The final publication is available at

http://dx.doi.org/10.1016/j.scitotenv.2015.11.147

Additional Information

Shifts in the suitable habitat available for brown trout (*Salmo trutta* L.) under short-term climate change scenarios.

R. Muñoz-Mas[a*], A, Lopez-Nicolas[b], F. Martínez-Capel[a], M. Pulido-Velazquez[b]

[a]Institut d'Investigació per a la Gestió Integrada de Zones Costaneres (IGIC), Universitat Politècnica de València, C/Paranimf 1, 46730 Grau de Gandia, València, Spain

[b]Research Institute of Water and Environmental Engineering (IIAMA), Universitat Politècnica de València, Camí de Vera s/n 46022, València, Spain.

* Corresponding author. Tel.: (+34) 622098521

E-mail addresses: pitifleiter@hotmail.com (R. Muñoz-Mas), anloni@cam.upv.es (A, Lopez-Nicolas), fmcapel@dihma.upv.es (F. Martínez-Capel), mapuve@hma.upv.es (M. Pulido-Velazquez)

1

ABSTRACT

The impact of climate change on the habitat suitability for large brown trout (*Salmo trutta* L.) was studied in a segment of the Cabriel River (Iberian Peninsula). The future flow and water temperature patterns were simulated at a daily time step with M5' models' trees (NSE of 0.78 and 0.97 respectively) for two short-term scenarios (2011-2040) under the Representative Concentration Pathways (RCP 4.5 and 8.5). An ensemble of five strongly regularized machine learning techniques (generalized additive models, multilayer perceptron ensembles, random forests, support vector machines and fuzzy rule base systems) was used to model the microhabitat suitability (depth, velocity and substrate) during summertime and to evaluate several flows simulated with River2D$^{©}$. The simulated flow rate and water temperature were combined with the microhabitat assessment to infer Bivariate Habitat Duration Curves (BHDCs) under historical conditions and climate change scenarios using either the Weighted Usable Area (WUA) or the Boolean-based Suitable Area (SA). The forecasts for both scenarios jointly predicted a significant reduction in the flow rate and an increase in water temperature (mean rate of change of *ca.* -25 % and +4 % respectively). The five techniques converged on the modelled suitability and habitat preferences; large brown trout selected relatively high flow velocity, large depth and coarse substrate. However, the model developed with support vector machines presented a significantly trimmed output range (Max.: 0.38), and thus its predictions were banned from the WUA-based analyses. The BHDCs based on the WUA and the SA broadly matched, indicating an increase in the number of days with less suitable habitat available (WUA and SA) and/or with higher water temperature (trout will endure impoverished environmental conditions *ca.* 82 % of the days). Finally, our results suggested the potential extirpation of the species from the study site during short time spans.

Keywords: Bivariate habitat duration curves, machine learning, Mediterranean river, models' ensemble, *Salmo trutta*, water temperature.

2

List of acronyms:

ANN – Artificial Neural Network

CART – Classification And Regression Tree

CCI – Correctly Classified Instances

CORDEX – Coordinated Regional Climate Downscaling Experiment

FRBS – Fuzzy Rule Base System

GA – Genetic Algorithm

GAM – Generalized Additive Model.

BHDC – Bivariate Habitat Duration Curve

HMU – Hydro-Morphological Unit.

HTS – Habitat Time Series

IPCC – Intergovernmental Panel on Climate Change

MLP – Multi-Layer Perceptron

MLPE – Multi-Layer Perceptron Ensemble

MSE – Mean Squared Error

NSE – Nash-Sutcliffe Efficiency

RCP – Representative Concentration Pathway

RF – Random Forests

SA – Suitable Area

Sn – Sensitivity

Sp – Specificity

SRESs – Special Report on Emissions Scenarios

SSP – Shared Socioeconomic Pathway

SVM – Support Vector Machine

TSK – zero-order presence-absence Takagi-Sugeno-Kang fuzzy rule base system

TSS – True Skill Statistic

WUA – Weighted Usable Area.

1 Introduction

Worldwide, climate change poses a major threat to freshwater biodiversity (Dudgeon, 2014). Recent and long-term changes in climate have altered mean annual temperatures and precipitation patterns, which may have a suite of negative effects for fluvial ecosystems such as an altered flow regime or increased water temperature (Ficke *et al.*, 2007; Palmer *et al.*, 2009). Furthermore, the effects of climate change may interact with other anthropogenic impacts such as land use changes or the introduction of non-indigenous species; thus, different scenarios consistently indicate that biodiversity will continue to decline over the 21st century (Ficke *et al.*, 2007). Consequently, it is a priority to identify which types of stressor combinations are likely to negatively affect species under current and forecasted climate scenarios.

Predicting the impacts of climate change requires proactive management responses if valuable riverine ecosystems are to be preserved (sensu Palmer *et al.*, 2008). Proactive management requires sound monitoring and predicting capabilities at the local scale (Palmer *et al.*, 2009) because, for every river of the world, there is a natural flow regime representative of the unaltered landscape, precipitation and temperature (Poff *et al.*, 1997). In accordance to these specific combinations of factors, some rivers would be consequently very stable, while others would be especially prone to alteration (Schneider *et al.*, 2013). For instance, the Eastern Mediterranean region of the Iberian peninsula is likely to be characterized by a continued decrease in water yield (Chirivella Osma *et al.*, 2014; Salmoral *et al.*, 2015), which may create a bottleneck for species survival.

Experimental studies have provided scientists with the necessary knowledge about the key environmental factors and biotic processes to plan adequate mitigation and conservation actions

4

(Filipe *et al.*, 2013). More specifically, the use of models to run scenarios that capture the spectrum of possible outcomes has become an invaluable tool for anticipating changes in hydrology and water temperatures (Palmer *et al.*, 2008). Among the different available approaches, the physical habitat simulation approach (Bovee *et al.*, 1998) has emerged as a powerful tool to quantify variations in the suitable habitat available for different river biota. Thus, this method can evaluate the effect of different management alternatives, restoration actions (Mouton *et al.*, 2007) and the potential effects of climate change (Almodóvar *et al.*, 2012; Hauer *et al.*, 2013; Papadaki *et al.*, 2016; Viganò *et al.*, 2015). The physical habitat simulation approach to evaluate the impact of climate change in unregulated river segments typically couples three different modelling components:

(i)   the hydrologic and water temperature models component

(ii)   the hydraulic model component

(iii)   the habitat suitability model component

Regarding the first component, process-based (*i.e.* physically-based) models have traditionally been preferred (Van Vliet *et al.*, 2012), and there is a myriad of capable models available for hydrologic simulation (Zhuo *et al.*, 2015), some with specific modules dedicated to water temperature forecasting (e.g. Luo *et al.*, 2013). However, machine learning-based approaches (data-based) such as neuro-fuzzy rule base models (Lohani *et al.*, 2012) or Artificial Neural Networks (ANNs) (Rabi *et al.*, 2015) have recently gained appreciable visibility (Zhuo *et al.*, 2015). Moreover, the Intergovernmental Panel on Climate Change (IPCC) has recently rendered a new set of climate change scenarios (IPCC, 2014), the so-called Representative Concentration Pathways (RCPs) scenarios, which superseded the Special Report on Emissions Scenarios projections published in 2000 (Nakicenovic *et al.*). The main advantage of the new RCP over previous scenarios relies on the consideration of international policies to control greenhouse emissions. Actually, climate projections have been downscaled through the Coordinated Regional Climate Downscaling Experiment (CORDEX) (Christensen *et al.*, 2014; Jones *et al.*, 2011), and thus they are ready to be applied at the

5

local scales. The use of the latest scenarios established for climate change studies represents an update of previous forecasts (Santiago *et al.*, 2015) since they have been demonstrated to play a paramount role in the final predictions about the impact of climate change in the flow regime (Döll and Schmied, 2012). Consequently, selecting the appropriate modelling approach for the hydrologic component of a climate change study represents a challenge by itself.

Regarding the hydraulic model, there is also a large set of dedicated software packages for hydraulic modelling. In physical habitat studies, typically one- or two-dimensional hydraulic models are used (Jowett and Duncan, 2012) to simulate spatial and temporal variations to water depth, flow velocity, substrate and cover (Conallin *et al.*, 2010).

Finally, the last component is the habitat suitability model for the target species. This component, unlike the hydrological and water temperature component, has been typically addressed by machine learning approaches, and examples include, Classification And Regression Trees (CARTs) (Fukuda *et al.*, 2014), Generalized Additive Models (GAMs) (Jowett and Davey, 2007), ANNs (Fukuda *et al.*, 2014; Muñoz-Mas *et al.*, 2014b), Fuzzy Rule Base Systems (FRBSs) (Fukuda *et al.*, 2013; Muñoz-Mas *et al.*, 2012), and Support Vector Machines (SVMs) (Fukuda *et al.*, 2014; Lin *et al.*, 2015). However, no consensus has been reached on the optimal machine learning technique since each one has its own unique structure and merits (Lin *et al.*, 2015). Moreover, the strengths and weaknesses of every machine learning technique must be considered; otherwise, the selection of an inappropriate habitat suitability model may result in erroneous predictions (Lin *et al.*, 2015). One method to overcome these limitations relies on the use of models' ensembles (Bouska *et al.*, 2015). Ensemble modelling consists of learning several models, each developed with a unique technique or with several different techniques, and combining their individual forecasts into a single prediction. Thus, almost every single technique has their ensemble counterpart; however, only Random Forests (RF) (Breiman, 2001), which is the CART ensemble counterpart, can be considered widespread in physical habitat simulation studies (e.g. Fukuda *et al.*, 2014; Lin *et al.*, 2015; Vezza *et al.*, 2015).

Riparian vegetation, macroinvertebrates and fish species are effective indicators of ecological status (Magdaleno and Fernández, 2011); however, fish species present some advantages that make them especially suited for that purpose. Fish occupy high trophic levels, are relatively easy to sample and to identify, and generally are known to react quickly to instream habitat constraints (Lorenz *et al.*, 2013). Furthermore, in the Iberian context, previous studies suggested that cold-water species, such as brown trout (*Salmo trutta* L. 1758), may suffer severe reductions in their distribution areas (Almodóvar *et al.*, 2012; Santiago *et al.*, 2015). Therefore, these species are an attractive ultimate target in climate change studies, particularly when the species occurs at the edge of its distribution area (Santiago *et al.*, 2015).

In the Iberian Peninsula, water temperature has been suggested to be the main driver of brown trout distribution (Almodóvar *et al.*, 2012; Santiago *et al.*, 2015), whereas in other alpine regions streamflow has also been stressed to potentially impact salmonids' distributions (Papadaki *et al.*, 2016; Tisseuil *et al.*, 2012; Viganò *et al.*, 2015). Previous studies suggest that present-day droughts cause relatively small and transient changes to Mediterranean stream fish assemblages, whereas longer and more severe droughts – as those expected in the streams of the Jucar River basin (Chirivella Osma *et al.*, 2014; Estrela *et al.*, 2012; Salmoral *et al.*, 2015) – may result in declines or local extinctions of the most sensitive species (Magalhães *et al.*, 2007).

The present study analysed the shifts in the suitable habitat available for the large brown trout (*Salmo trutta* L.) – body length > 20 cm – under two short-term (2011-2040) climate change scenarios (the Representative Concentration Pathways 4.5 and 8.5). The study was carried out in a segment of the Cabriel River (Iberian Peninsula) applying the physical habitat simulation approach. Hydrological modelling and water temperature modelling were carried out at a daily time step with M5' predicting the flow rate and water temperature along the historical and climate change scenarios. Several flows were simulated with River2D© and were assessed with a models' ensemble of presence-absence habitat suitability models that encompassed five different machine learning

7

techniques: generalized additive models, multi-layer perceptron ensembles, random forests, support vector machines and zero-order presence-absence Takagi-Sugeno-Kang fuzzy rule base systems. Standard habitat-flow curves were developed and simulated flows were transduced to habitat time series. Finally, the frequency and magnitude of the extreme events (*i.e.* droughts and temperature peaks) were analysed by means of bivariate habitat duration curves.

2 Materials and methods

2.1 The Cabriel River

The Cabriel River is an unregulated tributary of the Jucar River (eastern Iberian Peninsula). Its river basin has been affected by a marked human depopulation which is likely to continue through 2050 (INE, 2012), allowing us to neglect any potential change to the ecological status based on anthropogenic actions (*e.g.,* future regulation, water abstraction or increased pollution). In total, the river is 220 km long, the catchment elevation ranges from 490 to 1790 m a.s.l. and its drainage area covers 4750 km^2. The upper part of the basin, where the study site is located, has an average riverbed slope of 1.1% and the land cover – from the CORINE Land Cover classification – (Bossard *et al.*, 2000) mainly consists of forested areas (86%) and crops (12%). The study area has a typical mesomediterranean climate with a mean annual precipitation of *ca.* 500 mm, resulting in low flows and high evapotranspiration in summer and high flows in spring and autumn (Vezza *et al.*, 2015). However, it has suffered a significant reduction in mean flow since the 1980's (Estrela *et al.*, 2012), which may be exacerbated if climate change scenarios are met. Furthermore, at the study site the Cabriel River does not contain any invasive species (Vezza *et al.*, 2015), thus future changes in the ecosystem should be driven principally by changes in the flow regime and water temperature.

