

Mediterranean Marine Science

Vol 17, No 1 (2016)

VOL 17, No 1 (2016)



Modelling critical factors affecting the distribution of the vulnerable endemic Eastern Iberian barbel (*Luciobarbus guiraonis*) in Mediterranean rivers

E. J. OLAYA-MARIN, F. MARTÍNEZ-CAPEL, R. GARCÍA-BARTUAL, P. VEZZA

doi: [10.12681/mms.1351](https://doi.org/10.12681/mms.1351)

To cite this article:

OLAYA-MARIN, E. J., MARTÍNEZ-CAPEL, F., GARCÍA-BARTUAL, R., & VEZZA, P. (2015). Modelling critical factors affecting the distribution of the vulnerable endemic Eastern Iberian barbel (*Luciobarbus guiraonis*) in Mediterranean rivers. *Mediterranean Marine Science*, 17(1), 264–279. <https://doi.org/10.12681/mms.1351>

Modelling critical factors affecting the distribution of the vulnerable endemic Eastern Iberian barbel (*Luciobarbus guiraonis*) in Mediterranean rivers

E. JULIA OLAYA-MARÍN¹, F. MARTÍNEZ-CAPEL², R. GARCÍA-BARTUAL³ and P. VEZZA^{2,4}

¹ Grupo de investigaciones en Sistemas y Recursos Ambientales Sostenibles. Universidad Manuela Beltrán. Av. Circunvalar 60-00. Bogotá - Colombia

² Research Institute for Integrated Management of Coastal Areas (IGIC), Universitat Politècnica de València, C/ Paranimf, 1, 46730 Grau de Gandia (València), Spain

³ Research Institute of Water and Environmental Engineering (IIAMA), Universitat Politècnica de València, Spain.

⁴ The International Centre for Ecohydraulics Research (ICER), Faculty of Engineering and Environment, University of Southampton, University Rd, Southampton SO17 1BJ, UK

Corresponding author: esther.olaya@docentes.umb.edu.co

Handling Editor: Stamatis Zogaris

Received: 13 April 2015; Accepted: 1 August 2015; Published on line: 29 March 2016

Abstract

Luciobarbus guiraonis (Eastern Iberian barbel) is an endemic fish species restricted to Spain, mainly distributed in the Júcar River Basin District. Its study is important because there is little knowledge about its biology and ecology. To improve the knowledge about the species distribution and habitat requirements, nonlinear modelling was carried out to predict the presence/absence and density of the Eastern Iberian barbel, based on 155 sampling sites distributed throughout the Júcar River Basin District (Eastern Iberian Peninsula). We used multilayer feed-forward artificial neural networks (ANN) to represent nonlinear relationships between *L. guiraonis* descriptors and variables regarding the physical habitat and biological components (macroinvertebrates, fish, riparian forest). The gradient descent algorithm was implemented to find the optimal model parameters; the importance of the ANN's input variables was determined by the partial derivatives method. The predictive power of the model was evaluated with the Cohen's kappa (k), the correctly classified instances (CCI), and the area under the curve (AUC) of the Receiver Operating Characteristic (ROC) plots. The best model predicted presence/absence with a high performance ($k=0.66$, CCI= 87% and AUC= 0.85); the prediction of density was moderate (CCI = 62%, AUC=0.71 and $k=0.43$). The fundamental variables describing the presence/absence were: solar radiation (the highest contribution was observed between 2000 and 4200 WH/m²), drainage area (with the strongest influence between 3000 and 5.000 km²), and the proportion of exotic fish species (with relevant contribution between 50 and 100%). In the density model, the most important variables were the coefficient of variation of mean annual flows (relative importance of 50.5%) and the proportion of exotic fish species (24.4%). The models provide important information about the relation of *L. guiraonis* with biotic and abiotic variables, this new knowledge can help develop future studies and management plans for the conservation of this species in the Júcar River Basin District and, potentially, for the conservation of other endemic fish species of *Barbus* and *Luciobarbus* in Mediterranean rivers.

Keywords: Eastern Iberian Barbel, *Luciobarbus*, hydromorphology, Species distribution model, Mediterranean rivers, Fish habitat, Artificial Neural Networks.

Introduction

Ecological models are useful to understand the effects of environmental variables on species distribution and abundance and to assess the alteration of ecological patterns as a response to environmental changes (Jopp *et al.*, 2011). These models can be integrated in decision support systems for ecological restoration, impact assessment, reserve design and conservation planning (Pearce & Ferrier, 2000; Guisan & Thuiller, 2005; Franklin, 2010; Drew *et al.*, 2011; Olaya-Marín *et al.*, 2012; Olaya-Marín *et al.*, 2013). Moreover, ecological modelling has become an important tool for learning about the implications of stressors like climate change, hydrological regime alteration, water pollution, and

invasive species introduction on freshwater ecosystems (Drew *et al.*, 2011). This constitutes an important line of research in Mediterranean rivers (Olaya-Marín *et al.*, 2013; Vezza *et al.*, 2015) because 56% of the endemic freshwater fish species are threatened with extinction (Smith and Darwall, 2006). From a conservation point of view, models of species presence/absence and density may help in the search for better conservation and river restoration policies (Costa *et al.*, 2012; Muñoz-Mas *et al.*, 2012; Zarkami *et al.*, 2012).

Machine learning (ML) techniques have been seen as a promising discipline to advance current knowledge of ecological processes and patterns (Olden *et al.*, 2008; Drew *et al.*, 2011; Jopp *et al.*, 2011; Leclere *et al.*, 2011). These techniques have been developed from

artificial intelligence and applied in several disciplines of environmental sciences (Hsieh, 2009), owing to their ability to model nonlinear processes (Olden *et al.*, 2008; Hsieh, 2009). This ML feature allows us to derive better predictions and improve the effectiveness of decision making in environmental management (Evans & Cushman, 2009).

Currently, ML is a cornerstone and one of the most active research areas in the field of artificial intelligence (Jopp *et al.*, 2011). Artificial neural networks (ANN) are one of the most effective ML techniques to capture nonlinearities in ecological problems (Lek *et al.*, 2005; Franklin, 2010; Feio & Poquet, 2011), the use of ANN in freshwater studies demonstrates this affirmation. Such models have been successful in predicting the density and biomass of *Salmo trutta* (Baran *et al.*, (1996), the presence/absence of other fish species (Mastrorillo *et al.*, (1997), and also in predicting fish and decapod presence (Joy & Death, (2004). Fish presence/absence is well classified by ANN, both in temperate zones (Tirelli *et al.*, 2009; Tirelli & Pessani, 2011) and tropical ones (Hauser-Davis *et al.*, 2010), even when modelling with unbalanced datasets (in terms of prevalence). Olaya-Marín *et al.* (2012) found satisfactory results in modelling native fish richness with ANN; their approach provided an evaluation of the effects of hydromorphological changes and river restoration actions (weir removal) in three Spanish Mediterranean rivers.

This study focused on the Eastern Iberian barbel (*Luciobarbus guiraonis*) because it is a Mediterranean endemic freshwater fish mainly distributed in the Júcar River Basin, in Spain (Doadrio, 2001; Jiménez *et al.*,

2002), and its ecology and biology is poorly known (Doadrio, 2001); moreover, this species is facing a high risk of extinction and its population is estimated to decline by 30% in the next ten years (Crivelli, 2006). The ichthyofauna of the Iberian Peninsula has a high endemism (Doadrio & Aldeguer, 2007); therefore, the study of aquatic ecosystems in this area should be supported by reliable models to predict future ecological changes, in order to understand potential alterations and avoid their occurrence through the implementation of restoration actions (Clark *et al.*, 2001; Drew *et al.*, 2011; Olaya-Marín *et al.*, 2012).

In this paper, we evaluate the ability of ANN to identify the factors affecting the density and presence/absence of *L. guiraonis* in the Júcar River Basin District at the basin scale. The aims of this study were (i) to identify relevant environmental variables affecting the presence/absence and abundance of *L. guiraonis* and (ii) to assess the ranking of variable importance in the estimation of fish presence/absence and abundance.

Materials and Methods

Study area and data collection

The research was conducted in the main channels of the Júcar, Cabriel and Turia rivers (Eastern Iberian Peninsula). This area is characterised by torrential storms in the humid seasons, severe droughts in the dry seasons and a hydrological response controlled by the rainfall regime, which is typical in Mediterranean

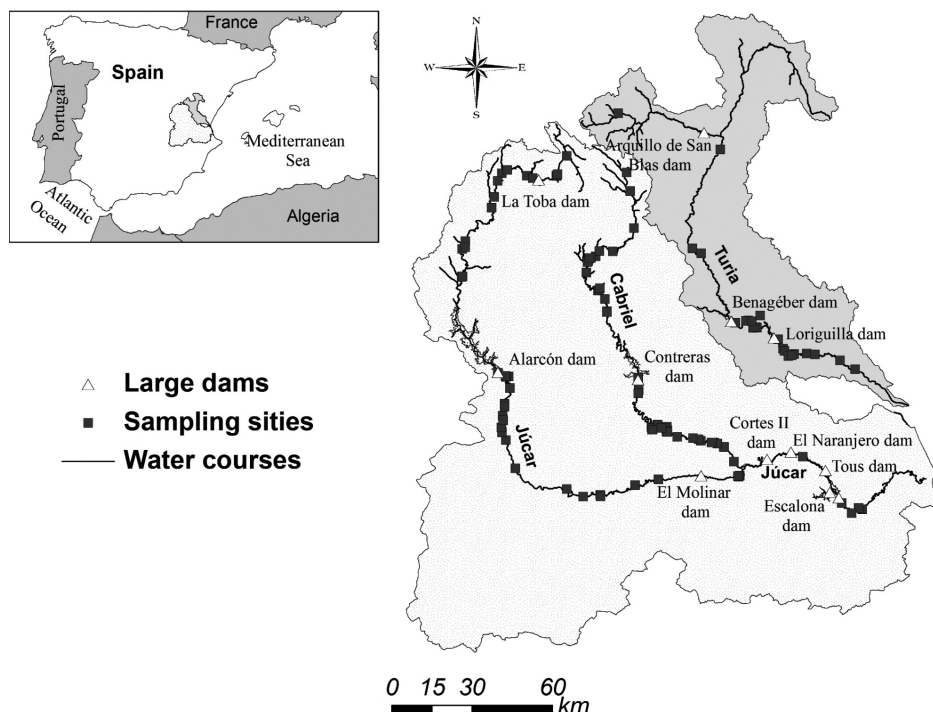


Fig. 1: Study area showing the distribution of the 145 sampling sites in the three rivers (Júcar, Cabriel and Turia rivers).

environments (Granado-Lorencio, 1996; Vila-Gispert *et al.*, 2005). Temporal irregularities of rainfall cause particular variability, expressed by interannual differences in discharge (Granado-Lorencio, 1996; Gasith & Resh, 1999). The natural hydrological regime in these rivers has been altered by dams and weirs, used for hydropower generation and water consumption for agricultural, industrial and domestic activities. The most altered habitats are in the middle and lower parts of the catchments, in which there is a poor development of riparian vegetation caused by agricultural pressures and wastewater discharges (Martínez-Capel *et al.*, 2008). This situation is mainly evidenced in the stretches located between the Alarcón Dam and the mouth of the Júcar river (Fig. 1).

