

# Development of Generalized Additive Models (GAMs) for *Salmo rizeensis* Endemic to North-Eastern Streams of Turkey

Reza Rezaei<sup>1,\*</sup> , Hatice Sengül<sup>1</sup>

<sup>1</sup>Hacettepe University Department of Environmental Engineering, 06800 Beytepe, Ankara, Turkey.

## Article History

Received 10 November 2017  
Accepted 13 February 2018  
First Online 18 February 2018

## Corresponding Author

Tel.: +90.507 5854933  
E-mail: rezarezaei2008@gmail.com

## Keywords

*Salmo rizeensis*  
Generalized Additive Models  
Habitat analysis  
Micro-habitat variables

## Abstract

Generalized Additive Models (GAMs) are widely used statistical models for species distribution in habitat and environmental management modeling as they enable incorporation of non-linearity. The objective of this study is to develop relationships between habitat variables, fish absence/presence (logistic GAM) and abundance (Poisson GAM) for the *S. rizeensis* using on-site observations in Solaklı River. Logistic GAMs correctly predicted absence/presence of adult and juvenile *S. rizeensis* at 79.3% and 74% of sampling areas respectively. Response curves of logistic GAMs results show differences in probability of finding the adult and juvenile fish. Adult *S. rizeensis* were mostly found in higher velocities (0.5-0.8 m/s) than in juvenile fish ( $\approx 0$  m/s). Also, the highest number of adult fish were recorded in deeper habitats ( $\approx 0.4$  m) rather than in juvenile fish ( $\approx 0$  m). Curves show adult fish are more independent to the presence of cover compared to juvenile fish. Velocity shelter and bedrock formation were the most common cover types. Chi-square test results of predicted values showed that developed Poisson GAMs of adult and juvenile *S. rizeensis* could not accurately represent fish abundance. The results show that while logistic GAM is applicable, Poisson GAM model is not applicable for the area.

## Introduction

The Black Sea coast of Turkey is home to two trout species from *Salmonidae* family: *Salmo coruhensis* and *Salmo rizeensis*. *Salmo rizeensis*, which was previously reported as *Salmo trutta labrax* and *Salmo trutta macrostigma* respectively by Zengin and Aksungur (2008) and Verep *et al.* (2016), is an endemic freshwater trout species, which prefers highly elevated large river systems (Verep *et al.*, 2016). On-site visits revealed that the population of freshwater trout in Solaklı Basin is declining due to stream channel modifications for various purposes. Maintaining a viable species population necessitate the collection of information about their habitat conditions to ensure favorable conditions. Population status of high-level organisms in the aquatic food chain such as *S. rizeensis* is an indicator of ecosystem health. For *S. rizeensis*, there is a lack of knowledge on its favorable micro-habitat conditions.

Both biotic and abiotic habitat variables are important in long-term sustainability of the species. To date, several studies have focused on the effect of abiotic factors, including water depth, velocity and

substrate composition on aquatic organism's populations and distribution. Jacobson *et al.*, (2015) investigated the link between river management and pallid sturgeon population dynamics for the Missouri River, USA emphasizing the importance of environmental variables such as flow and sediment regime on fish behavior and population at different life stages of the fish. Freeman, Bowen, Bovee and Irwin (2001) studied the effect of environmental variables and flow rate on juvenile fish abundance in Tallapoosa River. They found that the juvenile fish abundance is strongly related to habitat variables and flow rate. Abundance was more frequently correlated with persistence<sup>1</sup> of shallow-water habitats than with habitat availability and value of flow extremes.

Englund and Krupa (2000) studied the effect of water depth as an environmental variable, besides the biotic factors and predator presence on crayfish habitat use. They found that the water depth is a determinant factor in habitat use by small and large crayfish. Fox, Brien, Collas, and Nash (2000) modeled egg production patterns and egg density of three marine fish species in the Irish Sea using environmental variables. Location,

depth, time, depth-integrated sea-temperature and salinity were incorporated into the presence/absence and egg production models. Grossman, Ratajczak, Crawford and Freeman (1998) classified the stream fish species in Coweeta Creek, North Carolina, USA based on environmental variables such as mean velocity, depth and substrate using statistical analysis. Hydrologic variation (annual flow variability) was noted as a key determinant, having a stronger impact than seasonal or annual cycles of habitat availability.

Ecological models are widely used for modeling and analysis of habitat conditions of species to improve our understanding of the relationship between habitat variables and the species abundance or inhabit probability. To date, several species distribution models have been developed for this purpose. Two types of statistical models are more commonly applied to investigate the effect of habitat conditions on population status and dynamics: Generalized Linear Models (GLMs) and Generalized Additive Models (GAMs). GLMs, developed by Nelder and Baker (1972), are being used to predict the species' response to different ecological variables. As environmental patterns are generally non-linear, linear models like GLMs falls short of describing ecological conditions (Yee & Mitchell, 1991). However, there are examples of successful prediction using GLMs (Young, Iampietro, Kvittek, & Garza, 2010).

To overcome the linearity assumption of GLMs, Generalized Additive Models (GAMs) were developed by Hastie and Tibshirani (1990). GAMs are effective models in establishing a relationship between predictor variables and the response. GAMs allow depiction of a wider range of response curves than GLMs.

Comparison of ecological models show that GAMs have a higher performance or at least as well as other models such as MARS, CART and ANN (Moisen & Frescino, 2002; Walsh & Kleiber, 2001). As a more flexible model, GAMs are able to show nonlinear relationship between a group of predictor variables and a response, and estimate the response based on these non-parametric functions (Ahmadi-Nedushan *et al.*, 2006; T. Hastie & Tibshirani, 2010).

GAMs have been widely used in ecological research such as forest biota and vegetation distribution modeling (Austin, 2002; Drexler & Ainsworth, 2013; Moisen *et al.*, 2006; Yee & Mackenzie, 2002), marine fisheries (Bergstrom, Sundblad, Downie, Snickars, & Lindgarth, 2013; Denis, Lejeune, & Robin, 2002; Elith *et al.*, 2006; França & Cabral, 2015), and freshwater fisheries (Alexander, 2016; Jowett, Parkyn, & Richardson, 2008; Leathwick, Elith, & Hastie, 2006). The results of ecological models such as GAMs can be crucial for decision making for environmental management.

There are several GAMs distribution functions including commonly applied additive logistic regression (Binomial), log-additive model (Poisson), gamma and negative-binomial distributions (Yee & Mackenzie,

2002). Logistic GAMs aim to discover the relationship between selected variables and binary response (i.e. absence/presence) whereas logarithmic (Poisson) GAMs establish the relationship between predictor variables and abundance of response variable. Logistic and logarithmic GAMs can be presented by Equations 1 and 2, respectively (Hastie & Tibshirani, 2010).

$$g(\mu) = \alpha + \sum_{j=1}^p f_j(x_j) \quad (1)$$

$$g(\mu) = \log(\mu) = \sum_{j=1}^p f_j(x_j) \quad (2)$$

Where,  $\mu$  is the mean of the response variable,  $\alpha$  is a constant,  $f_j$  are regression coefficients and  $X_j$  are measured values for predictor variables.

