Generalized additive and fuzzy models in environmental flow assessment: a comparison employing the West Balkan trout (*Salmo farioides*)

R. Muñoz-Mas[a,*], Ch. Papadaki[b], F. Martínez-Capel[a], S. Zogaris[b], L. Ntoanidis[c] and E. Dimitriou[b]

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

[a]Research Institute for Integrated Management of Coastal Areas (IGIC), Universitat Politècnica de València. C/ Paranimf 1, 46730 Grau de Gandia. València. Spain, e-mail: pitifleiter@hotmail.com & fmcapel@dihma.upv.es

[b]Hellenic Centre for Marine Research, Institute of Marine Biological Resources and Inland Waters, 46.7 km of Athens - Sounio Ave., 19013, Anavissos Attica, Greece, e-mail: chrispap@hcmr.gr, zogaris@gmail.com & elias@hcmr.gr

[c]D. Argyropoulos & associates, Tinou 2, 15562 Cholargos, Athens, Greece, e-mail: l.ntoanidis@d-argyropoulos.gr

ABSTRACT

Human activities have altered flow regimes resulting in increased pressures and threats on river biota. Physical habitat simulation has been established as a standard approach among the methods for Environmental Flow Assessment (EFA). Traditionally, in EFA, univariate habitat suitability curves have been used to evaluate the habitat suitability at the microhabitat scale whereas Generalized Additive Models (GAMs) and fuzzy logic are considered the most common multivariate approaches to do so. The assessment of the habitat suitability for three size classes of the West Balkan trout (*Salmo farioides*; Karaman, 1938) inferred with these multivariate approaches was compared at three different levels. First the modelled patterns of habitat selection were compared by developing partial dependence plots. Then, the habitat assessment was spatially explicitly compared by calculating the fuzzy kappa statistic and finally, the habitat quantity and quality was compared broadly and at relevant flows under a hypothetical flow regulation, based on the Weighted Usable Area (WUA) *vs.* flow curves. The GAMs were slightly more accurate and the WUA-flow curves demonstrated that they were more optimistic in the habitat assessment with larger areas assessed with low to intermediate suitability (0.2-0.6). Nevertheless, both approaches coincided in the habitat assessment (the optimal areas were spatially coincident) and in the modelled patterns of habitat selection; large trout selected microhabitats with low flow velocity, large depth, coarse substrate and abundant cover. Medium sized trout selected microhabitats with low flow velocity, middle-to-large depth, any kind of substrate but bedrock and some elements of cover. Finally small trout selected

microhabitats with low flow velocity, small depth, and light cover only avoiding bedrock substrate. Furthermore, both approaches also rendered similar WUA-flow curves and coincided in the predicted increases and decreases of the WUA under the hypothetical flow regulation. Although on an equal footing, GAMs performed slightly better, they do not automatically account for variables interactions. Conversely, fuzzy models do so and can be easily modified by experts to include new insights or to cover a wider range of environmental conditions. Therefore, as a consequence of the agreement between both approaches, we would advocate for combinations of GAMs and fuzzy models in fish-based EFA.

Keywords: fuzzy kappa, habitat suitability, microhabitat, physical habitat simulation, *Salmo farioides*, Takagi-Sugeno-Kang fuzzy models

1 Introduction

Human activities such as water withdrawals (Benejam *et al.*, 2010), storing for irrigation purposes (Costa *et al.*, 2012) and hydropeaking (Yao *et al.*, 2015), directly alter river flow regimes in regulated streams impacting freshwater biota (Döll *et al.*, 2009). Moreover, indirectly, human activities have significantly modified precipitation patterns by altering climate (Kalogeropoulos and Chalkias, 2013) and land use (Döll *et al.*, 2009) thus flow regimes in unregulated streams are not exempt of anthropogenic impacts (Li *et al.*, 2015). To evaluate the threats posed by such phenomena the development of scientifically sophisticated tools has now become a fundamental area of research within the scientific community (Arthington *et al.*, 2006). The methods addressed to evaluate river flows were classified into four different categories (Tharme, 2003), namely: hydrological methods (*e.g.* Mathews and Richter, 2007), hydraulic methods (*e.g.* Lamouroux and Souchon, 2002), physical habitat methods (e.g. Muñoz-Mas *et al.*, 2014) and holistic methods (e.g. McClain *et al.*, 2014).

The hydrological methods rely on statistical analysis of hydrological data whereas the hydraulic methods analyse changes in simple hydraulic variables, such as wetted perimeter or maximum depth, as proxies of limiting factors for freshwater biota. Physical habitat methods assess the quantity and suitability of the physical habitat for the target species or assemblages under different flows on the basis of integrated hydrological, hydraulic and biological data (Maddock, 1999). The lasts approach typically encompasses a hydrodynamic model, in order to simulate spatial and temporal variations in critical hydraulic parameters; depth, flow velocity, substrate and cover (Boavida *et al.*, 2014) and a habitat suitability model usually developed at the microhabitat scale for target species thus overstepping the simplicity of the hydraulic methods at the expense of increasing the cost rates (Lamouroux and Souchon, 2002). Finally, several components of the riverine ecosystems as well as social and economic modules are incorporated under the framework of the holistic approaches for basin-scale evaluation.

Nowadays, legislative frameworks in many countries reflect modern societal needs for improved ecological conditions in regulated rivers including the implementation of environmental flow regimes (Katopodis, 2012). However, the requirements and the methods for their determination strongly depend on the considered jurisdiction (Tharme, 2003). For instance, Spanish legislation requires the development of physical habitat studies (Muñoz-Mas *et al.*, 2012) whereas environmental flow recommendations in Greece are based on simplified hydrological methods (Ministry of Environment, Energy and Climate Change, 2011). Hydrological methods have been criticized because they have often been simplified to flow rules that neglect natural system complexity (Arthington *et al.*, 2006). Avoiding this oversimplification, the physical habitat simulation has been identified by some practitioners as the most scientifically and legally defensible methodology for Environmental Flow Assessment (EFA) (Tharme, 2003). Therefore it has demonstrated to be adequate in evaluating the effect of different management alternatives (Yao *et al.*, 2015), restoration actions (Mouton *et al.*, 2007) and potential effects of climate change (Belgiorno *et al.*, 2013).

Regarding the habitat suitability model component in the physical habitat simulation, Waters (1976) suggested the application of continuous curves representing a suitability index (ranging from 0 to 1) for each variable (*e.g.* velocity or depth) instead of binary criteria; with one meaning maximum suitability and zero totally unsuitable. Since then the use of the so-called Habitat Suitability Curves (HSCs) became by far the most common approach in studies involving the physical habitat simulation (Muñoz-Mas *et al.*, 2012). The sum of the areas (*i.e.* cells or pixels) weighed by the inferred suitability within the entire domain of the hydrodynamic model correspond to the Weighted Usable Area (WUA) (Bovee *et al.*, 1998). The WUA is the most renowned general indicator of habitat quality and quantity and is usually calculated for every of the simulated flows thus becoming the WUA-flow curve (Boavida *et al.*, 2014). Upon the WUA-flow curve further calculations should be made for the EFA; for instance the comparison of alternative flow regimes and/or scenario analysis via habitat time series (Milhous *et al.*, 1990).

However, the variables within the aforementioned approach are treated independently for the estimation of the HSCs even though interactions among them were expected (Orth and Maughan, 1982). Consequently, there are examples of multivariate approaches (*e.g.* logistic regression) that demonstrated a greater ability in the determination of the presence or absence of some species (Guay *et al.*, 2000). Between the multivariate approaches those who have received increasing attention are the Generalized Additive Models (GAMs) (Hastie and Tibshirani, 1990) and those based on fuzzy logic (Zadeh, 1965). Although different in nature, the structure of the GAMs could be considered the natural succession of the HSCs because the effect of the set of inputs is simultaneously modelled with smooths curves that resemble the HSCs. On the other hand the popularity of fuzzy logic relies in its capability to mimic human reasoning (Muñoz-Mas *et al.*, 2016). Fuzzy logic describes the input space in linguistic terms (*e.g.* Low velocity or High depth), without loss of accuracy (Castro, 1995), and articulates their different combinations in a comprehensive rules set (Mouton *et al.*, 2008). Further, the mathematics behind are simple enough to be inspected, used and modified by human experts using expert knowledge or new insights to cover a wider range of environmental conditions (Mouton *et al.*, 2008), which emphasizes the usefulness of fuzzy logic to deal with impoverished or extirpated populations. Thus, Jowett and Davey (2007) have developed GAMs for large brown trout (*Salmo trutta*; Linnaeus, 1758) in New Zealand rivers, whereas Muñoz-Mas *et al.* (2012) developed the fuzzy counterpart for medium size individuals in Iberian rivers. Accordingly to that increasing interest, both techniques are actually implemented in commercial software packages; GAMs have been implemented in SEFA (Payne and Jowett, 2012) whereas CASiMiR allows the use of fuzzy models (Jorde, 1997; Schneider, 2001). Limited knowledge exists on the comparison of these two approaches in respect to the simulated habitat suitability (Fukuda *et al.*, 2013) and, as far as we know, there is no example of comparison of such models (developed upon the same database) in EFA.

Different taxa can be targeted in EFA studies. However, fish species can occupy high trophic levels (Sánchez-Hernández and Amundsen, 2015), they are relatively easy to sample and to identify, and generally are known to indicate in-stream habitat constraints (Lorenz *et al.*, 2013). Furthermore, fish are mobile species compared to other aquatic organism groups, *e.g.* benthic invertebrates, and often undergo ontogenetic shifts in their habitat selection (Ayllón *et al.*, 2010). Thus, to complete their life cycle, all the required habitats must be present. Consequently the state of fish populations and fish habitats has served as indicators of aquatic ecosystem health (Katopodis, 2012). Among fish species, salmonids play a crucial role in cold-water food webs and in the generation of ecosystem services (Schindler *et al.*, 2010). The West Balkan (W.B.) trout (*Salmo farioides*; Karaman, 1938) is a poorly studied Balkan endemicity (Delling, 2010) restricted to upland streams between Montenegro and western Greece (Kottelat and Freyhof, 2007) and is assessed as vulnerable in a state-wide conservation evaluation (Zogaris and Economou, 2009). Only some general hints about the optimal habitat for this trout are known, such as the typical salmonids' requirements for cold and fast flowing waters. However, until now there has been no investigation concerning the species' specific habitat preferences at the microhabitat scale.