Mediterranean fish communities are known for their low species richness, high endemism (Ferreira *et al.*, 2007) and the predominance of Cyprinids, which possess high specific diversity, are exclusive to epicontinental water bodies, and have typical morphofunctional and physiological adaptations to fluctuating environments (Granado-Lorencio, 1996; Granado-Lorencio, 2000). Knowledge of the specific ecological requirements of Mediterranean freshwater fish inhabiting the Iberian Peninsula is scarce (Ferreira *et al.*, 2007; Aparicio *et al.*, 2011; Maceda-Veiga & De Sostoa, 2011). Although some studies about the habitat requirements of cyprinids at the microhabitat and mesohabitat scale have been made in the Iberian Peninsula (Veza *et al.*, 2015), studies at the basin scale are limited. *L. guiraonis* is a cyprinid species and it has been poorly studied; consequently, there is little information about its biology and ecology, most of it having been inferred from similar species like *Luciobarbus graellsii* and *Luciobarbus bocagei* (Doadrio, 2001). However, these extrapolations could lead to erroneous conclusions about the ecology of *L. guiraonis* and the design of inadequate restoration measurements and management actions (Aparicio *et al.*, 2011).

The density and presence/absence of *Luciobarbus*, the proportion of exotic fish species, the total density of invasive fish and fish predator density were calculated from field data. These data were collected by single-pass electrofishing during the spring, summer and autumn from 2004 to 2010, in 145 sampling sites in the main channels of the Júcar, Cabriel and Turia rivers. The limits of the sampling sites were open (no netting) and the minimum length of the sites was 50 m.

Response and predictive variables

The sampling dataset consists of 145 sites along the main channels of the three Mediterranean rivers (Fig. 1). Fish density and presence/absence were the dependent variables for the models in function of biological and habitat variables. Species density was recorded as the number of fish caught per m².

Twenty-seven environmental variables were used

as potential predictors in the ANN models (Table 1), these variables were selected by taking into account their ecological importance for the fish life cycle (Oberdorff *et al.*, 1995; Granado-Lorencio, 1996; Jackson *et al.*, 2001; Bernardo *et al.*, 2003; Costa *et al.*, 2012). These environmental data were obtained from three sources (Table 1): in situ (fieldwork), GIS analysis and official monitoring networks (MN) of stream flow and biological variables (Confederación Hidrográfica del Júcar, Government of Spain). Geographical variables (*i.e.* altitude, distance from the source, length of channel without artificial barriers, and others) were computed in ArcGIS 9.3.1 with a 5-metre resolution digital elevation model supplied by the National Geographical Institute of Spain. Altitude, water temperature, and the longitudinal natural channel slope play an important role in the distribution of fish communities throughout the rivers, since they influence flow velocity, water oxygenation and the magnitude of sediments transported by the stream (Jackson *et al.*, 2001; De Sostoa, 2002); factors that affect the development of different habitats for fish life (Bernardo *et al.*, 2003; Costa *et al.*, 2012). The length of channel without artificial barriers, number of tributaries between artificial barriers and drainage area between artificial barriers were included in the research because dams and weirs are physical obstacles that limit fish migration along the river (García de Jalón & González del Tánago, 2007; García de Jalón *et al.*, 2007); moreover, longitudinal connectivity restoration is critical to reestablish the natural dynamics of freshwater ecosystems (Lake *et al.*, 2007). Other geographical and hydrological variables, such as distance from the source, drainage area, potential insolation or solar radiation, magnitude and variability of river discharge, are key factors for Mediterranean ichthyofauna conservation (Granado-Lorencio, 2000; Filipe *et al.*, 2010; Hermoso & Clavero, 2011).

It is well known that water quality affects the distribution and composition of fish communities (Schlosser, 1991; Jackson *et al.*, 2001). We considered water quality indices such as the Iberian Biomonitoring Working Party (IBMWP) and the general physico-chemical water quality index (ICGp), because they give an integrative estimation of environmental quality and helped us to reduce dimensionality. The IBMWP is a modification of the Biological Monitoring Working Party score system (1978), adapted to the Iberian Peninsula by Tercedor and Sánchez-Ortega (1988; 1996). This index assesses biological quality in water bodies based on macroinvertebrates; IBMWP values are distributed in five ranges of water quality: 101: very clean water; 100-61: unpolluted or not appreciably altered water; 60-36: partially polluted water with some evident effects; 35-16: very polluted water; 15: heavily polluted water.

The ICGp is a variation of the general quality index (Martínez-Muro, 2003; CHJ, 2008), originally developed by Provencher & Lamontagne (1977); it results from the

Table 1. Potential environmental variables selected to build the presence/absence and density predictive models.

Variable	Code	Method	Range	Mean
Channel length without artificial barriers (km)	CWB	GIS	0.8-79.0	43.8
Number of tributaries between artificial barriers	TAB	GIS	0.0-54.0	28.7
Altitude (m a.s.l)	ALT	GIS	28.0-1286.0	553.8
Drainage area (km ²)	DRA	GIS	95.0-18296.0	4189.0
Drainage area between artificial barriers	DAB	GIS	3.0-4624.0	866.0
Distance from the source (km)	DHS	GIS	20.5-383.7	168.0
Natural slope of the channel (%)	NSL	GIS	0.0-6.8	0.44
Solar radiation (WH/m ²)	SOR	GIS	1153.0-6298.0	3915.6
Water temperature (°C)	WAT	MN	5.8-19.9	13.6
Mean Annual flow rate (m ³ /s)	FMA	MN	0.03- 12.22	5.1
Mean monthly flow (Two year before sampling) (m ³ /s)	MMF		0.10- 13.63	5.13
Inter-annual mean flow (5 years before sampling)	FIA	MN	0.11- 12.36	6.02
Coefficient of variation of mean monthly flows (5 years before sampling)	FIM	MN	0.23-1.09	0.65
Coefficient of variation of mean annual flows (5 years before sampling)	FCV	MN	0.15-0.91	0.36
Inter-monthly flow variation of the mean monthly flows (5 years before sampling)	FVM	MN	0.36-3.37	0.83
Maximum monthly flow during spawning (April to June before sampling) (m ³ /s)	MaxMF	MN	0.01- 26.38	7.82
Mean monthly flow during spawning (April to June before sampling) (m ³ /s)	MeanMF	MN	0.01- 21.02	6.3
Minimum monthly flow during spawning (April to June before sampling) (m ³ /s)	MinMF	MN	0.00-7.87	2.73
Mean monthly low flow (of the two months with the lowest monthly flow for the year before sampling) (m ³ /s)	MeanLMF	MN	0.00- 8.18	2.9
Sampling year	FIY	n/a	2004-2010	n/a
River name	RN	n/a	1-3	n/a
Proportion of exotic fish species (%)	PEF	<i>In situ</i>	0.0-100.0	37.6
Total density of invasive fish (Fish/m ²)	DIF	<i>In situ</i>	0.000-0.008	0.002
Fish predator density (Fish/m ²)	FPD	<i>In situ</i>	0.0000-0.0035	0.0005
Index of Riparian Habitat Quality	QBR	MN	10-100	73.28
Iberian Biomonitoring Working Party	IBMWP	MN	61.0-260.0	124.3
Physicochemical Index of water quality	ICGp	MN	67.6-87.0	80.5

combination of 11 parameters and scores range from 0 (heavily polluted water) to 100 (very good quality). Biological processes, specifically the interactions between species, influence fish distribution (Fitzpatrick *et al.*, 2007; Broennimann *et al.*, 2012), but these kinds of variables are commonly neglected in species distribution modelling (Davis *et al.*, 1998; Fitzpatrick *et al.*, 2007). To deal with this issue, we have included in our analysis some variables related to the fish community and species interactions; specifically, the proportion of exotic fish species (%), total density of invasive fish (Fish/m²), and fish predator density (Fish/m²), because exotic species are one of the main threats for endemic ichthyofauna in Mediterranean rivers (Granado-Lorencio, 1996; Doadrio, 2001; Smith and Darwall, 2006). The exotic and invasive species were quantified separately, in accordance with the terminology by Copp *et al.* (2005). These authors defined an exotic species as a species that enters an ecosystem from outside the historic range of that species, and they define an invasive species in an ecosystem as

a non-indigenous species that may cause damage to the economy, environment, human health, recreation or public welfare.

At ungauged locations, the mean monthly flow was calculated through a linear interpolation based on the relationship between flow in natural conditions and the accumulated drainage area between gauged sites (Leopold & Maddock, 1953; Leopold *et al.*, 1964; Caissie & El-Jabi, 1995; Caissie, 2006a). River discharge and its pattern of variability define the lifecycle traits of Mediterranean fish species (Ferreira *et al.*, 2007); therefore, several variables regarding the magnitude and variability of river flows (at monthly scale) during the previous years and in the spawning season before the fish sampling (from April to June) were considered. The riparian habitat quality index (QBR, Munné *et al.*, 2003) was introduced in the study to assess the morphological conditions of the sampling sites. QBR integrates four components and synthesises various qualitative features that describe the conservation status of the riparian

area: total vegetation cover, vegetation cover structure, vegetation cover quality, and river channel morphological alterations. Each feature can be evaluated from 0 to 25 and the total valuation can range from 0 to 100. QBR is accepted as a good approximation of riparian quality in several regions, including Mediterranean environments (Aguilella *et al.*, 2005; Garófano-Gómez *et al.*, 2011). Finally, the fish sampling year and river name were used as potential predictive variables in order to discern if the consecutive years of data (timeline) or the river were relevant in predicting the density and presence/absence of *L. guiraonis*.

Artificial neural networks modelling

Artificial neural networks (ANN) are mathematical models inspired by the structure of the nervous system, the neuron being its fundamental building block (Lek *et al.*, 2005; Olden *et al.*, 2008). ANN are valuable tools in ecological studies because they have proved to be flexible, robust and generalisable (Lek *et al.*, 2005; Olden *et al.*, 2008; Alpaydm, 2010). The most prevalent type of ANN in ecological applications is the multilayer perceptron (MLP) (Özesmi *et al.*, 2006), which has been successfully applied in problems of pattern recognition, forecasting, signal processing and modelling of complex nonlinear systems (Lek *et al.*, 2005; Goethals *et al.*, 2007).

An MLP network basically consists of a set of sensory units in the input layer, one or more hidden layers of active nodes (artificial neurons) and an output layer of computational nodes. All input to a neuron in a particular layer comes from preceding layers through unidirectional connections characterised by their strength. The weights quantify the connection strength, thus the training (learning) phase of the ANN can be interpreted as the process of obtaining the best calibrated weight values, which are updated during the training process to reduce the differences between the observed and the predicted outputs. This process ends when a given stopping criterion is reached, which is important to determine whether the model has been optimally or sub-optimally trained. The use of a cross-validation set during the process allows the application of the so called *early stopping criteria*, which implies stopping the training phase when cross-validation error increases in a specified number of iterations. Therefore, further training may produce an overfitting of the training data without improving the generalisation performance of the network.

