed features and the hydrological regime, such as altitude, slope, rainfall and run-off. Overall, the fish community went from assemblages dominated by *S. trutta*, followed by assemblages with a dominance of small cyprinids, such as *S. pyrenaicus*, *S. alburnoides* and *C. oligolepis*, to assemblages characterised by the larger cyprinids, *B. bocagei* and *C. duriense*. Our results from the network of minimally disturbed sites

generally corroborated the organisation of freshwater fish communities in western Iberia that had been revealed in some earlier works (Godinho *et al.* 1997; Filipe *et al.* 2002; Santos *et al.* 2004a) and in other geographical areas (Lyons 1989; Rahel & Hubert 1991). Assemblage composition was also related to local-scale features, such as cover, abundance of macrophytes, substrate coarseness and shading – variables that other studies (Godinho *et al.* 2000; Pires *et al.* 2004; Santos *et al.* 2006) have found to be important. Cover may be important for resting and hiding, and can be related to invertebrate production (particularly submerged and emergent species) and structural habitat complexity (Moyle & Cech 1996). The availability of coarse substrate is important in that it provides protection against widespread mammals (such as the otter *Lutra lutra* and beech marten *Martes foina*) and birds (such as kingfisher *Alcedo atthis*, grey heron *Ardea cineria* and little egret *Egretta garzetta*) and prevents fish from being washed away by high flow events. Abundance of macrophytes was also found to be essential. In fact, cyprinids, such as *B. bocagei* and *C. polylepis*, eat large amounts of plant material (Magalhães 1992; Collares-Pereira *et al.* 1996), and *S. alburnoides* is a phytophil spawner (Fame Consortium 2004). Regional and local-scale factors were thus both significant correlates of the fish assemblage structure. However, there may be other factors influencing species distribution, including population dynamics, species interactions and historical factors, such as isolation arising from the formation and evolution of the basin systems (Carmona *et al.* 1999). Furthermore, our analyses should be viewed with caution as they represent an ecological snapshot, and interannual or more sampling-intensive studies would help to clarify the environmental correlates that are subject to temporal variability (Lohr & Fausch 1997; Wiley *et al.* 1997; Magalhães *et al.* 2002).

The generally high fit and classification accuracy of the logistic regression models support the basic hypothesis of our modelling approach that significant relationships exist between local habitat features, spatial variables and the occurrence of fish species in western Iberian rivers. However, several factors could

have been responsible for the lower accuracy classification that occurred with some of the models, particularly in the case of *S. alburnoides*, *S. carolitertii* and *C. polylepis*. There may also be habitat measures that are missing from this analysis and could increase predictive power. For example, additional measures often used in predictive models for fish at a local scale include dissolved oxygen, pH and conductivity (Bain & Robinson 1988; Fausch *et al.* 1988), and these variables could be important in this system as well. There may be other factors influencing the presence of fishes, including site-based population dynamics, species interactions and historical factors (Guisan & Zimmermann 2000). Fish are also constrained by their historically defined drainage basins (Williams *et al.* 2003). Unlike other authors (Hinch *et al.* 1994), who found that spatial autocorrelation might strongly affect inferences about large-scale distributional patterns, in our study this variable was not a determinant for species models, possibly due to the larger distance between neighbours (30 km) when compared with other studies (Magalhães *et al.* 2002).

Our models clearly emphasised the importance of examining species–habitat relationships at different spatial scales (local and regional). Altitude remained a major ecological factor for some species: *B. bocagei* and *C. oligolepis* were negatively related to this variable, whereas *S. trutta* showed a positive association. It is known that altitude influences species occurrence through water temperature (Huet 1949), thus indirectly governing fish population density via growth and fecundity (Lobón-Cerviá *et al.* 1996). Slope was also a regional key variable selected by the models, as it displayed significant negative associations for four cyprinid species (*B. bocagei*, *S. carolitertii*, *C. polylepis* and *C. oligolepis*). Slope makes a major contribution to the erosive force acting on substrate and bed scour in a given area (Cobb *et al.* 1992). Weaker slopes within our study area are associated with higher habitat heterogeneity, thus favouring higher species richness (e.g., coexistence of cyprinid species).