To improve our understanding of the relationships between environmental variables and species distribution, this paper presents results of GAMs development for the endemic *S. rizeensis* using data collected from the Haldizen (Uzungöl) branch of Solaklı Stream. The binomial and Poisson GAMs of adult and juvenile *S. rizeensis* were developed and compared for its accuracy in predicting species distribution.

## Methodology

### The Study Site

Figure 1 presents the location of the sampling area located in the Northeastern part of Turkey. Data collection was carried out in Haldizen (Uzungöl) branch of the Solaklı River, which had a relatively higher abundance of trout than Karaçam branch. The mainstream, on the other hand, has largely been affected by the ongoing modifications of the stream bed, and untreated wastewater discharge from small-scale nearby residential establishments and presence of multiple hydroelectric power plants.

In addition to anthropogenic effects, because of the preference of *S. rizeensis* of high elevations (Verer *et al.*, 2016), it is normal to see low numbers of *S. rizeensis* in the mainstream, which has a much lower elevation than Öğene (Karaçam) and Haldizen (Uzungöl) branches.

Sampling was carried out during medium annual flow rate of the stream in October of 2016. Two survey areas (each one about 300 meters in length) were selected for fish sampling on the stream.

The most dominant habitat types of the stream were rapid and run while the least commonly observed habitat types were glide and pool. Cobble, boulder and sand were the most dominant substrate types. Velocity shelter and bedrock formation were the most common cover types. There is not a complex cover composition at the study area, due to the scarcity of other types of cover such as undercut boulder, log jam, root wad and aquatic vegetation.

### Fish Sampling Method

The research team included four people that were actively involved in on-site data collection. The group electro-fished randomly selected points (quadrates) using a backpack electro fishing set, in the daytime. One fisherman used the electro fishing machine (SAMUS725MS) and held the fishing net in the downstream of the sampling point. A biologist measured fish length, a hydrologist gauged the velocity and depth at the sampling point, another person recorded the measured values and fish abundance for each quadrate. To reduce the effect of water turbulence caused by researchers, sampling points were adjusted to 5-6 meter intervals. A total of 150 sampling points was studied and the data on substrate composition, depth, mean velocity, distance to cover and the distance to the stream's edge, whether any fish was captured or not were recorded.

### Determination of Sampling Point Characteristics

Microhabitat variables were collected from quadrates of about 1 m<sup>2</sup> in the area. In addition to the measurement of mean velocity (at 0.6 times of depth

at each point or at 0.2 and 0.8 times of depth if the depth exceeds 1 m) and depth of quadrates, the percent of each substrate category in each quadrate were estimated visually following the criteria of Baker, Jowett, & Allibone (2003). The cover was classified according to Bovee (1986). For each quadrate, the distance to the nearest cover and stream edge were recorded.

The trout population was classified as juvenile and adult to eliminate the negative effect of possible differences in habitat selection behavior of juvenile and adult fish on GAM. The classification was based on the fish length. As reported by Zengin and Aksungur (2008) and Verep *et al.* (2016) in their studies of trout population in the same region, the minimum length of adolescence in *S. rizeensis* was 14 cm.

According to the length distribution of the sample from a total of 150 quadrates (Figure 2), 72 fish in the sample belong to the juvenile group and 39 fish to the adult fish group. From the sampled 150 quadrates, 20% of quadrates were occupied by adult and 30.6% by juvenile fish. Verep *et al.* (2016) reported 250 mm as the maximum length of *S. rizeensis*.

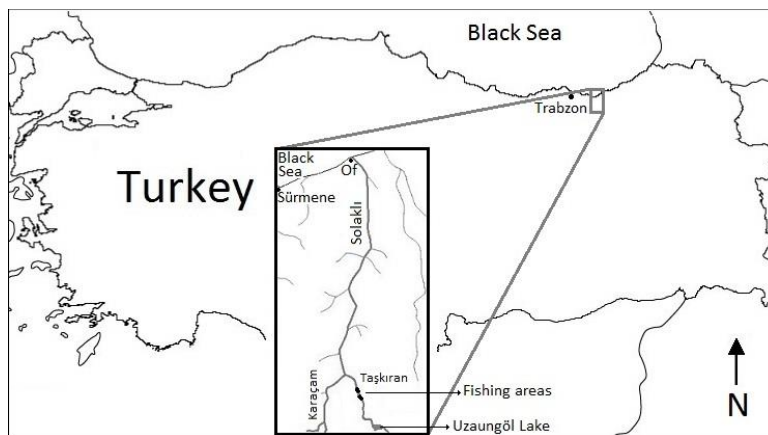


Figure 1. Sampling site location.

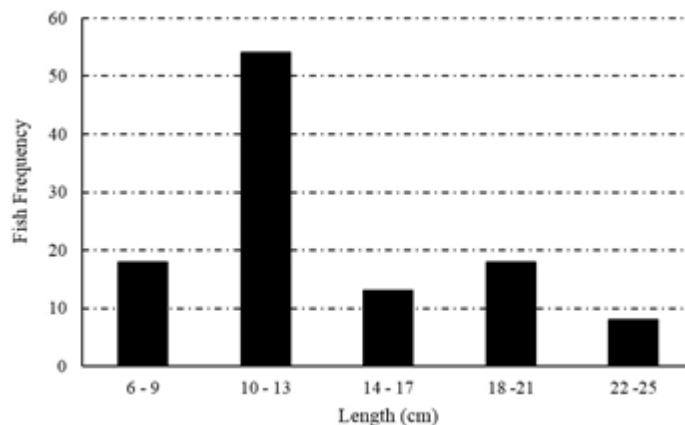


Figure 2. Length distribution of the sample.

## Development of GAM Models

In this study, physical habitat variables and developed logistic and logarithmic GAMs for adult and juvenile *S. rizeensis* were added to the models step-by-step. For all variables of Binomial and Poisson GAMs, the maximum df values were adjusted to 3. The cubic spline test was made to determine the non-linearity of model variables.

Model development were performed by using HSC development module of SEFA program (System for Environmental Flow Assessment) version 1.2 build 30 (Aquatic Habitat Analysts, 2012).

## The Sample Size

Accuracy of ecological models are sensitive to sample size (Cumming, 2000; Hernandez, Graham, Master, & Albert, 2006; Kadmon, Farber, & Danin, 2003; McPHERSON, JETZ, & Rogers, 2004; Stockwell & Peterson, 2002). Development of ecological models with low samples may result in unrepresented response (Carroll & Pearson, 1998; Pearce & Ferrier, 2000). There is also evidence that small, but adequate size of the sample does not affect model prediction performance. For example, Elith *et al.* (2006) did not find the evidence of sample size effect on GAM model prediction. They concluded that this finding may be related to the existence of adequate sampling points. According to Wisz *et al.* (2008), a sample size of 100 is adequate.