MLP regressions have been used in this research to predict the presence/absence and density of *L. guiraonis*, this means that simulated values were rounded to integer values in order to generate the final prediction of each model. Presence/absence is treated as a binary variable, where presence was denoted by 1 and absence was represented by 0. Density was converted to an ordinal variable, according to the number of fish per square metre; class 1 means a density of zero, class 2 comprises

densities between 0.001 and 0.019, and class 3 comprises densities in the range from 0.020 to 0.066 ind/m² (Table 2). These categories were created according to data frequency, by aggregating a similar number of records in the three classes.

The generalisation capacity of ANN can be restricted by the distribution pattern and magnitudes of the original data. Data pre-processing is highly recommended before building the ANN models (Goethals *et al.*, 2007). Accordingly, the input environmental variables were first transformed to ranges of similar order of magnitude; they were proportionally scaled between 0 and 1 in their range of values (e.g. Olden & Jackson, 2002; Park *et al.*, 2008; Tirelli & Pessani, 2009; Qin *et al.*, 2010). Another aspect playing a crucial role in the overall performance of the resulting network is the process of input variables selection; a two-step methodology was used herein. Firstly, an exploratory analysis to identify collinearity among the potential predictors was carried out by hierarchical cluster analysis using squared Spearman correlations (ρ^2) as a similarity measure. In the case of highly correlated variables ($\rho^2 > 0.8$), only the one with the highest ecological interpretability was chosen (Olaya-Marín *et al.*, 2012). Secondly, a forward stepwise method was conducted to eliminate irrelevant inputs, thus reducing network architecture complexity (Gevrey *et al.*, 2003).

Several MLP models were built and tested, in order to establish (by systematic trial and error) the optimal number of neurons in the hidden layer and the optimal transfer function in the hidden and output layers. Commonly, transfer functions are nonlinear; they transform the weighted sum of inputs into an output signal (Zhang *et al.*, 1998; Isa *et al.*, 2010) and it is typical to use the same transfer function in hidden and output layers (Lek *et al.*, 2005; Goethals *et al.*, 2007). MLP results are very sensitive to the implemented transfer functions in their layers (Piekiewicz & Rybicki, 2004; Isa *et al.*, 2010). Generally, the selection of a transfer function is based on the best performance by trial and error (Isa *et al.*, 2010), by comparing different transfer functions in the hidden and output layers. In this work, two transfer function combinations (hidden layer/output layer) were tested: hyperbolic tangent/linear, and logistic/linear; the combination offering the best performance was selected (Isa *et al.*, 2010; Olaya-Marín *et al.*, 2012).

The ANN models were designed with a single hidden

Table 2. Range and total number of data classes used to build the presence/absence and density predictive models.

Model	Class	Range	Number of data
Presence/ Absence	Presence	1-1	102
	Absence	0-0	43
Density	Class 1	0-0	43
	Class 2	0.001-0.019	52
	Class 3	0.020-0.066	50

layer and the number of neurons optimised by trial and error. Bishop (1996) has shown that a single hidden layer is sufficient for statistical applications with reasonable computation requirements. Moreover, the use of a single hidden layer is comparable to results using multiple hidden layers (Kurková, 1992). The dataset was randomly divided into three sections (training: 60%, validating: 20% and testing sets: 20%), these percentages are frequently used in the literature (Ryan *et al.*, 2004; Qin *et al.*, 2010). The optimisation method used to train the networks (Demuth *et al.*, 2010) was the gradient descent with momentum and adaptive learning rate backpropagation algorithm (*traingdx* function in Matlab), with a momentum constant of 0.9 and a learning rate of 0.01.

Three efficiency indices were used to evaluate the predictive capacity of the models: the percentage of Correctly Classified Instances (CCI) (Buckland & Elston, 1993; Fielding & Bell, 1997), Cohen's kappa (Cohen, 1960) calculated upon the confusion matrix (Table 3), and the area under the curve (AUC) of the Receiver Operating Characteristic (ROC) plots (Hanley & McNeil, 1982). Cohen's kappa measures the proportion of correctly classified points after accounting for the probability of chance agreement (Drew *et al.*, 2011). Kappa ranges from 0 to 1. According to previous studies (Koch *et al.*, 1977; Manel *et al.*, 2001; Drew *et al.*, 2011) the index can be valued as poor (0.00 to 0.39), moderate (0.40 – 0.59), substantial (0.60 to 0.79) or excellent (0.80 to 1). AUC is calculated as the area under the ROC curve (Franklin, 2010) and it is applicable only to binary variables (e.g., presence/absence). AUC is interpreted as the probability of correctly classifying a pair of randomly selected subjects, one from the presence group and the other from the absence group (Fielding & Bell, 1997; Franklin, 2010). AUC ranges from 0.50 to 1: a value between 0.50 and 0.70 indicates a low discrimination (poor model performance); from 0.70 to 0.90, moderate discrimination; and higher than 0.90, high discrimination (Swets, 1988; Manel *et al.*, 2001).

The selection of a parsimonious model is important in order to find a robust model, for example when different options of model structure or variables can be used upon

Table 3. Error matrix used to calculate the percentage of correctly classified instances (CCI) and Cohen's Kappa.

Predicted	Observed	
	Presence	Absence
Presence	a (true positive)	b (false positive)
Absence	c (false negative)	d (true negative)

the same database. We took into account the number of parameters using the Akaike Information Criterion (AIC). This performance metric considers the fitting error and the number of variables used to reach that error, and is useful to assess the relationship between performance and neural network size. A smaller AIC means a better performance in relation to the number of parameters used by the model. Thus, choosing the model with the smallest AIC implies selecting the simplest one, with the least inputs and hidden neurons. This criterion is valuable because models with fewer variables are much easier to interpret, have a lower level of prediction uncertainty (Jorgensen and Fath, 2011), and an optimal ANN architecture is the simplest model that adequately captures the relationships in the training data (D'Heygere *et al.*, 2006).

Finally, to evaluate the importance of input variables in each model, the partial derivatives method (PaD) was implemented (Dimopoulos *et al.*, 1995; Lek *et al.*, 1995). PaD method can be used to analyse the output changes as a response of small variations in each input variable, thus to estimate the relative importance of the input variables to predict presence/absence and density.

Results

The correlation matrix of the input variables indicated that altitude (ALT) and drainage area (DRA) were strongly correlated (Fig. 2). Following the literature, DRA has a higher ecological importance for fish and is broadly used to explain variations among aquatic communities (Ibarra *et al.*, 2003; Olaya-Marín *et al.*, 2012); thus, ALT was removed as potential predictive variable. Minimum monthly flow (MinMF, April-June)

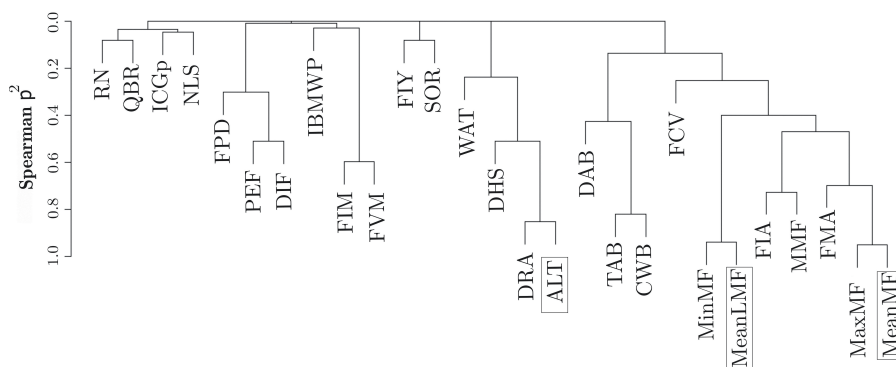


Fig. 2: Hierarchical clustering using squared Spearman correlation (ρ^2) of environmental variables. Nomenclature is shown in Table 1. The variables in boxes were excluded in further analyses.

was highly correlated ($\rho^2 = 0.97$) with the mean monthly flow of the 2 lowest flows (MeanLMF); MeanLMF was discarded because MinMF acts as a critical threshold for habitat availability, mainly in terms of wetted area and water depth, with a direct effect on the recruitment and density of the youngest fish cohorts. Maximum monthly flow (MaxMF) has a correlation of 0.98 with mean monthly flow (MeanMF); MaxMF is important because is related to high flows and floods, which are important to maintain habitat diversity and may occur between April-May, during the *L. guiraois* spawning period (Doadrio, 2001), we excluded MeanMF from further analyses. According to Figure 2, the number of tributaries between artificial barriers (TAB) had a strong correlation with the channel length without artificial barriers (CWB); however, we preserved both of them for the input selection through the forward stepwise method, given their importance for upstream and downstream migration, and the creation and availability of spawning habitats for the Mediterranean fish (Granado-Lorencio, 2000; Olaya-Marín *et al.*, 2012).

The hyperbolic tangent in the hidden layer and linear in output layer were the best transfer functions to predict presence/absence of *L. guiraois*. Figure 3a shows that Cohen's kappa increases at a high rate, from 1 to 8 predictive variables, and the rate of improvement is dramatically reduced from 8 to 19 variables; using 20 variables results in another significant increment (maximum $k=0.85$). The density model (Fig. 3b) reached the best performance using logistic and linear transfer functions; the best performance with a small number of inputs was reached by the model with 5 input variables. By using 10 or 15 variables, there was a potential for better performance; nevertheless, we selected the model with 5 predictors because a model with a reduced number of input variables is more interpretable and applicable (Drew *et al.*, 2011).