Physical habitat and channel morphology descriptors also played an important role in species occurrence. Stream width was positively related to the probability of the presence of the larger cyprinids, *B. bocagei* and *C. duriense*. This pattern (i.e., larger fish associated with wider and deeper habitat) has been found elsewhere in other Iberian rivers (Rincon *et al.* 1990; Godinho *et al.* 1997; Carmona *et al.* 1999; Pires *et al.* 1999; Santos *et al.* 2004b). Substrate coarseness and shading were also important correlates of fish species occurrence. The availability of coarse substrate for small-sized species, such as *S. pyrenaicus* and *S. alburnoides*, is important in that it provides cover and concealment and also relatively

better habitat for macroinvertebrates that provide them with a source of food (Doadrio 2001). In contrast, *C. paludica*'s negative response to substrate coarseness is consistent with some studies on this species (Perdices & Doadrio 1997; Valladolid & Przybylski 2003; Clavero *et al.* 2005). Shading was positively related to probability of occurrence of *S. pyrenaicus*, *C. polylepis* and *S. trutta*. The importance of shading, and thus canopy cover, to fishes has been demonstrated in a number of papers (Schlosser 1991; Schiemer *et al.* 1995; Stauffer *et al.* 2000). Larger fish, such as *C. polylepis* and *S. trutta*, have been found to favour areas with a high canopy cover, and this could be due to a combination of factors including physical stresses, food availability, predation risk or competitive interactions (Schlosser 1991; Collares-Pereira *et al.* 1995; Mellina *et al.* 2005; Santos *et al.* 2006). Shading is also important to small-sized *S. pyrenaicus*, as it is a major component of instream cover that provides refuge in well-conserved upper reaches and small Iberian streams (Corbacho & Sánchez 2001; Clavero *et al.* 2005). But our study also provided insights into the occurrence of less well-known Iberian species. For example, *C. oligolepis* was found to be negatively associated to large-scale descriptors, such as altitude, slope and catchment area – i.e., it inhabits mainly small coastal rivers and the lower reaches of main tributaries.

Both local and regional attributes were significant correlates of fish species distribution in western Iberia. The purely local attributes accounted for much of the model variation of *C. duriense* and *S. alburnoides*. Previous authors (Godinho *et al.* 1997; Pires *et al.* 1999; Santos *et al.* 2006) had pointed out that the distribution of these species is governed first by the suitability of physical habitat and hydraulic conditions, such as cover, substrate and velocity. On the other hand, factors that operate on a regional scale (temperature, altitude and catchment area) possessed a greater explanatory power for the two species inhabiting the smallest streams (*S. trutta* and *C. oligolepis*). Our results seem to demonstrate that the responses of species that inhabit small-order streams differed from those occurring in larger rivers. Also, the largest proportion of explained variation for some of the species (*S. pyrenaicus* and *B. bocagei*) was due to the interactions between both sets of attributes. Moreover, species that were highly associated with regional variables (*S. trutta*, *C. oligolepis* and *C. paludica*) also had strong shared effects, thus suggesting that some of the local attributes may be regional or spatially structured. This overlapping could be a consequence of local data acting partially as a synthetic descriptor of unmeasured underlying regional processes. Future research should be extended to the influence of other factors, in an attempt to further elucidate the

mechanisms governing the occurrence of Iberian fish species.

Knowing how fish species respond to habitat heterogeneity on multiple scales can be useful to fishery managers, watershed restoration groups and others with an interest in conservation (Fausch *et al.* 2002; Filipe *et al.* 2002; Allan 2004). Species that respond primarily to local factors, such as *S. alburnoides*, *S. pyrenaicus* and *C. duriense*, may have a patchy distribution and hence be more vulnerable to immediate disturbances, but their populations are nonetheless more easily recovered, while species that are principally governed by regional factors, such as *S. trutta* and *C. oligolepis*, would be especially affected by broad-scale disturbances, such as climate change and river discontinuities, and therefore less responsive to management actions. Nonetheless, the high rate of endemism of Iberian native ichthyofauna (Clavero *et al.* 2004) should deserve special concern for the conservation of inland waters. All studied species, particularly *S. alburnoides* and *S. pyrenaicus*, are currently threatened by deterioration or destruction of habitats, both by pollution and intense modifications (i.e., damming and channelisation) and the introduction of exotic species (Cabral *et al.* 2006). The GLM models provided insights into how both landscape and reach-scale factors affect fish species occurrence, and are thus useful for setting or refining management goals and evaluating the effectiveness of species management and conservation. Subsequent efforts that include ecological attributes of fish species, such as reproductive and trophic guilds (Oberdorff *et al.* 1993; Belliard *et al.* 1997) and also colonisation abilities (Grenouillet *et al.* 2004), and others that focus on variables related to dispersal success (Dunham & Rieman 1999), may improve predictions of the occurrence of native Iberian fish fauna.

