

Modelling stream fish species distribution in a river network: the relative effects of temperature versus physical factors

Buisson L, Blanc L, Grenouillet G. Modelling stream fish species distribution in a river network: the relative effects of temperature versus physical factors.
Ecology of Freshwater Fish. © 2007 The Authors Journal compilation © 2007 Blackwell Munksgaard

Abstract – The relative influence of temperature versus local physical factors on the spatial distribution of riverine fish species was investigated in a large watershed of south-western France. Using generalised additive models and hierarchical partitioning, we modelled the ecological responses of 28 fish species to a set of five environmental predictors, and we quantified the independent effect of each predictor. The spatial distribution of fish species was primarily determined by both mean temperature and position along the upstream–downstream gradient. However, responses to these environmental factors varied according to the species considered. Fish species with strong thermal requirements (e.g., common carp, black bullhead, Atlantic salmon) were mainly sensitive to temperature whereas longitudinal gradient was of primary importance for downstream species (e.g., common bream, largemouth bass, pike perch). Both the statistical methods used gave concordant results and appeared complementary. This dual-approach, quantifying the relative contribution of each environmental factor, appears particularly useful to understand the spatial distribution of stream fish species. Separating the effects of temperature versus habitat factors is crucial to accurately predict species distribution modifications in the current context of global change.

**L. Buisson¹, L. Blanc²,
G. Grenouillet¹**

¹Laboratoire Evolution et Diversité Biologique, UMR CNRS-Université Paul Sabatier, Toulouse cedex, France, ²Office National de l'Eau et des Milieux Aquatiques, Délégation régionale Aquitaine & Midi-Pyrénées, Toulouse, France

Key words: stream fish; species distribution; statistical modelling; GAM; hierarchical partitioning

L. Buisson, Laboratoire Evolution et Diversité Biologique, UMR CNRS-Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse cedex 4, France; e-mail: buisson@cict.fr

Accepted for publication September 24, 2007

Introduction

Species distribution modelling has always been a central issue in ecology (reviewed in Guisan & Thuiller 2005). In recent years, an increasing number of studies in ecology, biogeography, or conservation biology have tried to build predictive models of species distribution aiming at a better protection and management of natural resources and ecosystems. Generally, most of them use static modelling techniques which relate the currently observed distribution of species to a pool of available environmental variables (Guisan & Zimmermann 2000; Thuiller 2003).

In stream fish ecology, many studies have been conducted to relate fish assemblage structure to their

environment (reviewed in Matthews 1998). Fish assemblage structure has been frequently examined either at the scale of one single stream and its tributaries (e.g., Meador & Matthews 1992; Humpl & Pivnicka 2006; Taylor et al. 2006) or at the scale of several large hydrographical units (e.g., Rathert et al. 1999; Oberdorff et al. 2001; Pont et al. 2005; Lamouroux & Cattaneo 2006; Hoeinghaus et al. 2007). Most of these papers conclude that fish species distribution is influenced by a wide range of environmental factors.

Climate is known to be a major determining factor operating at large-scale for many taxa (e.g., Pearson & Dawson 2003). For stream fish, temperature appears to be one of the main determinant factors of spatial distribution (e.g., Shuter & Post 1990; Heino 2002; Brazner et al. 2005). Indeed freshwater fish are

particularly sensitive to temperature because they are ectothermic animals (Magnuson et al. 1979) in which both survival and growth are dependent on temperature. Temperature has an effect on their metabolism (Gillooly et al. 2001), breeding (Mills & Mann 1985), development and growth (Mann 1991; Wolter 2007) and behaviour (Taniguchi et al. 1998). Each fish species typically selects thermal habitats in which its growth rate is near maximum (Jobling 1981) and which maximises the metabolic power available for growth, activity and reproduction (Kelsch 1996).

However, other local environmental factors have been shown to determine fish assemblage structure as well. For instance, hydraulic conditions may also influence fish assemblages (Marchetti & Moyle 2001; Lamouroux & Cattaneo 2006), as do physical conditions such as substratum type (Humpl & Pivnicka 2006) or riparian vegetation (Maridet et al. 1998; Grown et al. 2003) and the chemical environment (Matthews et al. 1992; Taylor et al. 1993; Lappalainen & Soininen 2006). The distributional range of freshwater fish species seems therefore to be greatly affected by a combination of environmental factors acting at different scales (Matthews 1998; Jackson et al. 2001; Ferreira et al. 2007).

Stream fish species distribution is also influenced by the atypical structure of the watershed they live in. Indeed, one of the key features of river systems is the network structure of the drainage basin. A whole watershed offers a large variability of environmental conditions ranging from small headwater streams in mountain areas to large lowland channels downstream (Vannote et al. 1980). Each of the previously detailed environmental factors (e.g., temperature, current velocity, discharge, etc.) is thus organised according to a longitudinal upstream–downstream gradient. Fish can disperse throughout this gradient but prefer the environments that are the most adapted to their biology and needs. In European freshwater ecosystems, there has been described a fish species zonation along this longitudinal gradient (Huet 1959) characterised by successive replacements of species from upstream to downstream. Although previous studies emphasised the influence of tributaries' spatial position within the watershed on fish assemblage richness and structure (Osborne & Wiley 1992; Grenouillet et al. 2004), the upstream–downstream gradient remains the most well-known pattern in stream fish assemblages (reviewed in Matthews 1998). Thus, we propose that changes in fish assemblages are due to a gradient in environmental conditions if no other disturbance occurs (e.g., anthropic pollution).

Recently, several studies have focused on developing predictive models of fish assemblage structure according to a large diversity of environmental predictors (e.g., Oberdorff et al. 2001; Kennard et al.

2006) with the purpose that they could be used in streams bioassessment.

However, to our knowledge, despite this vast literature, studies aiming at separating the effects of climate from local habitat factors on stream fish species distributions are scarce. In this context, the aim of the present study was to determine the relative influence of both temperature and physical variables on the fish species distribution using a data set of French riverine fish species occurring in contrasting streams of the Adour–Garonne basin. Temperature was analyzed as a climate surrogate even if numerous other climatic factors may have a great influence. The detailed objectives were: (i) to model the response of fish species to several environmental factors, (ii) to determine the independent effect of each environmental predictor and to separate temperature from physical effects and (iii) to discuss the relative influence of temperature and local habitat on spatial distribution of stream fish assemblage.

Study area

The study area was the Adour–Garonne drainage basin in south-western France. This hydrographic network comprises 120,000 km of flowing waters draining a total area of 116,000 km². Six hydrographic sub-basins (Adour, Charente, Dordogne, Garonne, Lot, Tarn-Aveyron) form this large watershed which covers 1/5 of France. This area is characterised by much contrasted regions in terms of climate, geology and relief ranging from small headwater streams in the Pyrenean mountains to large rivers at the mouth of the main channels.

Material and methods

Fish data

We extracted data from the Office National de l'Eau et des Milieux Aquatiques (ONEMA) database covering a period of 12 years from 1994 to 2005. The ONEMA is the national fisheries organisation in charge of the protection and conservation of freshwater ecosystems in France. To regularly monitor fish assemblages in French rivers, the ONEMA samples a network of sites every year to collect fisheries data. In the Adour–Garonne watershed 141 sites were sampled annually between 1994 and 2005 (Fig. 1) leading to a set of 1557 capture sessions. Each site was consequently sampled several times but each *site* × year combination has been considered as one independent observation in the following analyses. We are aware that this sampling scheme may have potentially led to pseudoreplication, carrying statistical consequences which will be discussed. The selected sites were typical of the

Temperature versus physical determinants of stream fish distribution

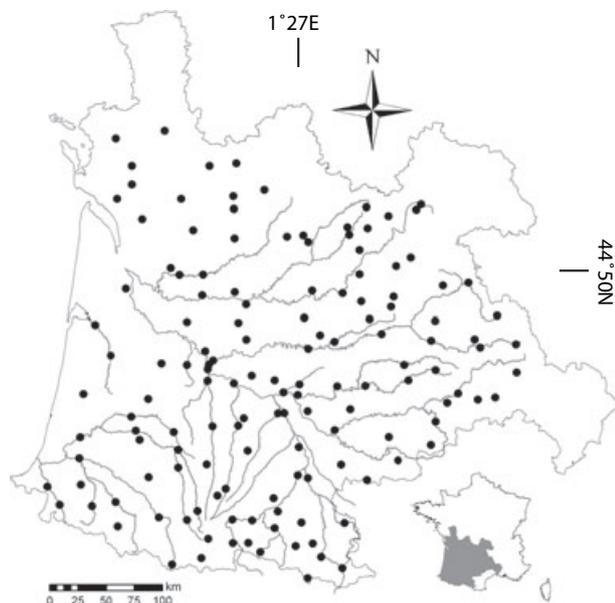


Fig. 1. Geographical distribution of fish sampling sites in the Adour-Garonne basin, France.

different kinds of rivers that can be found in the whole basin and covered a large diversity of environmental influences. Sites stream order ranged from 1 to 7 (median = 4). Sampling sites corresponded to whole stream reaches including several pool-riffle sequences. The prospected areas ranged from 80 to 4882 m² (mean = 856.2 ± 486.9 m²). Fish sampling was conducted with standardised electrofishing methods during low-flow periods (May–October). Most sites were sampled by two-pass removal electrofishing. Two different sampling methods were used depending on river depth and width: smaller rivers were sampled by wading and larger ones by boat. Fishing material was a towed Héron 180–1000 V (1–4 A) direct current apparatus (Dream électronique, Pessac, France). Caught fish were identified to the species level, the number of individuals per species counted, and then the fish were released back into the river. Only presence–absence data are used here.