2.2 Large brown trout data collection

The large brown trout data were retrieved from a previous study (Muñoz-Mas *et al.*, 2014b) performed during summertime (2007-2009) in several rivers of the Tagus (Guadiela and Cuervo), Jucar and Senia River basins (eastern Iberian Peninsula) (Fig. 1).
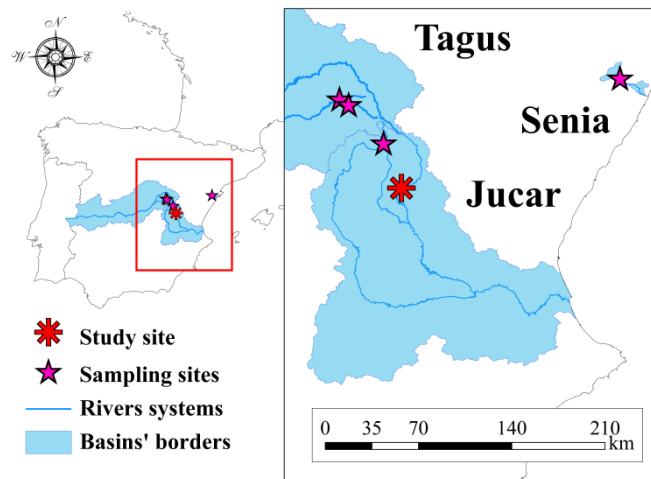

Fig. 1. Location of the large brown trout (*Salmo trutta* L.) sampling sites and location of the site where the shifts in the suitable habitat available under short-term (2011-2040) climate change scenarios were studied.

The sampling unit was the microhabitat, which is defined as a localized area (from zero to few m^2) – having relatively homogeneous conditions of depth, velocity, substrate, and cover – used by an aquatic organism for specific behaviours (*e.g.,* feeding) (Bovee *et al.*, 1998). The large brown trout sampling was undertaken by snorkelling because it proved to be superior to direct observation from the shoreline as well as electrofishing sampling. At this scale, direct observation tends to underestimate the number of individuals (Heggenes *et al.*, 1990), whereas electrofishing is characterized by coarser resolution and introduces a bias related to displacement caused by galvanotaxis (Brosse *et al.*, 2001). Furthermore, neither direct observation nor electrofishing allows the easy observation of fish behaviour, thus database could become noisy when including different activities (Heggenes *et al.*, 1990). In this study, only large brown trout that were 'feeding' or 'holding a feeding position' were considered because it is assumed that they are occupying such positions because they are the most energetically profitable (Rincón and Lobón-Cerviá, 1993).

9

Microhabitat sampling was performed following a modification of the equal effort sampling approach (Johnson, 1980). The river segment was stratified in Hydro-Morphological Units (HMUs) (*i.e.* pool, glide, riffle and rapid) (Muñoz-Mas *et al.*, 2012) and similar areas of slow and fast water HMUs – grouping pools with glides (slow) and riffles with rapids (fast) – were sampled. The diver carefully moved from downstream to upstream – ignoring disturbed fish – placing a weighted marker where trout were spotted. The hydraulic conditions (absence data) were measured along the selected HMUs in cross-sections whereas the presence data were measured at the markers. In the end, 98 presence cases and 1457 absence cases were sampled resulting in a relatively low prevalence dataset (*i.e.* ratio of presence cases over the entire dataset of 0.06).

Three variables were measured: mean flow velocity of the water column, depth and substrate. Velocity (m/s) was measured with an electromagnetic current meter (Valeport®, United Kingdom) and depth (m) was measured with a wading rod to the nearest cm. The percentage of each substrate class (*i.e.* bedrock, boulders, cobbles, gravel, fine gravel, sand, silt and vegetated soil) was visually recorded following the guidelines used in previous studies (Muñoz-Mas *et al.*, 2012), and the data were converted into the substrate index (-) (Mouton *et al.*, 2011) that typically ranges from zero to eight (Table 1 & Fig. 2).

Table 1. Summary of the principal statistics of the presence (large brown trout) and the absence datasets.

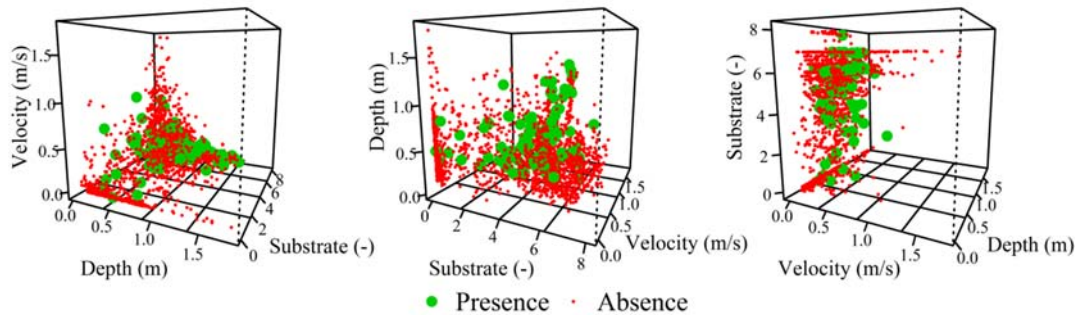|		Velocity (m/s)	Depth (m)	Substrate index (-)
Large brown trout	Min.	0.01	0.21	0.00
	1st Qu.	0.13	0.47	4.00
	Median	0.22	0.65	5.30
	Mean	0.26	0.69	4.92
	3rd Qu.	0.34	0.88	6.00
	Max.	1.00	1.45	8.00
Absences	Min.	0.00	0.03	0.00
	1st Qu.	0.04	0.31	2.80
	Median	0.12	0.50	5.60
	Mean	0.22	0.55	4.50
	3rd Qu.	0.29	0.74	6.40
	Max.	1.76	1.78	8.00

Fig. 2. Distributions of the three input variables; velocity (m/s), depth (m) and substrate (-), discriminated between the presence and absence classes.

2.3 Habitat suitability modelling

Model reliability and transparency is a major concern in habitat suitability modelling (Austin, 2007). Therefore, there are several common elements that must be considered in order to ensure the development of sound models. The first issue to avoid is the excessive parameterization, which tends to over-fit the data (reducing models' generalization) and may overestimate further predictions (Wenger and Olden, 2012). Another issue is low prevalence, as the over-representation of the absence category may shift the balance of errors, increasing the number of serious mistakes (*i.e.,* falsely predicted absences) (Platts *et al.*, 2008). Each modelling technique presents different parameter settings – the number of neurons in Artificial Neural Networks (ANNs) or the number of

11

fuzzy sets in Fuzzy Rule Base Systems (FRBSs) – that condition the propensity of the technique to over-fitting (Merow *et al.*, 2014), which may in turn also favour the assumption of the uncertain absence data as true negatives. To deal with such drawbacks, the number of parameters are limited and optimized by means of cross-validation approaches (Fukuda *et al.*, 2013). This optimization usually takes place through performance criteria irrespective of prevalence – such as the True Skill Statistic (TSS [-1, 1]) (Mouton *et al.*, 2009) and stimulating over-prediction by favouring a larger proportion of presence predictions in comparison with absence Sensitivity (Sn) > Specificity (Sp) (Fukuda *et al.*, 2013).

Following Austin's suggestions (2007), the habitat suitability models in this study were regularized as much as possible to render unimodal or monotonic responses. For the modelling techniques that maximize the classificatory capability (*e.g.*, RF), the parameter settings were adjusted to maximize the TSS (Sn > Sp), whereas the weighted Mean Squared Error (MSE) was minimized for those techniques (*e.g.*, ANN) that are trained on the basis of the overall error. Finally, partial dependence plots (Friedman, 2001) implemented in the R package *randomForest* (Liaw and Wiener, 2002) were used to graphically characterise the relationship between the input variables and the presence of large brown trout. Partial dependence plots yield easily interpretable univariate plots by depicting the influence of a given input on the output and account for the remaining variables within the model by averaging their effects.

2.3.1 *Generalized Additive Models – GAMs*

The Generalized Additive Model (GAM) (Hastie and Tibshirani, 1990) was implemented in *R* (R Core Team, 2015) with the *mgcv* package (Wood, 2004). A single tensor product was used to optimize the smooth curves because they are especially useful for representing covariate functions measured in different units (Wood, 2006). The maximum number of knots was limited to three, and the link function was set to *binominal* (*i.e.* logit link). Data prevalence has proved to significantly

12

impact the classificatory capability of GAMs (Platts *et al.*, 2008). The way to deal with this limitation has typically focused on the modification of the classification threshold, but some authors have suggested that weighting cases may yield similar results (Platts *et al.*, 2008). Consequently, in order to maximize GAMs' performance, but keeping the discriminant threshold at 0.5, the case *weights* implemented in the function *gam* were adjusted following a 3 fold cross-validation scheme. The objective function searched the maximization of the TSS while stimulating over-prediction. The absence cases were iteratively weighted around the prevalence value (0.06) from 1 % to 10 % at an increment rate of 1 %, whereas the presence cases were inversely weighted to sum up to 100 %. Once the optimal case weights were obtained, a single GAM was trained with the entire dataset and used in the subsequent analysis.

2.3.2   *Multi-Layer Perceptron Ensembles – MLPE*

Multi-Layer Perceptron Ensembles (MLPE) (Hansen and Salamon, 1990) are the ensemble counterpart of ANNs. Its development typically follows the overproduce-and-choose approach (Partridge and Yates, 1996). First, a large pool of MLPs is trained, usually by bagging (resampling with replacement) the complete dataset, and the combination of MLPs that produces the best mean prediction (arithmetic mean) on the out-of-bag dataset is sought.

During the overproduce phase, the MLP components were trained in *R* (R Core Team, 2015) with the package *monmlp* (Cannon, 2012) by minimising the Mean Squares Error (MSE). The number of nodes was restricted to two since it was demonstrated to be sufficient in habitat suitability modelling, and the hyperbolic tangent and logistic functions were used in the hidden and the output layers respectively (Muñoz-Mas *et al.*, 2014a). MLPs have proved sensitive to data prevalence (Olden *et al.*, 2002); therefore, to train every MLP, 66 % of the presences (*i.e.* 65 cases) and the same number of absences were randomly selected. The sample size did not allow training for all possible

combinations with a prevalence of 0.5. Therefore, 2000 MLPs were pooled in order to ensure that every presence case was faced to every absence case several times.

The MLPs selection was performed with a Genetic Algorithm (GA) (Holland, 1992). GAs are search-and-optimization algorithms based on the process of natural selection (Olden *et al.*, 2008). Several combinations of MLPs are encoded in bit strings (1 meaning selected and 0 discarded) corresponding to the population. The GA then searched for the minimization of the weighted MSE (*i.e.* 0.94 for presence cases and 0.06 for absence cases) in the out-of-bag dataset by evolving that population. The evolution takes place by more frequently reproducing (crossing) those individuals with better fitness (*i.e.* low weighted MSE) but regularly altering (mutating) the remainder of the individuals to properly sample the searching space. The selected GA was from the package *rgenoud* (Mebane Jr and Sekhon, 2011), and it considers nine operators driving the optimization which correspond to cloning, uniform mutation, boundary mutation, non-uniform mutation, polytope crossover, simple crossover, whole non-uniform mutation, heuristic crossover and local-minimum crossover. In the end, the whole set of operators were: 0.20, 0.75, 0.00, 0.00, 0.15, 0.75, 0.15, 0.35 and 0.00 respectively, whereas the population size and the number of generations were both set to 2000. Finally, the optimal MLPE was used to develop the partial dependence plots and to perform the habitat assessment.

2.3.3   Random Forests – RF

Random Forests (RF) (Breiman, 2001) is an ensemble machine learning technique based on the aggregation of a large set of Classification And Regression Trees (CARTs) (Breiman *et al.*, 1984). CARTs recursively split the training dataset into two parts, maximizing the homogeneity in the two child nodes until the leaves are homogeneous or the maximum number of allowed partitions is achieved (Vezza *et al.*, 2015). RF corrects many of the unpleasant issues of CARTs, such as over-fitting, by aggregating awkward models. Thus, in every split, not all input variables are tested but a

14

few of them (Breiman, 2001). During the prediction phase, every CART performs its prediction and the Boolean output of RF is calculated by majority vote but also probabilistic outputs are possible. Despite these aforementioned improvements, they have not automatically turned RF into a technique immune to issues of data prevalence (Freeman *et al.*, 2012).

Consequently, to optimize the performance of RFs, we applied a resampling strategy favouring the presence class in the *R* package *randomForests* (Liaw and Wiener, 2002). The presence sample size (*i.e.* 98 cases) was used as the reference value, and all possible combinations between 5 % and 95 % of that value were stepwise tested (at 5 % intervals) for the presence and absence sample sizes. Thus, the number of presence cases outnumbered absences in several trials. The number of tree components was kept sufficiently large (5000) to ensure convergence, and the selected sample sizes were those that maximized the TSS in the out-of-bag dataset (Sn > Sp). The RF trained with the optimal sample sizes was selected for subsequent analyses.

2.3.4   Support Vector Machines – SVMs

Support Vector Machines (SVMs) (Vapnik, 1995) are a non-probabilistic classificatory technique based on the construction of discriminant hyperplanes typically adjusted with radial basis kernel functions (Huang and Wang, 2006). However, in addition of the Boolean output, SVMs are actually able to render probabilistic outputs through the implementations of the Platt's (2000) approach. SVMs produce very competitive results when compared with other benchmarking classification techniques by adjusting only few parameter settings (Fukuda *et al.*, 2013).

The SVM was developed in *R* (R Core Team, 2015) with the function *svm* implemented in the *e1071* package (Dimitriadou *et al.*, 2011). The parameters C and γ were optimized in a real value search using the aforementioned GA that maximized the TSS (Sn > Sp) in a 3 fold cross-validation scheme. The tested ranges were based on Huang and Wang (2006), who modelled a large range of different datasets, with C and γ ranging from 0 to 300 and from 0 to 10 respectively. In addition, to

15

improve generalization during the training of each SVM, 3 fold cross-validation was internally performed (*i.e.* in the end a 3×3 cross $-$ validation scheme was followed). Data prevalence also affects the SVMs' classification capability (Osuna *et al.*, 1997). Therefore, the training cases were weighted accordingly to data prevalence: 0.06 for the absence class and 0.94 for presence. In this case, the population size and the number of generations were set to 100; the GA's operators were set to 0.25, 0.75, 0.15, 0.10, 0.15, 0.75, 0.0, 0.35 and 0.0 respectively, disabling the use of derivatives in the searching process. Finally, once the optimal parameters were obtained, a SVM was trained with the optimal parameter setting, and the entire dataset and this model were used to develop the partial dependence plots and to perform the habitat assessment.