### Acknowledgements

The authors would like to thank Paulo Pinheiro, Luís Lopes, António Albuquerque and Patricia Gonzalez for help in the field and GIS work. We are also indebted to the anonymous reviewer, whose suggestions greatly improved the manuscript.

### References

- Allan, J.D. 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. *Annual Review of Ecology, Evolution and Systematics* 35: 257–284.
- Augustin, N.H., Mugglstone, M.A. & Buckland, S.T. 1996. An autologistic model for the spatial distribution of wildlife. *Journal of Applied Ecology* 33: 339–347.
- Bain, M.B. & Robinson, C.L. 1988. Structure, performance, and assumptions of riverine habitat suitability index models. Alabama Cooperative Fish and Wildlife Research Unit, Aquatic Resources Research Series 88-3, Auburn.
- Beja, P. & Alcazar, R. 2003. Conservation of Mediterranean temporary pond under agricultural intensification: an evaluation using amphibians. *Biological Conservation* 114: 317–326.
- Belliard, J., Boët, P. & Tales, E. 1997. Regional and longitudinal patterns of fish community structure in the Seine River basin, France. *Environmental Biology of Fishes* 50: 133–147.
- Borcard, D., Legendre, P. & Drapeau, P. 1992. Partialling out the spatial component of ecological variation. *Ecology* 73: 1045–1055.
- Bovee, K.D. 1982. A guide to stream habitat analysis using the instream flow incremental methodology. *Instream Flow Inf. Pap.* 12. Fort Collins, CO: US Fish and Wildlife Service.
- Cabral, M.J., Almeida, J., Almeida, P.R., Delinger, T., Ferrand de Almeida, N., Oliveira, M.E., Palmeirim, J.M., Queiroz, A.L., Rogado, L. & Santos – Reis, M. 2006. Livro vermelho dos vertebrados de Portugal, 2nd edn. Lisboa: Instituto da Conservação da Natureza/Assírio & Alvim (Portuguese Red Data Book).
- Carmona, J.A., Doadrio, I., Márquez, A.L., Real, R., Hugueny, B. & Vargas, J.M. 1999. Distribution patterns of indigenous freshwater fishes in the Tagus River basin, Spain. *Environmental Biology of Fishes* 54: 371–387.
- Clavero, M., Blanco-Garrido, F. & Prenda, J. 2004. Fish fauna in Iberian Mediterranean river basin: biodiversity, introduced species and damming impacts. *Aquatic Conservation: Marine and Freshwater Ecosystems* 14: 575–585.
- Clavero, M., Blanco-Garrido, F. & Prenda, J. 2005. Fish–habitat relationships and fish conservation in small coastal streams in southern Spain. *Aquatic Conservation: Marine and Freshwater Ecosystems* 15: 415–426.
- Cobb, D.G., Galloway, T.D. & Flannagan, J.F. 1992. Effects of discharge and substrate stability on density and species composition of stream insects. *Canadian Journal of Fisheries and Aquatic Sciences* 49: 1788–1795.
- Collares-Pereira, M.J., Magalhães, M.F., Geraldes, A.M. & Coelho, M.M. 1995. Riparian ecotones and spatial variation of fish assemblages in Portuguese lowland streams. *Hydrobiologia* 303: 93–102.
- Collares-Pereira, M.J., Martins, M.J., Pires, A.M., Geraldes, A.M. & Coelho, M.M. 1996. Feeding behaviour of *Barbus bocagei* assessed under a spatial–temporal approach. *Folia Zoologica* 45: 65–76.
- Corbacho, C. & Sánchez, J.M. 2001. Patterns of species richness and introduced species in native freshwater fish faunas of a Mediterranean-type basin: the Guadiana River (southwest Iberian Peninsula). *Regulated Rivers: Research and Management* 17: 699–707.
- Doadrio, I. 2001. Atlas y libro rojo de los peces continentales de España. Madrid: Museo Nacional de Ciencias Naturales.
- Dunham, J.B. & Rieman, B.E. 1999. Metapopulation structure of bull trout: influences of physical, biotic, and geometrical landscape characteristics. *Ecological Applications* 9: 642–655.
- Fame Consortium. 2004. Manual for the application of the European fish index – EFI. A fish-based method to assess the ecological status of European rivers in support of the water framework directive. Version 1.1, January 2005.