In this study habitat suitability models for three size classes of the W.B. trout were developed by means of GAMs and fuzzy models. These models were used to infer the habitat suitability (spatially distributed and summarized in the WUA) in a study site in the mountainous part of the Acheloos River (Western Greece). Then the assessed suitability was spatially explicitly compared by calculating the fuzzy kappa statistic. Finally the WUA-flow curves were visually compared and the WUA values derived from the natural flow regime and those derived from the hypothetical extraction of the maximum amount of water legally permitted were numerically compared. The implications of the Greek legislation in EFA were discussed.

2 Materials and methods

2.1 Study site and data collection

The W.B. trout data collection was conducted, at the microhabitat scale, during summer 2014 in the Voidomatis River, north-western Greece; a reference river within the Northern Pindos National Park (Fig. 1). As a consequence the period of strict validity of the developed model would encompass only that season. The mean annual precipitation in the study area typically ranges between 1100 and 1700 mm, yielding a mean daily flow of 13 m^3/s (Woodward *et al.*, 2008); while during the period of data collection (July 2014) it presented a flow rate of 6.29 m^3/s.

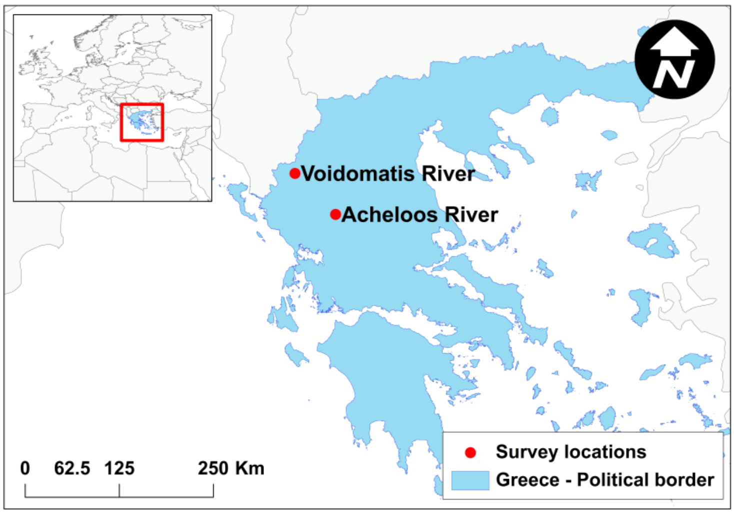

Fig. 1 Location of the site where microhabitat data of West Balkan trout were collected (Voidomatis River) and location where the physical habitat simulation was performed (Acheloos River).

A modification of the equal effort approach (Johnson, 1980) was applied in the selection of the surveyed area. This approach reduces the bias derived from the unbalanced fast- and slow-waters sampling (Muñoz-Mas *et al.*, 2012). Therefore, the river stretch was stratified in Hydro-Morphological Units (HMU) classified as pool, glide, run, riffle and rapid; then several HMUs were selected balancing the areas of slow (*i.e.* pool and glide) and fast (*i.e.* run, riffle and rapid) flow habitats. According to common procedures (Martínez-Capel *et al.*, 2009; Muñoz-Mas *et al.*, 2012), the microhabitat study was conducted by underwater observation (snorkelling) during daylight, classifying the observed individuals in three size classes; large (>20 cm), medium (20–10 cm) and small (<10 cm). The main purpose of the habitat suitability models in the physical habitat simulation approach is to determine habitat in an ecosystem that is best suited for a particular species life history, rather than for determining species abundance and diversity, as do population models (Tomsic *et al.*, 2007); consequently data were collected following a presence-absence scheme. The study focused on individuals that were ‘feeding’ or 'holding a feeding position' because it is assumed that they are occupying such positions as the most energetically profitable (Rincón and Lobón-Cerviá, 1993).

The absences were sampled along each HMU in 4 cross-sections uniformly distributed with 5 point sampling along each cross section, whereas the presences (*i.e.* W.B. trout observations) were measured at the corresponding locations. Depth [m] was measured with a wading rod to the nearest cm and the mean flow velocity of the water column (hereafter velocity [m/s]) was measured with a propeller current meter (OTT®). The percentage of each substrate class was visually estimated around the sampling point or fish location. The substrate classification was simplified from the American Geophysical Union size scale: bedrock, boulders (>256), cobbles (64–256 mm), gravel (8–64 mm), fine gravel (2–8 mm), sand (62 μm–2 mm), silt (< 62 μm) similarly to previous works (Martínez-Capel *et al.*, 2009; Muñoz-Mas *et al.*, 2012). Substrate composition was converted into a single value through the Substrate index [-], by summing the weighted percentages of each substrate type as follows: Substrate index = 0.08 · Bedrock % + 0.07 · Boulder % + 0.06 · Cobble % + 0.05 · Gravel % + 0.04 · Fine Gravel % + 0.03 · Sand % (Mouton *et al.*, 2011).

In addition, the abundance of 5 different cover types was also recorded. Namely, aquatic vegetation, undercut banks, woody debris, shade and large boulders. These cover types corresponded to the most commonly used by other salmonids (Heggenes *et al.*, 1999; Zika and Peter, 2002; Strakosh *et al.*, 2003); while they also summarize the concept of structural cover (*e.g.*, boulders, log jams) (Bovee *et al.*, 1998) and escape cover (*e.g.* vegetation, undercut banks) (Raleigh *et al.*, 1986). As they were written down the cover was scored with three values as follows; easy observation of the fish from the shore (1), observation of the fish possible by underwater observation from distant locations (2) and underwater observation of fish only from close locations (3). Finally, the cover types and their scores were summarized in a cover index [-] by summing the different scores at each location (*e.g.* none = 0, boulders 3 + undercut banks 1 = 4, etc.). In the end, 103 large, 73 medium and 69 small W.B. trout were recorded, whereas the hydraulic conditions in the surrounding area were measured at 241 sites (Fig. 2).

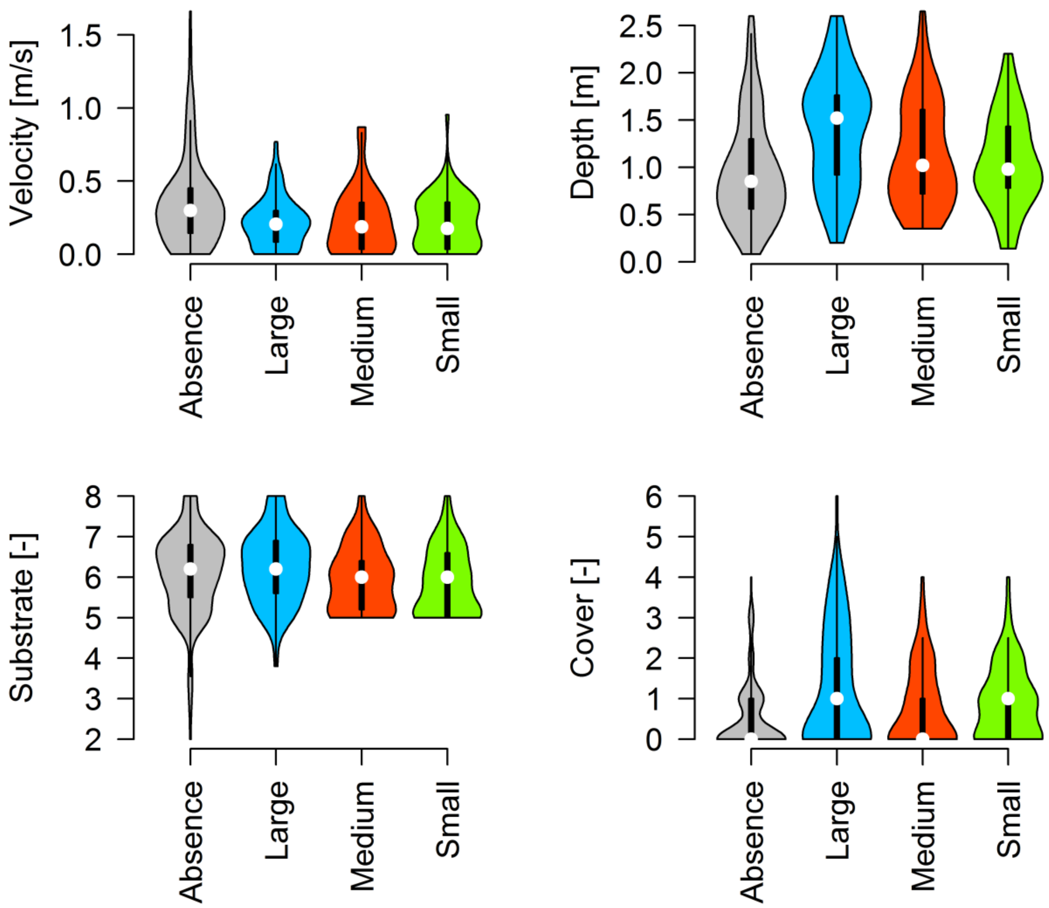

Fig. 2 Violin plots of the data collected in the Voidomatis River. They appear stratified by size class of West Balkan trout and the absences.

2.2 Habitat suitability modelling

2.2.1 Generalized Additive Models (GAMs)

The ecological gradient theory states that species responses to environmental variables are likely to be unimodal and often skewed although, straight-lines are adjusted without any justification (Austin, 2007). In this regard GAMs (Hastie and Tibshirani, 1990) are semi-parametric models, indicated to deal with nonlinearity, since they do not presuppose any specific type of distribution of the input variables applying smooth functions with different degree and number of curvatures (*i.e.* the s_i in equation 1) to simultaneously model their effects (Jowett and Davey, 2007).

$$g(E(y)) = \beta_0 + s_1(x_1) + s_2(x_2) + \dots + s_i(x_i) \quad (1)$$

where g is the link function, E is the expected value, β_0 is the intercept, x_i correspond to the input variables and s_i are the smooth functions.