2.3.5 *Zero-order Takagi-Sugeno-Kang fuzzy rule base systems – TSK*

Fuzzy rule base systems (FRBSs) correspond to a specific part of fuzzy logic (Zadeh, 1965) that mimics human reasoning. Variables are discretized in fuzzy sets (*e.g.*, Low velocity, Medium velocity, High velocity etc.) and due to the fuzzy nature of these sets a given value may belong with different proportions to two adjacent sets. The degree of membership of a given case to these fuzzy set is mathematized by means of membership functions, usually of trapezoidal or triangular shape (Fukuda, 2013; e.g. Muñoz-Mas *et al.*, 2012). Additionally, the relationships between the different combinations of fuzzy sets are articulated in fuzzy rules (IF-THEN sequences) such as: IF velocity is Low and depth is High and substrate is High THEN the habitat suitability is High. Zero-order Takagi-Sugeno-Kang fuzzy rule base systems (TSK) (Takagi and Sugeno, 1985) are a special case within the TSKs where the consequents of the fuzzy rules correspond to singletons (*i.e.* real numbers).

The TSK was implemented in *R* (R Core Team, 2015) with the help of the *frbs* package (Riza *et al.*, 2015) and restricted the consequents to a dichotomous output: zero or one (*i.e.* absence or presence). Nevertheless, the TSKs provide probability-like outputs. The input variables were

discretized in three fuzzy sets with triangular membership functions. The vertices of these triangular fuzzy sets were located in concordance with variables' quantiles (*i.e.* minimum, median and maximum values) because it helps to define the system over the entire universe of discourse (Muñoz-Mas *et al.*, 2012). The fuzzy rules were optimized three times (*i.e.* a zero or a one was assigned to every consequent) by maximizing the TSS (Sn > Sp). The optimisation was performed with the hill-climbing algorithm (see Mouton *et al.*, 2008 for further details) and the corresponding consequent (*i.e.* 0 or 1) was assigned by rounding up the mean value obtained through the three iterations. Finally, a value of 0.5 was assigned to the untrained fuzzy rules.

The classificatory capability of the TSK was inspected by performing a 3 fold cross-validation and the ultimate consequents for the TSK were calculated by rounding up the mean value obtained in the nine iterations. Then, the partial dependence plots were developed and the habitat assessment was performed.

2.4 Hydraulic modelling

The hydraulic simulation was retrieved from Muñoz-Mas et al. (2012; 2014b). It encompassed an approximately 300 m long reach of the Cabriel River inserted within a catch and release fishing preserve that met different requirements such as abundance of large brown trout, habitat heterogeneity and representativeness. Furthermore, it proved to have a stable channel for more than a decade (Muñoz-Mas *et al.*, 2014b). Data collection and model development followed standard procedures (Jowett and Duncan, 2012). The topographic data of the river channel and banks were collected using a Leica© Total Station with an average area of *ca.* 2 m^2 per topographic measurement, and the substrate composition was collected co-ordinately as described above. The hydrometry was performed in 11 cross-sections, with depth and velocity measured along these sections and the resulting information used to gauge the flow rate. Measurements were performed at three different flow rates (0.54, 1.04 and 2.75 m^3/s) and were used to calibrate the model. The 2D

17

hydraulic simulation was performed with River2D$^©$ (University of Alberta, 2002) by altering the bed roughness to adjust the outcomes to the measurements obtained during the hydrometric campaigns. Thirty-four different flows were simulated, ranging from 0.05 to 6.5 m^3/s; in all cases the water level was below the bankfull stage of the river channel.

2.5   Hydrology and water temperature modelling under the climate change scenarios

2.5.1   *Rainfall-runoff and water temperature modelling – M5'*

For effective management of water and freshwater ecosystems, estimates of flow rate and water temperature at high temporal resolution, preferably on daily basis, are required (Van Vliet *et al.*, 2012). Process-based distributed models often render better performance, although they can be overly complicated requiring excessive data inputs, which may lead to over-fitting (Zhuo *et al.*, 2015). To cope with such limitations, the rainfall-runoff and the water temperature models were developed in *R* (R Core Team, 2015) with the M5' algorithm (Quinlan, 1992) implemented in the package *Cubist* (Kuhn *et al.*, 2014). This technique has previously been demonstrated to proficiently model daily flows (Solomatine and Dulal, 2003; Taghi Sattari *et al.*, 2013) and presents interesting capabilities. The M5' algorithm is encompassed in the group of recursive partition machine learning techniques (*e.g.,* CARTs or RF) that divide the data space into smaller subspaces (Hothorn *et al.*, 2006). However, M5' fits multi-linear regression models to the leaves of the tree instead of the average value provided by Breiman's *et al.* CARTs (1984). The complete tree is effectively a piecewise linear model resulting in a model that can be seen as a committee of linear models, with each member specialized on particular subsets of the input space (Taghi Sattari *et al.*, 2013). These linear models at the leaves confer M5' the capability for extrapolation, in contrast to some other machine learning techniques that demonstrated a poor or null capacity to do so (Hettiarachchi *et al.*, 2005). The M5' algorithm implemented in the *Cubist* package presents some ancillary functionalities that may improve generalization and accuracy. *Cubist* allows the development of M5' models'

18

ensembles by recursively boosting the subsequent models' trees (*i.e.* the subsequent models are trained more intensely on the data that presented larger error). Predictions can be further assisted with similar training observations (*i.e. neighbours*), which may reduce the prediction of unreliable flows or temperatures, and it allows the control of the number of leaves in the ultimate model (*i.e.* the number of models in which the input space is eventually divided).

Following previous studies, the rainfall-runoff model was trained using lagged data (rainfall and temperature) as input variables (Solomatine and Dulal, 2003; Taghi Sattari *et al.*, 2013). Three sets of different lagged data encompassing the period 1971-2000 (9,873 days in total) were calculated using, daily, monthly and quarterly lags. These three groups of variables were intended to reflect the causes of peak, normal and base flows. Daily variables included the rainfall and temperature from the current day to the 7th preceding day (eight variables in total). The monthly variables were calculated using the moving average for the 12 previous months (twelve variables in total) and the previous year was calculated from the moving average for the 13th to the 24th previous months in a quarterly time step (four variables in total). Consequently, only the daily variables overlapped with the current month variable. No variable selection was performed, and thus correlated variables were included in the training dataset (Solomatine personal communication).

The input data for the water temperature model encompassed daily data from two different periods – 2006-2008 and 2010-2011 (1,109 days in total) – that were measured at the study site with a HOBO Water Temp ProV2® (USA) data logger. Input variables were the air temperature and flow rate on the current day and air temperature on the previous day because they have been shown to develop accurate models (Santiago *et al.*, 2015).

In order to improve models' generalization, the number of allowed leaves (from 1 to 25 or let M5' to choose the optimal number), the number of M5' models in the committee (from one to ten) and the number of *neighbours* data assisting in the ultimate prediction (from 0 to 5) were optimized with the GA included in the package *rgenoud* (Mebane Jr and Sekhon, 2011). In both cases, the

19

optimization sought the maximization of the Nash-Sutcliffe Efficiency (NSE) index (feasible range from $-\infty$ to 1) in a 3 fold cross-validation scheme. The optimization sought among combinations of integers then, the population size and the number of generations were set to 10 and 25 respectively, and the GA's operators were set to 0.10, 0.75, 0.15, 0.10, 0.15, 0.75, 0.0, 0.35 and 0.0. Once the optimal parameters were obtained, single M5' models were developed to predict the flow rate and, subsequently, the water temperature under the historical and the climate change scenarios.

2.5.2 *Flow regime and water temperature under the short-term (2011-2040) climate change scenarios*

Two different climate change scenarios on the extremes of the available spectrum of Representative Concentration Pathways (RCPs) were used to feed the rainfall-runoff model and subsequently the temperature model: the RCP 4.5 (cautious) and RCP 8.5 (pessimistic). The RCP 4.5 is a scenario in which total radiative forcing is stabilized shortly after 2100 without ever exceeding that value (Thomson *et al.*, 2011). Conversely, the RCP 8.5 is characterized by increasing greenhouse gas emissions over time, thus it is representative of scenarios in the literature that lead to the highest greenhouse gas concentration levels (Riahi *et al.*, 2007).

The historical data were retrieved from Herrera *et al.* (2015), and the future rainfall and temperature time series (RCP 4.5 and RCP 8.5 scenarios) were obtained from simulations described in CORDEX (Christensen *et al.*, 2014; Jones *et al.*, 2011). Specifically, these simulations were based in the global circulation model CCCma-CanESM2 (regional circulation model RCA4) and the ICHEC-EC-EARTH (regional circulation model DMI-HIRHAM5). However, in order to correct any bias related with the climate models, the projected rainfall and temperature data were adjusted in *R* (R Core Team, 2015) following the quantile-quantile approach (*fitQmapRQUANT* function) implemented in the package *qmap* (Gudmundsson *et al.*, 2012). Once the predictions were adjusted to the historical time series, the two series per scenario (RCP 4.5 or 8.5) were used to feed the

20

rainfall runoff model and, subsequently, the temperature model (Table 2). Finally, the comparison was restricted to summertime (*i.e.* from July to September) to fit the biological data and the two series per scenario treated as single realizations.

Table 2 Mean monthly inter-annual rainfall (mm) and temperature (ºC) in the historical time series (1971-2000) and the Representative Concentration Pathways (RCP) 4.5 and 8.5 (2011-2040). The corresponding percentages of variation appear between brackets.

	Rainfall (mm)			Temperature (ºC)		
	Historical	RCP 4.5	RCP 8.5	Historical	RCP 4.5	RCP 8.5
Oct	2.64	1.83 (-30.7%)	1.81 (-31.4%)	5.19	7.04 (35.6%)	6.97 (34.3%)
Nov	3.11	1.49 (-52.1%)	1.68 (-46.0%)	6.60	7.57 (14.7%)	7.83 (18.6%)
Dec	2.1	1.79 (-14.8%)	1.6 (-23.8%)	9.35	9.31 (-0.4%)	9.71 (3.9%)
Jan	2.93	1.79 (-38.9%)	1.82 (-37.9%)	10.65	12.42 (16.6%)	12.51 (17.5%)
Feb	2.35	1.64 (-30.2%)	1.78 (-24.3%)	14.83	16.44 (10.9%)	15.84 (6.8%)
Mar	1.15	1.45 (26.1%)	1.29 (12.2%)	19.23	21.32 (10.9%)	21.18 (10.1%)
Apr	0.27	0.52 (92.6%)	0.69 (155.6%)	23.22	25.90 (11.5%)	25.65 (10.5%)
May	0.40	0.83 (107.5%)	1.02 (155.0%)	22.61	24.66 (9.1%)	24.60 (8.8%)
Jun	1.03	1.43 (38.8%)	1.65 (60.2%)	19.21	20.53 (6.9%)	20.19 (5.1%)
Jul	2.25	1.62 (-28.0%)	1.46 (-35.1%)	14.00	15.05 (7.5%)	15.04 (7.4%)
Aug	2.80	1.70 (-39.3%)	1.72 (-38.6%)	9.15	10.14 (10.8%)	10.32 (12.8%)
Sep	3.40	2.16 (-36.5%)	1.92 (-43.5%)	6.52	7.71 (18.3%)	7.92 (21.5%)

2.6 Models' selection and impact of climate change on large brown suitable habitat

The analysis of potential shifts in the suitable habitat available was performed by developing Bivariate Habitat Duration Curves (BHDCs), an extension of the traditional HDCs (Milhous *et al.*, 1990). In its original conception, the HDCs depict the probability of exceedance of a given quantity of habitat. Therefore, such curves are very useful to evaluate the frequency and duration of suitable habitat, allowing conservationists or managers to discern the long term impact of different scenarios and possible bottlenecks for aquatic organisms (Parasiewicz, 2008).

The BHDCs were built in a sequential procedure. First, for every simulated flow, a general indicator of the habitat suitability for the target species was calculated, all the tested techniques were able to render Boolean and probabilistic outputs, and consequently two different BHDCs were built. The first BHDC was based on the Weighted Usable Area (WUA) (Bovee *et al.*, 1998). The WUA is the most renowned general indicator of aquatic habitat quality and quantity and is calculated as the

sum of the suitability (*i.e. probability of presence*) calculated for every set of hydraulics (velocity and depth) and the substrate multiplied by the corresponding area. The second BHDC was based on the Boolean outputs and corresponded to the sum of the areas where the models predicted the presence of the large brown trout, hereafter Suitable Area (SA). These general indicators (WUA and SA) were then gathered to develop the corresponding habitat-flow curves (*i.e.* WUA-flow and SA-flow curves). The following step corresponded to the development of the Habitat Time Series (HTS), which are the interpolation of the different running flows (historical and RCP scenarios) provided by the M5' rainfall-runoff model upon the WUA-flow and SA-flow curves. Finally, the BHDCs were built by calculating the probability of exceedance of the different values of WUA and SA interpolated from the HTS (only values from July to September) and the corresponding water temperature. In addition, in order to illustrate the potential deviations of the water temperature, two different thresholds were considered throughout the analysis. The first one (18.7 ºC), encompasses the thermal niche and was inferred in the central part of the Iberian Peninsula (Santiago *et al.*, 2015), whereas the second one (25 ºC), approximately determines the physiological niche (threshold of incipient lethality) and was determined indoors (Elliott and Elliott, 2010).

Traditionally in ensemble modelling, no models' selection is usually performed, and thus all predictions are gathered in the ultimate forecast (Bouska *et al.*, 2015). However, the different nature of the considered techniques – some of them principally of classificatory nature (*e.g.*, SVM) and others principally addressed for regression (*e.g.*, MLPE) – made it necessary to check significant differences in the predictions. Therefore, prior to developing the BHDCs, the bivariate boxplots for the values of WUA and SA included in the HTS (July to September) were compared using the package *boxplotdbl* (Tomizono, 2013), and thus only those predictions that were similar were kept for subsequent analyses. The *boxplotdou* function considers that two distributions are similar when the square which is surrounded by the first and third quartile has an overlap.