To predict the presence/absence of *L. guiraois*, the best neural network architecture found had three layers (i.e. 8→6→2) with eight nodes in the input layer, six neurons in the hidden layer, and two neurons in the output layer (Fig. 5a). The model with 12-inputs (12→9→2) has a kappa of 0.77 (Fig. 3 and Table 4), but it was not selected because it has almost twice the number of parameters of the 8-input model. This involves a test-AIC of 243.88 in the 12-input model and a test-AIC of 113.4 (68 neural network parameters) in the 8-input model (Table 5 and Fig. 4). Moreover, both models have a kappa coefficient classified as substantial. The eight predictive variables in the selected presence/absence model were (in order of importance): solar radiation (SOR), drainage area (DRA), proportion of exotic fish species (PEF), mean annual flow rate (FMA), coefficient of variation of mean monthly flow (FIM), number of tributaries between artificial barriers (TAB), natural slope of the channel (NSL) and the Iberian Biomonitoring Working Party (IBMWP). This network possesses a high percentage of CCI (87%), a substantial kappa coefficient (0.66) and a

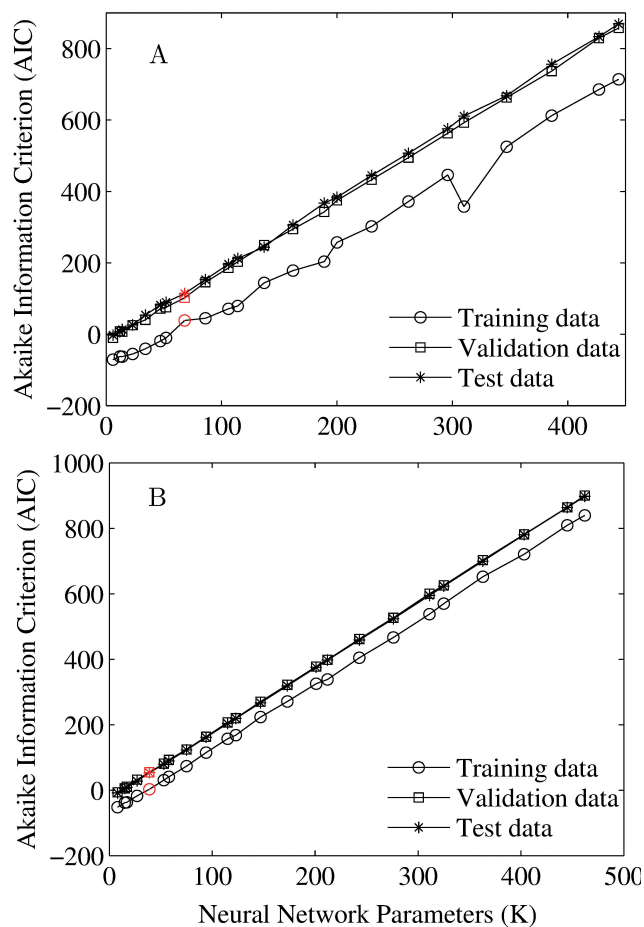


Fig. 3: Influence of the number of input variables and transfer functions (hidden layer/output layer) in models' performance. a) presence/absence of *L. guiraois*. b) Density of *L. guiraois*. Cohen's Kappa was calculated for the entire data set.

Table 4. Predictive results of ANN models (CCI= Percentage of correctly classified instances; Cohen's kappa; AUC= Area Under the Curve). ANN structure indicates the number of nodes in the three layers (input/hidden/output).

Model	ANN-structure	CCI	Cohen's kappa	AUC
Presence/absence	8→6→2	87%	0.66	0.85
	12→9→2	90%	0.77	0.93
Density	5→4→3	62%	0.43	0.71
	15→11→3	72%	0.57	0.83

good performance, evidenced by a value of 0.85 in the area under the ROC curve, which indicate that this model discriminates well (Manel *et al.*, 2001; Gabriels *et al.*, 2007).

For fish density, the best parsimonious model had 5 inputs (Table 5), namely; coefficient of variation of mean annual flows (FCV), proportion of exotic fish species (PEF), minimum monthly flow from April to June before sampling (MinMF), sampling year (FIY) and IBMWP. The optimal network topology for this case consisted of five nodes in the input layer, four neurons in the hidden and three nodes in the output (i.e., 5→4→3), and an

Table 5. Akaike Information Criterion (AIC) calculated in presence/absence and density models (RMSE= Root mean square error, n = sampling size, k = neural network parameters).

Model	Phase	RMSE	n	K	AIC	
Presence/Absence	ANN (8→6→2)	Training	0.327	87		38.76
		Validation	0.310	29	68	102.08
		Test	0.466	29		113.84
	ANN (12→9→2)	Training	0.224	87		143.8
		Validation	0.433	29	137	249.73
		Test	0.354	29		243.88
Density	ANN (5→4→3)	Training	0.422	87		2.92
		Validation	0.452	29	39	54.95
		Test	0.439	29		54.17
	ANN (15→11→3)	Training	0.375	87		338.56
		Validation	0.419	29	212	398.75
		Test	0.400	29		397.45

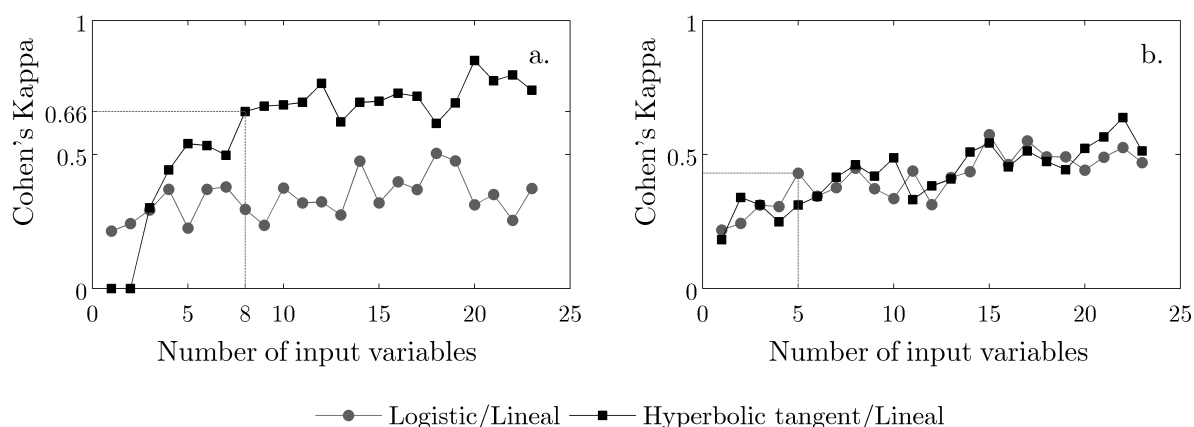


Fig. 4: Akaike Information Criterion (AIC) as a function of the number of neural network parameters, indicating the positive trend in model complexity. a) presence/absence of *L. guiraois*. b) Density of *L. guiraois*.

output layer which classifies density into three classes (Fig. 5b). This model presented a lower performance than the presence/absence model (Table 4); with CCI of 62%, a moderate kappa coefficient (0.43) and an AUC of 0.71. These values indicate a moderate efficiency of the model (e.g. Tirelli *et al.*, 2009).

Models with more input variables also had a moderate efficiency (Table 5). For example, the density model with 15 inputs (15→11→3) presents a CCI of 72%, a moderate kappa coefficient (0.57) and a moderate AUC (0.83). However, this model is clearly over dimensioned, including five times the number of free parameters (weights) as the selected model. This is clearly reflected in the AIC index, as shown in table 5 and Figure 4.

The implementation of the partial derivatives algorithm revealed that the most influential variables to predict the presence/absence of *L. guiraois* were: solar radiation (SOR), with a relative importance of 27.8%; drainage area (DRA), with 24.53%; and the proportion of exotic fish species (PEF), with 13.60% (Fig. 5a). Partial derivatives of each of these variables were plotted against the corresponding input values (e.g. Brosse *et al.*, 2003). Positive values in the y axis (Fig. 6) indicate a

positive relationship between the input and the output variable; conversely, negative values express an inverse relationship (Gevrey *et al.*, 2003; Olaya-Marín *et al.*, 2012). The PaD analysis (Fig. 6a) led to the following observations:

- The highest partial derivatives with respect to drainage area (DRA) occur in the range between 3000 to 5000 km^2 . This means that the increase of DRA is related to an expected increment of *L. guiraois* presence, but this relation is weaker for DRA values over 5000 km^2 .
- For PEF, in the range [50% - 100%] the highest negative partial derivatives are found. Therefore, values of PEF over 50% point towards a relevant reduction in the probability of presence for *L. guiraois*.
- The negative partial derivatives in relation to solar radiation (SOR) show that the increase in SOR contributes towards reducing the probability of presence, the largest negative contributions being found in the range from 2000 to 4200 WH/m^2 .

The most important variables in the density model were the coefficient of variation in the mean annual flow (FCV) with a relative importance of 50.5% (Fig. 5b), and the proportion of exotic fish species (PEF) with 24.40%.

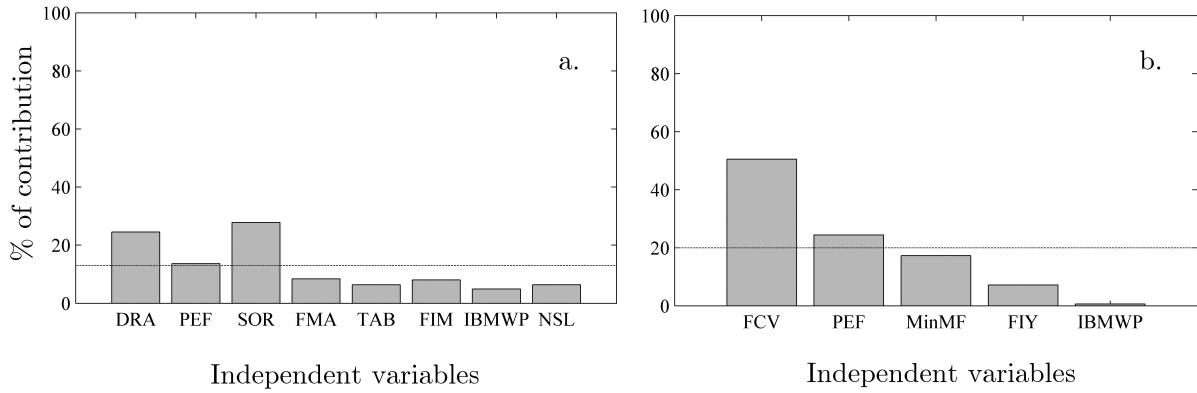


Fig. 5: Contribution of each independent variable predicting a) presence/absence (left) and b) density (right), based on the partial derivatives method (PaD). Dotted lines represents the level of significance (13 and 20% respectively) according to Brosse *et al.* (2003).

Given the moderate performance of this model, the PaD method was unable to support an ecological interpretation of the predictive variables. Partial derivatives with respect to FCV and PEF were positive and negative at the same time, without a precise tendency; therefore, it was

not possible to provide a robust mathematical conclusion about the effects of FCV and PEF on *L. guirao* density (Fig. 6b). The same situation occurred after performing PaD analysis with the model based on 15 input variables.