The expected value can be calculated as the direct aggregation of the effect derived from every variable (g = *gaussian link function*) or can be adjusted to pre-specified distributions such as *poisson* or *binomial,* constraining the outputs to the desired domain. The GAMs development was carried out in *R* (R Core Team, 2015) by means of the *mgcv* package (Wood, 2004). Tensor product smooths are especially useful for representing functions of covariates measured in different units (Wood, 2006). Therefore, instead of one *smooth spline* for each input variable a single *tensor product* was used for the optimization of the smooth curves. The maximum number of knots (*i.e.* the number of bends of every smooth curve) was restricted to three in order to obtain unimodal responses and due to the presence-absence nature of the collected data the selected *link function* was the *binominal*, which constraints the output to the range between 0 and 1. Data prevalence (*i.e.* the ratio of presence data within the entire dataset) was relatively low; 0.30, 0.23 and 0.22 for large, medium and small W.B. trout, respectively. In order to reduce the number of falsely predicted absences, the absence data were down-weighted accordingly to data prevalence because these values may impact the classification capability of habitat models (Maggini *et al.*, 2006; Platts *et al.*, 2008; Beakes *et al*., 2014). For instance the presence cases in the adult GAM were weighted by 0.70 and the absence by 0.30. No-variable selection was carried out so we avoided hypothesis tests in favour of global measures of model performance (Anderson *et al.*, 2000; Platts *et al.*, 2008). Consequently, input *p-values* (Wood, 2013) or *AIC* (Akaike, 1998) were not inspected. However a $3 \times 3\ fold\ cross\ validation$ scheme was followed to inspect the predictive capability of the developed GAMs calculating several performance criteria for every fold. Namely, overall accuracy or Correctly Classified Instances (CCI), Sensitivity (Sn) which corresponds to the ratio of presences correctly classified, Specificity (Sp) which corresponds to the ratio of absences correctly classified, Cohen's Kappa and the True Skill Statistics ($TSS = Sn + Sp - 1$) (Mouton *et al.*, 2010).

2.2.2 Takagi-Sugeno-Kang fuzzy models

The fuzzy logic approach, firstly introduced by Zadeh (1965), takes into account the inherent uncertainty of ecological variables by discretizing the inputs in fuzzy sets named using linguistic terms (*e.g.* Low velocity, Medium velocity, High velocity etc.). Owing to the fuzzy nature of these sets a given value may belong, (with different proportions), to more than one fuzzy sets. The degree of membership in each category is mathematized by means of membership functions usually of trapezoidal or triangular shapes (e.g. Muñoz-Mas *et al.*, 2012; Fukuda, 2013; Boavida *et al.*, 2014). For instance a triangular membership function is defined by three parameters (a_m, b_m and c_m); the membership degree linearly increases between a_m and b_m from zero to one and linearly decreases from one to zero between b_m and c_m (Fig. 3).

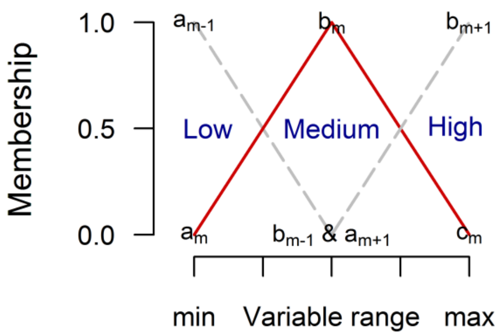

Fig. 3 Depiction and parameters defining triangular membership functions.

Furthermore the fuzzy logic approach allow modellers to express non-linear relations in an *interpretable* manner (Casillas *et al.*, 2005) because the relationship between the different combinations of fuzzy sets are articulated in IF-THEN sequences, which are known as fuzzy rules (Muñoz-Mas *et al.*, 2012). Different types of fuzzy models exist varying mostly in the nature of the consequent (*i.e.* the THEN part). Mamdani-Assilian fuzzy models (Mamdani, 1974) have their consequents defined also by fuzzy sets whereas Takagi-Sugeno-Kang (TSK) fuzzy models (Takagi and Sugeno, 1985) present linear functions (*e.g.* equation 2).

$$IF\ velocity\ is\ x_1\ and\ depth\ is\ x_2\ and\ substrate\ is\ x_3\ and\ cover\ is\ x_4\ THEN\ z = A_i \cdot x_1 + B_i \cdot x_2 + C_i \cdot x_3 + D_i \cdot x_4 + E_i \quad (2)$$

where i corresponds to the rule at hand and from A_i to E_i are the parameters of the consequent linear function. TSK fuzzy models were selected because they performed well in previous studies on habitat suitability modelling (Fukuda, 2013). These TSK fuzzy models were implemented in *R* (R Core Team, 2015) with the help of the *frbs* package (Riza *et al.*, 2015) developing zero order TSK fuzzy models ($i.e.\ A_i = B_i = C_i = D_i = 0$). Therefore, the consequent part corresponded to a dichotomous output, 0 or 1 (*i.e.* presence or absence). Each consequent is weighted by the fulfilment degree of the corresponding fuzzy rule (i) and summed. Thus, the TSK fuzzy model provided smooth outputs all along the feasible output range (from 0 to 1) in a similar way to the *binomial link function* selected for the GAMs. In order to match the ecological gradient theory (Austin, 2007) the complexity of the model was limited by considering three fuzzy sets with triangular shape per input variable (*e.g.* Low velocity, Medium velocity and High velocity). However if a given rule does not cover any input data it shall remain undetermined. To overcome such deficiency, a uniform distribution of the fuzzy sets over the variable range was implemented since it has been proved to reduce the number of untrained rules (Muñoz-Mas *et al.*, 2012). Consequently the vertices of the triangular fuzzy sets were placed in accordance with variables' quantiles. The fuzzy rules optimization was based on the TSS because its maximization usually renders models that balance the accuracy over the presence and absence classes (Mouton *et al.*, 2010). For every developed TSK fuzzy model the optimisation was performed nine times with the hill-climbing algorithm (see Mouton *et al.*, 2008 for further details) searching for the optimal value for every consequent (*i.e.* 0 or 1) and the ultimate consequent was assigned by rounding up the mean value obtained in the nine iterations. The 3×3 fold cross-validation scheme was also followed to inspect the predictive capability of the TSK fuzzy models over the same data subsets used in the GAMs section. Finally, the same performance criteria calculated for the GAMs were calculated.

2.3 Hydraulic modelling

A representative reach of the Acheloos River upstream of the Mesochora dam was selected in order to apply the hydraulic simulation. A topographic survey encompassing the main channel and banks, was carried out with a GPS/GNSS Geomax-Zenith 20 using geodesic references (*i.e.* GGRS '87 - Greek Geodetic Reference System) to improve the accuracy. Substrate percentages and cover types were co-ordinately recorded to match the requirements of the habitat suitability models. The topographic survey was then used to generate digital elevation models as a base for the hydraulic simulation.

HEC-RAS (Version 4.1) was used to perform a quasi-2D hydraulic simulation for several flows in regard to the mean monthly summer flows. The length of the representative reach was 390 m (Papadaki *et al.*, 2014); simulations were performed with 27 cross-sections along the river stretch placed in accordance with the general principles of 1D modelling (Jowett and Duncan, 2012). Manning's roughness coefficient was adjusted for model calibration by comparing the observed water surface elevations and velocities at 10 critical cross-sections and two surveyed flows (*i.e.* 4 m^3/s and 8.8 m^3/s) with the simulated model results.

For the quasi-2D hydraulic approach every cross-section was subdivided in 10 cells both in the main channel and the overbank area. Thereby, velocities were separately calculated for each cell of the simulated water stage. In the end, every pixel of the hydraulic model for each river flow presented a value for velocity, depth, substrate index and cover index on which the habitat assessment was then performed.

2.4 Comparison of the habitat suitability models and river habitat assessment

Model reliability and transparency is of major concern for ecological modelling (Austin, 2007). Unlike the analysis of GAMs, the analysis of TSK fuzzy models is straightforward. Thus, to concurrently characterize the relationship between the inputs variables and the outputs, the partial dependence plots (PDPs) implemented in the package *randomForests* (Liaw and Wiener, 2002) were developed allowing an easy comparison of the GAMs and the TSK fuzzy models. The PDPs depict the average of the outputs for an input variable and accounts for the effects of the remaining variables within the model by averaging their effect yielding interpretable univariate plots. However, as a consequence, the depicted output range may differ from the feasible one (*i.e.* from 0 to 1).

The outputs of the GAMs and the TSK fuzzy models match the range typically provided by the HSCs (*i.e.* from 0 to 1). Thus, in order to illustrate similitudes and differences in the habitat assessment (regardless the simulated flow or corresponding season), it was used to build the WUA-flow curves which were then visually compared. The WUA is the sum of the areas (in this case $1 x 1$ m pixels) weighted by the inferred suitability. Since a single WUA value per flow is inferred, similar values of WUA could dramatically differ in the spatial distribution of the assessed suitability. To overcome this limitation a spatially explicit pairwise comparison was performed by calculating the fuzzy kappa statistic (Hagen-Zanker *et al.*, 2005). Fuzzy kappa statistic is similar to the traditional Cohen's kappa and provides a meaningful index ranging from -1 to 1, with one corresponding to perfect agreement. The spatial explicit comparison was carried out with the Map Comparison Kit version 3.2.3 (Visser and De Nijs, 2006) by dividing the assessed suitability in 5 uniform intervals. This software allows performing the comparison with certain degree of tolerance between categories of the overlaid pixels and taking into account the surrounding area. However the extension of the area of influence affects the results obtained from the fuzzy kappa statistic; for instance, a large influence area has demonstrated to dramatically increase the values of the statistic (Rose *et al.*, 2009) thus providing awkward interpretation. Therefore it should be selected in accordance with grounded reasons such as known differences in map resolutions or the home range of the target species. Brown trout, a closely related species, has proved a home range of approximately 300 m (Ovidio *et al.*, 1998) a distance similar to the length of the area comprised in the hydraulic models. Therefore we calculated the fuzzy kappa by considering only the overlaying pixels ($1\ x\ 1\ m$) following the correspondence depicted in the similarity matrix where similarity between categories linearly decreases as the distance from the main diagonal increases (Table 1).

Table 1 Similarity matrix used in the calculation of the fuzzy kappa statistic. The similarity linearly decreases as the interval goes farther from the main diagonal.

		Suitability				
		0.0-0.2	**0.2-0.4**	**0.4-0.6**	**0.6-0.8**	**0.8-1.0**
Suitability	**0.0-0.2**	1.00	0.75	0.50	0.25	0.00
	0.2-0.4	0.75	1.00	0.75	0.50	0.25
	0.4-0.6	0.50	0.75	1.00	0.75	0.50
	0.6-0.8	0.25	0.50	0.75	1.00	0.75
	0.8-1.0	0.00	0.25	0.50	0.75	1.00

Greek legislation on environmental flows coincides with the period of strict validity of the developed models (*i.e.* summer). Consequently the developed models allowed the comparison in WUA terms of the hypothetical extractions of the largest flow legally permitted. Currently, Greek legislation establishes the minimum flow as a percentage of the natural flow according to the highest value of the following rules:

1. 30 % of the mean monthly flows of June, July and August

2. 50 % of the mean monthly flow of September.
3. 0.03 m^3/s.
4. 0.2 m depth at the thalweg if there is sensitive ichthyofauna present.