Our model is based on data collected in 150 points. 111 fish were captured at sampling points during the on-site study which meets an acceptable number of sample size. Partially homogeneous distribution of substrate structure, depth and habitat types along the stream and random sampling method reduce effects of probable sampling bias.

## Results

### Results of Logistic GAM

Table 1 presents logistic GAMs results for adult and juvenile fish. All variables used in the final GAM (last row on the left side) of adult fish (velocity, depth, substrate index, edge location and distance to cover) had a statistically significant effect on the model ( $P < 0.05$ ). The non-linearity test shows that the only statistically significant non-linear variable in adult fish logistic GAM is the depth (Table 2), with an optimum value of about 0.4 m (Figure 3).

The same variables for adult fish were used for developing logistic GAM of the juvenile fish group (Table 1). Edge location did not have a statistically significant effect on model, so it was not included in the final model (last row on the right side of Table 1). Also, the non-linearity test did not show a statistical significance for any of the variables (Table 2). The stepwise decreasing trend in residual deviance value at both adult and juvenile fish logistic GAMs in Table 1 shows the positive effects of variables on model

**Table 1.** Logistic generalized additive models of adult and juvenile *S. rizeensis*

Binomial model variables	Test for overall significance of adult fish logistic GAM			Binomial model variables	Test for overall significance of juvenile fish logistic GAM		
	Res. Dev.**	Res. df	P		Res. Dev.	Res. df	P
V*	145.62	146.14	0.2	V	171.11	146.07	0.01
V-D	136.15	143.5	0.02	V-D	166.25	143.21	0.01
V-D-Si	129.87	140.71	0.01	V-D-Si	162.24	140.29	0.02
V-D-Si-El	122.91	137.88	0.00	V-D-Si-Cd	158.29	137.56	0.03
V-D-Si-El-Cd	120.38	134.94	0.01				

\* V: Velocity, D: Depth, Si: Substrate index, El: Edge location and Cd: Distance to cover

\*\* Residual deviance

**Table 2.** Non-linearity test (cubic spline) results of logistic GAMs for adult and juvenile fish

Adult fish			Juvenile fish		
Variable	F	P	Variable	F	P
V*	0.96	0.38	V	1.07	0.34
D	6.11	0.00	D	0.37	0.67
Si	1.04	0.35	Si	1.02	0.36
El	0.59	0.55	Cd	1.05	0.34
Cd	1.32	0.27			

\* V: Velocity, D: Depth, Si: Substrate index, El: Edge location and Cd: Distance to cover

prediction.

Depth had the largest effect on fish presence of adult fish GAM response curves. According to the models result presented in Table 1, incorporation of depth to the model improved the model confidence level (from  $P=0.2$  to  $P=0.02$ ). As the distance of a location from the edge of the stream increases ( $> 1$  m), the probability of finding the adult fish gradually decreases. The logistic regression model correctly predicted absence-presence of adult fish at 79.3% of sampling areas (Cohen's kappa= 0.43 at 0.3 cut-off level) (Table 3).

For juvenile fish, velocity and depth were the most determinant micro-habitat variables affecting logistic GAM ( $P$  value=0.01) (Table 1). Logistic regression model correctly predicted absence-presence of juvenile fish at 74.0% of sampling areas (Cohen's kappa= 0.30 at 0.5 cutoff level).

There are differences in GAM response curves of juvenile and adult fish (Figure 3 and Figure 4). While the juvenile fish-finding probability decreases as the depth increase, there is a non-linear relationship between fish finding probability and increase in depth for adult fish. This indicates that the probability of adult fish finding increased as the depth increased up to about a depth of 0.4 m. Increase in the model

significance level ( $P$  value) with incorporation of the depth to the model confirms the importance of depth in the model (Table 1).

Juvenile fish logistic GAM results show that the optimum velocity for juvenile fish is zero (Figure 4). The sharp drop in finding probability curve related to velocity reveals the importance of this variable for juvenile fish. Contrarily, the adult fish presence increases moderately as velocity increases up to velocities of 0.5-0.8 m/s (Figure 3).

The distance to cover curves of adult and juvenile fish showed different patterns. As the distance to cover increases ( $> 1$  m) juvenile fish finding probability decreases, but there is an increase in adult fish finding probability as the distance to cover increases (up to about 2.5 m), as seen in Figure 3 and Figure 4 Also, as the substrate index value increases (at substrate index  $> 5$ ), adult and juvenile fish finding probability decreases.

### Results of Poisson GAM

In addition to logistic GAM, Poisson GAMs were developed as a fish abundance prediction model. Variables incorporated in the model included velocity, depth, substrate index, edge location and distance to

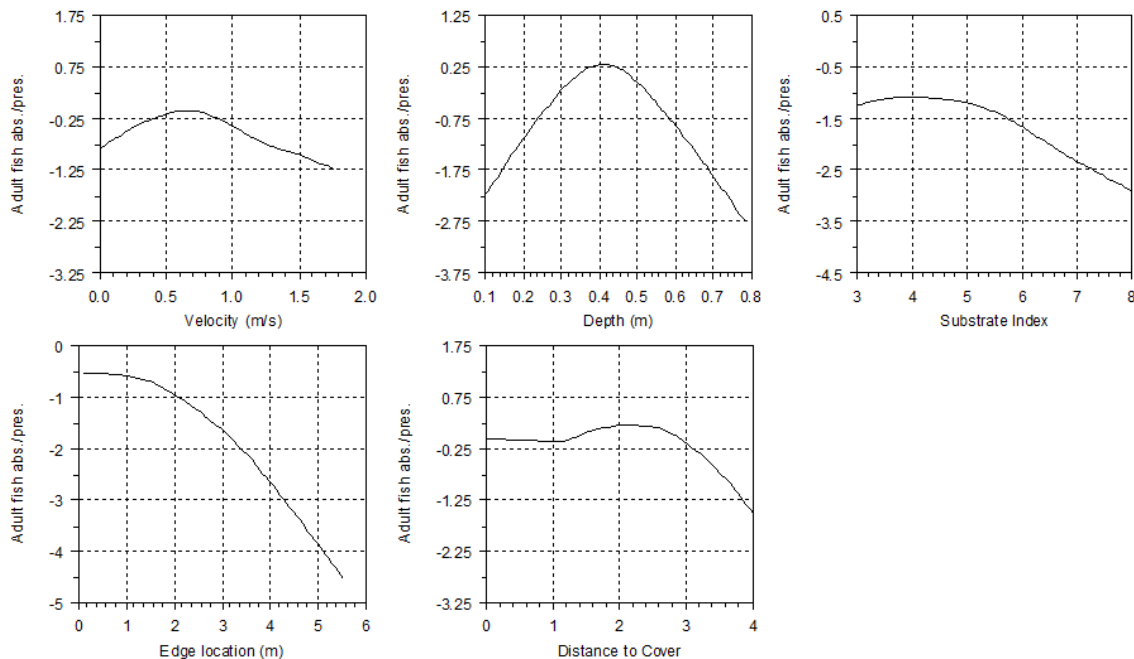


Figure 3. Response curves of variables of logistic GAM for adult *S. rizeensis*.