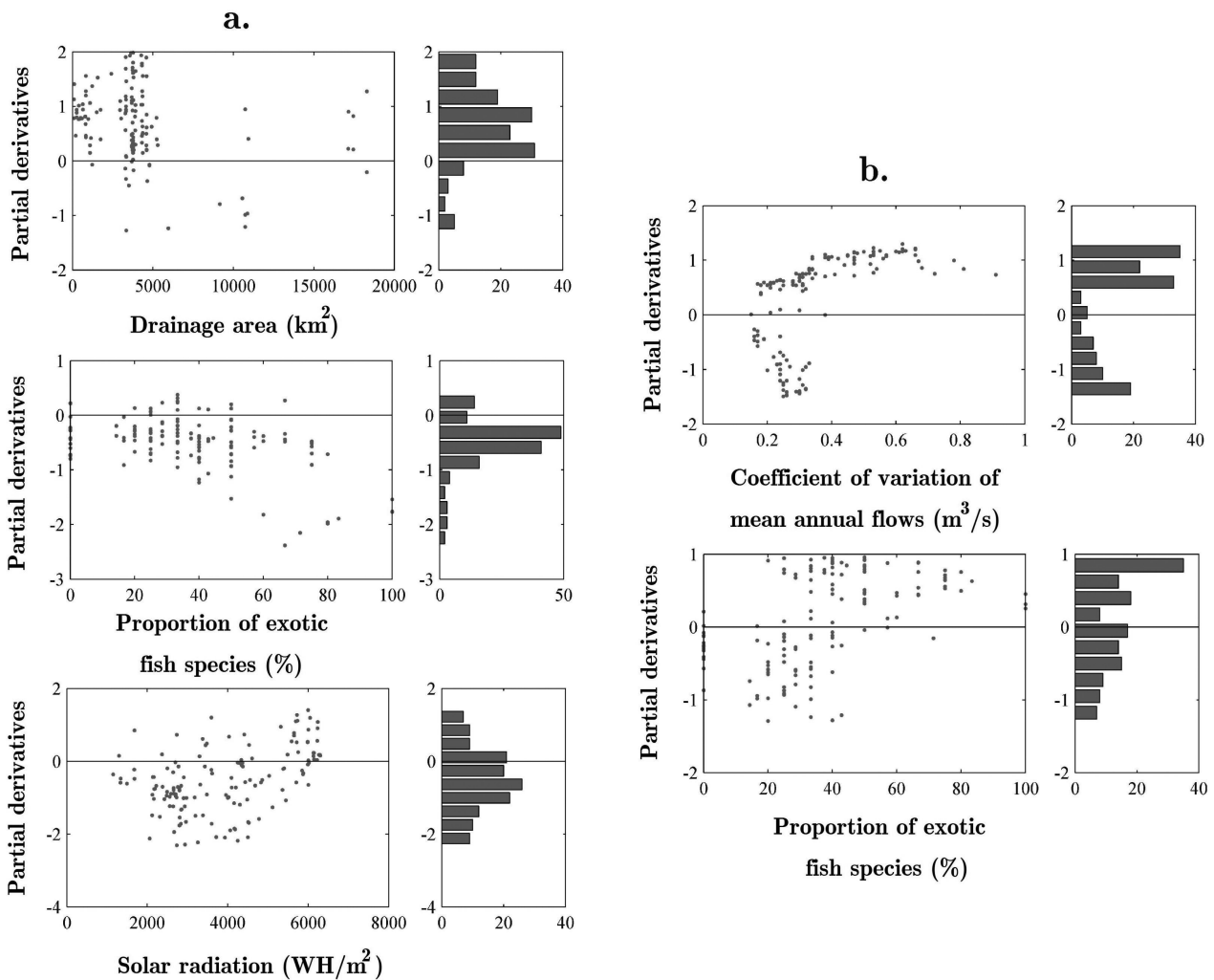


Fig. 6: Output partial derivatives of the most significant predictive variables, as a function of each environmental variable. Left a) Presence of *L. guirao*. Right b) Density of *L. guirao*.

Discussion

A modelling approach based on ANN has been presented herein to classify the presence/absence of the target fish species. It has shown to be efficient, based on a set of performance indices, and the variables involved constitute predictive variables with ecological relevance; moreover, the relationships revealed by the PaD method have ecological meaning, as discussed below. These features provide evidence to support the proposed model, illustrating that it can be used in further studies on the relationships between the environmental variables and *L. guiraonis*, and thus improve the understanding of the fluvial Mediterranean ecosystems.

Regarding the density model, it is important to highlight that the predictive variables are relevant for the target species, even though the best ANN model found had a moderate performance. In this aspect, it is important to remark that the model accuracy is sensitive to the thresholds used for the categorisation into three classes; in this study the categorisation was made in three classes according to frequency, thus each of the classes included between 43 and 51 samples. However, if we had a larger database, different ways of grouping could be tested, with potentially better results for density. Furthermore, we could test other specific modelling techniques for count data modelling to deal with many zeros (e.g. Lambert, 1992; Liu & Chan, 2010) or ANN ensembles (e.g. Muñoz-Mas *et al.*, 2015). Additionally, it was not possible to establish a clear relationship between the input variables and species density from PaD analysis (Fig. 6b); one of the reasons could be an interaction between FCV and PEF or another variable (e.g. Gevrey *et al.*, 2003), which probably indicates the difficulty in identifying a single input perturbation effect on species density (Gevrey *et al.*, 2003; Gevrey *et al.*, 2006). For this reason, a potential improvement in future analysis could be the use of PaD2 (see Gevrey *et al.*, 2003; Gevrey *et al.*, 2006, for details) to study the contribution of all the possible pair-wise combinations of input variables. Another limitation of the density model was the use of fish data collected by electrofishing with 1-pass. The collection of a relatively large database at a river basin scale involves difficulties in various aspects (logistics, budget, etc.) and for this study we could not obtain a database with multiple-passes; therefore, the algorithms to estimate fish abundance could not be applied. The fundamental problem in capturing data is that several factors affect the electrofishing efficiency (and thus the probability of fish capture), such as biological, environmental and technical factors (Zalewski & Cowx, 1990). Therefore, an assessment of the fish abundance at each site, based on the removal method (with two or three passes) can be more accurate in the assessment and modelling of fish abundance at the river basin scale. Nevertheless, this potential drawback was overcome with the classification of fish density into three classes.

In our opinion, the potential improvement in modelling with more accurate data would not compensate the high cost in fieldwork involved in such a sampling method.

The application of a multilayer perceptron artificial neural networks has been demonstrated in several studies to have the capability to model complex ecological patterns and processes with higher performances than traditional statistical approaches (Olden *et al.*, 2008; Franklin, 2010). Moreover, with other techniques it is difficult to represent dataset patterns and trends. Given the nonlinearities in ecological processes and patterns, linear modelling is not a promising field to develop predictive models. Despite the advantages of ANN, they have been categorised as *black-box* models due to the scant information given by the network about the relationship of each input variable and the dependent variables. This is explained by the fact that these relationships are implicit in the architecture of the MLP model. The black-box condition is the main disadvantage of MLP in contrast to traditional statistical approaches, through which we can quantify the influence of each independent variable in the modelling process and the level of confidence in the prediction.

Nevertheless, several methods have been developed to overcome this issue of MLP; one of these is the partial derivative method (PaD), which is used to assess the contribution of each input variable in the prediction. PaD have been considered the most useful method to identify the degree of contribution of input variables in ANN models (Park & Chon, 2007), but PaD is calculated in relation to one independent variable at a time. Thus, when one predictive variable interacts with another, it is difficult to explicitly represent the relationship, as was observed in PaD analysis of the density model in this work. Thus the improvement of the future models with the implementation of techniques like PaD2 is a relevant line of future research.

The partial derivatives method indicated that the predictive variables with the strongest contribution to predict presence/absence were drainage area (DRA), solar radiation (SOR) and the proportion of exotic fish species (PEF). Based on our results, *L. guiraonis* has a positive relation with drainage area in part of its range, and the species prefers the stretches located in the middle part of the watersheds (Fig 6a). Accordingly, Gortázar *et al.* (2007) found a low presence of *L. guiraonis* in the headwaters of the Cabriel River and a large population in the middle. Vezza *et al.* (2015) as well as Kottelat & Freyhof (2007), affirmed that the presence of *L. guiraonis* is more frequent in middle and lower river reaches, and related its decline with water abstraction and habitat modification; these alterations are present in the lower segments of the Júcar, Cabriel and Turia rivers. Martínez-Capel *et al.* (2008) discussed that the populations of the target species were declining in the Júcar River over the years, due to the large proportion of lentic habitats (produced by frequent weirs) and the high proportion of

fine-textured soils in the channel bed, which affects the survival of eggs. The same author revealed that the lack of recruitment could be the main cause of *L. guirao* declining, which is noticeable in the lower reaches of the Júcar and Turia rivers (Estrela *et al.*, 2004). Therefore, the weak relation between DRA and fish presence in sites with DRA > 5000 km² can be interpreted as a result of the hydromorphological and ecological alterations in those sites. In order to explore the interactions with the other two fundamental variables, their mean values (\pm SD) were calculated in sites with DRA > 5000 km²; the mean value of PEF (44.97 % \pm 29.60) was relatively high, suggesting negative effects of the exotic fish species, and the mean SOR (4544.5 WH/m² \pm 1310.7) higher than 4200 suggested a negative influence on *L. guirao* presence.

Concerning solar radiation (SOR), the sensitivity analysis indicated a negative influence of SOR on the presence of the target fish (Fig. 6b). Solar radiation is a critical environmental factor governing temperature change in fluvial systems (Brown & Krygier, 1970; Webb *et al.*, 2008; Isaak *et al.*, 2012). Water temperature is a key variable for fish survival because it affects their physiology and behaviour (Caissie, 2006b; Hrachowitz *et al.*, 2010); this environmental variable directly regulates dissolved oxygen concentration in the water, affecting spawning time, growing rates, and spatial-temporal distribution of species (Magnuson *et al.*, 1979; Jackson *et al.*, 2001; Baron *et al.*, 2002; Prchalová *et al.*, 2006). The negative influence of SOR on the target fish presence may be associated with high water temperatures and a deficit of oxygen during some periods of the year; specifically, within the range from 2000 to 4200 WH/m² the negative influence on probability could be also related with the interaction with PEF (36.71 % \pm 19.80) as an additional stressor for the native fish species. Furthermore, higher temperatures and the increase of evaporation can be associated with longer droughts in Mediterranean rivers, and water scarcity may lead to the reduction of water quality, which may pose severe risks to ecosystem integrity (Petrovic *et al.*, 2011).

Currently, there is a deficiency of knowledge about the effect of water temperature on *L. guirao* and more efforts are needed to investigate this relationship, given that climate change studies have indicated that Mediterranean rivers will experience an increase in the intensity and frequency of droughts (Bonada *et al.*, 2007; Mas-Martí *et al.*, 2010; Sabater & Tockner, 2010; Hermoso & Clavero, 2011), which could severely affect the establishment and survival of this species in the future. Specifically, the Júcar River is expected to suffer a future decline in water quantity (CEDEX, 2011) and quality, as a consequence of climate change; the wide presence of pesticides in water and fish, suggest potential severe effects on fish populations and other biota in future scenarios of climate change (Belenguer *et al.*, 2014). In different Mediterranean countries with a high percentage of endemic fish species, future research

is needed on the interaction between the most relevant environmental factors and multiple stressors, in order to help improve the fish populations and the resilience of the Mediterranean river ecosystems facing a future of water scarcity (Belenguer *et al.*, 2014). The negative effects of climate change can be more severe in countries with high endemism and where the knowledge of native fish species distributions are limited (Hermoso & Clavero, 2011). Therefore, it is very important to foster further research on the population status, distribution and habitat requirements (i.e. habitat suitability models) of the endemic fish species, especially in countries with a high fish diversity (see Barbieri *et al.*, 2015; Papadaki *et al.*, 2014), in order to prevent species extirpation and to facilitate the management of water resources compatible with ecological conservation.