- Fausch, K.D., Hawkes, C.L. & Parson, M.H. 1988. Models that predict standing crop of stream fish from habitat variables: 1950–85. Portland, OR: US Forest Service, Pacific Northwest Research Station, General Technical Report PNW-GTR-213.
- Fausch, K.D., Torgensen, C.E. & Li, H. 2002. Landscapes and riverscapes: bridging the gap between research and conservation of stream fishes. *BioScience* 52: 483–498.
- Fielding, A.H. & Bell, J.F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24: 38–49.
- Filipe, A.F., Cowx, I.G. & Collares-Pereira, M.J. 2002. Spatial modelling of freshwater fish in semi-arid river systems: a tool for conservation. *River Research and Applications* 18: 123–136.
- Filipe, A.F., Marques, T.A., Seabra, S., Tiago, P., Ribeiro, F., Moreira da Costa, L., Cowx, I.G. & Collares – Pereira, M.J. 2004. Selection of priority areas for fish conservation in Guadiana river Basin, Iberian Peninsula. *Conservation Biology* 18: 189–200.
- Godinho, F.N., Ferreira, M.T. & Cortes, R.V. 1997. Composition and spatial organisation of fish assemblages in the lower Guadiana basin, southern Iberia. *Ecology of Freshwater Fish* 6: 134–143.
- Godinho, F.N., Ferreira, M.T. & Santos, J.M. 2000. Variation in fish community composition along an Iberian river basin from low to high discharge: relative contributions of environmental and temporal variables. *Ecology of Freshwater Fish* 9: 22–29.
- Gore, J.A. & Nestler, J.M. 1988. Instream flow studies in perspective. *Regulated Rivers: Research and Management* 2: 93–101.
- Grenouillet, G., Pont, D. & Hérissé, C. 2004. Within-basin fish assemblage structure: the relative influence of habitat versus stream spatial position on local species richness. *Canadian Journal of Fisheries and Aquatic Sciences* 61: 93–102.
- Guisan, A. & Zimmermann, N.E. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* 135: 147–186.
- Harding, J.S., Benfield, E.F., Bolstad, P.V., Helfman, G.S. & Jones, E.B.D. III 1998. Stream biodiversity: the ghost of land use past. *Proceedings of National Academy of Sciences of United States of America* 95: 14843–14847.
- Heggenes, J., Northcote, T.G. & Peter, A. 1991. Seasonal habitat selection and preferences by cutthroat trout (*Oncorhynchus clarki*) in a small, coastal stream. *Canadian Journal of Fisheries and Aquatic Sciences* 48: 1364–1370.
- Hinch, S.G., Somers, K.M. & Collins, N.C. 1994. Spatial autocorrelation and assessment of habitat-abundance relationships in littoral zone fish. *Canadian Journal of Fisheries and Aquatic Sciences* 51: 701–712.
- Hoeinghaus, D.J., Winemiller, K.O. & Birnbaum, J.S. 2006. Local and regional determinants of stream fish assemblage structure: inferences based on taxonomic vs. functional groups. *Journal of Biogeography* 34: 324–338.
- Hosmer, D.W. Jr. & Lemeshow, S. 1989. Applied logistic regression. New York: John Wiley and Sons.
- Hubert, W.A. & Rahel, F.J. 1989. Relation of physical habitat to abundance of four non-game fishes in high-plain stream: a test of habitat suitability index models. *North American Journal of Fisheries Management* 9: 332–340.