Table 3. Estimate of adult and juvenile fish absence/presence

	Cutoff level	Cohen's kappa	Correct%
Adult fish	0.3	0.43	79.3
Juvenile fish	0.5	0.30	74.0

cover and fish abundance (Table 4) for adult and juvenile fish.

Similar to adult fish logistic GAM, depth is the only statistically significant non-linear variable in adult Poisson GAM (Table 5), with an optimum value of about 0.4 m (Figure 5). The probability of finding adult fish increased as the velocity increased up to 0.7 m/s. Also, as the distance to cover increased (> 1 m) the probability of finding adult fish decreased.

Similar variables were used in juvenile fish Poisson

GAM development. Velocity is a statistically significant non-linear variable.

The finding probability of juvenile fish decreases with an increase in velocity (Figure 6). As most of the juvenile fish were present in more stagnant water (velocity near zero), there is a sharp decline in the probability of fish finding curve with an increase in velocity. Also, the optimum value of depth for juvenile fish is about 0.3-0.4 m. Similar to logistic GAM, the probability of juvenile fish finding decreases as the

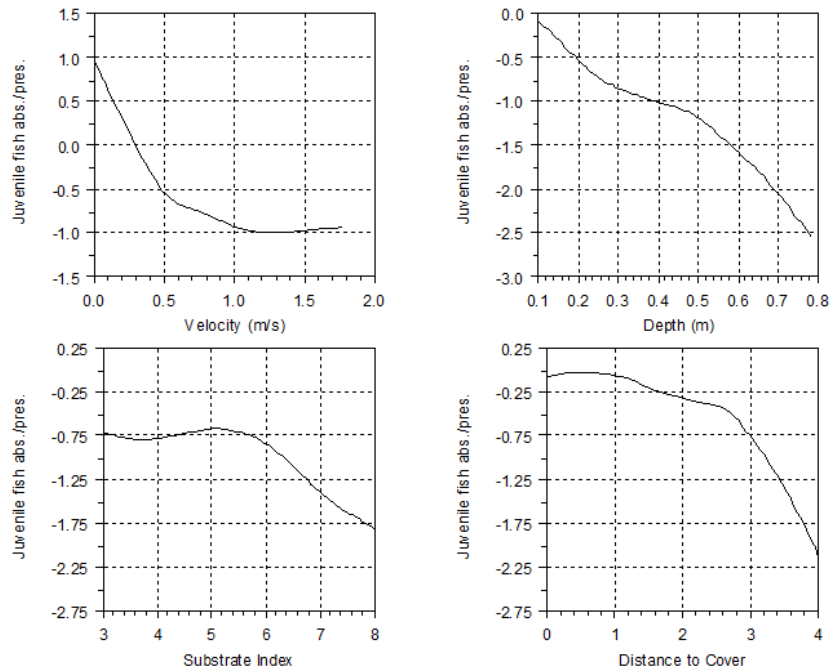


Figure 4. Response curves of variables contributed in the logistic GAM of juvenile *S. rizeensis*

Table 4. Poisson generalized additive models of adult and juvenile *S. rizeensis*

Poisson model variables	Test for overall significance of adult fish Poisson GAM			Poisson model variables	Test for overall significance of juvenile fish Poisson GAM		
	Res. Dev.	Res. df	P		Res. Dev.	Res. df	P
V	97.71	115.95	0.02	V	120.21	116.08	0.00
V-D	87.4	113.23	0.00	V-D	117.8	113.06	0.00
V-D-Si	80.19	110.36	0.00	V-D-Si	113.97	110.07	0.00
V-D-Si-El	77.72	107.56	0.00	V-D-Si-El	110.16	107.32	0.00
V-D-Si-El-Cd	74.85	104.71	0.00	V-D-Si-El-Cd	104.91	104.58	0.00

Table 5. Non-linearity test (cubic spline) results of Poisson GAMs for adult and juvenile fish

Adult fish			Juvenile fish	
Variable	F	P	Variable	F
V	2.50	0.09	V	3.33
D	9.02	0.00	D	0.81
Si	0.87	0.42	Si	2.02
El	0.81	0.44	El	1.65
Cd	0.91	0.40	Cd	1.93

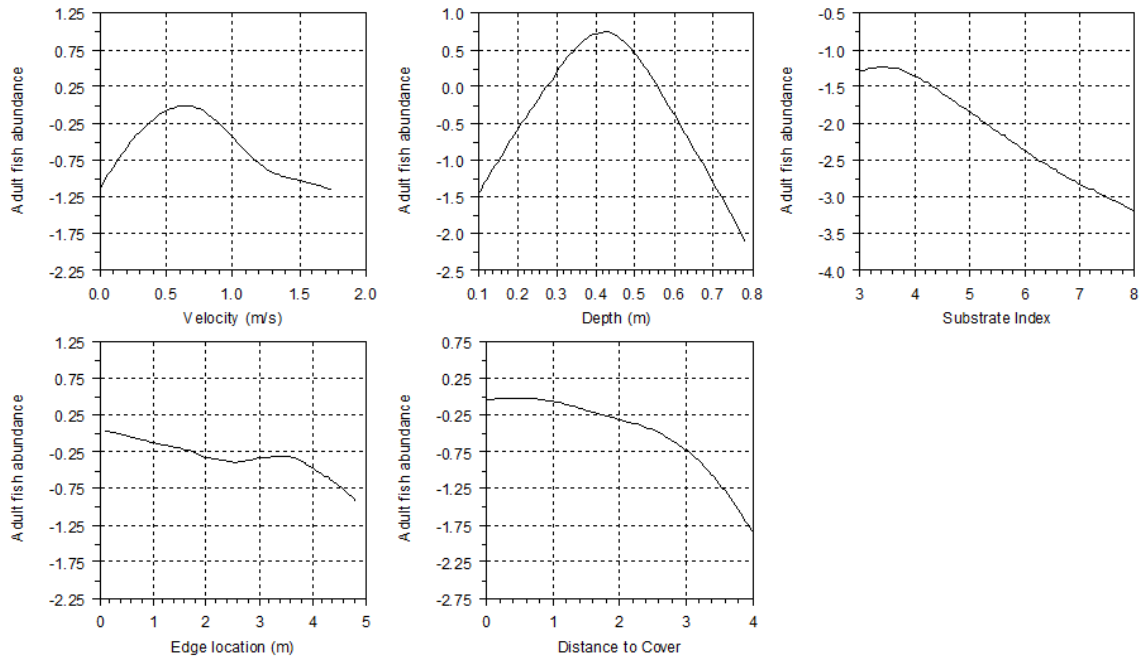


Figure 5. Response curves of variables contributed in the Poisson GAM of adult *S. rizeensis*.