Physical variables

Six environmental descriptors were available for each site (Table 1): distance from the headwater source DS (km), surface area of the drainage basin above the sampling site SDB (km²), elevation ELE (m), slope SLO (‰), mean stream width WID (m) and depth DEP (m) of the sampled rivers at sites. The last two variables, which were measured during the data collection, were used to describe the river size and could fluctuate from year to year whereas the first four variables, derived from 1:25,000-scale topographic maps, remained unchanged between 1994 and 2005.

Table 1. Characteristics of local habitat and climatic variables of the Adour-Garonne basin.

	SDB	DS	ELE	SLO	WID	DEP	T_m	T_a
Min	2.1	1	2	0.2	1.9	0.05	4.7	6.2
Mean	2114	75.7	304.5	8	24.4	0.67	12.7	10.5
Median	220.4	33.3	175	2.8	10.8	0.39	13.3	10.5
Max	49796	476	1800	154	167	3.8	15.3	13.1

SDB, surface of drainage basin above the site (km²); DS, distance from the source (km); ELE, elevation (m); SLO, slope (‰); WID, stream width (m); DEP, stream depth (m); T_m , mean air temperature (°C); T_a , air temperature range (°C).

Both DS and SDB describe the position of sites in the upstream–downstream gradient. Principal component analysis (PCA) was used to eliminate the colinearity between these two variables. The first axis of the PCA was kept as a synthetic variable describing the longitudinal gradient G :

$$G = -4.12 + 0.58 \log DS + 0.36 \log SDB \quad (1)$$

Thus, high G values corresponded to the most downstream sites.

WID, DEP and SLO influence water velocity and current characteristics. Following Oberdorff et al. (2001), a rough approximation of local velocity derived from the Chezy formula was calculated:

$$V = \log WID + \log DEP + \log SLO - \log(WID + 2DEP) \quad (2)$$

V was high for the most upstream rivers (steep slope, small width and depth).

ELE was not included in G or in V because it did not describe the longitudinal gradient or local hydraulic conditions: indeed, some small streams may be found at low elevations especially in the piedmont zone. ELE was log-transformed to check the hypothesis of normal distribution.

Climate variables

Streams are reasonably well-mixed water bodies that easily exchange heat with the atmosphere. It has already been found that air and river water temperatures are strongly positively correlated (Mohseni & Stefan 1999; Caissie 2006), especially for large-scale studies. Therefore air temperatures were used as a substitute for water temperatures.

The French Meteorological Service (Meteo France) provided us with daily minimum and maximum air temperatures between the 1st January 1994 and the 31st December 2005 for about 200 sites uniformly distributed within the Adour-Garonne watershed. From these data we first calculated mean annual minimum (maximum) temperature as the mean value of minimum (maximum) daily values for each year for each of the

200 meteorological sites. Then the mean annual minimum and maximum temperatures were interpolated by co-kriging for each year including elevation as a determinant factor. The R software (Ihaka & Gentleman 1996) and ‘gstat’ package (Pebesma 2004) were used to interpolate air temperatures. From this interpolation step the minimum and maximum temperatures (T_{\min} and T_{\max}) were predicted for the 141 fish-sampled sites for each year between 1994 and 2005. Two variables obtained from the minimum and maximum temperatures were used in this study, the mean annual air temperature (T_m) and the mean annual temperature range (T_a) at each of the 141 fish-sampling sites for each of the 12 years of study (Table 1):

$$T_m = \frac{T_{\min} + T_{\max}}{2} \quad (3)$$

$$T_a = T_{\max} - T_{\min} \quad (4)$$

T_m and T_a were both retained because they may be of great importance in freshwater fish biological requirements (Crisp 1996; Mann 1996). A different mean temperature and range was used for each *site* \times *year* combination although no significant trend between 1994 and 2005 was detected ($P = 0.48$, $R^2 = 0.05$).

Overall, three physical and two thermal variables were used to model the responses of freshwater fish species in the Adour–Garonne watershed. These five environmental descriptors were not completely independent but they were not highly correlated (Pearson’s coefficient values ranged from -0.72 to 0.65).

Statistical analyses

Species response modelling: generalised additive models

Generalised additive models are semi-parametric extensions of generalised linear models (GLM) (Hastie & Tibshirani 1990). They are designed to use the strengths of the GLMs without requiring the problematic step of postulating the shape of a response curve or a specific parametric function. GAMs are consequently more flexible and allow the combination of both linear and complex additive responses (e.g., non symmetrical responses) within the same model. GAMs were performed using the ‘gam’ library in the R software.

A key step in building GAMs is choosing the appropriate level of the ‘smoothness’ for a predictor (Guisan et al. 2002). In our study, we wanted to get simple responses interpretable in terms of fish biology, and showing only one single optimum. The smoothing parameter for each predictor was thus set to a value of two in equivalent degrees of freedom.

One purpose of this study was to select the most parsimonious model for each species, and thereby the model that best fits the data by using the smallest number of variables. The GAMs were built using a stepwise selection procedure to select both relevant explanatory variables and the level of complexity of the response shapes to each variable. First an initial model including all the environmental predictors in linear form was fitted (linear response). Then a series of models was fitted, each corresponding to a formula obtained by moving each of the variables one step up or down: each variable could be either dropped out of the model (no response to the variable), converted to quadratic form (bell-shaped response) using a second-order polynomial, or smoothed with two degrees of freedom (asymmetrical response) using a spline function. Once all the possible models had been fitted, the ‘best’ model in terms of the AIC statistic (Akaike Information Criterion) was selected (Akaike 1974; Burnham & Anderson 2002).

The response variable of fish species to environmental variability is a binary value (presence or absence of the species). Thus a binomial distribution of error was assumed and the probability of species occurrence was related to the set of environmental variables via a logistic link function.

$$\begin{aligned} \text{Species occurrence probability} = \\ \alpha + f_1(G) + f_2(V) + f_3(T_m) \\ + f_4(T_a) + f_5(ELE) + \varepsilon \end{aligned} \quad (5)$$

where f_i could be either absence of a predictor, the predictor in linear form, the predictor in quadratic form, or the smoothed predictor.

For statistical reasons (Gauch 1982) only species which had a mean occurrence (also called prevalence) $>5\%$ in our whole data set were studied.

Model evaluation

The original data set (1557 catches) was randomly separated into a model training set (987 catches = 63.4%) and a model validation set (570 catches = 36.6%) (split-sample approach). The training data set consisted of 7 years of samples for each site selected by randomisation. The other years were used for the validation. This partition was chosen to keep the same proportion of the 141 sites in the two data sets and to avoid a regional trend that could occur in one of the sets. The number of samples per site being equal in the training data set, model parameter estimates were considered unbiased.

The final series of models of fish species occurrence were fitted with the training set and then validated by applying them to the model validation set. Because the response variable of stream fish species represented a binary variable, a threshold value was set (Pearce &

Ferrier 2000) to compare the probability of occurrence of each fish species predicted for the validation set to the observed value (0 or 1). Below the threshold value, fish species was considered as absent and above as present. To calculate this value, we maximised the sum of two model performance measures (Manel et al. 2001; Luck 2002; Liu et al. 2005): sensitivity, which measures the ability of the model to predict the actual presence, and specificity, which measures the ability to predict the actual absence (Fielding & Bell 1997).