The hydrological data close to the study site of Mesochora are scarce with only two complete hydrologic cycles available (1986–1988) although, these data were used to infer the mean monthly flow (Fig. 4) which presented the minimum in August (1.44 m^3/s) and the maximum in April (32.91 m^3/s). Finally, the analysis focused on the months from June to September (*i.e.* 8.93, 4.92, 1.44, 1.85 m^3/s), and the same period but considering the worst scenario (*i.e.* 2.68, 1.48, 0.43, 0.92 m^3/s). The values of the WUAs for these flows (natural and hypothetically impacted) were interpolated from the corresponding WUA-flow curves for both models, GAM and TSK fuzzy, and the impact on the habitat suitability of Greek legislation was discussed.

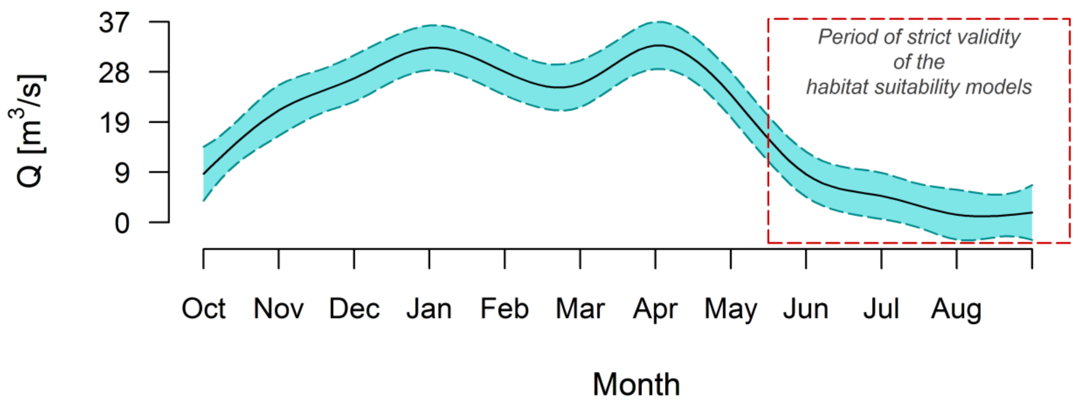

Fig. 4 Natural flow regime in the Acheloos River in the near vicinity of the Mesochora dam. Band width corresponds to the 0.95 confidence interval.

3 Results

Based on the results obtained during the $3 \times 3\ fold\ cross\ validation$ the GAMs would outperform the TSK fuzzy models (Table 2). The training of the ultimate models with the entire dataset mitigated such a trend and both models presented similar values of the performance criteria for the three size classes (Table 2 values between brackets).

Table 2 Accuracy or Correctly Classified Instances (CCI), Sensitivity (Sn), Specificity (Sp), Cohen's kappa (Kappa) and True Skill Statistics (TSS) for the developed models. The values for the ultimate models use in the habitat assessment (*i.e.* those without cross validation) appear between brackets

	Large		**Medium**		**Small**	
	GAM	**TSK**	**GAM**	**TSK**	**GAM**	**TSK**
CCI	0.67±0.05 (0.72)	0.45±0.35 (0.68)	0.64±0.05 (0.68)	0.46±0.35 (0.68)	0.63±0.07 (0.72)	0.48±0.36 (0.74)
Sn	0.66±0.07 (0.74)	0.49±0.35 (0.71)	0.64±0.06 (0.75)	0.48±0.29 (0.68)	0.55±0.14 (0.72)	0.43±0.24 (0.59)
Sp	0.68±0.08 (0.72)	0.46±0.32 (0.67)	0.64±0.08 (0.66)	0.47±0.34 (0.68)	0.66±0.08 (0.72)	0.51±0.36 (0.78)
Kappa	0.31±0.08 (0.41)	0.21±0.14 (0.34)	0.23±0.05 (0.32)	0.19±0.1 (0.29)	0.16±0.12 (0.35)	0.2±0.11 (0.33)
TSS	0.34±0.08 (0.46)	0.24±0.16 (0.38)	0.29±0.05 (0.42)	0.24±0.12 (0.37)	0.2±0.15 (0.44)	0.24±0.12 (0.37)

The PDPs showed similar pattern for both approaches basically differing in their smoothness degree with the TSK-Fuzzy model yielding piecewise rectilinear PDPs (Fig. 5).

In general, the large W.B. trout selected low flow velocity microhabitats with the largest depth, coarse-to-rocky substrates (cobble to bedrock) and abundant cover. The medium W.B. trout presented slight discrepancies between the GAM and the TSK model. The PDPs showed preference for low flow velocity with middle-to-large depth, whereas the substrate presented the largest discrepancy. The TSK fuzzy model placed the optimum for fine substrate whereas the GAM model did it for coarse substrate (gravel and cobble). Finally, the medium size class selected microhabitats with cover either scarce or abundant. The small W.B. trout also presented slight discrepancies between the GAM and the TSK fuzzy model. The PDPs coincided in the preference for microhabitats with low flow velocity but differed in regards to the optimal depth; the GAM stated as preferable deeper microhabitats. The small size class selected a wide range of substrate types from fine to coarse substrates and also selected microhabitats with either scarce or abundant cover.

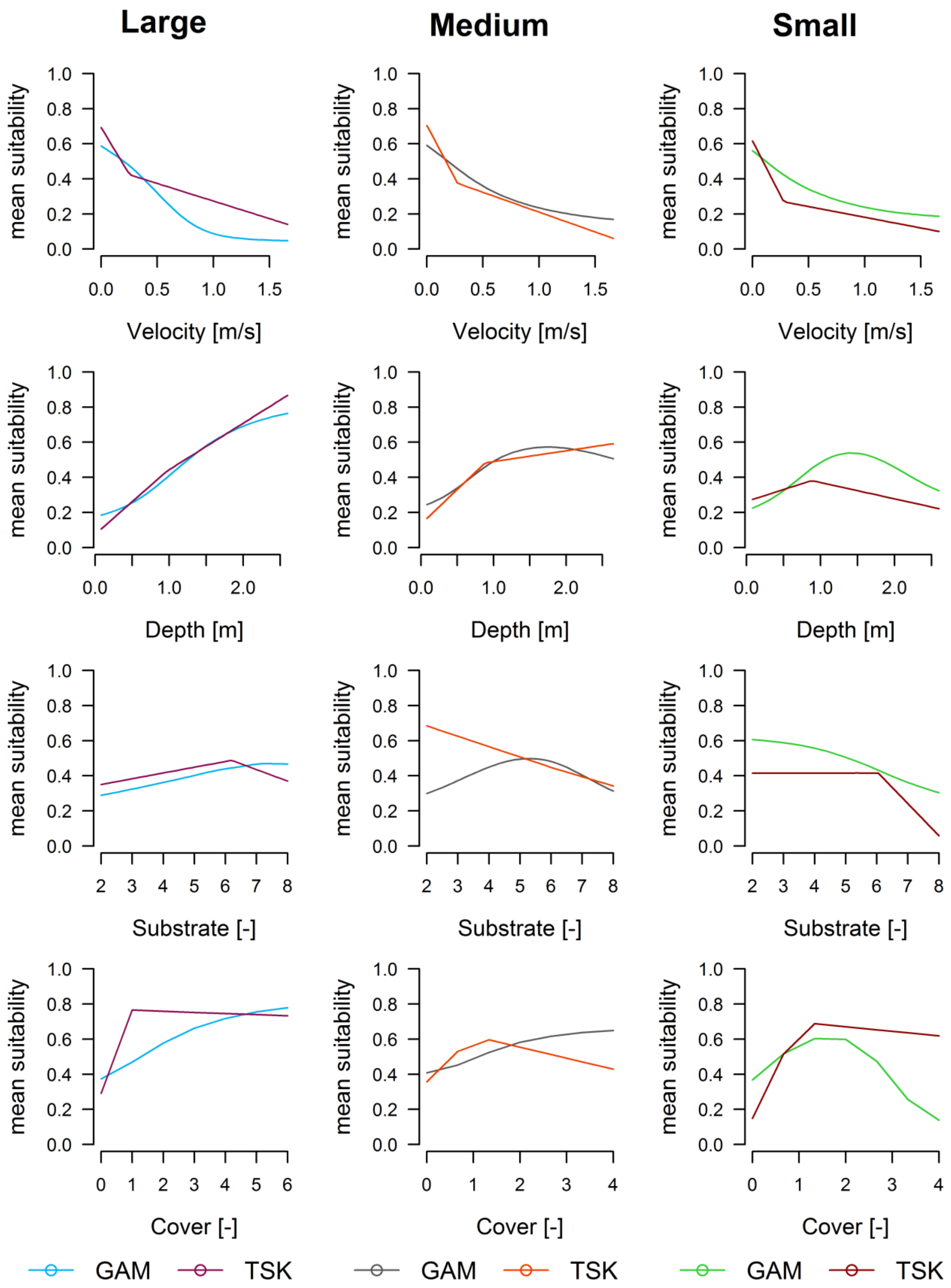

Fig. 5 Partial Dependence Plots (PDPs) calculated by means of the GAMs and the TSK fuzzy models for the three size classes of the West Balkan (W.B.) trout.

The study site at the Acheloos River presented low suitability for the W.B. trout thus the WUA-flow curves presented low values in comparison with the corresponding wetted area (Fig. 6). The TSK fuzzy models presented generally lower WUA values of the WUA than the GAM's counterparts but showing similar patterns. Though both curves presented a very gentle slope, only the WUA-flow curves for the large W.B. trout showed discrepant trends. Thus, the GAM-related curve presented a gentle decreasing trend and the TSK's an increasing one. The values of the fuzzy kappa were relatively low; nevertheless, in accordance with the concordant PDPs, the fuzzy kappa analysis suggested similar spatial distribution of the suitable and unsuitable microhabitats achieving the larger values for those flows with closer values of the WUA. Only the large W.B. trout presented an erratic pattern, especially for these flows between 0.5 and 5 m^3/s which presented the larger differences in terms of WUA but relatively high values of fuzzy kappa.