- Huet, M. 1949. Aperçu des relations entre la pente et les populations piscicoles des eaux courantes. *Revue Suisse d'Hydrologie* 11: 332–351.
- Leftwich, K.N., Angermeier, P.L. & Dolloff, C.A. 1997. Factors influencing behaviour and transferability of habitat models for a benthic stream fish. *Transactions of the American Fisheries Society* 126: 725–734.
- Legendre, L. & Legendre, P. 1998. Numerical ecology (developments in environmental modelling S.). Amsterdam: Elsevier.
- Leps, J. & Smilauer, P. 2003. Multivariate analysis of ecological data using canoco. UK: Cambridge University Press.
- Lobón-Cerviá, J., Dgebuadze, Y., Utrilla, C.G., Rincon, P.A. & Granado-Lorencio, C. 1996. The reproductive tactic of dace in central Siberia: evidence for temperature regulation of the spatio-temporal variability of its life history. *Journal of Fish Biology* 48: 1074–1087.
- Lohr, S.C. & Fausch, K.D. 1997. Multiscale analysis of natural variability in stream fish assemblages of a western Great Plain watershed. *Copeia* 4: 706–724.
- Lyons, J. 1989. Correspondence between the distribution of fish assemblages in Wisconsin streams and Omernik's ecoregions. *American Midland Naturalist* 122: 163–182.
- Magalhães, M.F. 1992. Feeding ecology of the Iberian cyprinid *Barbus bocagei* Steindachner, 1865 in a lowland river. *Journal of Fish Biology* 40: 123–133.
- Magalhães, M.F., Batalha, D.C. & Collares-Pereira, M.J. 2002. Gradients in stream fish assemblages across a Mediterranean landscape: contributions of environmental factors and spatial structure. *Freshwater Biology* 47: 1015–1031.
- Margalef, R. 1968. Perspectives in ecological theory. Chicago, IL: University of Chicago Press.
- Marsh-Matthews, E. & Matthews, W.J. 2000. Geographic, terrestrial and aquatic factors: which most influence the structure of stream fish assemblages in the Midwestern United States. *Ecology of Freshwater Fish* 9: 9–21.
- Mason, C.F. 1991. Longman scientific and technical. *Biology of Freshwater Pollution*. Longman Scientific and Technical. New York: John Wiley & Sons.
- Mellina, E., Hinch, S.G. & MacKenzie, K.D. 2005. Seasonal movement patterns of stream-dwelling rainbow trout in north-central British Columbia, Canada. *Transactions of the American Fisheries Society* 134: 1021–1037.
- Moreira, F., Beja, P., Morgado, R., Reino, L., Gordinho, L., Delgado, A. & Borralho, R. 2005. Effects of field management and landscape context on grassland wintering birds in southern Portugal. *Agriculture Ecosystems & Environment* 109: 59–74.
- Moyle, P.B. & Cech, J.J. 1996. Fishes: an introduction to ichthyology. NJ: Benjamin Cummings.
- Nagelkerke, N.J.D. 1991. A note on a general definition of the coefficient of determination. *Biometrika* 78: 691–692.
- Oberdorff, T., Guilbert, E. & Lucchetta, J.C. 1993. Patterns of fish species richness in the Seine River basin, France. *Hydrobiologia* 259: 157–167.
- Oberdorff, T., Pont, D., Hugueny, B. & Chessel, D. 2001. A probabilistic model characterizing fish assemblages of French rivers: a framework for environmental assessment. *Freshwater Biology* 46: 399–415.