3 Results

3.1 Habitat suitability modelling

3.1.1 Performance

In regards to the Boolean outputs, none of the considered techniques proved markedly superior to any other during the cross-validation phase. Support Vector Machines (SVMs) yielded the highest True Skill Statistic (TSS), whereas the zero-order Takagi-Sugeno-Kang fuzzy rule base system (TSK) yielded the lowest (Table 3). The TSK presented three untrained rules corresponding to the combinations of High velocity and High depth. The five techniques rendered varying Sensitivity (Sn) but always within the range from *ca.* 0.7 to *ca.* 0.8, whereas the range for Specificity (Sp) was narrower and, in any case, inferior to the Sn. The Multi-Layer Perceptron Ensemble (MLPE) optimization minimized the weighted Mean Squared Error (MSE) with eight MLPs, and thus it did not allow comparison. Conversely, the patterns on the performance of the ultimate models upon the complete dataset (*i.e.* the one without cross-validation) showed notable divergences. Random forests (RF) performed significantly better than any other technique, whereas the performance of the TSK was the worst. Regarding the probabilistic outputs, four out of five techniques rendered suitability for most of the feasible range (*i.e.* from zero to one), and only the SVM presented a significantly trimmed output range with a maximum value of 0.38.

Table 3 Summary of the Boolean-related performance criteria; True Skill Statistics (TSS), Sensitivity (Sn), and Specificity (Sp), and ranges of probabilistic outputs calculated during the training phase (cross-validation) and for the ultimate models.

	Cross-validation					Overall				
	TSS	*Sn*	*Sp*	Min	Max	TSS	Sn	Sp	Min	Max
GAM	0.45	0.80	0.64	0.00	1.00	0.48	0.86	0.62	0.00	1.00
MLPE	weighted Mean Squared Error = 0.18					0.51	0.81	0.70	0.01	0.91
RF	0.43	0.74	0.69	0.04	0.96	0.69	0.99	0.70	0.04	0.98
SVM	0.49	0.78	0.70	0.01	0.35	0.49	0.80	0.69	0.01	0.38
TSK	0.34	0.69	0.65	0.00	1.00	0.40	0.70	0.70	0.00	0.98

3.1.2 Partial dependence plots

The partial dependence plots reflected the internalities of every technique. The plots for TSK presented piecewise liner segments caused by the triangular membership functions and the plots for RF presented erratic patterns induced by the aggregation of the component random trees (Fig. 3). All of them coincided in assigning low suitability to low velocity, increasing up to *ca.* 0.5 m/s, but the TSK's achieved optimum suitability close to 1 m/s. For greater velocities, the techniques differed in the modelled suitability; SVM's and TSK's plots decreased whereas the other techniques persisted in high suitability. The partial dependence plots for depth showed a strong similarity across the techniques, but the SVM's showed an optimum plateau from 0.5 m to 1.5 m depth, after which suitability decreased. Conversely, the rest of them suggested a monotonic increase in habitat suitability. The partial dependence plots for substrate showed the optima around four (approximately fine gravel).
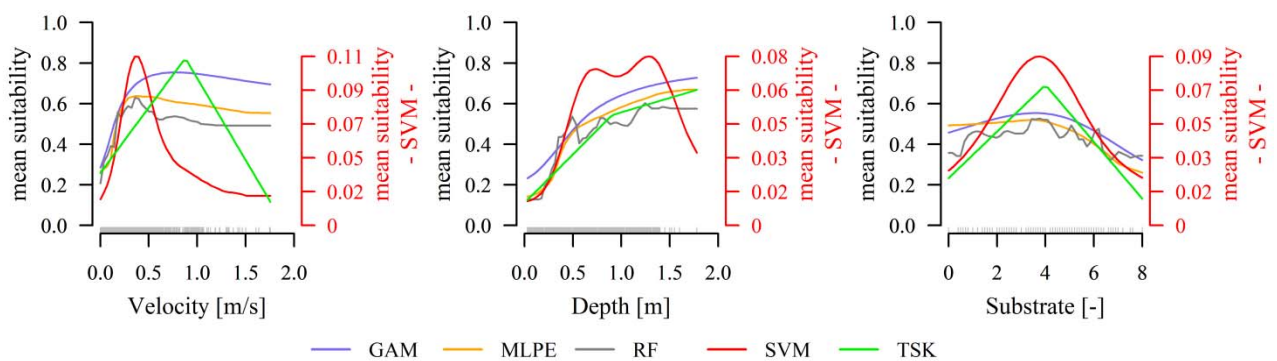

Fig. 3. Partial dependence plots calculated by means of the five techniques. The ones for the Support Vector Machine were depicted in a different scale because it rendered markedly inferior values. The corresponding axes appear depicted in the right side.

24

3.2 Hydrologic modelling

3.2.1 Rainfall-runoff and water temperature modelling – M5'

The M5' rainfall-runoff model optimization achieved the NSE of 0.78±0.14 with only nine multi-linear models (*i.e. leaves*), using a committee of ten M5' models and assisting the predictions with two neighbouring cases (Fig. 4).
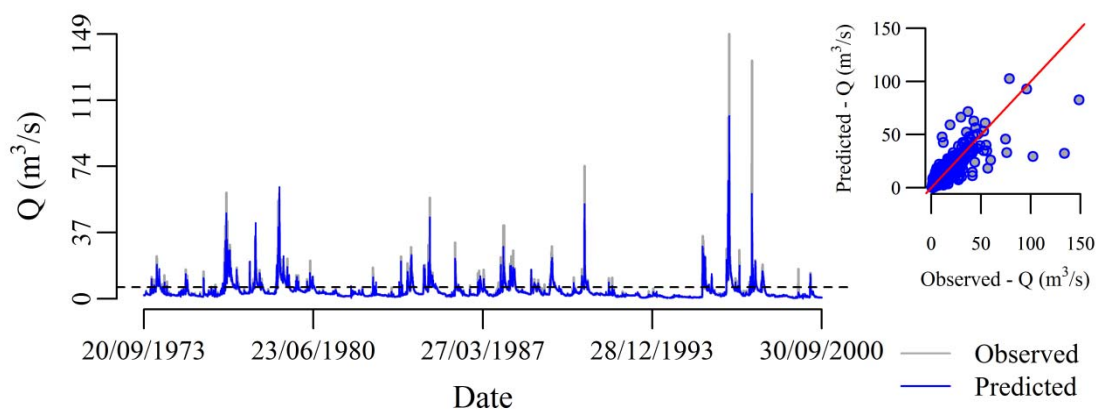

Fig. 4 Rainfall-runoff model performance (observed *vs.* predicted plot). Black dashed line depicts the highest simulated flow (6.5 m³/s).

The M5' water temperature model achieved almost perfect accuracy (NSE of 0.97 ± 0.00) for the three complete cycles (summer to spring) encompassed in the training data (Fig. 5). The ultimate model consisted of a committee of six members with 32 rules, whereas the predictions were assisted by nine neighbouring cases.
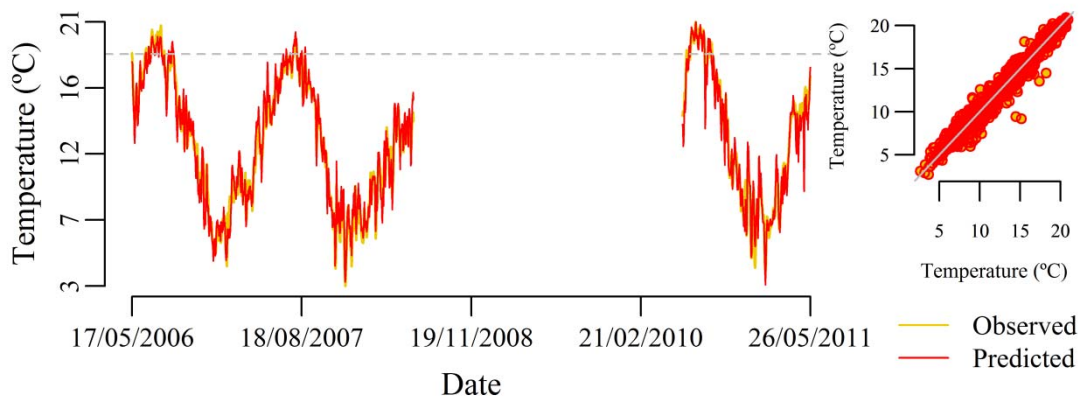
Fig. 5 Water temperature model performance (observed *vs.* predicted plot). Grey dashed line depicts the limit of the thermal niche (18.7 ºC).

3.2.2 *Summertime flow regime under the short-term (2011-2040) climate change scenarios*

The simulated daily flows for summertime (July to September) suggested a decrement of *ca.* -29 % for the RCP 4.5 and of *ca.* -20 % for the RCP 8.5, although the maximum flow predicted for the RCP 8.5 was larger than the maximum observed in the historical flow series (Table 4).**Error! Reference source not found.** For the historical time series, the RCP 4.5 and the RCP 8.5 scenarios, 2, 0.5 and 1 % of the predicted flows were beyond the largest simulated flow (*i.e.* 6.5 m^3/s) (Fig. 6). The two scenarios predicted the complete depletion of river flow for several non-consecutive days (Table 4). The complete flow and water temperature series appear depicted in the Appendix A.
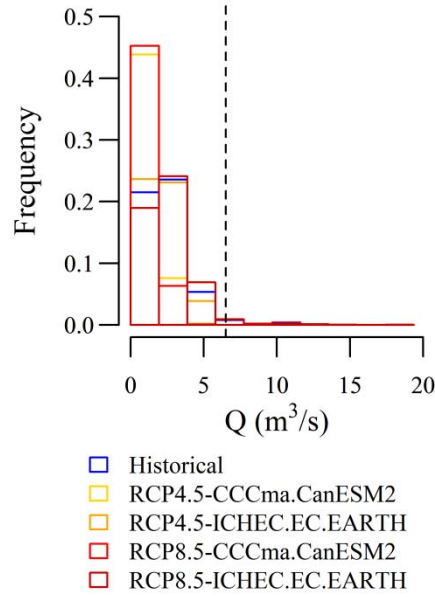
Fig. 6. Histogram of the summer daily flows for the historical time series and the Representative Concentration Pathways (RCP) 4.5 and 8.5. For the historical, the RCP 4.5 and the RCP 8.5 scenarios, 2, 0.5 and 1 % of the predicted flows were beyond the largest simulated flow (*i.e.* 6.5 m^3/s) – dashed line.

During summertime, the water temperature model predicted increases of *ca.* +4.0% for the RCP 4.5 and *ca.* +4.2% for the RCP 8.5. These predictions overtook the thermal niche threshold (18.7 ºC) more frequently than the historical temperature series, and even surpassed the physiological threshold (25 ºC) (Table 4).

Table 4 Principal statistics of the summer daily flows and water temperature for the historical time series and the Representative Concentration Pathways (RCP) 4.5 and 8.5 scenarios. Between brackets the percentages of variation.

	Flow (m^3/s)			Water temperature (ºC)		
	Historical	RCP 4.5	RCP 8.5	Historical	RCP 4.5	RCP 8.5
Min.	0.46	0.00 (-100.0%)	0.00 (-100.0%)	15.44	13.57 (-12.1%)	12.52 (-18.9%)
1st Qu.	1.52	0.98 (-35.5%)	0.97 (-36.2%)	18.44	18.80 (2.0%)	18.77 (1.8%)
Median	2.17	1.48 (-31.8%)	1.61 (-25.8%)	19.07	19.87 (4.2%)	19.77 (3.7%)
Mean	2.44	1.74 (-28.7%)	1.94 (-20.5%)	19.11	19.90 (4.1%)	19.81 (3.7%)
3rd Qu.	3.10	2.27 (-27.0%)	2.58 (-17.0%)	19.80	21.07 (6.4%)	20.95 (5.8%)
Max.	16.52	8.66 (-48.6%)	19.36 (14.9%)	24.67	25.50 (3.4%)	26.14 (6.0%)

3.3 Shifts in the suitable habitat for the large brown trout

3.3.1 Habitat-flow curves

The Weighted Usable Area (WUA) versus flow curves inferred with the five different techniques broadly coincided, except for SVM (Fig. 7). The WUA-flow curves predicted a parabolic increase of WUA in accordance with the flow rate, whereas the SVM's presented a decrease beyond *ca.* 2 m^3/s. The Suitable Area (SA) flow curves also exhibited a similar pattern. Therefore, nearly all techniques predicted a parabolic increase in SA with the flow rate, whereas the SVM predicted a significant decrease in SA for higher flows (*i.e.* beyond *ca.* 2 m^3/s). To assign a value of WUA and SA beyond the highest simulated flow (6.5 m^3/s), the curves were extrapolated using the *smooth.spline* function in *R* (R Core Team, 2015).
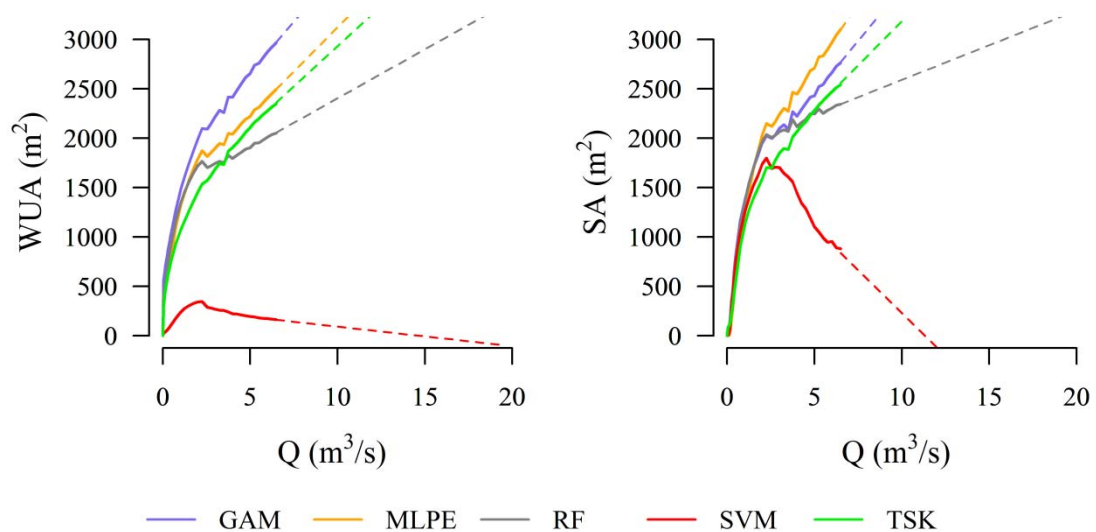

Fig. 7. Weighted Usable Area (WUA) *vs.* flow and Suitable Area (SA) *vs.* flow curves inferred by means of the five machine learning techniques. Dashed segments correspond to the extrapolation performed to assess the highest predicted flows (*i.e.* beyond 6.5 m^3/s).