The negative relationship between the proportion of exotic fish species (PEF) and the presence of *L. guirao* (Fig. 6a) is supported by Doadrio (2001), who described a general reduction of this species due to the introduction of exotic species. The adverse effects of exotic species on native fauna are well known, such as predation, competition, hybridisation, disease vector and habitat alteration (Granado-Lorenzo, 2000; Almeida & Grossman, 2012). In the Júcar, Cabriel and Turia rivers, these invasions have been favoured by the construction of dams and weirs, which have created a suitable habitat for the establishment of exotic species (Olaya-Marín *et al.*, 2012), changing from a lotic system to a lentic one, where some exotic species find the suitable habitat for spawning (e.g. *Micropterus salmoides* and *Lepomis gibbosus*). Moreover, these hydraulic structures have segregated *L. guirao* in isolated segments and interrupted upstream migration in the spawning season. Generally speaking, many exotic species cannot adapt well to the natural Mediterranean hydrological regime, characterised by prolonged drought and desiccation in summer and a torrential regime in autumn and winter (Doadrio & Aldeguer, 2007); hence, they are present in the studied area because the flow regulation, impounded waters and artificial obstacles have benefited the establishment of these species (Corbacho & Sánchez, 2001; Vila-Gispert *et al.*, 2005). The study and control of exotic species deserves special attention because of their impact on native fish communities (Clavero *et al.*, 2010).

Below the large Alarcón Dam (Júcar River), our database confirmed the relevant percentage of exotic fish with a mean value of 73.8 %. In the three river basins, *L. guirao* cohabits with the following exotic species: pumpkinseed (*Lepomis gibbosus*), largemouth bass (*Micropterus salmoides*), northern pike (*Esox lucius*), pikeperch (*Sander lucioperca*), pyrenean gudgeon (*Gobio lozanoi*), bleak (*Alburnus alburnus*), common carp (*Cyprinus carpio*), iberian straight mouth nase (*Pseudochondrostoma polylepis*), gudgeon (*Gobio gobio*) and rainbow trout (*Oncorhynchus mykiss*). Regarding the interactions of PEF with the other most

relevant variables, when PEF was higher than 50 % we found values of SOR unfavourable for the target fish (4411.4 WH/m² ± 1499.3) and DRA reached large values where the influence on fish presence was also negative (4623.8 km² ± 3187.0).

The predictive variables of the density model (FCV, PEF, MinMF, FIY and IBMWP) have a relevant influence on Mediterranean fish species from an ecological point of view (Granado-Lorenzo, 2000; Doadrio, 2001; Bernardo *et al.*, 2003; Olaya-Marín *et al.*, 2012). Previous studies have demonstrated that the coefficient of variation of the mean annual flow (FCV) is a fundamental variable in the hydrologic characterisation of the flow regime in Mediterranean rivers (Belmar *et al.*, 2011; Belmar *et al.*, 2012) and it is a critical variable to interpret and predict the richness of native fish species in the Júcar River Basin (Olaya-Marín *et al.*, 2012; Olaya-Marín *et al.*, 2013). The FCV is important for fish distribution in Mediterranean rivers because the lifecycle of native fish is well adapted to fluctuating discharges as a function of natural seasonality (Doadrio & Aldeguer, 2007). Moreover, stream flow is one of the main drivers of the fish population dynamics, as demonstrated in Mediterranean rivers and elsewhere (Alonso-González *et al.*, 2004; Lobón-Cervia & Mortensen, 2005).

As explained above, exotic species are a severe hazard for native freshwater fish in Mediterranean areas (Doadrio, 2001; Smith & Darwall, 2006; Hermoso & Clavero, 2011), thus the selection of PEF was coherent with previous ideas. The minimum monthly flow from April to June (MinMF) is important to *L. guiraonis* because this species migrates to spawn in these months (Kottelat & Freyhof, 2007). The IBMWP was found as a fundamental variable to predict fish distribution in Mediterranean rivers against other water-quality indices (Carballo *et al.*, 2009; Olaya-Marín *et al.*, 2012). The selection of sampling year (FIY) in this model was also interesting. In Mediterranean rivers, this variable could be interpreted as the random influence of the climate variability on the river flow, thus affecting the FCV; additionally, this variable might be related to the trend of increasing populations of exotic fish species. Some studies have analysed the expansion of such exotic species and, on average, the main river basins in the Iberian Peninsula have more invasive than native species (Clavero & García-Berthou, 2006). Although the PaD did not allow us to observe any temporal trend, we considered the increase of PEF throughout the years as a reasonable hypothesis. The variation of the mean PEF in the main stem of the Júcar River was coherent with such a hypotheses, with 40.2, 38.3, 57.6, 64.6 and 72.9 % (from 2005 to 2009; N = 42 data), but the trend was relatively stable in the Turia River, with 35.0, 25.3, 32.0, 30.4, 38.2, 32.6 and 31.1 % (from 2004 to 2010, N = 83 data). Therefore, a larger database with a better distribution regarding sampling years would be necessary to further investigate this aspect.

Apart from the specific limitations, the presence/absence model provides important information about the relation of *L. guiraonis* with biotic and habitat variables. This knowledge complements other models performed at the fish community level (Olaya-Marín *et al.*, 2012), and could be used to support future studies and practical decisions about the management and conservation of this species in the Júcar River Basin District. The density model did not permit the establishment of a clear relationship between the predictive variables and density of *L. guiraonis*. This should be considered in future studies seeking to advance understanding of ecological interactions in Mediterranean stream ecosystems and the critical factors for the conservation of the endemic fish population.

Acknowledgements

This study was partially funded by the MINECO, Spanish Ministry of Economy and Competitiveness, with the projects SCARCE (Consolider-Ingenio 2010 CSD2009-00065), IMPADAPT (CGL2013-48424-C2-1-R) and with Feder funds. There was also funding from the Project POTEVOL “Evaluación del Potencial Ecológico de Ríos Regulados por Embalses y Desarrollo de Criterios para su mejora según la Directiva Marco del Agua” (CGL2007-66412). In addition, the present research was developed in the frame of the EU-funded HolRiverMed project (Marie Curie Actions, Intra-European Fellowships for the mobility of European researchers). We thank the Confederación Hidrográfica del Júcar (Spanish Ministry of Agriculture, Food and Environment) for the data provided to develop this study and we also owe our gratitude to Juan Jiménez for the collaboration in building the first fish database for this research.

References

- Aguilella, A., Riera, J., Gómez-Serrano, M. A., Mayoral, O., Moreyra, E., 2005. *Evaluación del estado ecológico de los ríos de la cuenca hidrográfica del Júcar mediante el uso del índice QBR*, Jardí Botànic, Universitat de València, Valencia, 258 pp.
- Alba-Tercedor, A., 1996. *Macroinvertebrados acuáticos y calidad de las aguas de los ríos*. p. 203-213. In: *IV Simposio del Agua en Andalucía (SIAGA)*, Almería, Spain.
- Alba-Tercedor, J., Sánchez-Ortega, A., 1988. Un método rápido y simple para evaluar la calidad biológica de las aguas corrientes basado en el de Hellawell (1978). *Limnetica*, 4, 51-56.
- Almeida, D., Grossman, G. D., 2012. Utility of direct observational methods for assessing competitive interactions between non-native and native freshwater fishes. *Fisheries Management and Ecology*, 19 (2), 157-1.
- Alonso-González, C., García de Jalón, D., Gortázar, J., Baeza Sanz, D., 2004. Abiotic control of brown trout (*Salmo trutta* L.) population dynamics by highly variable stream

- flow regimes in a central Iberian mountain basin. P.12-17. In: 5th International Symposium on Ecohydraulics "Aquatic habitats: analysis & restoration". IAHR, Madrid.
- Alpaydin, E., 2010. *Introduction to Machine Learning*, 2nd ed, MIT Press, Massachusetts, 584 pp.
- Aparicio, E., Carmona-Catot, G., Moyle, P. B., García-Berthou, E., 2011. Development and evaluation of a fish-based index to assess biological integrity of Mediterranean streams. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 21 (4), 324-337.
- Baran, P., Lek, S., Delacoste, M., Belaud, A., 1996. Stochastic models that predict trout population density or biomass on a mesohabitat scale. *Hydrobiologia*, 337 (1), 1-9.
- Barbieri, R., Zogaris, S., Kalogianni, E., Stoumboudi, M., Chatzinikolaou, et al., 2015. *Freshwater Fishes and Lampreys of Greece*, Monographs on Marine Sciences No. 8. Hellenic Center for Marine Research, Athens, 134 pp.
- Baron, J. S., Poff, N. L., Angermeier, P. L., Dahm, C. N., Gleick, P. H. et al., 2002. Meeting ecological and societal needs for freshwater. *Ecological Applications*, 12 (5), 1247-1260.
- Belenguer, V., Martínez-Capel, F., Masiá, A., Picó, Y., 2014. Patterns of presence and concentration of pesticides in fish and waters of the Júcar River (Eastern Spain). *Journal of Hazardous Materials*, 265, 271-279.
- Belmar, O., Velasco, J., Martínez-Capel, F., 2011. Hydrological classification of natural flow regimes to support environmental flow assessments in intensively regulated Mediterranean Rivers, Segura River Basin (Spain). *Environmental Management*, 47 (5), 992-1004.
- Belmar, O., Velasco, J., Martínez-Capel, F., Peredo-Parada, M., Snelder, T., 2012. Do environmental stream classifications support flow assessments in Mediterranean basins? *Water Resources Management*, 26 (13), 3803-3817.
- Bernardo, J. M., Ilhéu, M., Matono, P., Costa, A. M., 2003. Interannual variation of fish assemblage structure in a Mediterranean river: implications of streamflow on the dominance of native or exotic species. *River Research and Applications*, 19 (5-6), 521-532.
- Bishop, C. M., 1996. *Neural Networks for Pattern Recognition*, University Press, Oxford, 504 pp.
- Bonada, N., Dolédec, S., Statzner, B., 2007. Taxonomic and biological trait differences of stream macroinvertebrate communities between Mediterranean and temperate regions: implications for future climatic scenarios. *Global Change Biology*, 13 (8), 1658-1671.
- Broennimann, O., Fitzpatrick, M. C., Pearman, P. B., Petitpierre, B., Pellissier, L. et al., 2012. Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, 21 (4), 481-497.
- Brosse, S., Arbuckle, C. J., Townsend, C. R., 2003. Habitat scale and biodiversity: influence of catchment, stream reach and bedform scales on local invertebrate diversity. *Biodiversity and Conservation*, 12 (10), 2057-2075.
- Brown, G. W., Krygier, J. T., 1970. Effects of Clear-Cutting on Stream Temperature. *Water Resources Research*, 6 (4), 1133-1139.
- Buckland, S. T., Elston, D. A., 1993. Empirical Models for the Spatial Distribution of Wildlife. *Journal of Applied Ecology*, 30 (3), 478-495.
- Caissie, D., 2006a. *River discharge and channel width relationships for New Brunswick rivers*. Canadian Technical Report of Fisheries and Aquatic Sciences, No 2637, 26 pp.
- Caissie, D., 2006b. The thermal regime of rivers: a review. *Freshwater Biology*, 51 (8), 1389-1406.
- Caissie, D., El-Jabi, N., 1995. Comparison and regionalization of hydrologically based instream flow techniques in Atlantic Canada. *Canadian Journal of Civil Engineering*, 22 (2), 235-246.
- Carballo, R., Cancela, J., Iglesias, G., Marín, A., Neira, X. et al., 2009. WFD indicators and definition of the ecological status of rivers. *Water Resources Management*, 23 (11), 2231-2247.
- CEDEX, 2011. *Evaluación del impacto del cambio climático en los recursos hídricos en régimen natural*. Centro de Estudios y Experimentación de Obras Públicas (CEDEX) por la Dirección General del Agua con la participación de la Oficina Española del Cambio Climático. Resumen ejecutivo, Madrid.
- Clark, J. S., Carpenter, S. R., Barber, M., Collins, S., Dobson, A., et al., 2001. Ecological Forecasts: An Emerging Imperative. *Science*, 293 (5530), 657-660.
- Clavero, M., García-Berthou, E., 2006. Homogenization dynamics and introduction routes of invasive freshwater fish in the Iberian Peninsula. *Ecological Applications*, 16 (6), 2313-2324.
- Clavero, M., Hermoso, V., Levin, N., Kark, S., 2010. Biodiversity research: Geographical linkages between threats and imperilment in freshwater fish in the Mediterranean Basin. *Diversity and Distributions*, 16 (5), 744-754.
- Cohen, J., 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20 (1), 37-46.
- 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 & Management*, 17 (6), 699-707.
- Costa, R. M. S., Martínez-Capel, F., Muñoz-Mas, R., Alcaraz-Hernández, J. D., Garófano-Gómez, V., 2012. Habitat suitability modelling at mesohabitat scale and effects of dam operation on the endangered Júcar nase, *Parachondrostoma arrigonis* (river Cabriel, Spain). *River Research and Applications*, 28, 740-752.
- Crivelli, A. J., 2006. *Luciobarbus guiraonis*. The IUCN Red List of Threatened Species, Version 2014.3. www.iucnredlist.org. (Accessed 07 February 2015).
- CHJ, 2008. *Explotación de la red de vigilancia de la calidad de las aguas, mediante índices bióticos, en el ámbito de la Confederación Hidrográfica del Júcar*. Confederación Hidrográfica del Júcar, Valencia. 105 pp.
- D'Heygere, T., Goethals, P. L. M., De Pauw, N., 2006. Genetic algorithms for optimisation of predictive ecosystems models based on decision trees and neural networks. *Ecological Modelling*, 195 (1-2), 20-29.
- Davis, A., Jenkinson, L., Lawton, J., Shorrocks, B., Wood, S., 1998. Making mistakes when predicting shifts in species range in response to global warming. *Nature*, 391, 783-786.
- De Sostoa, A., 2002. Las comunidades de peces en las cuencas mediterráneas: Caracterización y Problemática p. 51-56. In: *Atlas y Libro Rojo de los Peces Continentales de España*. I. Doadrio (Ed). CSIC y Ministerio del Medio Ambiente, Madrid.