For this threshold value the Kappa index was calculated. Kappa index measures the rate of correct classification after removing the probability of chance agreement. Values of the Kappa statistic can be classified as: 0–0.4 = poor; 0.4–0.6 = moderate; 0.6–0.8 = good; 0.8–1 = excellent (Landis & Koch 1977).

A second method of evaluating models performance was the area under the curve (AUC) of a receiver operating characteristic (ROC) plot (Hanley & McNeil 1982; Zweig & Campbell 1993; Pearce & Ferrier 2000). Unlike the Kappa index, this technique is a threshold-independent measure that relates relative proportions of correctly and incorrectly classified cells over a wide and continuous range of threshold levels. The AUC ranges between 0.5 for models that discriminate no better than random sorting would, and 1 for models that discriminate perfectly (Swets 1988).

Relative importance of the environmental variables: hierarchical partitioning

The multivariate regression methods described above produce the best model to predict values of the dependent variable (the probability of species occurrence). Nevertheless, these techniques do not highlight the most likely causal variables and the relative influence of each predictor on the dependent variable. The method of hierarchical partitioning developed by Chevan & Sutherland (1991) fills this gap. It has only recently been used in ecology (e.g., Heikkinen et al. 2005; Pont et al. 2005; Luoto et al. 2006). Hierarchical partitioning is recommended when the predictors are significantly intercorrelated because in this case the response of the dependent variable to a given descriptor may be an indirect effect of high correlation between that descriptor and other variables.

Hierarchical partitioning was performed using the 'hier.part' package which was run on the R statistical software. Logistic regression and log-likelihood as a measure of goodness-of-fit were used for the hierarchical partitioning of the probability of species occurrence.

Hierarchical partitioning involves computing the increase in fit for all models with a particular predictor compared with the equivalent model with-

out that predictor, and averaging the improvement in fit (here, the reduction of deviance) across all possible models with that predictor. The outcome is, for each explanatory predictor i , an estimate of the independent effect (I_i) and joint influence with all other variables (J_i). The relative independent contribution of each predictor i ($%I_i$) can thus be evaluated. According to Pont et al. (2005), it was considered that a predictor had high explanatory power when $%I_i$ was higher than $100/N$ (where N is the number of predictors, thus $%I_i$ is higher than 20% in this study).

Statistical significances of the relative independent contributions were tested by a randomisation routine, yielding Z-scores for the generated distribution of randomised independent contributions and a measure of statistical significance based on an upper 0.95 confidence limit (Mac Nally 2002). According to Mac Nally (2002), if the calculated independent contribution of a predictor is not significant, this predictor can be eliminated because it does not explain the dependent variable.

The ratio $|I/J|$ was also analyzed (Chevan & Sutherland 1991; Mac Nally 2000): if ratio $|I/J| < 1$, it indicates that the predictor variables are highly correlated.

Hierarchical partitioning had nonetheless a main drawback: only linear relationships could be assessed and the values of independent contribution were thus not meaningful for all the species responses.

Results

Overall 54 fish species were found in the Adour–Garonne basin between 1994 and 2005. Local species richness ranged from 1 to 22 species (mean = 8.2 ± 4.5). Eight species composed almost 90% of the captured individuals: Eurasian minnow (21.2%), gudgeon (20.5%), brown trout (15.1%), bleak (8%), stone loach (7.4%), roach (6.9%), chub (5.9%) and barbel (4%) (see Table 2 for scientific names). For some very rare species, <5 individuals were caught during the 12-year study period: twaite shad *Alosa fallax*, thinlip mullet *Liza ramado*, river lamprey *Lampetra fluviatilis*, sea trout *Salmo trutta trutta*, European plaice *Pleuronectes platessa* and goldfish *Carassius auratus*.

Twenty-eight fish species occurring in more than 5% of the whole data set were retained for the study (Table 2). Gudgeon, Eurasian minnow, brown trout, chub, stone loach and roach were present in more than 50% of the catches. The 28 selected species belonged to 10 families: Anguillidae (1), Balitoridae (1), Centrarchidae (2), Cottidae (1), Cyprinidae (14), Esocidae (1), Ictaluridae (1), Percidae (3), Petromyzontidae (1) and Salmonidae (3).

Code	Common name	Species	Family	Occurrence
Abb	Bream	<i>Abramis brama</i>	Cyprinidae	0.169
Ala	Bleak	<i>Alburnus alburnus</i>	Cyprinidae	0.407
Amm	Black bullhead	<i>Ameiurus melas</i>	Ictaluridae	0.096
Ana	European eel	<i>Anguilla anguilla</i>	Anguillidae	0.491
Bab	Stone loach	<i>Barbatula barbatula</i>	Balitoridae	0.590
Bar	Barbel	<i>Barbus barbus</i>	Cyprinidae	0.445
Blb	White bream	<i>Blicca bjoerkna</i>	Cyprinidae	0.097
Cac	Crucian carp	<i>Carassius carassius</i>	Cyprinidae	0.127
Cht	French nase	<i>Chondrostoma toxostoma</i>	Cyprinidae	0.158
Cog	Bullhead	<i>Cottus gobio</i>	Cottidae	0.226
Cyc	Common carp	<i>Cyprinus carpio</i>	Cyprinidae	0.134
Esl	Northern pike	<i>Esox lucius</i>	Esocidae	0.157
Gog	Gudgeon	<i>Gobio gobio</i>	Cyprinidae	0.712
Gyc	Ruffe	<i>Gymnocephalus cernua</i>	Percidae	0.065
Lap	Brook lamprey	<i>Lampetra planeri</i>	Petromyzontidae	0.203
Lec	European chub	<i>Leuciscus cephalus</i>	Cyprinidae	0.592
Leg	Pumpkinseed	<i>Lepomis gibbosus</i>	Centrarchidae	0.308
Lel	Common dace	<i>Leuciscus leuciscus</i>	Cyprinidae	0.285
Mis	Largemouth bass	<i>Micropterus salmoides</i>	Centrarchidae	0.078
Onm	Rainbow trout	<i>Onchorhynchus mykiss</i>	Salmonidae	0.077
Pef	Perch	<i>Perca fluviatilis</i>	Percidae	0.256
Php	Eurasian minnow	<i>Phoxinus phoxinus</i>	Cyprinidae	0.677
Rur	Roach	<i>Rutilus rutilus</i>	Cyprinidae	0.523
Sas	Atlantic salmon	<i>Salmo salar</i>	Salmonidae	0.055
Sat	Brown trout	<i>Salmo trutta fario</i>	Salmonidae	0.611
Sce	Rudd	<i>Scardinius erythrophthalmus</i>	Cyprinidae	0.128
Stl	Pike perch	<i>Stizostedion lucioperca</i>	Percidae	0.063
Tit	Tench	<i>Tinca tinca</i>	Cyprinidae	0.168

Table 2. Mean occurrence of the 28 fish species studied.

Species responses to the environmental predictors

All the specific models retained at least three out of five variables after the selection procedure (Table 3): six species selected only three variables, 13 species selected four variables and the other nine retained all five variables. Species that selected only three variables retained the three physical variables (G , V and ELE) in most cases and did not respond to the thermal variables (T_m and T_a). Overall, in the 28 species models, at least two out of the three physical variables were retained. G was selected in all models except for common carp and Atlantic salmon, V was only absent from the white bream model and ELE from the rainbow trout model. Thermal range (T_a) was dropped from the models of 15 out of 28 species, and mean temperature (T_m) was not included in the models of largemouth-bass, common bream, stone loach, rudd, pike perch and tench.

Variables in the form of smoothed predictors were the most common (about 47%), then as quadratic predictors (39.1%). Linear predictors were only selected in 13.9% of the cases.

The amount of deviance explained by the 28 specific models ranged from 8.7% (rainbow trout) to 56.6% (eel) with a mean of 31.3% ($\pm 13.2\%$) (Table 3).

AUC values derived from the validation process (Table 3) ranged from 0.755 (tench) to 0.933 (eel) (mean = 0.846 ± 0.065). Thus AUC was considered as fair (from 0.7 to 0.8) for eight species, good (from

0.8 to 0.9) for 11 species and excellent (from 0.9 to 1) for nine species.

Kappa index values gave similar results (Table 3) with values ranging from 0.194 (rainbow trout) to 0.748 (brown trout) (mean = 0.429 ± 0.188). Fifteen specific models had poor Kappa indexes, seven species moderate ones, and only six species models obtained good Kappa indexes.