3.3.2 Models' selection

The WUA values obtained (interpolated) for the SVM clearly differed from those provided by the other techniques (Fig. 8). The *boxplotdou* function did not aggregate this distribution,

although it showed overlap in the y-axis (SA), and consequently, it remained in subsequent analysis based on the SA.
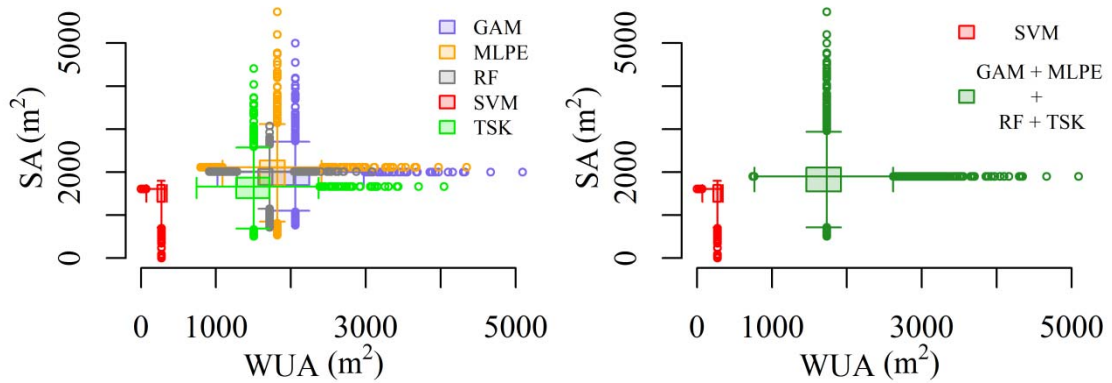

Fig. 8. Bivariate boxplots depicting the similitude between the interpolated Weighted Usable Area (WUA) and Suitable Area (SA) inferred using the five machine learning techniques. The depicted values are restricted to summertime (July to September).

3.3.3 Bivariate Habitat Duration Curves – BHDCs

The scatterplot of the inferred WUA and the corresponding water temperatures suggested that the two scenarios (RCP 4.5 and 8.5) may exert similar alteration to the available suitable habitat (Fig. 9). Both scenarios predicted an increment in the number of days with smaller WUA and/or with higher temperature. The RCP 4.5 scenario predicted a slightly larger increase in temperature whereas the RCP 8.5 predicted an increase in the frequency of lower WUA values. The Bivariate Habitat Duration Curves (BHDCs) predicted worse habitat conditions during 86 % and 79 % of the days for the RCP 4.5 and 8.5 respectively. The rainfall-runoff predicted the complete depletion of the river for several isolated days, and consequently the study site would be unsuitable in such conditions.
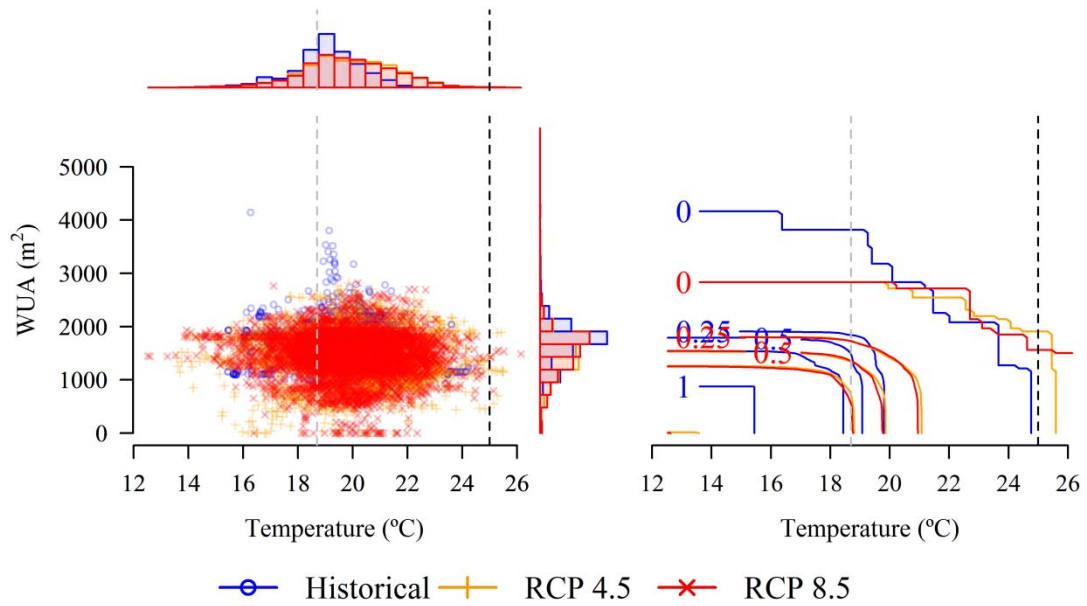
29

Fig. 9. Scatterplot of Weighted Usable Area (WUA) *vs.* water temperature (left) and Bivariate Habitat Duration Curves (BHDCs) inferred for the historical and the Representative Concentration Pathways (RCP) 4.5 and 8.5 (right). Iso-lines on the right plot represent the probability of exceedance. Grey dashed line depicts the limit of the thermal niche (18.7 ºC) and the black one the physiological niche (25 ºC).

The scatterplot of SA values and water temperature suggested that the two scenarios (RCP 4.5 and 8.5) may also exert similar alteration (increasing frequencies of low SA and high temperature) (Fig. 10), although the predicted SA values were smaller than their WUA counterparts. The RCP 4.5 predicted a slightly larger increase in temperature and the RCP 8.5 predicted lower values of SA. Consequently, the BHDCs predicted worse habitat conditions during 85 % (RCP 4.5) and 77 % (RCP 8.5) of the days on the basis of lower SA values and/or higher water temperature. In accordance with the Boolean nature of the SA, the study site was predicted to be more frequently unsuitable even without the complete depletion of river.

Under the historical conditions, the study site experienced extended periods of time with the water temperature above the thermal niche threshold (18.7 ºC), but it never surpassed the physiological niche threshold (25 ºC). Conversely, the predicted water temperature in the RCP scenarios would extend the period beyond the threshold of the thermal niche and even surpass the physiological niche threshold (threshold of incipient lethality).
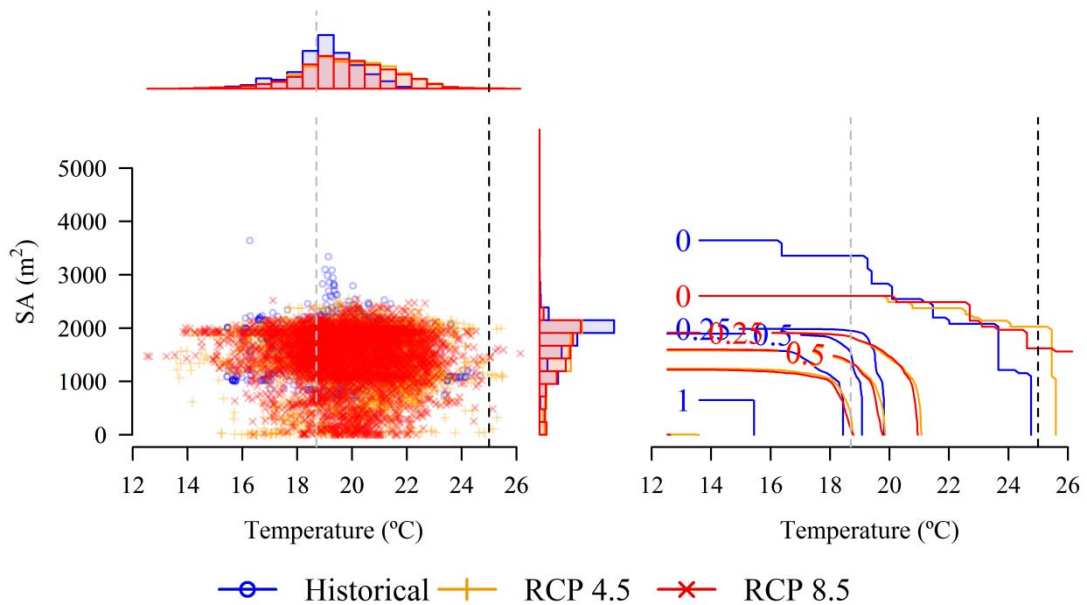
30

Fig. 10. Scatterplot of Suitable Area (SA) *vs.* water temperature (left) and Bivariate Habitat Duration Curves (BHDCs) inferred for the historical and the Representative Concentration Pathways (RCP) 4.5 and 8.5 (right). Iso-lines on the right plot represent the probability of exceedance. Grey dashed line depicts the limit of the thermal niche (18.7 ºC) and the black one the physiological niche (25 ºC).

4 Discussion

4.1 Habitat suitability models' reliability – Performance and ecological relevance

4.1.1 Performance

This study developed five regularized habitat suitability models that performed satisfactorily, and their predictive capability did not show significant differences. However, like in previous studies, Random Forests (RF) proved to be the most accurate approach and Fuzzy Rule Base Systems (FRBSs) – specifically the zero-order presence-absence Takagi-Sugeno-Kang FRBSs (TSK) – the least (Bouska *et al.*, 2015; Fukuda *et al.*, 2013).

There is a controversy about the generalization capability of RF therefore, it is possible to encounter positive (Bouska *et al.*, 2015; Vezza *et al.*, 2015) and negative (Wenger and Olden, 2012) evidences. The number of allowed splits and the number of cases in the terminal nodes can currently be constrained (Liaw and Wiener, 2002); however, our experience suggested that, due to the random

31

nature of RF training, such restrictions become meaningless after the aggregation of the individual predictions. The volatility of RF causes of the erratic patterns depicted in the partial dependence plots that disagree with the precepts of the ecological gradient theory (Austin, 2007). Consequently from our personal viewpoint, and despite the higher predictive accuracy, such a technique should not automatically relegate the other alternatives.

FRBSs underperformance is common (e.g. Fukuda *et al.*, 2013); however, they are not necessarily outperformed by any other technique (Mouton *et al.*, 2011). FRBSs such as the TSKs are potentially interpretable universal approximators (Castro, 1995; Fukuda *et al.*, 2011b), but increasing accuracy typically relies on either increasing the number of fuzzy sets and/or in the modification of their membership functions, although such modifications tend to be detrimental to their interpretability (Fukuda *et al.*, 2011b). Consequently, none of these modifications has been popularized, the number of fuzzy sets and membership functions are usually set beforehand, and the ultimate performance is based on that single trial (Fukuda *et al.*, 2013; e.g. Mouton *et al.*, 2009; Mouton *et al.*, 2011; Muñoz-Mas *et al.*, 2012). Nevertheless, a model's performance strongly depends on the problem at hand, and therefore testing the capabilities of FRBSs would be always worthwhile, especially in exploratory analysis for unknown species.

The Generalized Additive Model (GAM) and the Multi-Layer Perceptron Ensemble (MLPE) posed no major concern because the actual implementation of the packages – *mgcv* and *monmlp* – allow fully control of the models' complexity either by controlling the number of knots or the number of nodes. Their intermediate accuracy was in line with the larger regularization which rendered partial dependence plots ecologically sound. Setting aside the irregularities of RF', the GAM's and MLPE's partial dependence plots fitted well the RF' ones – with very little regularization – so they are likely to reflect the underlying relationships.

Finally, the Support Vector Machine (SVM) presented the largest discrepancies. The probabilistic output of the SVM was strongly affected by class overlapping and low prevalence, and Platt's (2000)

32

approach was unable to render comparable outputs. The use of class weights modifies the discriminant threshold, and consequently the Boolean outputs did not present such marked differences. Accordingly, the use of classificatory SVM in studies where a gradual (*i.e.* probabilistic) output is desired – for instance, to calculate Weighted Usable Area (WUA) – should be discouraged unless the different classes present an easy discrimination.

4.1.2 *Ecological relevance*

Regardless of the previous considerations, from an ecological viewpoint SVM rendered trustworthy results because several authors have suggested a pronounced decrease in habitat suitability beyond a water velocity of *ca.* 0.5 m/s (Ayllón *et al.*, 2010; Jowett and Davey, 2007). That decrease is certainly reasonable because of the increased energy costs of standing higher velocities (Rincón and Lobón-Cerviá, 1993), and thus the patterns of the SVM's and TSK's partial dependence plots were considered credible. In contrast, the remaining techniques suggested high habitat suitability beyond 0.5 m/s. This pattern was previously discussed by Muñoz-Mas *et al.* (2014b), and it was hypothesized that the higher water temperature was playing an important role by enhancing the natatorial capacity of brown trout. It could be argued that the strong regularization handicapped the proper modelling by tightening the partial dependence plots. However, RF lacks of effective regularization and the corresponding partial dependence plot matched the ones for the GAM and the MLPE. Thus, the potential influence of the strong regularization was ruled out. Data on high velocities were scarce within the training dataset, and it was considered that using a models' ensemble reduced the ultimate suitability assigned to high velocity, rendering predictions more consistent with the ecological knowledge about the species.