- Ormerod, S.J. 2003. Current issues with fish and fisheries: editor's overview and introduction. *Journal Applied Ecology* 40: 204–213.
- Pearce, J. & Ferrier, S. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling* 133: 225–245.
- Perdices, A. & Doadrio, I. 1997. Threatened fishes of the world: *Cobitis paludica* (De Buen, 1930) (Cobitidae). *Environmental Biology of Fishes* 49: 360.
- Pires, A.M., Cowx, I.G. & Coelho, M.M. 1999. Seasonal changes in fish community structure of intermittent streams in the middle reaches of the Guadiana basin, Portugal. *Journal of Fish Biology* 54: 235–249.
- Pires, A.M., Costa, L.M., Alves, M.J. & Coelho, M.M. 2004. Fish assemblage structure across the Arade basin (southern Portugal). *Cybiurn* 28: 357–365.
- Poff, N.L. & Allan, J.D. 1995. Functional organization of stream fish assemblages in relation to hydrological variability. *Ecology* 76: 606–627.
- Pont, D., Hugueny, B. & Oberdorff, T. 2005. Modelling habitat requirement of European fishes: do species have similar responses to local and regional environmental constraints? *Canadian Journal of Fisheries and Aquatic Sciences* 62: 163–173.
- Rahel, F.J. & Hubert, W.A. 1991. Fish assemblages and habitat gradients in a Rocky Mountain-Great Plains stream: biotic zonation and additive patterns of community change. *Transactions of the American Fisheries Society* 120: 319–332.
- Refcond. 2003. Guidance on establishing reference conditions and ecological status class boundaries for inland surface waters. Final version. CIS Working Group 2.3 – REFCOND.
- Reino, L. 2005. Variation partitioning for range expansion of an introduced species: the common waxbill *Estrilda astrild* in Portugal. *Journal of Ornithology* 146: 377–382.
- Ricciardi, A. & Rasmussen, J.B. 1999. Extinction rates of North American freshwater fauna. *Conservation Biology* 13: 1220–1222.
- Rincon, P.A., Vellasco, J.C., Gonzalez-Sanchez, N. & Pollo, C. 1990. Fish assemblages in small streams in western Spain: the influence of an introduced predator. *Archive fur hydrobiologie* 118: 81–91.
- Santos, J.M., Godinho, F.N., Ferreira, M.T. & Cortes, R.V. 2004a. The organisation of fish assemblages in the regulated Lima basin, Northern Portugal. *Limnologica* 34: 224–235.
- Santos, J.M., Godinho, F.N. & Ferreira, M.T. 2004b. Micro-habitat use by Iberian nase *Chondrostoma polylepis* and Iberian chub *Squalius carolitertii* in three small streams, north-west Portugal. *Ecology of Freshwater Fish* 13: 223–230.
- Santos, J.M., Ferreira, M.T., Pinheiro, A.N. & Bochechas, J. 2006. Effects of small hydropower plants on fish assemblages in medium-sized streams in Central and Northern Portugal. *Aquatic Conservation: Marine and Freshwater Ecosystems* 16: 373–388.
- Saunders, D.L., Meeuwing, J.J. & Vincent, C.J. 2002. Freshwater protected areas: strategies for conservation. *Conservation Biology* 16: 30–41.
- Schiemer, F., Zalewisk, M. & Thorpe, J. 1995. Land/inland water ecotones: intermediate habitats critical for conservation and management. *Hydrobiologia* 303: 259–264.
- Schlosser, I.J. 1991. Stream fish ecology: a landscape perspective. *Bioscience* 41: 704–712.
- Stanford, J.A., Ward, J.V., Liss, W.J., Frissell, C.A., Williams, R.N., Lichatowich, J.A. & Coutant, C.C. 1996. A general protocol for restoration of regulated rivers. *Regulated Rivers: Research and Management* 12: 391–413.
- Stauffer, J.C., Goldstein, R.M. & Newman, R.M. 2000. Relationship of wooded riparian zones and runoff potential to fish community composition in agricultural stream. *Canadian Journal of Fisheries and Aquatic Sciences* 57: 307–316.
- Swets, K.A. 1988. Measuring the accuracy of diagnostic systems. *Science* 240: 1285–1293.
- Taylor, C.M. 1996. Abundance and distribution within a guild of benthic stream fishes: local processes and regional patterns. *Freshwater Biology* 36: 385–396.
- Ter Braak, C.J.F. & Smilauer, P. 1998. *Canoco: reference manual and user's guide to canoco for Windows*. Ithaca, NY: Microcomputer Power.
- Ter Braak, C.J.F. & Verdonschot, P.F.M. 1995. Canonical correspondence analysis and related multivariate methods in aquatic ecology. *Aquatic Sciences* 57: 255–289.
- Valladolid, M. & Przybylski, M. 2003. Feeding ecology of *Cobitis paludica* and *Cobitis calderoni* in Central Spain. *Folia Biologica (Krakow)* 51(Suppl.):135–141.
- Wiley, M.J., Kohler, S.L. & Seelbach, P.W. 1997. Reconciling landscape and local views of aquatic communities: lessons from Michigan trout streams. *Freshwater Biology* 37: 133–148.
- Williams, L.R., Taylor, C.M., Warren, M.L. & Clingenpeel, J.A. 2003. Environmental variability, historical contingency and the structure of regional fish and macroinvertebrate fauna in Ouachita Mountain stream systems. *Environmental Biology of Fishes* 67: 203–216.