- Demuth, H., Beale, M., Hagan, M., 2010. *Neural network toolbox user's guide*, The MathWorks Inc, Natick, Massachusetts. 901 pp.
- Dimopoulos, Y., Bourret, P., Lek, S., 1995. Use of some sensitivity criteria for choosing networks with good generalization ability. *Neural Processing Letters*, 2 (6), 1-4.
- Doadrio, I., 2001. *Atlas y libro rojo de los peces continentales de España*, Ministerio de Medio Ambiente, Madrid. 358 pp.
- Doadrio, I., Aldeguer, M., 2007. *La invasión de especies exóticas en los ríos*, Ministerio de Medio Ambiente & Universidad Politécnica de Madrid, Madrid, 124 pp.
- Drew, C. A., Wiersma, Y., Huettmann, F., 2011. *Predictive species and habitat modeling in landscape ecology: concepts and applications*, Springer, New York. 328 pp.
- Estrela, T., Fidalgo, A., Fullana, J., Maestu, J., Pérez, M. A. et al., 2004. *Júcar Pilot River Basin, provisional article 5 report Pursuant to the Water Framework Directive*, Confederación Hidrográfica del Júcar, Valencia, 200 pp.
- Evans, J. S., Cushman, S., 2009. Gradient modeling of conifer species using random forests. *Landscape Ecology*, 24 (5), 673-683.
- Feio, M. J., Poquet, J. M., 2011. Predictive Models for Freshwater Biological Assessment: Statistical Approaches, Biological Elements and the Iberian Peninsula Experience: A Review. *International Review of Hydrobiology*, 96 (4), 321-346.
- Ferreira, T., Oliveira, J., Caiola, N., De Sostoa, A., Casals, F. et al., 2007. Ecological traits of fish assemblages from Mediterranean Europe and their responses to human disturbance. *Fisheries Management and Ecology*, 14 (6), 473-481.
- 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 (01), 38-49.
- Filipe, A. F., Magalhães, M. F., Collares-Pereira, M. J., 2010. Native and introduced fish species richness in Mediterranean streams: the role of multiple landscape influences. *Diversity and Distributions*, 16 (5), 773-785.
- Fitzpatrick, M. C., Weltzin, J. F., Sanders, N. J., Dunn, R. R., 2007. The biogeography of prediction error: why does the introduced range of the fire ant over-predict its native range? *Global Ecology and Biogeography*, 16 (1), 24-33.
- Franklin, J., 2010. *Mapping species distributions: spatial inference and prediction*, Cambridge University Press, New York, 338 pp.
- Gabriels, W., Goethals, P., Dedecker, A., Lek, S., De Pauw, N., 2007. Analysis of macrobenthic communities in Flanders, Belgium, using a stepwise input variable selection procedure with artificial neural networks. *Aquatic Ecology*, 41 (3), 427-441.
- García de Jalón, D., González del Tánago, M., 2007. *Obras hidráulicas y ecosistemas fluviales*, EOI, Madrid, 31 pp.
- García de Jalón, D., Sánchez Navarro, R., Serrano, J., 2007. *Ateraciones de los regímenes de caudales de los ríos*, Ministerio de Medio Ambiente, Madrid. 65 pp.
- Garófano-Gómez, V., Martínez-Capel, F., Peredo-Parada, M., Olaya-Marín, E. J., Muñoz-Mas, R. et al., 2011. Assessing hydromorphological and floristic patterns along a regulated Mediterranean river: The Serpis River (Spain). *Limnetica*, 30 (2), 307-238.
- Gasith, A., Resh, V. H., 1999. Streams in Mediterranean climate regions: Abiotic Influences and Biotic Responses to Predictable Seasonal Events. *Annual Review of Ecology and Systematics*, 30 (1), 51-81.
- Gevrey, M., Dimopoulos, I., Lek, S., 2003. Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecological Modelling*, 160 (3), 249-264.
- Gevrey, M., Dimopoulos, I., Lek, S., 2006. Two-way interaction of input variables in the sensitivity analysis of neural network models. *Ecological Modelling*, 195 (1-2), 43-50.
- Goethals, P., Dedecker, A., Gabriels, W., Lek, S., De Pauw, N., 2007. Applications of artificial neural networks predicting macroinvertebrates in freshwaters. *Aquatic Ecology*, 41 (3), 491-508.
- Gortázar, J., Alonso González, C., Iturriaga, C., Hernández, D., Baeza Sanz, D., García de Jalón, D., 2007. *Estudio hidrobiológico de la cuenca del río Cabriel en las provincias de Albacete y Cuenca*, Castilla la Mancha (Ecohidráulica SL-FUCOVASA) Madrid.
- Granado-Lorencio, C., 1996. *Ecología de peces*, Universidad de Sevilla, Sevilla, 353 pp.
- Granado-Lorencio, C., 2000. *Ecología de comunidades: el paradigma de los peces de agua dulce*, Universidad de Sevilla, 284 pp.
- Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8 (9), 993-1009.
- Hanley, J. A., McNeil, B. J., 1982. The meaning and use of the area under a receiver operating characteristics curve. *Radiology*, 143, 29-36.
- Hauser-Davis, R. A., Oliveira, T. F., Silveira, A. M., Silva, T. B., Ziolli, R. L., 2010. Case study: Comparing the use of nonlinear discriminating analysis and Artificial Neural Networks in the classification of three fish species: acaras (*Geophagus brasiliensis*), tilapias (*Tilapia rendalli*) and mullets (*Mugil liza*). *Ecological Informatics*, 5 (6), 474-478.
- Hellawell, J. M., 1978. *Biological surveillance of rivers*, Water Research Center, Stevenage, 332 pp.
- Hermoso, V., Clavero, M., 2011. Threatening processes and conservation management of endemic freshwater fish in the Mediterranean basin: a review. *Marine and Freshwater Research*, 62 (3), 244-254.
- Hrachowitz, M., Soulsby, C., Imholt, C., Malcolm, I. A., Tetzlaff, D., 2010. Thermal regimes in a large upland salmon river: a simple model to identify the influence of landscape controls and climate change on maximum temperatures. *Hydrological Processes*, 24 (23), 3374-3391.
- Hsieh, W. W., 2009. *Machine Learning Methods in the Environmental Sciences: Neural Networks and Kernels*, Cambridge University Press, Vancouver BC, 364 pp.
- Ibarra, A. A., Gevrey, M., Park, Y.-S., Lim, P., Lek, S., 2003. Modelling the factors that influence fish guilds composition using a back-propagation network: assessment of metrics for indices of biotic integrity. *Ecological Modelling*, 160 (3), 281-290.
- Isa, I. S., Omar, S., Saad, Z., Osman, M. K., 2010. Performance comparison of different multilayer perceptron network activation functions in automated weather classification. p. 71-75. In: *Proceedings of the 2010 Fourth Asia International Conference on Mathematical/Analytical Modelling and Computer Simulation*, Kota Kinabalu, 26-28 May 2010, Malaysia.