All three performance measures were significantly and positively cross-correlated ($P < 0.0001$). The prevalence was correlated with the Kappa index (Pearson's coefficient = 0.74, $P < 0.0001$) and the amount of explained deviance to a smaller extent (Pearson's coefficient = 0.50, $P = 0.0065$) as well.

Four types of response to the temperature and physical factors were observed (Fig. 2): (i) only one species (brown trout) showed a decreasing probability of occurrence all along the upstream–downstream gradient (negative sigmoidal response); (ii) some species (e.g., bullhead, stone loach, brook lamprey, French nase, Eurasian minnow and dace) preferred an intermediate longitudinal gradient and showed a bell-shaped response; (iii) other species (e.g., bleak, eel, barbel, bream, northern pike, chub, roach, gudgeon, perch, pumpkinseed or tench) revealed to be more often present at high longitudinal gradient scores in the downstream parts of the Adour–Garonne basin (positive sigmoidal response); (iv) it was difficult to model the responses of crucian carp, common carp, ruffe, black bullhead, Atlantic salmon and rainbow

Temperature versus physical determinants of stream fish distribution

Table 3. Results derived from the hierarchical partitioning and generalised additive models. The relative independent contribution of each environmental predictor is given as a percentage of the total independent contribution. For the models selected from the stepwise selection procedure, 'poly' and 's' correspond to polynomial and spline functions respectively. Given are percentage of explained deviance, threshold value, AUC (with a standard deviation calculated after 200 permutations) and the Kappa index, calculated from the validation data set (except the amount of explained deviance). See Methods for details and Table 2 for species codes.

Code	Hierarchical partitioning					Generalised additive models				
	G	V	T_m	T_a	ELE	Selected models	Explained deviance (%)	Threshold	AUC \pm SD	Kappa
Abb	47.49	7.72	16.96	2.65	25.18	s(G,2) + s(V,2) + poly(ELE,2)	0.420	0.342	0.921 \pm 0.013	0.579
Ala	46.59	14.11	21.40	3.37	14.53	s(G,2) + s(V,2) + s(T_m ,2) + poly(T_a ,2) + poly(ELE,2)	0.525	0.510	0.923 \pm 0.010	0.714
Amm	15.09	26.44	41.32	3.85	13.30	poly(G,2) + poly(V,2) + poly(T_m ,2) + poly(ELE,2)	0.276	0.136	0.793 \pm 0.023	0.198
Ana	12.98	11.39	19.35	2.90	53.38	s(G,2) + s(V,2) + s(T_m ,2) + s(ELE,2)	0.566	0.553	0.933 \pm 0.010	0.699
Bab	9.80	32.45	30.44	16.51	10.80	s(G,2) + s(V,2) + s(T_a ,2) + s(ELE,2)	0.168	0.656	0.763 \pm 0.018	0.428
Bar	47.64	12.81	22.79	6.19	10.57	poly(G,2) + s(V,2) + T_m + poly(T_a ,2) + s(ELE,2)	0.480	0.501	0.890 \pm 0.013	0.618
Bib	61.14	5.54	12.52	2.57	18.24	s(G,2) + T_m + poly(ELE,2)	0.362	0.072	0.894 \pm 0.016	0.270
Cac	15.99	27.45	36.11	4.34	16.10	poly(G,2) + V + s(T_m ,2) + poly(T_a ,2) + poly(ELE,2)	0.257	0.187	0.802 \pm 0.022	0.332
Cht	3.06	14.55	61.44	5.47	15.47	poly(G,2) + s(V,2) + poly(T_m ,2) + poly(T_a ,2) + s(ELE,2)	0.272	0.252	0.829 \pm 0.020	0.313
Cog	9.75	53.17	8.39	8.72	19.96	poly(G,2) + s(V,2) + s(T_m ,2) + poly(ELE,2)	0.220	0.245	0.773 \pm 0.023	0.301
Cyc	7.55	29.21	50.23	3.62	9.39	s(V,2) + T_m + poly(T_a ,2) + poly(ELE,2)	0.213	0.254	0.839 \pm 0.023	0.399
Esl	33.90	13.18	16.06	2.24	34.61	poly(G,2) + s(V,2) + poly(T_m ,2) + s(ELE,2)	0.263	0.171	0.804 \pm 0.022	0.236
Gog	23.59	22.52	30.98	10.47	12.44	poly(G,2) + s(V,2) + s(T_m ,2) + poly(ELE,2)	0.371	0.747	0.892 \pm 0.015	0.593
Gyc	61.67	2.67	7.32	9.87	18.47	s(G,2) + s(V,2) + T_m + T_a + poly(ELE,2)	0.281	0.121	0.913 \pm 0.016	0.360
Lap	29.22	5.39	9.92	26.64	28.83	poly(G,2) + s(V,2) + poly(T_m ,2) + s(T_a ,2) + s(ELE,2)	0.158	0.191	0.763 \pm 0.021	0.233
Lec	34.15	17.00	26.49	5.02	17.35	s(G,2) + poly(V,2) + s(T_m ,2) + T_a + poly(ELE,2)	0.493	0.549	0.912 \pm 0.014	0.719
Leg	26.97	9.71	16.91	3.27	43.14	s(G,2) + s(V,2) + T_m + poly(ELE,2)	0.311	0.251	0.844 \pm 0.017	0.473
Lel	38.91	32.16	14.62	5.38	8.93	poly(G,2) + s(V,2) + s(T_m ,2) + s(ELE,2)	0.246	0.351	0.808 \pm 0.019	0.426
Mis	60.91	2.97	15.66	3.35	17.11	G + s(V,2) + poly(ELE,2)	0.412	0.189	0.926 \pm 0.015	0.579
Onm	12.90	32.58	32.44	15.89	6.19	s(G,2) + s(V,2) + s(T_m ,2) + T_a	0.087	0.112	0.766 \pm 0.036	0.194
Pef	37.02	12.13	19.45	3.44	27.96	s(G,2) + V + T_m + poly(ELE,2)	0.190	0.195	0.762 \pm 0.022	0.305
Php	36.06	13.97	25.47	6.41	18.09	s(G,2) + poly(V,2) + poly(T_m ,2) + s(ELE,2)	0.289	0.696	0.823 \pm 0.019	0.518
Rur	26.48	24.04	18.36	3.08	28.04	s(G,2) + s(V,2) + T_m + poly(T_a ,2) + s(ELE,2)	0.517	0.525	0.921 \pm 0.011	0.744
Sas	2.05	34.24	41.81	4.47	17.43	s(V,2) + T_m + poly(T_a ,2) + poly(ELE,2)	0.257	0.082	0.844 \pm 0.033	0.159
Sat	16.71	25.24	27.42	3.15	27.49	poly(G,2) + poly(V,2) + s(T_m ,2) + poly(T_a ,2) + s(ELE,2)	0.516	0.526	0.927 \pm 0.011	0.748
Sce	33.79	11.21	25.38	3.80	25.82	G + s(V,2) + poly(ELE,2)	0.173	0.279	0.767 \pm 0.029	0.336
Stl	59.30	7.29	18.86	4.58	9.96	s(G,2) + V + poly(ELE,2)	0.291	0.140	0.912 \pm 0.018	0.285
Tit	26.42	23.57	25.26	3.58	21.16	poly(G,2) + s(V,2) + poly(ELE,2)	0.163	0.197	0.755 \pm 0.024	0.265

ELE, elevation (m); T_m , mean air temperature ($^{\circ}$ C); T_a , air temperature range ($^{\circ}$ C); AUC, area under the curve.

trout because their probability of occurrence was highly variable wherever along the gradient. Furthermore, it was noticed that the prevalences of these species were among the lowest and the Kappa values of these specific models were very low as well.

Relative effects of temperature versus physical variables on fish species distribution

In hierarchical partitioning, the independent contributions of all five variables included in the specific models were statistically significant for all species, except the longitudinal gradient G in the Atlantic salmon model. The other small independent contributions were surprisingly all significant.

Only 10 species out of the 28 studied species had larger independent contributions than joint effects (ratio $|I/J| > 1$), and this was especially true for stone loach, brook lamprey, rainbow trout and French nase.

On average, the longitudinal gradient G had the highest independent contribution (29.90%) followed by mean temperature T_m (24.76%) and elevation ELE (20.36%) (Table 3). The thermal range T_a had the

lowest independent contribution (6.24%) and rarely exceeded 20% (only for brook lamprey).