In natural rivers, large brown trout has been reported to prefer pools (Armstrong *et al.*, 2003; Ayllón *et al.*, 2010), and consequently several authors suggested a monotonic increase in suitability related to depth (Ayllón *et al.*, 2010). Others have suggested either a monotonic increase or slight

decrease beyond 3 m deep on the basis of several changes on the models' settings (Jowett and Davey, 2007). Such patterns perfectly fit the ones depicted in the depth partial dependence plots, further supporting our modelled preferences.

Finally, large brown trout has been observed occupying preferably medium-to-coarse substrates (Armstrong *et al.*, 2003; Ayllón *et al.*, 2010), which agrees with the patterns depicted in the partial dependence plots for substrate.

In conclusion, the models' ensemble in this study fit well with previous knowledge about brown trout. Moreover, the differences between the partial dependence plots and the consequent differences in the habitat suitability assessment would be attenuated by assembling the different predictions, highlighting the benefits of using models' ensembles and increasing the reliability of the predicted subsequent shifts in suitable habitat.

2.5 Flow regime alteration under the climate change scenarios

The optimal rainfall-runoff and water temperature model were built with a relatively low number of multi-linear models (*i.e.* members in the committee and leaves in the tree components) and therefore presented good generalization. The Nash-Sutcliffe Efficiency (NSE) of the rainfall-runoff model (0.78) was superior than the median value ($\widetilde{NSE}$: 0.70) of several approaches reported by Zhuo *et al.* (2015), whereas the water temperature model posed no major concern since the NSE was close to the maximum (NSE: 0.96).

Regardless of the cautious or pessimistic nature of the Representative Concentration Pathways (RCPs), both proposed similar decreases in flow rate as well as increases in water temperature. The RCPs scenarios are based on the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC, 2014). Therefore in accordance with Rogelj *et al.* (2012), the RCP 8.5 would better fit the A1FI and the RCP 4.5 the B1 scenarios of the Special Report on Emissions Scenarios (SRESs) (Nakicenovic *et al.*, 2000). In the Iberian Peninsula, studies focused on the SRESs have been focused

principally in the A2 and B2 scenarios (Chirivella Osma *et al.*, 2014; Estrela *et al.*, 2012). As a consequence, comparisons should be based on general trends rather than specific values.

Our results suggest similar decreases in the mean value of the running flow *ca.* -25 % (RCP 4.5: -28.7 % and RCP 8.5; -20.5 %) than the year-round value predicted for the entire Jucar River basin using the B2 scenarios of the SRESs (*ca.* -20 %), but disagree with the predictions based on the A2, which suggested an increase (*ca.* +25 %) (Chirivella Osma *et al.*, 2014). Unlike the SRESs that predicted divergent patterns of changes in rainfall (with increases in the A2 and decreases in the B2), the RCP scenarios broadly coincided with the predicted changes in rainfall and temperature. Thus, collectively, they should be viewed as the extension of the decreases in water yield observed in the Cabriel River basin, which has suffered a significant reduction in mean flow since 1980 (Estrela *et al.*, 2012). Historically, the strong influence of the Mediterranean coastal convective phenomena has been poorly reproduced by the general circulation models probably causing the divergent predictions (Chirivella Osma *et al.*, 2014). In this regard, our results would corroborate Santiago *et al.* (2015) statement that the use of the latest scenarios established for climate change studies represent a refinement over previous scenarios because they jointly predict a decrement of the flow rate.

Though performed in alpine streams, the predicted decrease in running flows broadly coincided with those predicted in several studies on the topic performed abroad (e.g. Hauer *et al.*, 2013; Papadaki *et al.*, 2016; Viganò *et al.*, 2015). Nevertheless, in a short time span, the RCP scenarios are going to be likewise superseded by the new Shared Socioeconomic Pathways (SSPs) (Kriegler *et al.*, 2012). The new SSPs consist of a framework for climate change research that combines the former RCPs with alternative pathways of socioeconomic development (O'Neill *et al.*, 2014). Consequently, the results presented herein could be also superseded in the near future.

4.2 Shifts in the suitable habitat for the large brown trout

The relationship between climate change, river flows and salmonid's survival has been previously suggested (Tisseuil *et al.*, 2012). However, in the Iberian context, the only available study that considered flow rate predicted stable carrying capacity and concluded that only temperature will affect the species distribution (Almodóvar *et al.*, 2012). Conversely, our results highlighted the crucial impact of flow alteration, since we predicted unsuitable condition and even the complete depletion of the river segment. Our results are in line with some of the aforementioned studies performed in mountainous areas (Hauer *et al.*, 2013; Papadaki *et al.*, 2016) where significant reductions in flow rate were accompanied by pronounced reductions in the suitable habitat available. Channel morphology plays a fundamental role in the impact of flow decrease (Turner and Stewardson, 2014), and therefore others also predicted stable carrying capacity during summer (Viganò *et al.*, 2015). Furthermore, the mesomediterranean region of the Iberian Peninsula is likely to experience a significantly larger reduction in water yield (Chirivella Osma *et al.*, 2014; Salmoral *et al.*, 2015) in comparison to those areas described by Almodóvar *et al* (2012), highlighting the need for detailed studies to combine the effect of climate change (*i.e.* flow rate) with river morphology.

The WUA-based assessment rendered slightly more optimistic predictions than the SA-based one. This pattern was caused by the mean nature of the WUA indicator, since every area with a WUA value greater than zero adds a small quantity to the final summed prediction. Conversely, the Boolean output used in the calculation of the SA rules out any area with suitability inferior to 0.5 – independently of the suitability – and thus, lower values appear in the corresponding Bivariate Habitat Duration Curve (BHDC). In this regard, the WUA has been criticized because a large area of relatively unsuitable conditions can render similar values than one smaller but largely suitable area, and thus, in some cases, is an ambiguous indicator of overall suitability. It is indubitable that more gradual predictions are significantly more informative that Boolean predictions, and thus resulting in the typical preference for abundance-related indices over dichotomous ones (Fukuda *et al.*, 2011a).

36

However, these differences were not as relevant in this study, and the ultimate shifts in the suitable habitat available should be between both BHDCs.

Fluctuations in relative abundances have been significantly correlated with variables reflecting the severity of summer droughts (*e.g.* minimum flow) (Magalhães *et al.*, 2007). Furthermore, it has been demonstrated that water temperature affects the feeding capability of brown trout (Elliott and Elliott, 2010). Therefore it can be asserted that large brown trout is going to experience an increasing number of days (*i.e. ca.* 82 %) were the suitable habitat – physical or, specifically, thermal – is going to be noticeably impoverished. Furthermore, on the basis of the predicted flow depletion, the species could be even extirpated from the study site. Brown trout has been successfully introduced due to its ecological flexibility and has a worldwide distribution (Klemetsen *et al.*, 2003), and thus, in regard to its applied thermal thresholds, we considered them only a general guideline for the extant population of brown trout (Muñoz-Mas *et al.*, 2014b), as the study site typically has a water temperature above the threshold of the thermal niche (*i.e.* 18.7 ºC). Moreover, the threshold of the physiological niche (threshold of incipient lethality) depends on the dissolved oxygen – neglected in this study – and the specimens' tolerance. Consequently, we could not suggest the species extirpation solely based on the increments in water temperature (mean increment of ca. +4 %). Previous studies may have different results by considering specific thermal thresholds (Almodóvar *et al.*, 2012) or by considering climate-induced changes in the flow regime (Santiago *et al.*, 2015).

A recommendable method to enhance the resistance and resiliency of brown trout is to undertake mitigation and restoration measures focused on creating or facilitating access to thermal and hydraulic refugia (Almodóvar *et al.*, 2012). Brown trout could search for refugia in deeper pools, as we confirmed their existence in previous studies (e.g. Vezza *et al.*, 2015), if naturally permeable barriers do not become impassable due to reduced flows. However, in the Cabriel River, thermal refugia with the sufficient capacity to harbour an extant population are unlikely to be available because they should be located in the headwaters which may suffer the largest effect of decreased

flows (Tisseuil *et al.*, 2012). In addition, brown trout cohabits with several cyprinid species (Costa *et al.*, 2012; Vezza *et al.*, 2015) that may compete with brown trout for space and trophic resources (Santiago *et al.*, 2015). Besides, brown trout is a territorial fish (Ayllón *et al.*, 2010) thus such a suite of inter- and intra-species competition for waning resources is likely to be fierce. Therefore, we can also assert that climate change, including global warming, could be particularly dramatic in the basins of the Iberian Peninsula flowing to the Mediterranean Sea due to the synergistic effects of warming and flow reduction and the consequent increase in the competition for resources (Santiago *et al.*, 2015).

In accordance with the limited capacity to manipulate the Cabriel River suggested by other authors (Palmer *et al.*, 2008), we would appeal to increase the grade of protection of the Cabriel River basin and the frequency and extension of monitoring programmes to properly anticipate ecosystem bottlenecks. We acknowledge that the negative effect of climate change is unlikely to be supressed by such actions, although intensified surveillance should allow the proper implementation of adequate management actions (*e.g.*, catch and release or re-stocking) to reduce the predicted impact of climate change in the Cabriel River.

5   Conclusions

The shifts in the suitable habitat available for the large brown trout (body length > 20 cm) under two short-term (2011-2040) climate change scenarios – Representative Concentration Pathways (RCPs) 4.5 and 8.5 – was studied in a segment of the Cabriel River following the physical habitat simulation approach. The use of a model's ensemble of five different machine learning techniques (generalized additive models, multilayer perceptron ensembles, fuzzy rule base systems, random forests and support vector machines) reduced discrepancies in modelled habitat preferences. The ensemble had a superior fit to the prior knowledge of brown trout habitat preferences – selecting relatively high flow velocity, large depth and coarse substrate – and increased the reliability of the

predicted climate-change shifts in suitable habitat. Support Vector Machines (SVMs) were strongly affected by class overlapping and low prevalence, and thus the probabilistic output was deemed inappropriate, and banned in the ensemble, to calculate the Weighted Usable Area (WUA). The flow rate and water temperature were modelled with M5' models trees achieving excellent performance (NSE of 0.78 and 0.97 respectively). M5' predicted similar alteration for both RPCs scenarios (4.5 and 8.5), mean summertime decrease in flow rate of *ca.* - 25 % and increase in water temperature of *ca.* +4 %. Furthermore the Bivariate Habitat Duration Curves (BHDCs) based on the WUA and the Suitable Area (SA) broadly matched, indicating an increase in the number of days with less suitable habitat and/or with higher water temperature. Large brown trout are predicted to endure worse habitat conditions *ca.* 82 % of the days; however, the trout in the Cabriel proved to tolerate higher temperatures than other extant populations in the Iberian Peninsula. Finally, in accordance with the forecasted depletion of river flow, we predicted a potential extirpation of the species from the study site during short time spans. This study represents a valuable example of the ultimate impacts of climate change in a cold-water species such as brown trout. Consequently, we would increases the grade of protection for the Cabriel River and the frequency and extent of the monitoring programs to anticipate ecosystem bottlenecks.

Acknowledgments

39

References

Almodóvar, A., Nicola, G.G., Ayllón, D. and Elvira, B., 2012. Global warming threatens the persistence of Mediterranean brown trout. Global Change Biol. 18 (5), 1549–1560. http://dx.doi.org/10.1111/j.1365-2486.2011.02608.x.

Armstrong, J.D., Kemp, P.S., Kennedy, G.J.A., Ladle, M. and Milner, N.J., 2003. Habitat requirements of Atlantic salmon and brown trout in rivers and streams. Fish. Res. 62 (2), 143-170. http://dx.doi.org/10.1016/S0165-7836(02)00160-1.

Austin, M., 2007. Species distribution models and ecological theory: A critical assessment and some possible new approaches. Ecol. Model. 200 (1-2), 1-19. http://dx.doi.org/10.1016/j.ecolmodel.2006.07.005.

Ayllón, D., Almodóvar, A., Nicola, G.G. and Elvira, B., 2010. Ontogenetic and spatial variations in brown trout habitat selection. Ecol. Freshw. Fish. 19 (3), 420–432. http://dx.doi.org/10.1111/j.1600-0633.2010.00426.x.

Bossard, M., Feranec, J. and Otahel, J., 2000. CORINE land cover technical guide - Addendum 2000 - 40, pp. 105.

Bouska, K.L., Whitledge, G.W. and Lant, C., 2015. Development and evaluation of species distribution models for fourteen native central U.S. fish species. Hydrobiologia 747 (1), 159-176. http://dx.doi.org/10.1007/s10750-014-2134-8.

Bovee, K.D., Lamb, B.L., Bartholow, J.M., Stalnaker, C.B., Taylor, J. and Henriksen, J., 1998. Stream habitat analysis using the instream flow incremental methodology Geological Survey - Information and Technology Report 1998-0004, Fort Collins, CO (USA), pp. 130.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5-32. http://dx.doi.org/10.1023/A:1010933404324.

Breiman, L., Friedman, J., Stone, C.J. and Olshen, R.A., 1984. Classification and Regression Trees. Chapman & Hall, 368 pp.

Brosse, S., Laffaille, P., Gabas, S. and Lek, S., 2001. Is scuba sampling a relevant method to study fish microhabitat in lakes? Examples and comparisons for three European species. Ecol. Freshw. Fish 10 (3), 138-146. http://dx.doi.org/10.1034/j.1600-0633.2001.100303.x.

Cannon, A.J., 2012. monmlp: Monotone multi-layer perceptron neural network. R package version 1.1.2.

Castro, J.L., 1995. Fuzzy logic controllers are universal approximators. IEEE Trans. Syst. Man Cybern. 25 (4), 629-635. http://dx.doi.org/10.1109/21.370193.

Chirivella Osma, V., Capilla Romá, J.E. and Pérez Martín, M.A., 2014. Modelling regional impacts of climate change on water resources: the Júcar basin, Spain. Hydrol. Sci. J. 60 (1), 30-49. http://dx.doi.org/10.1080/02626667.2013.866711.

40

Christensen, O., Gutowski, W., Nikulin, G. and Legutke, S., 2014. CORDEX Archive Design Version 3.1, March 2014, pp. 21.