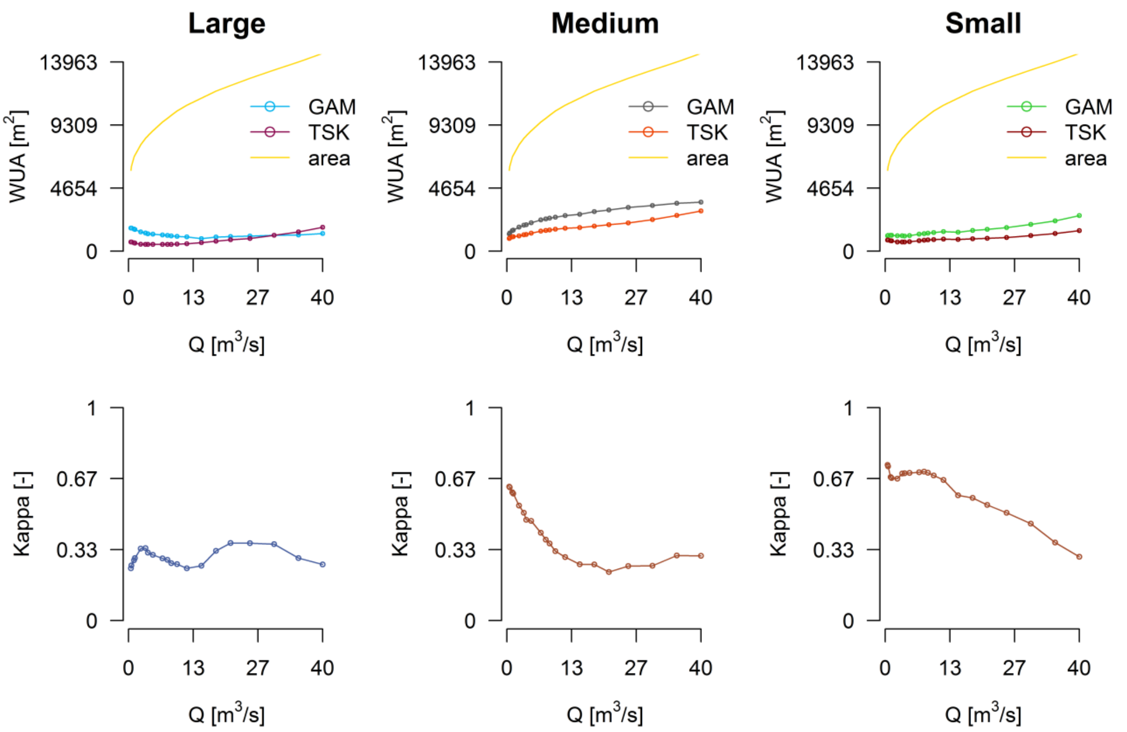

Fig. 6 Upper sequence; WUA-flow curves calculated with the GAM and the TSK-fuzzy. Lower sequence fuzzy kappa-flow curves for the three size classes of the West Balkan (W.B.) trout. The highest curve corresponds to the wetted area.

Generally, the GAMs demonstrated to be more optimistic in the habitat assessment by significantly increasing the pixels assessed with low to intermediate suitability (*i.e.* from 0.2 to 0.6). Fig. 7 depicts the habitat assessment for the flows with the most discrepant WUA; 0.6, 25 and 40 m^3/s for the large, the medium and the small W.B. trout respectively (0.5, 21 and 40 m^3/s considering the lowest values of fuzzy kappa). The regions assessed with high suitability were almost coincident but the areas assessed with low to intermediate suitability were larger for the GAMs which in accordance with the values stated in the similarity matrix, caused the relatively low values of the fuzzy kappa statistic.

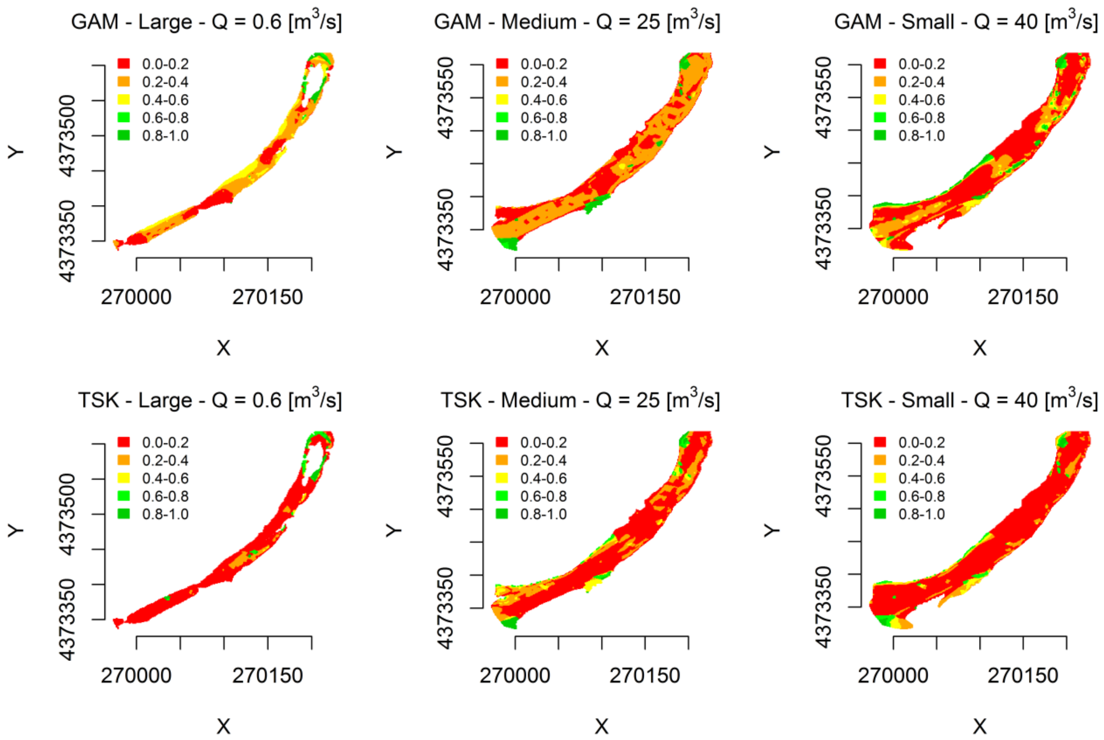

Fig. 7 General view of the habitat assessment for the flows with the most discrepant Weighted Usable Area.

In accordance with the patterns observed in the WUA-flow curves the hypothetical reduction of the running flows following legal minimum flow norms would present either positive or negative values. The large W.B. trout would experience an increase of the WUA in each of the analysed flows, regardless the considered model, GAM or TSK (Fig. 8). Conversely, the medium size class would experience a decrement of the WUA for every month and habitat suitability model. Finally the small W.B. trout was the only size class that mixed the trends. Both models suggested a decrement in June whereas the GAM yielded almost the same values of WUA for the natural and the hypothetically regulated counterparts. Conversely the TSK fuzzy model suggested a small increase of the WUA for the same period.

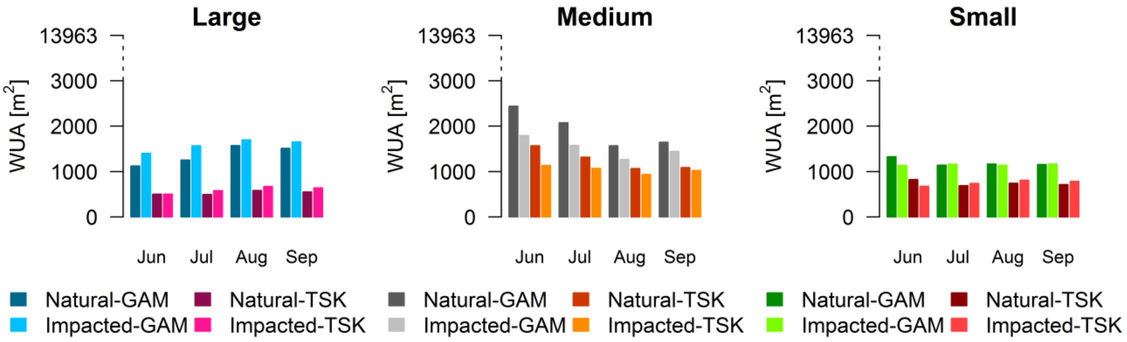

Fig. 8 Weighted Usable Area (WUA) for the mean monthly natural flow in the analysed period and values derived from a hypothetical extraction of the largest legal amount of water, 70 % in June, July and August and 50 % in September. Maximum in the y-axis correspond to the wetted area of the largest monthly mean flow (April; 32.91 m^3/s).

The WUA-flow curves inferred with the GAMs showed steeper shapes; consequently the per cent variation of WUA was larger than the ones for the TSK fuzzy models (Table 3). Only the TSK fuzzy model for small W.B. trout inverted this trend by showing smaller variations than the GAM. Even though the minimum hypothetical reduction in the running flow would be of 50 % that percentage of variation was never exceeded either positively or negatively; the largest predicted impact would affect largely the medium W.B. trout with a predicted reduction of *ca.* 37 % in June.

Table 3 Per cent variation of the WUA derived from the hypothetical extraction of the largest legal amount of water (release of only the minimum legal environmental flow during summer).

		June	July	August	September
Large	GAM	20%	20%	7%	8%
	TSK	0%	15%	13%	14%
Medium	GAM	-36%	-32%	-24%	-14%
	TSK	-38%	-23%	-14%	-6%
Small	GAM	-17%	2%	-2%	1%
	TSK	-22%	7%	8%	9%

4 Discussion

4.1 Models' comparison

In accordance with the calculated performances and the agreement between the PDPs, the entire set of habitat suitability models (GAMs and TSK fuzzy models) were considered adequate for EFA. Analysing the performance criteria of the ultimate models, both approaches presented similar values in magnitude, which practically coincided with those obtained in previous studies involving similar datasets of salmonids (Muñoz-Mas *et al.*, 2012; Muñoz-Mas *et al.*, 2014). In those studies, a Mandami-Asilian fuzzy model was developed for medium brown trout and another one for large brown trout by means of probabilistic neural networks achieving values of kappa and TSS near to 0.4. However analysing the performance criteria obtained through the cross-validation the GAMs proved a larger predictive capability and especially a larger stability (*i.e.* smaller standard deviation). Fuzzy models are universal approximators (Castro, 1995) therefore they can over-fit the data. The use of adequate datasets has proved fundamental in the development of proficient fuzzy models (Yi *et al.*, 2014); thus two thirds of the data from any of our datasets (*i.e.* the one for large, medium and small W.B. trout) training 81 rules, which corresponds to 3 fuzzy sets to the 4^{th} degree, have demonstrated to be insufficient to render generalizing models in every one of the nine trials. Consequently, some models poorly performed over the corresponding validation datasets. Accordingly, on the basis of selecting the most stable and accurate model, GAMs could be considered a slightly preferable option for EFA, especially taking into account that there was no validation with independent data.