G was the main contributing predictor for 14 species and ranged from 2.05% (Sas) to 61.67% (Gyc). G was also considered as a highly explanatory predictor ($\%I > 20\%$) for four extra species. Largemouth bass, white bream, ruffe and pike perch were very sensitive to the longitudinal gradient ($\%I_G > 59\%$) and for these four species all the other variables had little importance.

T_m was the main contributing variable for only six species and ranged from 7.32% (Gyc) to 61.44% (Cht). T_m was the unique main predictor for French nase ($\%I_{T_m} = 61.4\%$). However, T_m was a highly explanatory variable for nine extra species.

Elevation had the third mean highest independent contribution and ranged from 6.19% (Onm) to 53.38% (Ana). It was the most important variable for five species but five extra species also showed a strong independent effect of elevation. Eel was the only species that responded strongly only to elevation ($\%I_{ELE} = 53.4\%$) and for which the independent contributions of the four other variables were very low.

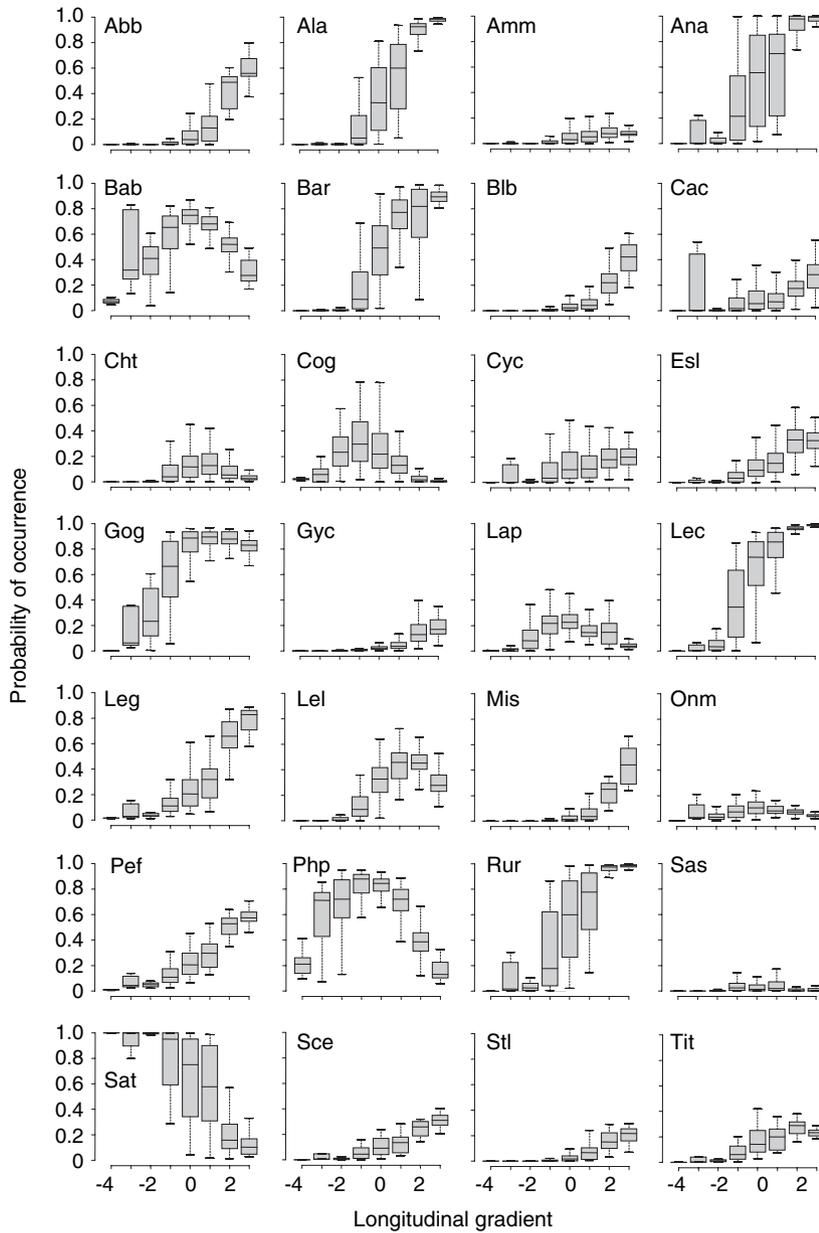


Fig. 2. Probability of occurrence of the 28 fish species for eight classes of longitudinal gradient G values. See Table 3 for model details and Table 2 for species codes.

The velocity measurement V was the main effect for only three species but was $>20\%$ for nine other species. Bullhead was sensitive to only this predictor ($\%I_V = 53.2\%$).

Ultimately, it was observed that most of the 28 species distributions studied were determined by one or two single environmental predictors.

Discussion

This study was a triple-purpose analysis of a large data set reflecting the spatial distribution of stream fish species in the Adour–Garonne watershed in southwestern France. The first aim was to describe the ecological responses of the 28 most common fish species to several environmental variables using

species-specific modelling techniques. Secondly, we attempted to determine the relative effects of both temperature and physical variables for the 28 fish species by hierarchical partitioning. The third goal was to compare the results obtained from both methods used and to discuss the techniques employed.

Fish species responses to the environmental variables

The responses of the 28 most common fish species in the Adour–Garonne to the selected environmental factors were analyzed by using both classical regression methods and hierarchical partitioning. The GAM models allowed us to describe the fish species responses to temperature and physical factors, whereas hierarchical partitioning was used to determine the

importance of each factor retained in the GAMs. These methods gave concordant results.

First, although the environmental factors chosen to study fish species distribution were simple, they were ecologically relevant. All the specific models effectively selected at least three variables out of five, and even if the amount of deviance explained by the models was about 30% on average, it reached almost 50% for six species. AUC values showed that model predictive quality was satisfactory as well. Furthermore, except for the longitudinal gradient in the Atlantic salmon model, all independent contributions calculated with the hierarchical partitioning method were significant. Nevertheless, we found that thermal range was not at all relevant in explaining fish species distribution. This variable was excluded from more than half of the species-models and its independent contribution was very low for almost all species. Thermal range was not found to be an important predictor in Pont et al. (2005) either. While mean water or air temperature has been extensively used as an indicator of fish ecological requirements (e.g., Wang et al. 2003; Lappalainen & Soininen 2006; Wolter 2007), very few authors have regarded the influence of the thermal amplitude on fish biology (e.g., Rathert et al. 1999), maybe because its interpretation is less obvious. Indeed, each fish species is characterised by a tolerance range of temperatures varying from a few degrees to more than 10° (Crisp 1996; Mann 1996). Although each species has a thermal preferendum, they are flexible and can adapt to a certain extent by physiological adjustments (Johnson & Kelsch 1998) or behavioural thermoregulation (e.g., Heggenes et al. 1993). Temperate species are used to seasonal variation in temperatures (Johnson & Kelsch 1998) and this may be one explanation of why most species did not respond to the thermal range in this study. Another reason why the thermal range may not appear important here could be its relatively restricted variability (6.2–13.1 °C but 50% of the sites between 10 and 11.1 °C).

The four other environmental variables studied appeared to have a strong influence on fish species distribution. Thus, both local habitat and mean temperature play an important role when looking at the relationship between species probability of occurrence and environmental conditions. Fish species distribution is influenced by a combination of both thermal and physical factors, but many other factors which have not been considered in this study are also of major importance for freshwater fish (Matthews 1998). This is probably the reason why the explanatory performances of our species-models were not excellent. Indeed, climate has been reduced to thermal variables whereas it is well-known that it is rather a complex interplay between thermal and hydrological

regimes. Factors such as precipitation or discharge could have been integrated in our models to better separate the effects of the whole climate component from the physical factors. In addition, other stream characteristics (e.g., chemical variables) have not been addressed in this study, although these factors are expected to act upon the fish species responses (e.g., Cooney & Allen 2006). Moreover, to prevent the influence of site disturbance, only reference sites (i.e., least-affected by human activity) could have been included in our data set (Kennard et al. 2006). Taking into account such relevant factors could thus improve the quality of our species distribution models.

Overall, the position along the upstream–downstream gradient obtained the highest mean independent contributions and was selected in almost all the specific models. The mean temperature was the second most explanatory variable followed by elevation. These results are in agreement with both the zonation concept of Huet (1959) and thermal niche preferences (Magnuson et al. 1979) that are known to be the two factors that best explain fish species distribution (reviewed in Matthews 1998).