- Isaak, D. J., Wollrab, S., Horan, D., Chandler, G., 2012. Climate change effects on stream and river temperatures across the northwest U.S. from 1980–2009 and implications for salmonid fishes. *Climatic Change*, 113 (2), 499-524.
- 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 (1), 157-170.
- Jiménez, J., Lacomba, I., Sancho, V., Risueño, P., 2002. *Peces continentales, anfibios y reptiles de la Comunidad Valenciana*, Generalitat Valenciana, Valencia, 271 pp.
- Jopp, F., Reuter, H., Breckling, B. (Eds.), 2011. *Modelling complex ecological dynamics: an Introduction into ecological modelling for students, Teachers and Scientists*. Springer-Verlag, Berlin, 413 pp.
- Jorgensen, S. E., Fath, B. D., 2011. *Fundamentals of ecological modelling: applications in environmental management and research. 4th ed.*, Elsevier, Amsterdam, 432 pp.
- Joy, M. K., Death, R. G., 2004. Predictive modelling and spatial mapping of freshwater fish and decapod assemblages using GIS and neural networks. *Freshwater Biology*, 49 (8), 1036-1052.
- Koch, G. G., Landis, J. R., Freeman, J. L., Freeman, D. H., Jr., Lehnen, R. G., 1977. A General Methodology for the Analysis of Experiments with Repeated Measurement of Categorical Data. *Biometrics*, 33 (1), 133-158.
- Kottelat, M., Freyhof, J., 2007. *Handbook of European freshwater fishes*, Kottelat, Cornol, Switzerland and Freyhof, Berlin, Germany, 646 pp.
- Kurková, V., 1992. Kolmogorov's theorem and multilayer neural networks. *Neural Networks*, 5 (3), 501-506.
- Lake, P. S., Bond, N., Reich, P., 2007. Linking ecological theory with stream restoration. *Freshwater Biology*, 52 (4), 597-615.
- Lambert, D., 1992. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, 34 (1), 1-14.
- Leclere, J., Oberdorff, T., Belliard, J., Leprieur, F., 2011. A comparison of modeling techniques to predict juvenile 0+ fish species occurrences in a large river system. *Ecological Informatics*, 6 (5), 276-285.
- Lek, S., Belaud, A., Dimopoulos, I., Lauga, J., Moreau, J., 1995. Improved estimation, using neural networks, of the food consumption of fish populations. *Marine and Freshwater Research*, 46 (8), 1229-1236.
- Lek, S., Scardi, M., Verdonshot, P., Descy, J. P., Park, Y. S. (Ed.), 2005. *Modelling community structure in freshwater ecosystems*. Springer-Verlag, Berlin, 518 pp.
- Leopold, L. B., Maddock, T., 1953. *The hydraulic geometry of stream channels and some physiographic implications*, U.S. Govt. Print. Off., Washington, 56 pp.
- Leopold, L. B., Wolman, M. G., Miller, J. P., 1964. *Fluvial processes in geomorphology*, W.H. Freeman, San Francisco, 544 pp.
- Liu, H., Chan, K., 2010. Introducing COZIGAM: an R package for unconstrained and constrained zero-inflated generalized additive model analysis. *Statistical Software*, 35 (11), 1-26.
- Lobón-Cerviá, J., Mortensen, E., 2005. Population size in stream-living juveniles of lake-migratory brown trout *Salmo trutta* L.: the importance of stream discharge and temperature. *Ecology of Freshwater Fish*, 14 (4), 394-401.
- Maceda-Veiga, A., De Sostoa, A., 2011. Observational evidence of the sensitivity of some fish species to environmental stressors in Mediterranean rivers. *Ecological Indicators*, 11 (2), 311-317.
- Magnuson, J. J., Crowder, L. B., Medvick, P. A., 1979. Temperature as an Ecological Resource. *American Zoologist*, 19 (1), 331-343.
- 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 (5), 921-931.
- Martínez-Capel, F., Costa, R., Muñoz-Mas, R., 2008. *Evaluación de las poblaciones de peces en el río Júcar bajo el embalse de Alarcón, en las comarcas de la Manchuela Conquense (Cuenca) y la Mancha Júcar-centro (Albacete)*, Universitat Politècnica de Valencia, Valencia.
- Martínez-Muro, 2003. *Control de calidad de aguas fluviales: diseño y puesta a punto de indicadores físico-químicos y biológicos*, Universitat de València, Facultat de Ciències Químiques, València, 380 pp.
- Mas-Martí, E., García-Berthou, E., Sabater, S., Tomanova, S., Muñoz, I., 2010. Comparing fish assemblages and trophic ecology of permanent and intermittent reaches in a Mediterranean stream. *Hydrobiologia*, 657 (1), 167-180.
- Mastrorillo, S., Lek, S., Dauba, F., Belaud, A., 1997. The use of artificial neural networks to predict the presence of small-bodied fish in a river. *Freshwater Biology*, 38 (2), 237-246.
- Munné, A., Prat, N., Solà, C., Bonada, N., Rieradevall, M., 2003. A simple field method for assessing the ecological quality of riparian habitat in rivers and streams: QBR index. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 13 (2), 147-163.
- Muñoz-Mas, R., Martínez-Capel, F., Schneider, M., Mouton, A. M., 2012. Assessment of brown trout habitat suitability in the Júcar River Basin (SPAIN): Comparison of data-driven approaches with fuzzy-logic models and univariate suitability curves. *Science of The Total Environment*, 440, 123-131.
- Muñoz-Mas, R., Martínez-Capel, F., Alcaraz-Hernández, J. D., Mouton, A. M., 2015. Can multilayer perceptron ensembles model the ecological niche of freshwater fish species? *Ecological Modelling*, 309-310, 72-81.
- Oberdorff, T., Guégan, J.-F., Hugueny, B., 1995. Global scale patterns of fish species richness in rivers. *Ecography*, 18 (4), 345-352.
- Olaya-Marín, E. J., Martínez-Capel, F., Soares Costa, R. M., Alcaraz-Hernández, J. D., 2012. Modelling native fish richness to evaluate the effects of hydromorphological changes and river restoration (Júcar River Basin, Spain). *Science of The Total Environment*, 440, 95-105.
- Olaya-Marín, E. J., Martínez-Capel, F., Vezza, P., 2013. A comparison of artificial neural networks and random forests to predict native fish species richness in Mediterranean rivers. *Knowledge and Management of Aquatic Ecosystems*, 409 (07) DOI: 10.1051/kmae/2013052.
- Olden, J. D., Jackson, D. A., 2002. Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks. *Ecological Modelling*, 154 (1-2), 135-150.
- Olden, J. D., Lawler, J. J., Poff, N. L., 2008. Machine learning methods without tears: A primer for ecologists. *The Quarterly Review of Biology*, 83 (2), 171-193.

- Özesmi, S. L., Tan, C. O., Özesmi, U., 2006. Methodological issues in building, training, and testing artificial neural networks in ecological applications. *Ecological Modelling*, 195 (1-2), 83-93.
- Papadaki, C., Ntoanidis, L., Zogaris, S., Martinez-Capel, F., Muñoz-Mas, R. *et al.*, 2014. Habitat hydraulic modelling for environmental flow restoration in upland streams in Greece. p. 385-392. In: *12th International Conference on Protection and Restoration of the Environment*, Skiathos, 29 June - 3 July 2014. University of Thessaly, Volos.
- Park, Y.-S., Lek, S., Chon, T.-S., Verdonshot, F. M., 2008. Evaluation of Environmental Factors to Determine the Distribution of Functional Feeding Groups of Benthic Macroinvertebrates Using an Artificial Neural Network. *Journal of Ecology and Field Biology*, 31 (3), 233-241.
- Park, Y. S., Chon, T. S., 2007. Biologically-inspired machine learning implemented to ecological informatics. *Ecological Modelling*, 203 (1-2), 1-7.
- Pearce, J., Ferrier, S., 2000. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, 133 (3), 225-245.
- Petrovic, M., Ginebreda, A., Acuña, V., Batalla, R. J., Elozegi, A. *et al.*, 2011. Combined scenarios of chemical and ecological quality under water scarcity in Mediterranean rivers. *Trends in Analytical Chemistry*, 30 (8), 1269-1278.
- Piekniewski, F., Rybicki, L., 2004. *Visual comparison of performance for different activation functions in MLP networks*. pp. 2947-2952. In: Proceedings of International Joint Conference on Neural Networks IJCNN '04, Budapest, Hungary, 25-29 July 2004, Hungary.
- Prchalová, M., Slavík, O., Bartoš, L., 2006. Patterns of cyprinid migration through a fishway in relation to light, water temperature and fish circling behaviour. *International Journal of River Basin Management*, 4 (3), 213-218.
- Provencher, M., Lamontagne, M. P., 1977. *Méthode de la détermination d'un indice d'appréciation de la qualité des eaux selon différentes utilisations*, Service de la qualité des eaux. Ministère des Richesses naturelles, Québec, 76 pp.
- Qin, Z., Su, G.-l., Zhang, J.-e., Ouyang, Y., Yu, Q. *et al.*, 2010. Identification of important factors for water vapor flux and CO₂ exchange in a cropland. *Ecological Modelling*, 221 (4), 575-581.
- Ryan, M., Müller, C., Di, H. J., Cameron, K. C., 2004. The use of artificial neural networks (ANNs) to simulate N₂O emissions from a temperate grassland ecosystem. *Ecological Modelling*, 175 (2), 189-194.
- Sabater, S., Tockner, K., 2010. Effects of Hydrologic Alterations on the Ecological Quality of River Ecosystems Water Scarcity in the Mediterranean. p. 15-39. In: *Water Scarcity in the Mediterranean: Perspectives Under Global Change*. Sabater, S., Barceló, D. (Eds). Springer, Berlin / Heidelberg.
- Schlosser, I. J., 1991. Stream fish ecology: a landscape perspective. *Bioscience*, 41, 704-712.
- Smith, K. G., Darwall, W. R. T., 2006. *The status and distribution of freshwater fish endemic to the mediterranean basin*, IUCN -The World Conservation Union, Gland, Switzerland/Cambridge, 41 pp.
- Swets, J., 1988. Measuring the accuracy of diagnostic systems. *Science*, 240 (4857), 1285-1293.
- Tirelli, T., Pessani, D., 2009. Use of decision tree and artificial neural network approaches to model presence/absence of *Telestes muticellus* in piedmont (North-Western Italy). *River Research and Applications*, 25 (8), 1001-1012.
- Tirelli, T., Pessani, D., 2011. Importance of feature selection in decision-tree and artificial-neural-network ecological applications. *Alburnus alburnus alborella: A practical example*. *Ecological Informatics*, 6 (5), 309-315.
- Tirelli, T., Pozzi, L., Pessani, D., 2009. Use of different approaches to model presence/absence of *Salmo marmoratus* in Piedmont (Northwestern Italy). *Ecological Informatics*, 4 (4), 234-242.
- Veza, P., Muñoz-Mas, R., Martinez-Capel, F., Mouton, A., 2015. Random forests to evaluate biotic interactions in fish distribution models. *Environmental Modelling & Software*, 67, 173-183.
- Vila-Gispert, A., Alcaraz, C., García-Berthou, E., 2005. Life-history traits of invasive fish in small Mediterranean streams. *Biological Invasions*, 7 (1), 107-116-116.
- Webb, B. W., Hannah, D. M., Moore, R. D., Brown, L. E., Nobilis, F., 2008. Recent advances in stream and river temperature research. *Hydrological Processes*, 22 (7), 902-918.
- Zalewski, M., Cowx, I. G., 1990. Factors affecting the efficiency of electric fishing. p. 89-111. In: *Fishing with Electricity: Applications in Freshwater Fisheries Management*. Cowx, I. G., Lamarque, P. (Eds). Fishing News Books, Oxford.
- Zarkami, R., Sadeghi, R., Goethals, P., 2012. Use of fish distribution modelling for river management. *Ecological Modelling*, 230, 44-49.
- Zhang, G., Eddy Patuwo, B., Y. Hu, M., 1998. Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, 14 (1), 35-62.