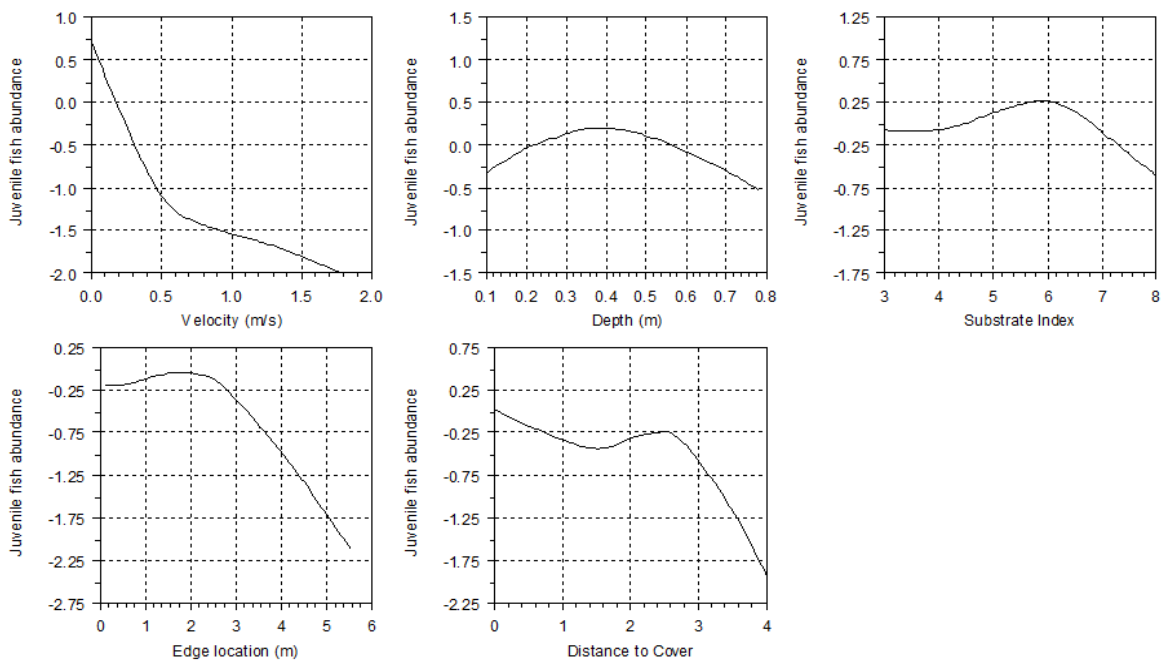


Figure 6. Response curves of variables contributed in the Poisson GAM of juvenile *S. rizeensis*

distance to cover increases.

Despite the overall abundance prediction value of the models being close to observed values (approximately, 25% error for the adult Poisson model and 11% error for the juvenile Poisson model), both Poisson models over-predicted the fish abundance at unoccupied quadrates and under-predicted the abundance at occupied quadrates (Figure 7). Hence, Poisson GAM is not suitable for estimating abundance.

Chi-square test results showed that the  $X^2$  value of the models (58.07 and 59.56 for adult and juvenile fish Poisson models respectively) were much higher than the maximum acceptable value (40.11, at  $\alpha=0.05$ ).

Contour plots of adult and juvenile fish occurrence with respect to depth and velocity are presented in Figure 8. The contour plot shows the velocity and depth interactions' effect on the fish presence, aforementioned above.

## Discussion

Logistic GAMs correctly predict the absence/presence of adult and juvenile fish at 79.3% and 74% of quadrat respectively (Table 3). The higher accuracy of juvenile fish logistic GAM than that of adult fish is related to the larger sample size of juvenile fish. In general, logistic and Poisson models show that hydraulic variables (depth and velocity) are determinant factors in predicting both adult and juvenile fish presence.

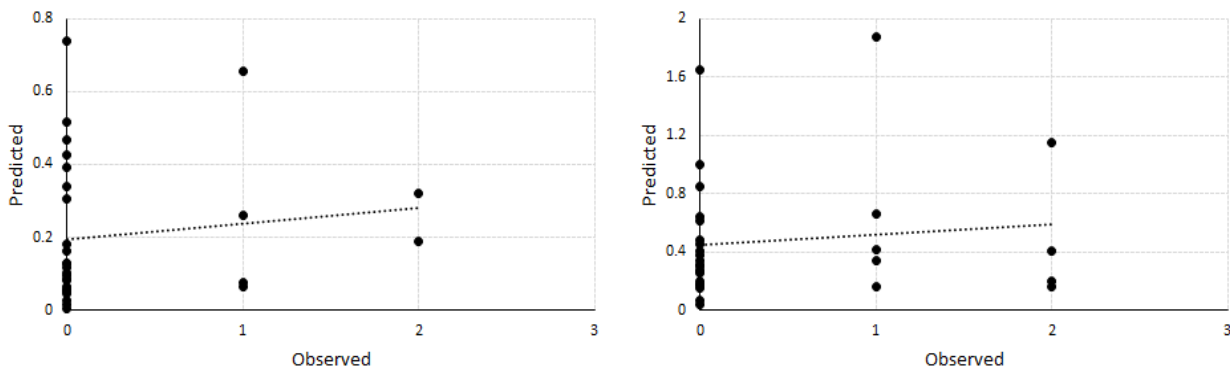
Velocity curves of adult and juvenile trout shows that the adult fish can better tolerate the physical stress of higher velocities compared to the juvenile fish. Logistic GAMs showed that the adult fish prefer deeper habitats than the juvenile fish.