Secondly, our results confirm those of some rare previous studies (e.g., Olden 2003; Pont et al. 2005) showing the need to use a species-specific modelling approach. Indeed, a different model was obtained for each species. Even if the selected variables were the same between two species-models, responses were not the same as model coefficients were different. The hierarchical partitioning approach led to quite the same conclusion because the independent and joint effects were different for each species, showing that the fish species distributions were not determined by the same environmental factors. To date, in spite of these rare studies, many other papers have studied fish community structure only at the species richness level (Angermeier & Winston 1998; Rathert et al. 1999; Kennard et al. 2006; Taylor et al. 2006) or with synthetic descriptors (e.g., guilds: Lamouroux & Cattaneo 2006). Multivariate statistical approaches were also used to group together species having common responses to one or more environmental factors (Taylor et al. 1993; Brazner et al. 2005; Humpl & Pivnicka 2006) but the individual response of each species has rarely been discussed. Although this species-specific approach gave interesting results, it could be criticised as it did not allow us to include some important biotic interactions which could greatly influence fish species distribution. Nonetheless, previous studies suggested that species interactions are negligible in structuring communities compared with the strong influence of regional processes (Oberdorff et al. 1998).

Therefore, although many other biotic or abiotic factors could have been included in the models, our

results are interesting with regard to fish species biology. Indeed, at the species level the longitudinal gradient was a highly explanatory variable for species with contrasted ecological requirements (e.g., feeding, reproduction, behaviour). Nevertheless, except for Eurasian minnow and brook lamprey, all these species generally occupy the intermediate or lower parts of streams. Furthermore, very high values of the relative independent contribution of G were found for the most downstream species (white bream, largemouth bass, ruffe, pike perch) indicating that these species are only present in the downstream parts of the largest watershed and do not have other strong ecological requirements. This result is concordant with the findings of Pont et al. (2005) who found that species that responded strongly to drainage basin area were species of the downstream parts of rivers. These authors found that the slope was of great importance to almost all species considered. In our study, the effect of slope was not analyzed as a single factor but was included in the water velocity index. Unlike this previous reference, velocity was found to have a strong effect for only two small benthic fishes (bullhead and stone loach) and two salmonid species (rainbow trout and Atlantic salmon). This is in agreement with the current literature (Zweimüller 1995; Crisp 1996; Mann 1996; Lamouroux et al. 1999; Keith & Allardi 2001; Lamouroux & Cattaneo 2006): these four species are known to prefer fast-flowing waters and require strong local hydraulic constraints, especially during the spawning period.

Mean temperature appeared to be the single climatic variable influencing species distribution. This was especially true for species which have pronounced thermal requirements, including some warm-water species such as non-native species (e.g., common and crucian carps, black-bullhead), but also tench or gudgeon, as well as some cold- or cool-water species such as French nase, Atlantic salmon, but also rainbow and brown trouts and their associated species, stone loach and Eurasian minnow. These results are in good agreement with the known thermal requirements of fish species (Crisp 1996; Mann 1996) as the non-indigenous species are the ones that prefer and support the highest temperatures whereas the salmonid species are found particularly in mountain streams where the water temperature can drop down to a few degrees above 0 °C (Heggenes et al. 1993). The geographic range of all these species could consequently be explained by temperature.

Lastly, for some species, the relative independent contributions of all the variables were almost equal (e.g., tench, brown trout, roach) indicating that the distribution of these species could not be explained by one or two single factors but was determined by a

combination of all the climatic and physical factors studied (Brazner et al. 2005).

Methodological aspects

A large part of the results we obtained with the two different statistical methods used are concordant with fish species biology. More importantly, these two methods were highly complementary (Heikkinen et al. 2005; Luoto et al. 2006). Hierarchical partitioning solves the problem of multicollinearity between the predictors as it calculates each effect independently of the others. A major shortcoming of this approach is that non-monotonous functions cannot be taken into account whereas nonlinear responses of fish to environmental factors are quite common (Lamouroux et al. 1999; Oberdorff et al. 2001; Pont et al. 2005). No information on the type of response is thus given. This may lead to spurious conclusions as a predictor can have a low independent contribution if considered as linear, and yet be of major importance when taken as a non-linear predictor. For instance, bullhead and stone loach had a bell-shaped response to the longitudinal gradient (Fig. 2) and we found that G had a low independent contribution (9.75 and 9.8%, respectively). This is not surprising as G was considered as a linear predictor in hierarchical partitioning for both species, whereas it was expressed in a quadratic and smoothed form in respective GAMs. While Mac Nally (2002) suggested selecting the most significant variables through a randomisation process, this method did not help us to select any of the variables for any of the species (except Atlantic salmon). Hence the influence of each factor could only be discussed on the basis of the independent contribution values. No predictive model can be generated by using only hierarchical partitioning. Generalised additive models and the associated selection procedure are a reasonably good alternative to solving this dual-problem. These models allow more flexibility in modelling species responses and it has already been shown that modelling techniques able to summarise non-linear and complex relationships have greater efficiency than linear approaches (Brosse & Lek 2000; Bio et al. 2002; Olden & Jackson 2002; Segurado & Araujo 2004; Araujo et al. 2005). The inclusion of smoothed predictors or second-order forms of a variable in GAMs allowed us to describe bell-humped or sigmoidal responses that are relatively coherent with the known responses of the species. Furthermore, in our results more than 86% of the selected variables were expressed in a non-linear form, confirming the advantage of the use of generalised additive models instead of linear regression settings.

Both of our statistical approaches are also concordant in that, when an environmental factor was

dropped from the model during the selection procedure, its independent contribution calculated by hierarchical partitioning was very low (in almost all cases), although the opposite was not always true. Indeed, the GAM selection procedure did not lead us to greatly reduce the number of variables for each species model. Therefore GAMs can be used to describe the species response to the most explanatory abiotic factors, whereas hierarchical partitioning confirms whether the predictors selected in GAMs are amongst the most likely causal variables (Epps et al. 2004; Gibson et al. 2004) and quantify the effects of the predictors when responses are linear.

To date, AUC is one of the most used performance measures in ecology (e.g., Manel et al. 1999, 2001; Guisan & Zimmermann 2000; Pearce & Ferrier 2000; Luck 2002; Gibson et al. 2004; Araujo et al. 2005; Luoto et al. 2006). Being an independent threshold measure of quality, it appears to be the strictest performance measure (Fielding & Bell 1997). Eel had the highest AUC value and was also the species for which the percentage of explained deviance was the greatest. In contrast, rainbow trout had both the lowest explained deviance and Kappa index, and its AUC value was one of the poorest. One explanation can be given to understand this result. Some species did not occur often in the Adour–Garonne watershed even though it was chosen to retain only the species occurring in more than 5% of the sampled sites. Rainbow trout was one of these comparatively rare species, which may explain why its model obtained such poor performance measures, given the high degree of correlation between prevalence and Kappa index scores. Furthermore rainbow trout is not an isolated case as the 15 species which had a poor Kappa index (below 0.4) also had the lowest occurrences (<0.256). This problem of sensitivity to prevalence has already been discussed by several authors (Manel et al. 2001; Liu et al. 2005; Virkkala et al. 2005) who obtained discordant results. Some of them obtained low Kappa values for species with both the lowest and highest prevalence (Virkkala et al. 2005) showing that low Kappa indexes occur when the size of one class clearly exceeds the other (Fielding & Bell 1997). However, this complete pattern was not found in our work. It was also noticed in our study that for most of the species with a low Kappa index, the response curve was very flat and it was difficult to classify these species as upstream, intermediate or downstream species. Therefore, to ensure that our results are independent of species prevalence, we could only focus on the AUC values. We think that our AUC values are strong enough as we calculated them with a separate data set that had not previously been used to build our models. Nonetheless, this validation set was spatially autocorrelated

with the calibration set as the same sites were used for both data sets (Legendre 1993), differing only in the dates. Our models' performance measures may consequently be optimistic estimates of their true predictive capability (Araujo et al. 2005). However, even if AUC values are slightly overestimated, 20 species had AUC values >0.8 indicating that 71% of our specific models showed good to excellent model performance. Another slight concern in our methodological approach is the pseudoreplication issue that occurred because we used several catches collected at different years for each site in the calibration set. Thus the repetition of each site could lead to underestimation of standard errors and consequently, to an increase of the chance of detecting statistically significant results by inflated Type I error rate (Hurlbert 1984; Millar & Anderson 2004). Species-models could therefore be slightly different leading to different predictions as well as explanatory and predictive performances.