Conallin, J., Boegh, E. and Jensen, J.K., 2010. Instream physical habitat modelling types: An analysis as stream hydromorphological modelling tools for EU water resource managers. Int. J. River Basin Manage. 8 (1), 93-107. http://dx.doi.org/10.1080/15715121003715123.

Costa, R.M.S., Martínez-Capel, F., Muñoz-Mas, R., Alcaraz-Hernández, J.D. and 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 Res. Appl. 28 (6), 740-752. http://dx.doi.org/10.1002/rra.1598.

Dimitriadou, E., Hornik, K., Leisch, F., Meyer, D. and Weingessel, D., 2011. e1071: Misc Functions of the Department of Statistics (e1071), TU Wien (Austria). R Package Version 1.5-25.

Döll, P. and Schmied, H.M., 2012. How is the impact of climate change on river flow regimes related to the impact on mean annual runoff? A global-scale analysis. Environ. Res. Lett. 7 (1), http://dx.doi.org/10.1088/1748-9326/7/1/014037.

Dudgeon, D., 2014. Threats to freshwater biodiversity in a changing world. In: Anonymous Global Environmental Change. Springer, pp. 243-253.

Elliott, J.M. and Elliott, J.A., 2010. Temperature requirements of Atlantic salmon Salmo salar, brown trout Salmo trutta and Arctic charr Salvelinus alpinus: Predicting the effects of climate change. J. Fish Biol. 77 (8), 1793-1817. http://dx.doi.org/10.1111/j.1095-8649.2010.02762.x.

Estrela, T., Pérez-Martin, M. and Vargas, E., 2012. Impacts of climate change on water resources in Spain. Hydrol. Sci. J. 57 (6), 1154-1167. http://dx.doi.org/10.1080/02626667.2012.702213.

Ficke, A.D., Myrick, C.A. and Hansen, L.J., 2007. Potential impacts of global climate change on freshwater fisheries. Rev. Fish Biol. Fish. 17 (4), 581-613. http://dx.doi.org/10.1007/s11160-007-9059-5.

Filipe, A.F., Lawrence, J.E. and Bonada, N., 2013. Vulnerability of stream biota to climate change in mediterranean climate regions: A synthesis of ecological responses and conservation challenges. Hydrobiologia 719 (1), 331-351. http://dx.doi.org/10.1007/s10750-012-1244-4.

Freeman, E.A., Moisen, G.G. and Frescino, T.S., 2012. Evaluating effectiveness of down-sampling for stratified designs and unbalanced prevalence in Random Forest models of tree species distributions in Nevada. Ecol. Model. 233 1-10. http://dx.doi.org/10.1016/j.ecolmodel.2012.03.007.

Friedman, J.H., 2001. Greedy function approximation: A gradient boosting machine. Ann. Stat. 29 (5), 1189-1232. http://dx.doi.org/10.1214/aos/1013203451.

Fukuda, S., 2013. Effects of data prevalence on species distribution modelling using a genetic takagi-sugeno fuzzy system. IEEE International Workshop on Genetic and Evolutionary Fuzzy Systems (GEFS), Singapore, 21-27.

Fukuda, S., De Baets, B., Waegeman, W., Verwaeren, J. and Mouton, A.M., 2013. Habitat prediction and knowledge extraction for spawning European grayling (*Thymallus thymallus* L.) using

a broad range of species distribution models. Environ. Model. Softw. 47 1-6. http://dx.doi.org/10.1016/j.envsoft.2013.04.005.

Fukuda, S., Mouton, A.M. and De Baets, B., 2011a. Abundance versus presence/absence data for modelling fish habitat preference with a genetic Takagi-Sugeno fuzzy system. Environ. Monit. Assess. 184 (10), 6159–6171. http://dx.doi.org/10.1007/s10661-011-2410-2.

Fukuda, S., Nakajima, J., De Baets, B., Waegeman, W., Mukai, T., Mouton, A.M. and Onikura, N., 2011b. A discussion on the accuracy-complexity relationship in modelling fish habitat preference using genetic Takagi-Sugeno fuzzy systems. Genetic and Evolutionary Fuzzy Systems (GEFS), 2011 IEEE 5th International Workshop on, Paris (France), 81-86.

Fukuda, S., Tanakura, T., Hiramatsu, K. and Harada, M., 2014. Assessment of spatial habitat heterogeneity by coupling data-driven habitat suitability models with a 2D hydrodynamic model in small-scale streams. Ecol. Informatics http://dx.doi.org/10.1016/j.ecoinf.2014.10.003.

Gudmundsson, L., Bremnes, J.B., Haugen, J.E. and Engen-Skaugen, T., 2012. Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations – A comparison of methods. Hydrol. Earth Syst. Sci. 16 (9), 3383-3390. http://dx.doi.org/10.5194/hess-16-3383-2012.

Hansen, L.K. and Salamon, P., 1990. Neural network ensembles. IEEE T. Pattern Anal. 12 (10), 993-1001. http://dx.doi.org/10.1109/34.58871.

Hastie, T.J. and Tibshirani, R.J., 1990. Generalized Additive Models. Chapman & Hall/CRC, London, (UK), 352 pp.

Hauer, C., Unfer, G., Holzmann, H., Schmutz, S. and Habersack, H., 2013. The impact of discharge change on physical instream habitats and its response to river morphology. Clim. Change 116 (3-4), 827-850. http://dx.doi.org/10.1007/s10584-012-0507-4.

Heggenes, J., Brabrand, Å and Saltveit, S., 1990. Comparison of Three Methods for Studies of Stream Habitat Use by Young Brown Trout and Atlantic Salmon. T. Am. Fish. Soc. 119 (1), 101-111. http://dx.doi.org/10.1577/1548-8659(1990)1192.3.co;2.

Herrera, S., Fernández, J. and Gutiérrez, J.M., 2015. Update of the Spain02 gridded observational dataset for EURO-CORDEX evaluation: Assessing the effect of the interpolation methodology. Int. J. Climatol. http://dx.doi.org/10.1002/joc.4391.

Hettiarachchi, P., Hall, M.J. and Minns, A.W., 2005. The extrapolation of artificial neural networks for the modelling of rainfall–runoff relationships. J. Hydroinform. 7 (4), 291-296.

Holland, J.H., 1992. Genetic algorithms. Sci. Am. 267 (1), 66-72.

Hothorn, T., Hornik, K. and Zeileis, A., 2006. Unbiased recursive partitioning: A conditional inference framework. J. Comput. Graph. Stat. 15 (3), 651-674. http://dx.doi.org/10.1198/106186006X133933.

Huang, C.-. and Wang, C.-., 2006. A GA-based feature selection and parameters optimization for support vector machines. Expert Sys. Appl. 31 (2), 231-240. http://dx.doi.org/10.1016/j.eswa.2005.09.024.

Instituto Nacional de Estadística (INE), 2012. Proyecciones de población a largo plazo. INE. 2015, 31/08/2015.

IPCC, 2014. Climate Change 2014: Summary for Policymakers. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY (USA), 1132 pp.

Johnson, H.D., 1980. The comparison of usage and availability measurement for evaluating resource preference. Ecology 61 65-71. http://dx.doi.org/10.2307/1937156.

Jones, C., Giorgi, F. and Asrar, G., 2011. The Coordinated Regional Downscaling Experiment: CORDEX–an international downscaling link to CMIP5. CLIVAR Exchanges 16 (2), 34-40. www.clivar.org/sites/default/files/imported/publications/exchanges/Exchanges_56.pdf.

Jowett, I.G. and Davey, A.J.H., 2007. A comparison of composite habitat suitability indices and generalized additive models of invertebrate abundance and fish presence-habitat availability. Trans. Am. Fish. Soc. 136 (2), 428-444. http://dx.doi.org/10.1577/t06-104.1.

Jowett, I.G. and Duncan, M.J., 2012. Effectiveness of 1D and 2D hydraulic models for instream habitat analysis in a braided river. Ecol. Eng. 48 92-100. http://dx.doi.org/10.1016/j.ecoleng.2011.06.036.

Klemetsen, A., Amundsen, P.A., Dempson, J.B., Jonsson, B., Jonsson, N., O'Connell, M.F., *et al*, 2003. Atlantic salmon *Salmo salar* L., brown trout *Salmo trutta* L. and Arctic charr *Salvelinus alpinus* (L.): A review of aspects of their life histories. Ecol. Freshw. Fish 12 (1), 1-59. http://dx.doi.org/10.1034/j.1600-0633.2003.00010.x.

Kriegler, E., O'Neill, B.C., Hallegatte, S., Kram, T., Lempert, R.J., Moss, R.H., *et al*, 2012. The need for and use of socio-economic scenarios for climate change analysis: A new approach based on shared socio-economic pathways. Global Environ. Change 22 (4), 807-822. http://dx.doi.org/10.1016/j.gloenvcha.2012.05.005.

Kuhn, M., Weston, S., Keefer, C. and Coulter, N., 2014. Cubist: Rule- and Instance-Based Regression Modeling - C code for Cubist by Ross Quinlan -. R package version 0.0.18.

Liaw, A. and Wiener, M., 2002. Classification and Regression by *randomForest*. R News 3 (2), 18-22.

Lin, Y.-., Lin, W.-. and Wu, W.-., 2015. Uncertainty in Various Habitat Suitability Models and Its Impact on Habitat Suitability Estimates for Fish. Water 7 (8), 4088-4107. http://dx.doi.org/10.3390/w7084088.

Lohani, A.K., Kumar, R. and Singh, R.D., 2012. Hydrological time series modeling: A comparison between adaptive neuro-fuzzy, neural network and autoregressive techniques. J. Hydrol. 442-443 23-35. http://dx.doi.org/10.1016/j.jhydrol.2012.03.031.

Lorenz, A.W., Stoll, S., Sundermann, A. and Haase, P., 2013. Do adult and YOY fish benefit from river restoration measures? Ecol. Eng. 61, Part A (0), 174-181. http://dx.doi.org/10.1016/j.ecoleng.2013.09.027.

Luo, Y., Ficklin, D.L., Liu, X. and Zhang, M., 2013. Assessment of climate change impacts on hydrology and water quality with a watershed modeling approach. Sci. Total Environ. 450–451 72-82. http://dx.doi.org/10.1016/j.scitotenv.2013.02.004.

Magalhães, M.F., Beja, P., Schlosser, I.J. and Collares-Pereira, M.J., 2007. Effects of multi-year droughts on fish assemblages of seasonally drying Mediterranean streams. Freshw. Biol. 52 (8), 1494-1510. http://dx.doi.org/10.1111/j.1365-2427.2007.01781.x.

Magdaleno, F. and Fernández, J.A., 2011. Hydromorphological alteration of a large Mediterranean river: Relative role of high and low flows on the evolution of riparian forests and channel morphology. River Res. Appl. 27 (3), 374-387. http://dx.doi.org/10.1002/rra.1368.

Mebane Jr, W.R. and Sekhon, J.S., 2011. Genetic optimization using derivatives: The rgenoud package for R. J. Stat. Softw. 42 (11), 1-26.

Merow, C., Smith, M.J., Edwards, T.C., Guisan, A., Mcmahon, S.M., Normand, S., et al, 2014. What do we gain from simplicity versus complexity in species distribution models? Ecography 37 (12), 1267-1281. http://dx.doi.org/10.1111/ecog.00845.

Milhous, R.T., Bartholow, J.M., Updike, M.A. and A.R., M., 1990. Reference manual for generation and analysis of Habitat Time Series - Version II Biological Report 90; 27, Washington DC, (USA), pp. 249.

Mouton, A.M., Alcaraz-Hernández, J.D., De Baets, B., Goethals, P.L.M. and Martínez-Capel, F., 2011. Data-driven fuzzy habitat suitability models for brown trout in Spanish Mediterranean rivers. Environ. Model. Softw. 26 (5), 615–622. http://dx.doi.org/10.1016/j.envsoft.2010.12.001.

Mouton, A.M., Jowett, I., Goethals, P.L.M. and De Baets, B., 2009. Prevalence-adjusted optimisation of fuzzy habitat suitability models for aquatic invertebrate and fish species in New Zealand. Ecol. Inform. 4 (4), 215-225. http://dx.doi.org/10.1016/j.ecoinf.2009.07.006.

Mouton, A.M., Schneider, M., Depestele, J., Goethals, P.L.M. and De Pauw, N., 2007. Fish habitat modelling as a tool for river management. Ecol. Eng. 29 (3), 305-315. http://dx.doi.org/10.1016/j.ecoleng.2006.11.002.

Mouton, A.M., Schneider, M., Peter, A., Holzer, G., Müller, R., Goethals, P.L.M., et al, 2008. Optimisation of a fuzzy physical habitat model for spawning European grayling (*Thymallus thymallus* L.) in the Aare river (Thun, Switzerland). Ecol. Model. 215 (1–3), 122–132. http://dx.doi.org/10.1016/j.ecolmodel.2008.02.028.

Muñoz-Mas, R., Alcaraz-Hernández, J.D. and Martínez-Capel, F., 2014a. Multilayer Perceptron Ensembles (MLP Ensembles) in modelling microhabitat suitability for freshwater fish. XVII Congreso Español sobre Tecnologías y Lógica Fuzzy (ESTYLF 2014), Zaragoza (Spain), 609-614.

Muñoz-Mas, R., Martínez-Capel, F., Garófano-Gómez, V. and Mouton, A.M., 2014b. Application of Probabilistic Neural Networks to microhabitat suitability modelling for adult brown trout (*Salmo trutta* L.) in Iberian rivers. Environ. Model. Softw. 59 (0), 30-43. http://dx.doi.org/10.1016/j.envsoft.2014.05.003.

Muñoz-Mas, R., Martínez-Capel, F., Schneider, M. and Mouton, A.M., 2012. Assessment of brown trout habitat suitability in the Jucar River Basin (SPAIN): Comparison of data-driven approaches with fuzzy-logic models and univariate suitability curves. Sci. Total Environ. 440 123-131. http://dx.doi.org/10.1016/j.scitotenv.2012.07.074.