Depth, substrate index and edge location have a positive effect on adult fish logistic GAM, as seen by decreasing  $P$  value of the model (Table 1). As the water depth gradually increases as the distance to the river edge increases due to river bed topography of the

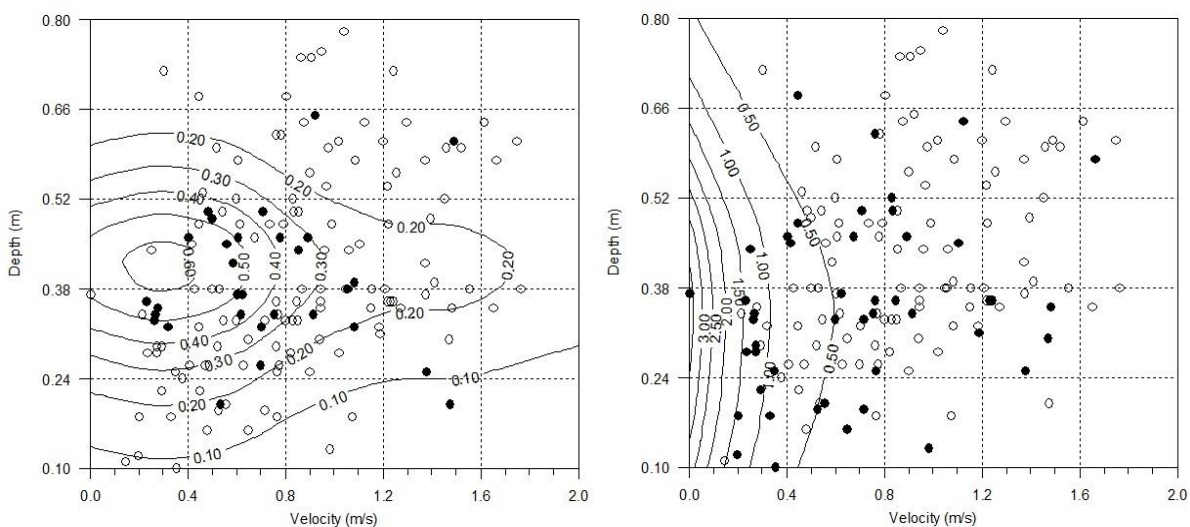
sampling area, we can assume that the distance to the edge is directly related to depth. Hence, the depth is the most important variable of adult fish logistic GAM.

Studies in different regions that also develop bivariate and multivariate models approve the controlling role of depth as a hydraulic variable (Ayllon, Almodovar, Nicola, & Elvira, 2009; Vismara, Azzellino, Bosi, Crosa, & Gentili, 2001). Munoz-Mas *et al.* (2016) study results on *Salmo farioides* showed that the large fish prefer low velocity and deep parts. The pool sections of the river are where the large fish could be found. Similar results were reported for large brown trout by Bovee (1978).

Supporting the findings of the study, the difference in water depth selection patterns between adult and juvenile fish was also reported by Ayllon *et al.* (2009). Their findings showed that as fish body sizes of trout species increase, they tend to select deeper and slower flowing sections of the stream. Selection of shallower areas by small size fish could be related to habitat competition between large and small fish.



**Figure 7.** Comparison between observed and predicted abundance of adult (left) and juvenile (right) *S. rizeensis* (X and Y axis indicate fish abundance).



**Figure 8.** Contour plots of adult (left) and juvenile (right) fish abundance using depth and velocity (filled and blank dots indicate occupied and unoccupied quadrates, respectively)



Because the study area is located in a mountainous region, the stream velocity is relatively high, except the shallow parts in the river bank and there are a few pools. Also, there are rare amounts of cover elements such as instream vegetation, undercut banks and log jams in the stream. Adult fish can only use water turbulence and depth, which has higher velocity, as cover elements due to these habitat characteristics. Deeper areas provide a form of overhead cover for fish (Benndorf, 1989; Bovee *et al.*, 1998).

The preference of shallow waters by juvenile fish is also related to its higher motivation for foraging (Ayllon *et al.*, 2009). The juvenile fish uses shallow parts of the stream which usually have low velocities to minimize energy consumption due to swimming. Low velocity enhances food deposition and therefore increases food availability. The optimum value of velocity for juvenile trout abundance is about zero, which is in agreement with findings of Munoz-Mas *et al.* (2016) and Jowett *et al.* (2008).

As mentioned in the methodology section, there is not a complex cover structure in the study area. About 75% of covers in the sampling areas are velocity shelters. Presence of velocity shelter is a more important determinant for juvenile fish compared to adult fish. Juvenile fish cannot tolerate high velocities. The need for the presence of velocity shelters for energy preservation for growth of juvenile trout has been emphasized in several studies (Fausch, 1984; Hughes & Dill, 1990).

Our model supports the need for the presence of cover for juvenile fish. Model results show that the juvenile *S. rizeensis* prefer cobble and gravel substrates. Similar results have been reported for small and medium trout in literature (Ayllon *et al.*, 2009; Bovee, 1978). Decrease in adult fish finding probability by increase in substrate index is in contrast with previous studies on trout species (Ayllon *et al.*, 2009; Bovee, 1978; Munoz-Mas *et al.*, 2016).

Chi-square test result shows Poisson GAMs of adult and juvenile fish could not correctly predict the fish abundance at quadrates, although there are similarities between logistic and Poisson GAMs curves. As there is no change in significance level (*P* value) of the Poisson GAMs prediction by addition of more variables in subsets (Table 4), it can be inferred that the models were unable to explain the statistical significance difference between the variables. It is likely that low population of fish is the major cause of lack of sufficient accuracy of Poisson models.

The close similarity between velocity, depth and substrate index curves of logistic and Poisson GAMs of adult fish is associated with low frequency of adult fish in each quadrate (mostly equal to 1). In other words, abundance data is very close to binomial (absence/presence) data.

## Conclusion

Habitat plays an important role on species population (Love, Schroeder, Lenarz, & Cochrane, 2006). The low population of *S. rizeensis* in the studied area necessitates taking protective actions, including habitat modification strategies by respective regulatory agencies. The lack of knowledge on habitat conditions and preferences of aquatic organisms can lead to misapplications that can negatively affect population dynamics. Channelizing of a river, for example, is one of the most commonly encountered misapplications. As channelization of a river has many habitat implications such as decrease in food productivity, limiting pool areas, increase in water velocity and elimination of natural covers in the river body, resulting in an overall decrease in area suitable for fish and decreases growth and survival conditions (Jacobson *et al.*, 2015; Jowett *et al.*, 2008; Pretty *et al.*, 2003). Several correct habitat improvement actions may be taken (Jacobson *et al.*, 2015; Kondratieff & Richer, 2014; Santos, Arau, & Brotto, 2008), among which include:

1. change in river width and flow management; which is strongly related to stream morphology,
2. increase in cover number by deploying artificial structures, and
3. excavation of pools.

In order to establish effective regulatory mechanisms, a reliable overview of favorable habitat conditions of each species is needed. Using habitat based ecological models selection and design of conservation areas or re-evaluation of existing areas can be made (Love *et al.*, 2006; Pittman, Christensen, Caldwell, Menza, & Monaco, 2007). Although some of these models are not able to represent habitat needs of all species inhabiting a specific area, they can be developed for ecologically and/or economically valuable species (Lindeman, Pugliese, Waugh, & Ault, 2000). GAMs have proved their ability of disclosing non-linear relationship patterns between variables and response, with their great flexibility (Venables & Ripley, 2002).