The main reason for the GAM outperformance is its greater flexibility in responses adjustment and the only way to increase the flexibility of the TSK fuzzy model is the increase of the amount of fuzzy sets (increasing granularity) and/or testing different membership functions. There are several approaches to simultaneously optimize the number and/or the shapes of the membership functions simultaneously with the optimization of the consequents (e.g. Casillas *et al.*, 2005; Alcalá-Fdez *et al.*, 2009). However, these approaches tend to be detrimental to the *interpretability* (*i.e.* the capability to express the behaviour of the real system in a comprehensible way), which is a fundamental advantage of fuzzy logic based models (Casillas *et al.*, 2005). The membership functions (in this case triangular) condition the transitions between the suitability assigned to the different regions of the universe of discourse (*i.e.* the ones described in the fuzzy rules) and thus linear membership functions turned in linear PDPs.

Despite specific studies demonstrated that there is not an optimal membership function applicable to every problem (Mitaim and Kosko, 2001) in most of the cases the studies addressed to EFA skipped the analysis of different alternatives (e.g. Muñoz-Mas *et al.*, 2012; Yi *et al.*, 2014; Boavida *et al.*, 2014). Gaussian or bell-shaped membership functions could produce rounded and smooth PDPs (Mitaim and Kosko, 2001) however, triangular membership functions present remarkable advantages; they are defined by few parameters which can be easily tuned (Alcalá-Fdez *et al.*, 2009) and the sum of membership for each data is always one. As a consequence, triangular membership functions still are being used in the development of novel modelling approaches (e.g. Casillas *et al.*, 2005; Alcalá-Fdez *et al.*, 2009) and thus we considered them an adequate choice. On the other hand increasing the number of fuzzy sets may increase models' accuracy. However, it also increases the possibility of over-fitting the data and the ratio of undetermined rules (Mouton *et al.*, 2008). In our study, the PDPs of both approaches markedly matched, in contrast with previous studies where they differed (Fukuda *et al.*, 2013). Therefore, it was considered that the prior constraint by limiting the amount of knots and of fuzzy sets up to three allowed the development of sound models that fitted well with the ecological gradient theory (Austin, 2007) and thus, the differences in models' performance were insufficient to trigger the search of additional improvements.

4.2 West Balkan trout habitat selection

The PDPs for the W.B. trout closely resembled the habitat selection patterns observed in other salmonids of mountain streams, especially the brown trout. Large W.B. trout selected habitats with low velocity, large depth, coarse substrate, even bedrock, and abundant cover. Such patterns practically coincided with Bovee's (1978) HSCs for large brown trout with the only difference appearing in the selection related to bedrock substrate. Likewise, Ayllón *et al.* (2010) and Muñoz-Mas *et al.* (2014) also reported the use of large depth and coarse (also bedrock) substrates; however in those warmer Mediterranean rivers brown trout selected faster microhabitats, most probably because those rivers presented higher summer water temperatures (22 ºC) – enhancing the natatorial capacity – than the ones observed in the Voidomatis River (10-12 ºC summer temperature). Although, none of these studies independently considered cover a variable that can be also influencing such differences. Cover is a more difficult variable to identify and quantify what may explain its absence from many habitat studies (Heggenes *et al.*, 1999). Nevertheless Strakosh *et al.* (2003) studied the patterns of cover selection of medium-to-large brown trout (body length > 17 cm) finding that the most important cover types were the undercut banks, vegetation, log jams, water turbulence and depth; whereas overhanging canopy and shade proved to be of lesser importance. We summarized the available cover in a single index although we can asseverate that the most used cover coincided with those detailed above; but the shaded area was profusely used by the large W.B. trout.

The PDPs of medium W.B. trout for velocity, depth and substrate also matched those patterns of habitat selection described by Bovee’s (1978) HSCs. Both PDPs coincided with Bovee’s (1978) HSCs by stating 0 m/s as the most suitable flow velocity and a gentle decrement of the suitability in comparison with the more abrupt decrement observed in the PDPs for the large W.B. trout. In addition medium W.B. trout selected shallower microhabitats than the large counterpart. Such differences typical of salmonids (Gibson, 1993) have been also reported in Iberian rivers where juvenile brown trout occupied smaller depth than the adults (Ayllón *et al.*, 2010; Muñoz-Mas *et al.*, 2012; Muñoz-Mas *et al.*, 2014). Conversely the PDPs for substrate differed from the suitability described within the aforementioned literature; several authors suggested acute HSCs with the optimum at cobbles (Bovee, 1978; Ayllón *et al.*, 2010; Muñoz-Mas *et al.*, 2012), whereas the PDPs for the medium W.B. trout suggested a wider optimal range from fine to coarse substrate. The GAM stated the optimum at cobbles whereas the TSK fuzzy model displaced it to silt and sand. However we cannot rule out that these differences are only caused by the number of degrees of freedom set up during the development of these HSCs which were significantly larger in comparison with the GAMs herein developed (Bovee, 1978; Muñoz-Mas *et al.*, 2012). Habitat selection in salmonids is based on their competitive abilities and the profitability of territories in terms of both potential net energy intake rate and predation risk (Ayllón *et al.*, 2009). The existing literature stated a weaker over-selection of microhabitats with cover for medium size brown trout in comparison with the large counterpart (Vismara *et al.*, 2001) which would be concordant with the patterns described in the PDPs for medium

W.B. trout. Cover was summarized in a single predictor although, it would be plausible that lighter cover provided enough shelter for these smaller individuals. However, trout species have a territorial behaviour, consequently the distribution of younger individuals could be also affected by older fish through intercohort competition (Ayllón *et al.*, 2009) displacing the smaller and weaker individuals from optimal microhabitats which could also possibly explain these differences.

The differences between the PDPs for the small W.B. trout and the small brown trout literature were larger. Our results showed that small W.B. trout tended to occupy near-bank microhabitats with low flow velocity (optimum at 0 m/s) and lower depth than their larger counterparts, whereas the literature lacks consensus about the most suitability habitats for the small brown trout. For instance Bovee (1978) and Ayllón *et al.* (2009) suggested a wider optimal range for velocity than the medium counterpart but a more restricted one for depth, (0.3 m to 0.5 m), which is significantly shallower than the one depicted in the corresponding PDPs. The PDPs for substrate do not fit better the patterns of substrate selection described in brown trout literature. Thereby, while our results suggested a wide range of suitable substrates, brown trout studies restricted the suitable substrates to gravel and cobble (Bovee, 1978; Ayllón *et al.*, 2009). Nevertheless, we considered the modelled suitability plausible since it was similar to observations in some Iberian rivers (Muñoz-Mas *et al.* unpublished). Larger brown trout tended to occupy areas with deeper water and more cover than did yearling brown trout (Heggenes, 1988a) apparently because small brown trout easily shelter in the cobble-boulder substrate's interstitial spaces (Heggenes, 1988b). Such a pattern of cover use could explain the differences observed in the PDPs however, likewise the medium size case, they could be caused by the aforementioned intercohort competition. To sum up we conclude that W.B. trout habitat selection certainly resemble those described for brown trout but the abundance and types of the available microhabitats (Rincón and Lobón-Cerviá, 1993) and the modelling technique (Fukuda *et al.*, 2013) could have influenced the inferred preferences. Therefore we acknowledge that this comparison should be cautiously interpreted as it might need further verification.

4.3 Environmental flow assessment

The populations of the W.B. trout in the Acheloos River have declined during the last decade; thus, W.B. trout is currently rather scarce in the main-stem of the river system as we confirmed during the summer sampling. Such a phenomenon has been suggested to be caused by severe overfishing involving illegal spear fishing and electrofishing since instream and riparian conditions in this stretch of river are not degraded (Zogaris *et al.*, 2009). This section of the Acheloos River is dominated by low populations of cyprinids (Economou *et al.*, 2007) thus the extensive shallow braided channel may not suit dense trout populations. Consequently, the scarcity of W.B. trout did not allow performing any validation of the developed habitat suitability models. However, interviewed anglers stated that the large W.B. trout were always found in the large and deep pools. These comments, together with the aforementioned similarities with brown trout habitat selection patterns, enhance the credibility of the low values of WUA calculated for most of the simulated flows and the subsequent comparison.

Nowadays a common approach to overcome the possible bias of using a given modelling technique is the use of models' ensembles, based on a single technique (Muñoz-Mas *et al.*, 2015) or combining the predictions of several techniques (Muñoz-Mas *et al.*, 2016). Nevertheless the coincidences between the PDPs, the patterns of the WUA-flow curves and, especially in the effects of the hypothetical extraction of the maximum amount of water legally permitted (fairly coincident, positively or negatively) suggested this approach, though recommendable, unnecessary. Certainly, the relatively low values of the fuzzy kappa statistic suggested low similarity. Although the most discrepant flows (*i.e.* 0.6, 25 and 40 m^3/s for large, medium and small W.B. trout respectively) presented the optimal areas in the same locations as well as any other pair of flows did in accordance with the increasing values of the fuzzy kappa statistic. Such low values of the fuzzy kappa statistic have been caused by the more classificatory character of the TSK fuzzy models (*i.e.* they tended to provide lopsided values either towards zero or one). However, another reason that could be playing a significant role for such a low values is the well documented dependence of the kappa statistic on data *prevalence* (Allouche *et al.*, 2006). As a consequence we cannot discard that these low values of the fuzzy kappa have been exacerbated by the bias on the categories of the assessed suitability, since the TSK fuzzy models assessed most of the pixels within the category from 0 to 0.2 and very few to the remaining categories. The study site resembled a deep run, a morphology characterized by relatively high flow velocity which tends to increase with the increase of the flow rate. Therefore in accordance with the modelled habitat requirements we concluded that the resulting low suitability of the site is certainly plausible and thus both approaches, GAMs and TSK fuzzy models, should be considered almost equal for EFA though the per cent reduction in the WUAs slightly varied. We referred the increase or decrease on WUA to the WUA in natural flow regime however, environmental flow legislation typically refers it to a specific WUA value (*e.g.* the maximum WUA) to facilitate the proper comparison (Muñoz-Mas *et al.*, 2012). As a consequence, the effects of the hypothetical water abstraction, which varied regarding the flow and size, should be viewed as illustrative of the changing trends in the suitable habitat available and the absolute per cent differences ignored. Likewise previous studies (Li *et al.*, 2015), the reduction of the flow rate can have a positive effect as it had for the large individuals but also negatives as it demonstrated for medium and, to a lesser extent, for small W.B. trout. Accordingly to these divergent effects the shifts in the WUA proved insufficient to evaluate either positively or negatively the Greek provisions for the minimum flow; habitat time series analysis (Milhous *et al.*, 1990) should be performed in the near future to ascertain its properness. Nevertheless we considered hard to believe that a reduction of 70 % of the flow rate can be innocuous for the inhabiting biota.