To conclude, our results revealed that the spatial distribution of stream fish species could be modelled in a simple manner, and that modelling stream fish distribution using both temperature and physical habitat descriptors could enhance our understanding of fish assemblage structure.

In the current context of global change, this approach reveals interesting perspectives with regards to aquatic ecosystems management. Conservation biologists and environmental managers are indeed very interested in these kinds of models as they appear to be a simple and useful tool to not only explain current species distribution, but also to predict what could happen under global change scenarios (Guisan & Thuiller 2005). As it has been discussed in this study, stream fish could be highly vulnerable to broad-scale environmental change because they are fairly responsive to temperature and only show restricted dispersal ability because of the drainage network structure (Tonn 1990; Sinokrot et al. 1995; Eaton & Sheller 1996). Therefore, isolating the effects of climate from habitat is one of the main challenging issues to accurately predict the impacts of climate change on fish species distribution in streams. To date, different scenarios of climate change are available (Gibelin & Déqué 2003) and could be applied to the specific models developed in this study. Nevertheless, to be applied in the context of climate change, climate cannot be restricted to temperature and other climatic variables for which climate scenarios are available (e.g., precipitation, hydrological cycle) should be included in our fish species distribution models. Thus, testing the effects of these climatic scenarios would allow us to forecast the future range distribution of these common French fish species and also to predict potential modifications of fish assemblages that could have serious consequences for aquatic ecosystems functioning.

Acknowledgements

This research was part of the EU project Euro-limpacs (contract number GOEC-CT-2003-505540). We are indebted to the Office National de l'Eau et des Milieux Aquatiques for providing fish data. We are grateful to L. Tudesque for his instrumental help with thematic maps and GIS support. We also thank F. Leprieur and F. Sanspiche, as well as three anonymous referees, for their insightful comments on the early draft of this manuscript. C. Rogers and C. Shinn gave linguistic advice.

References

- Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19: 716–723.
- Angermeier, P.L. & Winston, M.R. 1998. Local vs. regional influences on local diversity in stream fish communities of Virginia. *Ecology* 79: 911–927.
- Araujo, M.B., Pearson, R.G., Thuillier, W. & Erhard, M. 2005. Validation of species–climate impact models under climate change. *Global Change Biology* 11: 1504–1513.
- Bio, A.M.F., De Becker, P., De Bie, E., Huybrechts, W. & Wassen, M. 2002. Prediction of plant species distribution in lowland river valleys in Belgium: modelling species response to site conditions. *Biodiversity and Conservation* 11: 2189–2216.
- Brazner, J.C., Tanner, D.K., Detenbeck, N.E., Batterman, S.L., Stark, S.L., Jagger, L.A. & Snarski, V.M. 2005. Regional, watershed, and site-specific environmental influences on fish assemblage structure and function in western Lake Superior tributaries. *Canadian Journal of Fisheries and Aquatic Sciences* 62: 1254–1270.
- Brosse, S. & Lek, S. 2000. Modelling roach (*Rutilus rutilus* L.) microhabitat using linear and non-linear techniques. *Freshwater Biology* 44: 441–452.
- Burnham, K.P. & Anderson, D.R. 2002. Model selection and multimodel inference: a practical-theoretic approach, 2nd edn. New York: Springer-Verlag.
- Caissie, D. 2006. The thermal regime of rivers: a review. *Freshwater Biology* 51: 1389–1406.
- Chevan, A. & Sutherland, M. 1991. Hierarchical partitioning. *The American Statistician* 45: 90–96.
- Cooney, P.B. & Allen, M.S. 2006. Effects of introduced groundwater on water chemistry and fish assemblages in Central Florida lakes. *Hydrobiologia* 556: 279–294.
- Crisp, D.T. 1996. Environmental requirements of common riverine European salmonid fish species in freshwater with particular reference to physical and chemical aspects. *Hydrobiologia* 323: 201–221.
- Eaton, J.G. & Sheller, R.M. 1996. Effects of climate warming on fish thermal habitat in streams of the United States. *Limnology and Oceanography* 41: 1109–1115.
- Epps, C.W., Mac Cullough, D.R., Wehausen, J.D., Bleich, V.C. & Rechel, J.L. 2004. Effects of climate change on population persistence of desert-dwelling mountain sheep in California. *Conservation Biology* 18: 102–113.
- Ferreira, M.T., Sousa, L., Santos, J.M., Reino, L., Oliveira, J., Almeida, P.R. & Cortes, R.V. 2007. Regional and local environmental correlates of native Iberian fish fauna. *Ecology of Freshwater Fish*, doi:10.1111/j.1600-0633.2007.00241.x.
- 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.
- Gauch, H.G. 1982. Multivariate analysis in community ecology. New York: Cambridge University Press, pp. 298.
- Gibelin, A.-L. & Déqué, M. 2003. Anthropogenic climate change over the Mediterranean region simulated by a global variable resolution model. *Climate Dynamics* 20: 327–339.
- Gibson, L.A., Wilson, B.A., Cahill, D.M. & Hill, J. 2004. Spatial prediction of rufous bristlebird habitat in a coastal heathland: a GIS-based approach. *Journal of Applied Ecology* 41: 213–223.
- Gillooly, J.F., Brown, J.H., West, G.B., Savage Van, M. & Charnov, E.L. 2001. Effects of size and temperature on metabolic rate. *Science* 293: 2248–2251.
- 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.
- Growns, I., Gehrke, P.C., Astles, K.L. & Pollard, D.A. 2003. A comparison of fish assemblages associated with different riparian vegetation types in the Hawkesbury-Nepean River system. *Fisheries Management and Ecology* 10: 209–220.
- Guisan, A. & Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8: 993–1009.
- Guisan, A. & Zimmermann, N.E. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* 135: 147–186.
- Guisan, A., Edwards, T.C. & Hastie, T. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling* 157: 89–100.
- Hanley, J.A. & McNeil, B.J. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143: 29–36.
- Hastie, T. & Tibshirani, R. 1990. Generalized additive models. London: Chapman and Hall.
- Heggenes, J., Krog, O.M.W., Lindås, O.R., Dokk, J.G. & Bremnes, T. 1993. Homeostatic behavioural responses in a changing environment: brown trout (*Salmo trutta*) become nocturnal during winter. *Journal of Animal Ecology* 62: 295–308.
- Heikkinen, R.K., Luoto, M., Kuussaari, M. & Pöyry, J. 2005. New insights into butterfly–environment relationships using partitioning methods. *Proceedings of the Royal Society B* 272: 2203–2210.
- Heino, J. 2002. Concordance of species richness patterns among multiple freshwater taxa: a regional perspective. *Biodiversity and Conservation* 11: 137–147.
- Hoeinghaus, D.J., Winemiller, K.O. & Birnbaum, J.S. 2007. Local and regional determinants of stream fish assemblage structure: inferences based on taxonomic vs. functional groups. *Journal of Biogeography* 34: 324–338.
- Huet, M. 1959. Profiles and biology of Western European streams as related to fish management. *Transactions of the American Fisheries Society* 88: 156–163.