Further studies may focus on interactions such as food availability and velocity. Availability of food has a controlling role in fish habitat selection. Analyses of the food content of each quarter may prove to be challenging, but it may improve the prediction accuracy of the model.

## Acknowledgements

Funding for this research was provided by Hacettepe University Research Fund, Scientific Research Projects Coordination Unit (project number: 5364). We would like to thank Prof. Dr. Aydın Akbulut

for his valuable assistance during fieldwork. We also thank DSİ 22nd Regional Directorate personnel (Cihan Ziya Bayraktar, Mehmet Aksoy and Muhammet Aydoğan), for helping us with gauging of water velocity within quadrates during on-site observation. We also greatly appreciate the editor and two anonymous reviewers for their valuable comments.

## References

- Ahmadi-Nedushan, B., St-Hilaire, A., Berube, M., Robichaud, E., Thiemonge, N., & Bobee, B. (2006). A review of statistical methods for the evaluation of aquatic habitat suitability for instream flow assessment. *River Research and Applications*, 22(5), 503–523. <https://dx.doi.org/10.1002/rra.918>
- Alexander, R. E. (2016). *A comparison of GLM, GAM, and GWR modeling of fish distribution and abundance in Lake Ontario* (MSc Thesis). University of Southern California, Los Angeles, California, USA.
- Aquatic Habitat Analysts, I. (2012). SEFA program (System for Environmental Flow Assessment). Retrieved from <http://sefa.co.nz/>
- Austin, M. P. (2002). Spatial prediction of species distribution: An interface between ecological theory and statistical modelling. *Ecological Modelling*, 157(2–3), 101–118. [https://dx.doi.org/10.1016/S0304-3800\(02\)00205-3](https://dx.doi.org/10.1016/S0304-3800(02)00205-3)
- Ayllon, D., Almodovar, A., Nicola, G. G., & Elvira, B. (2009). Interactive effects of cover and hydraulics on Brown Trout habitat selection patterns. *River Research and Applications*, 25, 1051–1065. <https://dx.doi.org/10.1002/rra>
- Baker, C. F., Jowett, I. G., & Allibone, R. M. (2003). Habitat use by non-migratory Otago galaxiids and implications for water management. Science for Conservation Press, (221), 33 pp.
- Benndorf, J. (1989). Predation: direct and indirect impacts on aquatic communities. *International Review of Hydrobiology*, 74(3), 351–352.
- Bergstrom, U., Sundblad, G., Downie, A., Snickars, M., & Lindgarth, M. (2013). Evaluating eutrophication management scenarios in the Baltic Sea using species distribution modelling. *Journal of Applied Ecology*, 680–690. <https://dx.doi.org/10.1111/1365-2664.12083>
- Bovee, K. D. (1978). Probability of use criteria for the family Salmonidae (Report No. 4). Missoula, USA, Cooperative Instream Flow Service Group, 90 pp.
- Bovee, K. D. (1986). Development and evaluation of habitat suitability criteria for use in the instream flow incremental methodology (Report No. 21). Washington DC, USA, National Ecology Center, 263 pp.
- Bovee, K. D., Lamb, B. L., Bartholow, J. M., Stalnaker, C. B., Taylor, J., Henriksen (1998). Stream habitat analysis using the instream flow incremental methodology (Report No. USGS/BRD/ITR-1998-0004). Fort Collins, USA, U.S. Geological Survey, Biological Resources Division Information and Technology, 131 pp.
- Carroll, S. S., & Pearson, D. L. (1998). The effects of scale and sample size on the accuracy of spatial predictions of Tiger Beetle (Cicindelidae) species richness. *Ecography*, 21(4), 401–414. Retrieved from <http://www.jstor.org/stable/3683175>.
- Cumming, G. S. (2000). Using between-model comparisons to fine-tune linear models of species ranges. *Journal of Biogeography*, 27(2), 441–455. <https://dx.doi.org/10.1046/j.1365-2699.2000.00408.x>
- Denis, V., Lejeune, L., & Robin, J. P. (2002). Spatio-temporal analysis of commercial trawler data using General Additive models: patterns of Loliginid abundance in the north-east Atlantic. *ICES Journal of Marine Science*, 59, 633–648. <https://dx.doi.org/10.1006/jmsc.2001.1178>
- Drexler, M., & Ainsworth, C. H. (2013). Generalized additive models used to predict species abundance in the Gulf of Mexico: an ecosystem modeling tool. *PLoS ONE*, 8(5). <https://dx.doi.org/10.1371/journal.pone.0064458>
- Elith, J., Graham, C. H., Anderson, R. P., Dudik, M., Ferrier, S., Guisan, A., & Zimmermann, N. E. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29(2), 129–151. <https://dx.doi.org/10.1111/j.2006.0906-7590.04596.x>
- Englund, G., & Krupa, J. J. (2000). Habitat use by crayfish in stream pools: influence of predators, depth and body size. *Freshwater Biology*, 43, 75–83.
- Fausch, K. D. (1984). Profitable stream positions for salmonids: relating specific growth rate to net energy gain. *Canadian Journal of Zoology*, 62(3), 441–451. <https://dx.doi.org/10.1139/z84-067>
- Fox, C. J., Brien, C. M. O., Collas, M. D., & Nash, R. D. M. (2000). Patterns in the spawning of cod (*Gadus morhua* L.), sole (*Solea solea* L.) and plaice (*Pleuronectes platessa* L.) in the Irish Sea as determined by generalized additive modelling. *Fisheries Oceanography*, 9(1), 33–49.
- França, S., & Cabral, H. N. (2015). Environmental modelling & software Predicting fish species richness in estuaries: Which modelling technique to use? *Environmental Modelling and Software*, 66(2015), 17–26. <https://dx.doi.org/10.1016/j.envsoft.2014.12.010>
- Freeman, M. C., Bowen, Z. H., Bovee, K. D., & Irwin, E. R. (2001). Flow and habitat effects on juvenile fish abundance in natural and altered flow regimes. *Ecological Applications*, 11(1), 179–190. <https://dx.doi.org/10.2307/3061065>.
- Grossman, G. D., Ratajczak, R. E., Crawford, M., & Freeman, M. (1998). Assemblage organization in stream fishes: effects of environmental variation and interspecific interactions. *Ecological Monographs*, 68(3), 395–420. <https://dx.doi.org/10.2307/2657245>.
- Hastie, T. J., & Tibshirani, R. J. (1990). Generalized additive models. New York, USA, CRC Press, 352 pp.
- Hastie, T., & Tibshirani, R. (2010). Generalized Additive Models. *Statistical Science*, 1(3), 297–310. <https://dx.doi.org/10.1214/ss/1177013604>
- Hernandez, P. A., Graham, C., Master, L. L., & Albert, D. L. (2006). The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, 29(5), 773–785. <https://dx.doi.org/10.1111/j.0906-7590.2006.04700.x>
- Hughes, N. F., & Dill, L. M. (1990). Position choice by drift-feeding salmonids: model and test for Arctic grayling (*Thymallus arcticus*) in subarctic mountain streams, interior Alaska. *Canadian Journal of Fisheries and Aquatic Sciences*, 47(10), 2039–2048. <https://dx.doi.org/10.1139/f90-228>.
- Jacobson, R. B., Parsley, M. J., Annis, M. L., Colvin, M. E., Welker, T. L., & James, D. A. (2015). Development of Conceptual Ecological Models Linking Management of the Missouri River to Pallid Sturgeon Population