4.4 Models' selection

For the foregoing we considered that the GAMs and the TSK fuzzy counterparts quite similar models. In this case the only element that could tip the balance between GAMs or TSK fuzzy models was the accuracy and the stability, which was superior in the GAMs since the PDPs were ecologically relevant and fitted well each other and the habitat selection patterns of other salmonids. However, GAMs need sound training datasets and, in their very basic implementation, do not consider variables interactions. Conversely, the mathematics behind the zero order TSK fuzzy models are simple enough to allow their modification or their development by means of experts (*e.g.* following Ahmadi-Nedushan *et al.*, 2008) which upholds their validity for EFA, especially, dealing with impoverished populations. In addition fuzzy models will be specially suited to do exploratory analysis when interactions between variables are suspected to exist. As a consequence we would not advocate for one or the other approach rather for combinations of them in accordance with the necessities and limitations of the problem at hand.

5 Conclusions

GAMs outperformed TSK fuzzy models due to greater flexibility in modelling habitat suitability. The PDPs for the GAMs and the TSK fuzzy models suggested similar habitat selection. Large W.B. trout selected slow flowing microhabitats with the greatest depth, coarse and bedrock substrates and abundant cover. The medium-sized W.B. trout mostly selected microhabitats with low flow velocity but they proved more versatile by tolerating higher flow velocity. In terms of depth, substrate and cover they occupied deep areas with coarse substrate but were not as restrictive regarding the abundance of cover than the large counterpart. Finally the small W.B. trout selected shallow microhabitats with low flow velocity and fine-to-coarse substrate. Apparently the small W.B. trout used the interstitial space of the coarse substrate for concealment thus proved a weaker preference for microhabitats with abundant cover. The habitat selection patterns as well as the ontogenetic shift in the habitat preferences resembled those observed for the brown trout. In accordance with the similarities observed in the PDPs both approaches yielded similar habitat suitability assessment. The study site in the Acheloos River indicted a low suitability for the W.B. trout although the GAMs provided more optimistic results. The TSK models presented generally values of the WUA slightly lower than the GAM's but the shape of the PDPs, the habitat assessment (optimal microhabitats) and the shape of the WUA-flow curves largely matched. Therefore, the predicted variation in the WUA exerted by the hypothetical flow reduction was similar for both modelling approaches. However, the sign of the hypothetical change in the WUA varied, being positive for the large W.B trout and negative for the remaining size classes. Thus, in accordance with these divergent effects it has not been possible to evaluate the Greek state-legislated requirements for the minimum flow. Nevertheless, as a consequence of the agreement between the modelling approaches, we would advocate for combinations of GAMs and TSK fuzzy models in environmental flow assessment.

ACKNOWLEDGMENTS

This study was supported by the ECOFLOW project funded by the Hellenic General Secretariat of Research and Technology in the framework of the NSRF 2007-2013. We are grateful for field assistance of Dimitris Kommatas, Orfeas Triantafillou and Martin Palt and to Alcibiades N. Economou for assistance in discussions on trout biology and ecology.

REFERENCES

Ahmadi-Nedushan, B., St-Hilaire, A., Bérubé, M., Ouarda, T.B.M.J. and Robichaud, É, 2008. Instream flow determination using a multiple input fuzzy-based rule system: a case study. River Res. Appl. 24 (3), 279–292. http://dx.doi.org/10.1002/rra.1059.

Akaike, H., 1998. Information Theory and an Extension of the Maximum Likelihood Principle. In: Parzen, E., Tanabe, K. and Kitagawa, G.(ed.), Selected Papers of Hirotugu Akaike. Springer, New York, (USA), pp. 199-213.

Alcalá-Fdez, J., Alcalá, R., Gacto, M.J. and Herrera, F., 2009. Learning the membership function contexts for mining fuzzy association rules by using genetic algorithms. Fuzzy Sets Syst. 160 (7), 905-921. http://dx.doi.org/10.1016/j.fss.2008.05.012.

Allouche, O., Tsoar, A. and Kadmon, R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS).

J. Appl. Ecol. 43 (6), 1223–1232. http://dx.doi.org/10.1111/j.1365-2664.2006.01214.x.

Anderson, D.R., Burnham, K.P. and Thompson, W.L., 2000. Null hypothesis testing: Problems, prevalence, and an alternative. J. Wildl. Manage. 64 (4), 912-923. http://dx.doi.org/10.2307/3803199.

Arthington, A.H., Bunn, S.E., Poff, N.L. and Naiman, R.J., 2006. The challenge of providing environmental flow rules to sustain river ecosystems. Ecol. Appl. 16 (4), 1311-1318. http://dx.doi.org/10.1890/1051-0761(2006)016[1311:TCOPEF]2.0.CO;2.

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., 2009. Interactive effects of cover and hydraulics on brown trout habitat selection patterns. River Res. Appl. 25 (8), 1051-1065. http://dx.doi.org/10.1002/rra.1215.

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.

Beakes, M.P., Moore, J.W., Retford, N., Brown, R., Merz, J.E. and Sogard, S.M., 2014. Evaluating statistical approaches to quantifying juvenile Chinook salmon habitat in a regulated California river. River Res. Appl. 30 (2), 180-191. http://dx.doi.org/10.1002/rra.2632.

Belgiorno, V., Naddeo, V., Scannapieco, D., Zarra, T. and Ricco, D., 2013. Ecological status of rivers in preserved areas: Effects of meteorological parameters. Ecol. Eng. 53 (0), 173-182. http://dx.doi.org/10.1016/j.ecoleng.2012.12.039.

Benejam, L., Angermeier, P.L., Munné, A. and García-Berthou, E., 2010. Assessing effects of water abstraction on fish assemblages in Mediterranean streams. Freshw. Biol. 55 (3), 628-642. http://dx.doi.org/10.1111/j.1365-2427.2009.02299.x.

Boavida, I., Dias, V., Ferreira, M.T. and Santos, J.M., 2014. Univariate functions versus fuzzy logic: Implications for fish habitat modeling. Ecol. Eng. 71 533-538. http://dx.doi.org/10.1016/j.ecoleng.2014.07.073.

Bovee, K.D., 1978. Probability of use criteria for the family Salmonidae. Instream Flow Information Paper No. 4. FWS/OBS-78/07 Washington DC, (USA), pp. 53.

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.

Casillas, J., Cordón, O., del Jesus, M.J. and Herrera, F., 2005. Genetic tuning of fuzzy rule deep structures preserving interpretability and its interaction with fuzzy rule set reduction. IEEE Trans. Fuzzy Syst. 13 (1), 13-29. http://dx.doi.org/10.1109/TFUZZ.2004.839670.

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.

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.

Delling, B., 2010. Diversity of western and southern Balkan trouts, with the description of a new species from the Louros River, Greece (Teleostei: Salmonidae). Ichthyol. Explor. Freshwaters 21 (4), 331-344.

Döll, P., Fiedler, K. and Zhang, J., 2009. Global-scale analysis of river flow alterations due to water withdrawals and reservoirs. Hydrol. Earth Syst. Sci. 13 (12), 2413-2432. http://dx.doi.org/10.5194/hess-13-2413-2009.

Economou, A.N., Zogaris, S., Chatzinikolaou, Y., Tachos, V., Giakoumi, S., Kommatas, D., Koutsikos, N., Vardakas, L., Blasel, K. and Dussling, U., 2007. Development of an ichthyological multimetric index for ecological status assessment of Greek mountain streams and rivers Athens (Greece), pp. 166.

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.

Gibson, R.J., 1993. The Atlantic salmon in fresh water: spawning, rearing and production. Rev. Fish Biol. Fisher. 3 (1), 39-73. http://dx.doi.org/10.1007/bf00043297.

Guay, J.C., Boisclair, D., Rioux, D., Leclerc, M., Lapointe, M. and Legendre, P., 2000. Development and validation of numerical habitat models for juveniles of atlantic salmon (*Salmo salar*). Can. J. Fish. Aquat. Sci. 57 (10), 2065-2075. http://dx.doi.org/10.1139/cjfas-57-10-2065.

Hagen-Zanker, A., Straatman, B. and Uljee, I., 2005. Further developments of a fuzzy set map comparison approach. Int. J. Geogr. Inf. Sci. 19 (7), 769-785. http://dx.doi.org/10.1080/13658810500072137.

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

Heggenes, J., 1988a. Effects of short-term flow fluctuations on displacement of, and habitat use by, brown trout in a small stream. Trans. Am. Fish. Soc. 117 (4), 336-344. http://dx.doi.org/10.1577/1548-8659(1988)1172.3.CO;2.

Heggenes, J., 1988b. Substrate preferences of brown trout fry (*Salmo trutta*) in artificial stream channels. Can. J. Fish. Aquat. Sci. 45 (10), 1801-1806. http://dx.doi.org/10.1139/f88-211.

Heggenes, J., Baglinière, J.L. and Cunjak, R.A., 1999. Spatial niche variability for young Atlantic salmon (*Salmo salar*) and brown trout (*S. trutta*) in heterogeneous streams. Ecol. Freshw. Fish 8 (1), 1-21. http://dx.doi.org/10.1111/j.1600-0633.1999.tb00048.x.

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.

Jorde, K., 1997. Ökologisch begründete, dynamische Mindestwasserregelungen bei Ausleitungskraftwerken. Universität Stuttgart, Stuttgart (Germany), pp. 158.

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.

Kalogeropoulos, K. and Chalkias, C., 2013. Modelling the impacts of climate change on surface runoff in small Mediterranean catchments: Empirical evidence from Greece. Water Environ. J. 27 (4), 505-513. http://dx.doi.org/10.1111/j.1747-6593.2012.00369.x.

Katopodis, C., 2012. Ecohydraulic approaches in aquatic ecosystems: Integration of ecological and hydraulic aspects of fish habitat connectivity and Suitability. Ecol. Eng. 48 (0), 1-7. http://dx.doi.org/10.1016/j.ecoleng.2012.07.007.