Nakicenovic, N., Davidson, O., Davis, G., Grübler, A., Kram, T., La-Rovere, E.L.*, et al*, 2000. Emissions Scenarios: A Special Report of Working Group III of the IPCC.Cambridge (United Kingdom), pp. 1-27.

Olden, J.D., Jackson, D.A. and Peres-Neto, P.R., 2002. Predictive models of fish species distributions: A note on proper validation and chance predictions. T. Am. Fish. Soc. 131 (2), 329-336. http://dx.doi.org/10.1577/1548-8659(2002)1312.0.co;2.

Olden, J.D., Lawler, J.J. and Poff, N.L., 2008. Machine learning methods without tears: A primer for ecologists. Q. Rev. Biol. 83 (2), 171-193. http://dx.doi.org/10.1086/587826.

O'Neill, B.C., Kriegler, E., Riahi, K., Ebi, K.L., Hallegatte, S., Carter, T.R.*, et al*, 2014. A new scenario framework for climate change research: The concept of shared socioeconomic pathways. Clim. Change 122 (3), 387-400. http://dx.doi.org/10.1007/s10584-013-0905-2.

Osuna, E., Freund, R. and Girosi, F., 1997. Training support vector machines: An application to face detection. Proceedings of the 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Juan, PR (USA), 130-136.

Palmer, M.A., Lettenmaier, D.P., Poff, N.L., Postel, S.L., Richter, B. and Warner, R., 2009. Climate change and river ecosystems: Protection and adaptation options. Environ. Manage. 44 (6), 1053-1068. http://dx.doi.org/10.1007/s00267-009-9329-1.

Palmer, M.A., Reidy Liermann, C.A., Nilsson, C., Flörke, M., Alcamo, J., Lake, P.S.*, et al*, 2008. Climate change and the world's river basins: Anticipating management options. Front. Ecol. Environ. 6 (2), 81-89. http://dx.doi.org/10.1890/060148.

Papadaki, C., Soulis, K., Muñoz-Mas, R., Martinez-Capel, F., Zogaris, S., Ntoanidis, L.*, et al*, 2016. Potential impacts of climate change on flow regime and fish habitat in mountain rivers of the south-western Balkans. Sci. Total Environ. 540 418-428. http://dx.doi.org/10.1016/j.scitotenv.2015.06.134.

Parasiewicz, P., 2008. Habitat time series analysis to define flow augmentation strategy for the Quinebaug River, Connecticut and Massachusetts, USA. River Res. Appl. 24 (4), 439-452. http://dx.doi.org/10.1002/rra.1066.

Partridge, D. and Yates, W.B., 1996. Engineering Multiversion Neural-Net Systems. Neural Comp. 8 (4), 869-893. http://dx.doi.org/10.1162/neco.1996.8.4.869.

Platt, J., 2000. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In: Smola, A.J. and Bartlett, P.J.(ed.), Advances in Large Margin Classifiers. MIT Press, Cambridge, MA (USA), pp. 61-74.

Platts, P.J., McClean, C.J., Lovett, J.C. and Marchant, R., 2008. Predicting tree distributions in an East African biodiversity hotspot: model selection, data bias and envelope uncertainty. Ecol. Model. 218 (1-2), 121-134. http://dx.doi.org/10.1016/j.ecolmodel.2008.06.028.

Poff, N.L., Allan, J.D., Bain, M.B., Karr, J.R., Prestegaard, K.L., Richter, B.D., *et al*, 1997. The natural flow regime: A paradigm for river conservation and restoration. BioScience 47 (11), 769-784. http://dx.doi.org/10.2307/1313099.

Quinlan, J.R., 1992. Learning with continuous classes. The 5th Australian Joint Conference on AI, Sydney (Australia), 343-348.

R Core Team, 2015. R: A language and environment for statistical computing. Version 3.2.1.

Rabi, A., Hadzima-Nyarko, M. and Šperac, M., 2015. Modelling river temperature from air temperature: case of the River Drava (Croatia). Hydrol. Sci. J. 60 (9), 1490-1507. http://dx.doi.org/10.1080/02626667.2014.914215.

Riahi, K., Grübler, A. and Nakicenovic, N., 2007. Scenarios of long-term socio-economic and environmental development under climate stabilization. Technol. Forecast. Soc. Change 74 (7), 887-935. http://dx.doi.org/10.1016/j.techfore.2006.05.026.

Rincón, P.A. and Lobón-Cerviá, J., 1993. Microhabitat use by stream-resident brown trout: bioenergetic consequences. T. Am. Fish. Soc. 122 (4), 575-587. http://dx.doi.org/10.1577/1548-8659(1993)1222.3.CO;2.

Riza, L.S., Bergmeir, C., Herrera, F. and Benítez, J.M., 2015. frbs: Fuzzy rule-based systems for classification and regression in R. J. Stat. Software 65 (6), 1-30. http://dx.doi.org/10.18637/jss.v065.i06.

Rogelj, J., Meinshausen, M. and Knutti, R., 2012. Global warming under old and new scenarios using IPCC climate sensitivity range estimates. Nat. Clim. Change 2 (4), 248-253. http://dx.doi.org/10.1038/nclimate1385.

Salmoral, G., Willaarts, B.A., Troch, P.A. and Garrido, A., 2015. Drivers influencing streamflow changes in the Upper Turia basin, Spain. Sci. Total Environ. 503-504 258-268. http://dx.doi.org/10.1016/j.scitotenv.2014.07.041.

Santiago, J.M., García de Jalón, D., Alonso, C., Solana, J., Ribalaygua, J., Pórtoles, J., *et al*, 2015. Brown trout thermal niche and climate change: Expected changes in the distribution of cold-water fish in central Spain. Ecohydrology http://dx.doi.org/10.1002/eco.1653.

Schneider, C., Laizé, C.L.R., Acreman, M.C. and Flörke, M., 2013. How will climate change modify river flow regimes in Europe? Hydrol. Earth Syst. Sci. 17 (1), 325-339. http://dx.doi.org/10.5194/hess-17-325-2013.

Solomatine, D.P. and Dulal, K.N., 2003. Model trees as an alternative to neural networks in rainfall-runoff modelling. Hydrol. Sci. J. 48 (3), 399-412. http://dx.doi.org/10.1623/hysj.48.3.399.45291.

Taghi Sattari, M., Pal, M., Apaydin, H. and Ozturk, F., 2013. M5 model tree application in daily river flow forecasting in Sohu Stream, Turkey. Water Resour. 40 (3), 233-242. http://dx.doi.org/10.1134/S0097807813030123.

Takagi, T. and Sugeno, M., 1985. Fuzzy identification of systems and its applications to modeling and control. IEEE Trans. Syst. Man Cybern. 15 (1), 116-132. http://dx.doi.org/10.1109/TSMC.1985.6313399.

Thomson, A.M., Calvin, K.V., Smith, S.J., Kyle, G.P., Volke, A., Patel, P., *et al*, 2011. RCP4.5: A pathway for stabilization of radiative forcing by 2100. Clim. Change 109 (1), 77-94. http://dx.doi.org/10.1007/s10584-011-0151-4.

Tisseuil, C., Vrac, M., Grenouillet, G., Wade, A.J., Gevrey, M., Oberdorff, T., *et al*, 2012. Strengthening the link between climate, hydrological and species distribution modeling to assess the impacts of climate change on freshwater biodiversity. Sci. Total Environ. 424 193-201. http://dx.doi.org/10.1016/j.scitotenv.2012.02.035.

Tomizono, S., 2013. boxplotdbl: Double Box Plot for Two-Axes Correlation. R package version 1.2.2.

Turner, M. and Stewardson, M., 2014. Hydrologic indicators of hydraulic conditions that drive flow-biota relationships. Hydrolog. Sci. J. 59 (3-4), 659-672. http://dx.doi.org/10.1080/02626667.2014.896997.

Van Vliet, M.T.H., Yearsley, J.R., Franssen, W.H.P., Ludwig, F., Haddeland, I., Lettenmaier, D.P., *et al*, 2012. Coupled daily streamflow and water temperature modelling in large river basins. Hydrol. Earth Syst. Sci. 16 (11), 4303-4321. http://dx.doi.org/10.5194/hess-16-4303-2012.

Vapnik, V.(eds.), 1995. The nature of statistical learning theory. Springer-Verlag New York, New York, NY (USA), 314 pp.

Vezza, P., Muñoz-Mas, R., Martinez-Capel, F. and Mouton, A., 2015. Random forests to evaluate biotic interactions in fish distribution models. Environ. Model. Softw. 67 173-183. http://dx.doi.org/10.1016/j.envsoft.2015.01.005.

Viganò, G., Confortola, G., Fornaroli, R., Cabrini, R., Canobbio, S., Mezzanotte, V., *et al*, 2015. Effects of Future Climate Change on a River Habitat in an Italian Alpine Catchment. J. Hydrol. Eng. 04015063. http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0001293.

Wenger, S.J. and Olden, J.D., 2012. Assessing transferability of ecological models: An underappreciated aspect of statistical validation. Methods Ecol. Evol. 3 (2), 260-267. http://dx.doi.org/10.1111/j.2041-210X.2011.00170.x.

47

Wood, S.N., 2004. Stable and efficient multiple smoothing parameter estimation for generalized additive models. J. Am. Stat. Assoc. 99 (467), 673-686. http://dx.doi.org/10.1198/016214504000000980.

Wood, S.N., 2006. Generalized Additive Models: An introduction with R. Chapman and Hall/CRC Press, London, (UK), 410 pp.

Zadeh, L.A., 1965. Fuzzy sets. Inform. Control 8 (3), 338-353. http://dx.doi.org/10.1016/S0019-9958(65)90241-X.

Zhuo, L., Dai, Q. and Han, D., 2015. Meta-analysis of flow modeling performances-to build a matching system between catchment complexity and model types. Hydrol. Processes 29 (11), 2463-2477. http://dx.doi.org/10.1002/hyp.10371.

Flow regime under the short-term (2011-2040) climate change scenarios

The predicted flow time series under Representative Concentration Pathways (RCP) scenarios based on the ICHEC-EC-EARTH (RCP 4.5 and 8.5) showed low variability, whereas the CCCma-CanESM2 roughly mimicked the historical time series but with lower flow rates (Fig. A 1).
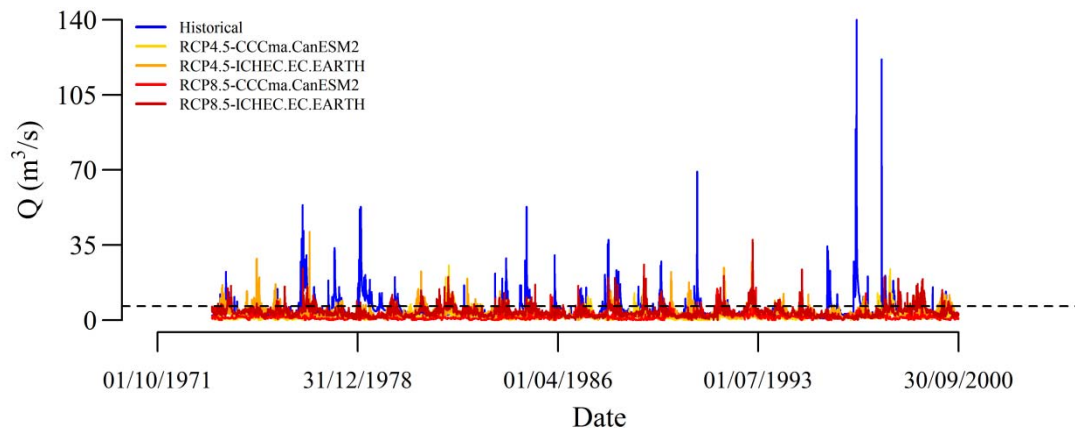

Fig. A 1 Flow time series for the historical and Representative Concentration Pathways (RCP) 4.5 and 8.5 scenarios. Black dashed line depicts the highest simulated flow (*i.e.* 6.5 m^3/s).

Water temperature was determined by the current season (Fig. A 2), and therefore the water temperature predicted under the RCP scenarios (4.5 and 8.5) largely mimicked the historical time series but with higher temperature. The largest increase was predicted in winter whereas the predicted increase during summertime was of inferior magnitude. These predictions overtook the thermal niche threshold (18.7 ºC) more frequently than the historical temperature series, and even surpassed the physiological threshold (25 ºC).
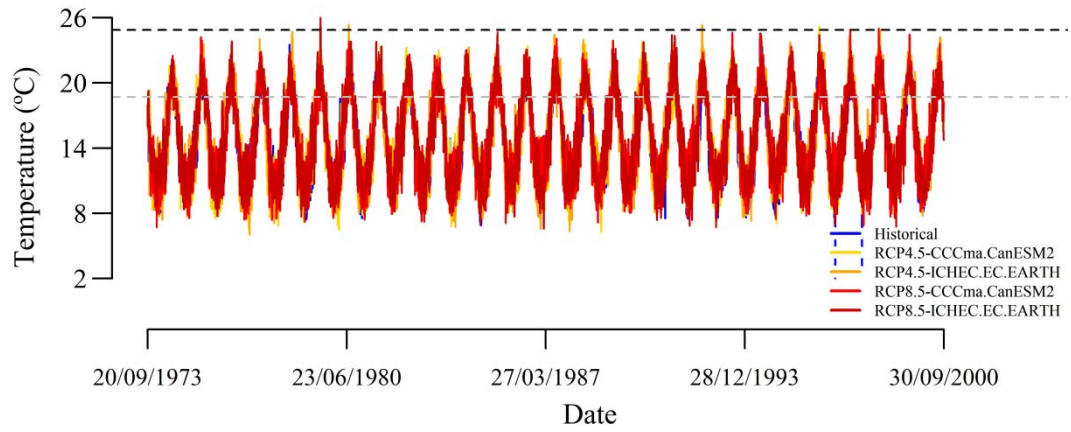
Fig. A 2 Water temperature time series for the historical and Representative Concentration Pathways (RCP) 4.5 and 8.5 scenarios. Grey dashed line depicts the limit of the thermal niche (18.7 ºC) and the black one the physiological niche – threshold of incipient lethality – (25 ºC).