Temperature versus physical determinants of stream fish distribution

- Humpl, M. & Pivnicka, K. 2006. Fish assemblages as influenced by environmental factors in streams in protected areas of the Czech Republic. *Ecology of Freshwater Fish* 15: 96–103.
- Hurlbert, S.H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54: 187–211.
- Ihaka, R. & Gentleman, R. 1996. R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics* 5: 99–314.
- Jackson, D.A., Peres-Neto, P.R. & Olden, J.D. 2001. What controls who is where in freshwater fish communities – the roles of biotic, abiotic, and spatial factors? *Canadian Journal of Fisheries and Aquatic Sciences* 58: 157–170.
- Jobling, M. 1981. Temperature tolerance and the final preferendum – rapid methods for the assessment of optimum growth temperatures. *Journal of Fish Biology* 19: 439–455.
- Johnson, J.A. & Kelsch, S.W. 1998. Effects of evolutionary thermal environment on temperature-preference relationships in fishes. *Environmental Biology of Fishes* 53: 447–458.
- Keith, P. & Allardi, J. 2001. Atlas des poissons d'eau douce de France. *Patrimoines Naturels* 47: 387.
- Kelsch, S.W. 1996. Temperature selection and performance by bluegills: evidence for selection in response to available power. *Transactions of the American Fisheries Society* 125: 948–955.
- Kennard, M.J., Pusey, J.B., Arthington, A.H., Harch, B.D. & Mackay, S.J. 2006. Development and application of a predictive model of freshwater fish composition to evaluate river health in eastern Australia. *Hydrobiologia* 572: 33–57.
- Lamouroux, N. & Cattaneo, F. 2006. Fish assemblages and stream hydraulics: consistent relations across spatial scales and regions. *River Research and Applications* 22: 727–737.
- Lamouroux, N., Capra, H., Pouilly, M. & Souchon, Y. 1999. Fish habitat preferences in large streams of southern France. *Freshwater Biology* 42: 673–687.
- Landis, J. & Koch, G. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33: 159–174.
- Lappalainen, J. & Soinen, J. 2006. Latitudinal gradients in niche breadth and position – regional patterns in freshwater fish. *Naturwissenschaften* 93: 246–250.
- Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? *Ecology* 74: 1659–1673.
- Liu, C., Berry, P.M., Dawson, T.P. & Pearson, R.G. 2005. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* 28: 385–393.
- Luck, G.W. 2002. The habitat requirements of the rufous tree creeper. 2. Validating predictive habitat models. *Biological Conservation* 105: 395–403.
- Luoto, M., Heikkinen, R.K., Pöyry, J. & Saarinen, K. 2006. Determinants of the biogeographical distribution of butterflies in boreal regions. *Journal of Biogeography* 33: 1764–1778.
- Mac Nally, R. 2000. Regression and model-building in conservation biology, biogeography and ecology: the distinction between – and reconciliation of – ‘predictive’ and ‘explanatory’ models. *Biodiversity and Conservation* 9: 655–671.
- Mac Nally, R. 2002. Multiple regression and inference in ecology and conservation biology: further comments on identifying important predictor variables. *Biodiversity and Conservation* 11: 1397–1401.
- Magnuson, J.J., Crowder, L.B. & Medvick, P.A. 1979. Temperature as an ecological resource. *American Zoologist* 19: 331–343.
- Manel, S., Dias, J.M., Buckton, S.T. & Ormerod, S.J. 1999. Alternative methods for predicting species distribution: an illustration with Himalayan river birds. *Journal of Applied Ecology* 36: 734–747.
- Manel, S., Williams, H.C. & Ormerod, S.J. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology* 38: 921–931.
- Mann, R.H.K. 1991. Growth and production. In: Winfield, I.J. & Nelson, J.S., eds *Cyprinids fishes – systematics, biology and exploitation*. London: Chapman and Hall, pp. 457–482.
- Mann, R.H.K. 1996. Environmental requirements of European non-salmonid fish in rivers. *Hydrobiologia* 323: 223–235.
- Marchetti, M.P. & Moyle, P.B. 2001. Effects of flow regime on fish assemblages in a regulated California stream. *Ecological Applications* 11: 530–539.
- Maridet, L., Wasson, J.G., Philippe, M., Amoros, C. & Naiman, R.J. 1998. Trophic structure of three streams with contrasting riparian vegetation and geomorphology. *Archiv für Hydrobiologie* 144: 61–85.
- Matthews, W.J. 1998. *Patterns in freshwater fish ecology*. New York: Chapman and Hall, Norwell.
- Matthews, W.J., Hough, D.J. & Robison, H.W. 1992. Similarities in fish distribution and water quality patterns in streams of Arkansas: congruence of multivariate analyses. *Copeia* 1992: 296–305.
- Meador, M.R. & Matthews, W.J. 1992. Spatial and temporal patterns in fish assemblage structure of an intermittent Texas stream. *American Midland Naturalist* 127: 106–114.
- Millar, R.B. & Anderson, M.J. 2004. Remedies for pseudoreplication. *Fisheries Research* 70: 397–407.
- Mills, C.A. & Mann, R.H.K. 1985. Environmentally-induced fluctuations in year-class strength and their implications for management. *Journal of Fish Biology* 27: 209–226.
- Mohseni, O. & Stefan, H.G. 1999. Stream temperature/air temperature relationship: a physical interpretation. *Journal of Hydrology* 218: 128–141.
- Oberdorff, T., Hugueny, B., Compin, A. & Belkessam, D. 1998. Non-interactive fish communities in the coastal streams of north-western France. *Journal of Animal Ecology* 67: 472–484.
- 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.
- Olden, J.D. 2003. A species-specific approach to modeling biological communities and its potential for conservation. *Conservation Biology* 17: 854–863.
- Olden, J.D. & Jackson, D.A. 2002. A comparison of statistical approaches for modelling fish species distributions. *Freshwater Biology* 47: 1976–1995.
- Osborne, L.L. & Wiley, L.J. 1992. Influence of tributary spatial position on the structure of warmwater fish communities. *Canadian Journal of Fisheries and Aquatic Sciences* 49: 671–681.

- Pearce, J. & Ferrier, S. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling* 133: 225–245.
- Pearson, R.G. & Dawson, T.P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimatic envelope models useful? *Global Ecology and Biogeography* 12: 361–371.
- Pebesma, E.J. 2004. Multivariable geostatistics in S: the gstat package. *Computers & Geosciences* 30: 683–691.
- Pont, D., Hugueny, B. & Oberdorff, T. 2005. Modelling habitat requirement of European fishes: do species have similar responses to local and regional constraints? *Canadian Journal of Fisheries and Aquatic Sciences* 62: 163–173.
- Rathert, D., White, D., Sifneos, J.C. & Hughes, R.M. 1999. Environmental correlates of species richness for native freshwater fish in Oregon, USA. *Journal of Biogeography* 26: 257–273.
- Segurado, P. & Araujo, M.B. 2004. An evaluation of methods for modelling species distributions. *Journal of Biogeography* 31: 1555–1568.
- Shuter, B.J. & Post, J.R. 1990. Climate, population variability, and the zoogeography of temperate fishes. *Transactions of the American Fisheries Society* 119: 314–336.
- Sinokrot, B.A., Stefan, H.G., McCormick, J.H. & Eaton, J.G. 1995. Modeling of climate change effects on stream temperature and fish habitats below dams and near groundwater inputs. *Climatic Change* 30: 182–200.
- Swets, K. 1988. Measuring the accuracy of diagnostic systems. *Science* 240: 1285–1293.
- Taniguchi, Y., Rahel, F.J., Novingen, D.C. & Gerow, K.G. 1998. Temperature mediation of competitive interactions among three fish species that replace each other along longitudinal stream gradients. *Canadian Journal of Fisheries and Aquatic Sciences* 55: 1894–1901.
- Taylor, C.M., Winston, M.R. & Matthews, W.J. 1993. Fish species-environment and abundance relationships in a Great Plains river system. *Ecography* 16: 16–23.
- Taylor, C.M., Holder, T.L., Fiorillo, R.A., Williams, L.R., Thomas, R.B. & Warren, M.L. 2006. Distribution, abundance, and diversity of stream fishes under variable environmental conditions. *Canadian Journal of Fisheries and Aquatic Sciences* 63: 43–54.
- Thuiller, W. 2003. BIOMOD – optimizing predictions of species distributions and projecting potential future shifts under global change. *Global Change Biology* 9: 1353–1362.
- Tonn, W.M. 1990. Climate change and fish communities: a conceptual framework. *Transactions of the American Fisheries Society* 119: 337–352.
- Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R. & Cushing, C.E. 1980. The river continuum concept. *Canadian Journal of Fisheries and Aquatic Sciences* 37: 130–137.
- Virkkala, R., Luoto, M., Heikkinen, R.K. & Leikola, N. 2005. Distribution patterns of boreal marshland birds: modelling the relationships to land cover and climate. *Journal of Biogeography* 32: 1957–1970.
- Wang, L., Lyons, J., Rasmussen, P., Seelbach, P., Simon, T., Wiley, M., Kanehl, P., Baker, E., Niemela, S. & Stewart, P.M. 2003. Watershed, reach, and riparian influences on stream fish assemblages in the Northern Lakes and Forest Ecoregion, USA. *Canadian Journal of Fisheries and Aquatic Sciences* 60: 491–505.
- Wolter, C. 2007. Temperature influence on the fish assemblage structure in a large lowland river, the lower Oder River, Germany. *Ecology of Freshwater Fish*, doi:10.1111/j.1600-0633.2007.00237.x.
- Zweig, M.H. & Campbell, G. 1993. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry* 39: 561–577.
- Zweimüller, I. 1995. Microhabitat use by two small benthic stream fish in a second order stream. *Hydrobiologia* 303: 125–137.