Kottelat, M. and Freyhof, J., 2007. Handbook of European Freshwater Fishes. Kottelat & Freyhof Publishing, Cornol (Switzerland) & Berlin (Germany), 646 pp.

Lamouroux, N. and Souchon, Y., 2002. Simple predictions of instream habitat model outputs for fish habitat guilds in large streams. Freshw. Biol. 47 (8), 1531-1542. http://dx.doi.org/10.1046/j.1365-2427.2002.00880.x.

Li, R., Chen, Q., Tonina, D. and Cai, D., 2015. Effects of upstream reservoir regulation on the hydrological regime and fish habitats of the Lijiang River, China. Ecol. Eng. 76 (0), 75-83. http://dx.doi.org/10.1016/j.ecoleng.2014.04.021.

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

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.

Maddock, I., 1999. The importance of physical habitat assessment for evaluating river health. Freshwat. Biol. 41 (2), 373-391. http://dx.doi.org/10.1046/j.1365-2427.1999.00437.x.

Maggini, R., Lehmann, A., Zimmermann, N.E. and Guisan, A., 2006. Improving generalized regression analysis for the spatial prediction of forest communities. J. Biogeogr. 33 (10), 1729-1749. http://dx.doi.org/10.1111/j.1365-2699.2006.01465.x.

Mamdani, E.H., 1974. Application of fuzzy algorithms for control of simple dynamic plant. Proceedings of the Institution of Electrical Engineers 121 (12), 1585-1588. http://dx.doi.org/10.1049/piee.1974.0328.

Martínez-Capel, F., García De Jalón, D., Werenitzky, D., Baeza, D. and Rodilla-Alamá, M., 2009. Microhabitat use by three endemic Iberian cyprinids in Mediterranean rivers (Tagus River Basin, Spain). Fisheries Manag. Ecol. 16 (1), 52-60. http://dx.doi.org/10.1111/j.1365-2400.2008.00645.x.

Mathews, R. and Richter, B.D., 2007. Application of the indicators of hydrologic alteration software in environmental flow setting. J. Am. Water Resour. Assoc. 43 (6), 1400-1413. http://dx.doi.org/10.1111/j.1752-1688.2007.00099.x.

McClain, M.E., Subalusky, A.L., Anderson, E.P., Dessu, S.B., Melesse, A.M., Ndomba, P.M., Mtamba, J.O.D., Tamatamah, R.A. and Mligo, C., 2014. Comparing flow regime, channel hydraulics, and biological communities to infer flow-ecology relationships in the Mara River of Kenya and Tanzania. Hydrol. Sci. J. 59 (3-4), 801-819. http://dx.doi.org/10.1080/02626667.2013.853121.

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.

Ministry of Environment, Energy and Climate Change, 2011. 196978/2011. FEK 518/B/5/04/2011.

Mitaim, S. and Kosko, B., 2001. The shape of fuzzy sets in adaptive function approximation. IEEE Trans. Fuzzy Syst. 9 (4), 637-656. http://dx.doi.org/10.1109/91.940974.

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., De Baets, B. and Goethals, P.L.M., 2010. Ecological relevance of performance criteria for species distribution models. Ecol. Model. 221 (16), 1995-2002. http://dx.doi.org/10.1016/j.ecolmodel.2010.04.017.

Mouton, A.M., De Baets, B., Peter, A., Holzer, G., Müller, R., Goethals, P.L.M., 2008. Entropy-based fuzzy set optimisation for reducing ecological model complexity. Proceedings of the iEMSs Fourth Biennial Meeting: International Congress on Environmental Modelling and Software, Barcelona, (Spain), 1750-1757.

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. and De Pauw, N., 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., Lopez-Nicolas, A., Martínez-Capel, F. and Pulido-Velazquez, M., 2016. Shifts in the suitable habitat available for brown trout (*Salmo trutta* L.) under short-term climate change scenarios. Sci. Total Environ. 544 686–700. http://dx.doi.org/ 10.1016/j.scitotenv.2015.11.147.

Muñoz-Mas, R., Martínez-Capel, F., Alcaraz-Hernández, J.D. and Mouton, A.M., 2015. Can multilayer perceptron ensembles model the ecological niche of freshwater fish species? Ecol. Model. 309–310 (0), 72-81. http://dx.doi.org/10.1016/j.ecolmodel.2015.04.025.

Muñoz-Mas, R., Martínez-Capel, F., Garófano-Gómez, V. and Mouton, A.M., 2014. 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.

Orth, D.J. and Maughan, O.E., 1982. Evaluation of the incremental methodology for recommending instream flows for fishes. Trans. Am. Fish. Soc. 111 (4), 413–445. http://dx.doi.org/10.1577/1548-8659(1982)1112.0.co;2.

Ovidio, M., Baras, E., Goffaux, D., Birtles, C. and Philippart, J.C., 1998. Environmental unpredictability rules the autumn migration of brown trout (*Salmo trutta* L.) in the Belgian ardennes. Hydrobiologia 371-372 263-274. http://dx.doi.org/10.1023/A:1017068115183.

Papadaki, C., L. Ntoanidis, L., Zogaris, S., Martinez-Capel, F., Muñoz-Mas, R., Evelpidou, N. and Dimitriou, E., 2014. Habitat hydraulic modelling for environmental flow restoration in upland streams in Greece. 12th International Conference on Protection and Restoration of the Environment, Skiathos island, (Greece).

Payne, T.R. and Jowett, I.G., 2012. SEFA - Computer Software System for Environmental Flow Analysis based on the Instream Flow Incremental Methodology. 9th International Symposium on Ecohydraulics, Vienna, (Austria).

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. 10.1016/j.ecolmodel.2008.06.028.

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

Raleigh, R.F., Zuckerman, L.D. and Nelson, P.C., 1986. Habitat suitability index models and instream flow suitability curves: brown trout, revised FWS/OBS - 82/10.71, Washington DC, (USA), pp. 65.

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.

Rose, K.A., Roth, B.M. and Smith, E.P., 2009. Skill assessment of spatial maps for oceanographic modeling. J. Mar. Syst. 76 (1-2), 34-48. http://dx.doi.org/10.1016/j.jmarsys.2008.05.013.

Sánchez-Hernández, J. and Amundsen, P., 2015. Trophic ecology of brown trout (*Salmo trutta* L.) in subarctic lakes. Ecol. Freshwat. Fish 24 (1), 148-161. http://dx.doi.org/10.1111/eff.12139.

Schindler, D.E., Hilborn, R., Chasco, B., Boatright, C.P., Quinn, T.P., Rogers, L.A. and Webster, M.S., 2010. Population diversity and the portfolio effect in an exploited species. Nature 465 (7298), 609-612. http://dx.doi.org/10.1038/nature09060.

Schneider, M., 2001. Habitat - und Abflussmodellierung für Fließgewässer mit unscharfen Berechnungsansätzen. Universität Stuttgart, Stuttgart (Germany), pp. 180.

Strakosh, T.R., Neumann, R.M. and Jacobson, R.A., 2003. Development and assessment of habitat suitability criteria for adult brown trout in southern New England rivers. Ecol. Freshw. Fish. 12, 4265-274. http://dx.doi.org/10.1046/j.1600-0633.2003.00022.x.

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.

Tharme, R.E., 2003. A global perspective on environmental flow assessment: Emerging trends in the development and application of environmental flow methodologies for rivers. River Res. Appl. 19 (5-6), 397-441. http://dx.doi.org/10.1002/rra.736.

Tomsic, C.A., Granata, T.C., Murphy, R.P. and Livchak, C.J., 2007. Using a coupled eco-hydrodynamic model to predict habitat for target species following dam removal. Ecol. Eng. 30 (3), 215-230. http://dx.doi.org/10.1016/j.ecoleng.2006.11.006.

Vismara, R., Azzellino, A., Bosi, R., Crosa, G. and Gentili, G., 2001. Habitat suitability curves for brown trout (*Salmo trutta fario* L.) in the River Adda, Northern Italy: Comparing univariate and multivariate approaches. Regul. River. 17 (1), 37–50. http://dx.doi.org/10.1002/1099-1646(200101/02)17:13.0.CO;2-Q.

Visser, H. and De Nijs, T., 2006. The map comparison kit. Environ. Model. Softw. 21 (3), 346-358. http://dx.doi.org/10.1016/j.envsoft.2004.11.013.

Waters, B.F., 1976. A Methodology for Evaluating the Effects of Different Streamflows on Salmonid Habitat. Proceedings of the Symposium and Specialty Conference on Instream Flow Needs, Bathesda, MD (USA), 13.

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

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., 2013. On p-values for smooth components of an extended generalized additive model. Biometrika 100 (1), 221-228. http://dx.doi.org/10.1093/biomet/ass048.

Woodward, J.C., Hamlin, R.H.B., Macklin, M.G., Hughes, P.D. and Lewin, J., 2008. Glacial activity and catchment dynamics in northwest Greece: Long-term river behaviour and the slackwater sediment record for the last glacial to interglacial transition. Geomorphology 101 (1-2), 44-67. http://dx.doi.org/10.1016/j.geomorph.2008.05.018.

Yao, W., Rutschmann, P. and Sudeep, 2015. Three high flow experiment releases from Glen Canyon Dam on rainbow trout and flannelmouth sucker habitat in Colorado River. Ecol. Eng. 75 (0), 278-290. http://dx.doi.org/10.1016/j.ecoleng.2014.11.024.

Yi, Y., Cheng, X., Wieprecht, S. and Tang, C., 2014. Comparison of habitat suitability models using different habitat suitability evaluation methods. Ecol. Eng. 71 (0), 335-345. http://dx.doi.org/10.1016/j.ecoleng.2014.07.034.

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

Zika, U. and Peter, A., 2002. The introduction of woody debris into a channelized stream: effect on trout populations and habitat. River Res. Appl. 18 (4), 355-366. http://dx.doi.org/10.1002/rra.677.

Zogaris, S. and Economou, A.N., 2009. West Balkan Trout, *Salmo fariodes*. In: Legakis, A. and Maragou, P.(ed.), Red data book of threatened animals of Greece. Hellenic Zoological Society, Athens (Greece), pp. 141-143.

Zogaris, S., Chatzinikolaou, Y. and Dimopoulos, P., 2009. Assessing environmental degradation of montane riparian zones in Greece. J. Environ. Biol. 30 (5), 719-726.