- Dynamics (Report No. 2015–1038). Virginia, USA, U.S. Geological Survey, 54 pp.
- Jowett, I. G., Parkyn, S. M., & Richardson, J. (2008). Habitat characteristics of crayfish (*Paraneohorsos planifrons*) in New Zealand streams using generalised additive models (GAMs). *Hydrobiologia*, 596(1), 353–365. <https://dx.doi.org/10.1007/s10750-007-9108-z>.
- Kadmon, R., Farber, O., & Danin, A. (2003). A systematic analysis of factors affecting the performance of climatic envelope models. *Ecological Applications*, 13(3), 853–867. [https://dx.doi.org/10.1890/1051-0761\(2003\)013\[0853:ASAOFA\]2.0.CO;2](https://dx.doi.org/10.1890/1051-0761(2003)013[0853:ASAOFA]2.0.CO;2).
- Kondratieff, M. C., & Richer, E. E. (2014). Stream Habitat Investigations and Assistance (Progress Report For Project Number: F-161-R-20). Colorado, USA, Colorado Division of Wildlife, Aquatic Wildlife Research Section, 73 pp.
- Leathwick, J. R., Elith, J., & Hastie, T. (2006). Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions, 9, 188–196. <https://dx.doi.org/10.1016/j.ecolmodel.2006.05.022>
- Lindeman, K. C., Pugliese, R., Waugh, G. T., & Ault, J. S. (2000). Developmental patterns within a multispecies reef fishery : management applications for essential fish habitats and protected areas. *Bulletin of Marine Science*, 66(3), 929–956.
- Love, M. S., Schroeder, D. M., Lenarz, B., & Cochrane, G. R. (2006). Gimme shelter: The importance of crevices to some fish species inhabiting a deeper-water rocky outcrop in Southern California. *CalCOFI Rep*, 47, 119–126.
- McPherson, J. M., Jetz, W., & Rogers, D. J. (2004). The effects of species' range sizes on the accuracy of distribution models: ecological phenomena or statistical artefact?. *Journal of Applied Ecology*, 41(5), 811–823. <https://dx.doi.org/10.1111/j.0021-8901.2004.00943.x>
- Moisen, G. G., Freeman, E. A., Blackard, J. A., Frescino, T. S., Zimmermann, N. E., & Edwards, T. C. (2006). Predicting tree species presence and basal area in Utah: a comparison of stochastic gradient boosting, generalized additive models, and tree-based methods. *Ecological Modelling*, 199(2), 176–187. <https://dx.doi.org/10.1016/j.ecolmodel.2006.05.021>.
- Moisen, G. G., & Frescino, T. S. (2002). Comparing Five Modelling Techniques for Prediction Forest Characteristics. *Ecological Modelling*, 157, 209–225. [https://dx.doi.org/10.1016/S0304-3800\(02\)00197-7](https://dx.doi.org/10.1016/S0304-3800(02)00197-7)
- Munoz-Mas, R., Papadaki, C., Martínez-capel, F., Zogaris, S., Ntoanidis, L., & Dimitriou, E. (2016). Generalized additive and fuzzy models in environmental flow assessment : A comparison employing the West Balkan trout ( *Salmo farioides* ; Karaman , 1938 ). *Ecological Engineering*, 91, 365–377. <https://dx.doi.org/10.1016/j.ecoleng.2016.03.009>
- Nelder, J. A., & Baker, R. J. (1972). Generalized linear models. *Journal of the Royal Statistical Society*, 135(3), 370-384. <https://dx.doi.org/10.2307/2344614>
- Pearce, J., & Ferrier, S. (2000). An evaluation of alternative algorithms for fitting species distribution models using logistic regression. *Ecological Modelling*, 128(2), 127–147. [https://dx.doi.org/10.1016/S0304-3800\(99\)00227-6](https://dx.doi.org/10.1016/S0304-3800(99)00227-6)
- Pittman, S. J., Christensen, J. D., Caldow, C., Menza, C., & Monaco, M. E. (2007). Predictive mapping of fish species richness across shallow-water seascapes in the Caribbean. *Ecological Modelling*, 204(1–2), 9–21. <https://dx.doi.org/10.1016/j.ecolmodel.2006.12.017>
- Pretty, J. L., Harrison, S. S. C., Shepherd, D. J., Smith, C., Hildrew, A. G., & Hey, R. D. (2003). River rehabilitation and fish populations : assessing the benefit of instream structures. *Journal of Applied Ecology*, 40, 251–265. <https://dx.doi.org/10.1046/j.1365-2664.2003.00808.x>.
- Santos, L. N., Arau, F. G., & Brotto, D. S. (2008). Artificial structures as tools for fish habitat rehabilitation in a neotropical reservoir. *Aquatic Conservation: Marine And Freshwater Ecosystems*, 18(6), 896–908. <https://dx.doi.org/10.1002/aqc.931>
- Stockwell, D. R. ., & Peterson, A. T. (2002). Effects of sample size on accuracy of species distribution models. *Ecological Modelling*, 148(1), 1–13. [https://dx.doi.org/10.1016/S0304-3800\(01\)00388-X](https://dx.doi.org/10.1016/S0304-3800(01)00388-X)
- Venables, W. N., & Dichmont, C. M. (2004). GLMs, GAMs and GLMMs: An overview of theory for applications in fisheries research. *Fisheries Research*, 70(2–3 SPEC. ISS.), 319–337. <https://dx.doi.org/10.1016/j.fishres.2004.08.011>
- Verep, B., Turan, D., Bilgin, S., Terzi, E., Kaya, C., & Mutlu, T. (2016). Abundance of Natural Trout Stocks, Interaction of Anthropogenic Processes and Hydro Power Plants in the River Basin of Rize City. *Journal of Anatolian Environmental&Animal Sciences*, 2016(2), 56–63.
- Vismara, R., Azzellino, A., Bosi, R., Crosa, G., & Gentili, G. (2001). Habitat suitability curves for brown trout (*Salmo trutta fario* L.) in the River Adda, Northern Italy: comparing univariate and multivariate approaches. *Regulated Rivers: Research & Management*, 17(1), 37–50. [https://dx.doi.org/10.1002/1099-1646\(200101/02\)17:1<37::AID-RRR606>3.0.CO;2-Q](https://dx.doi.org/10.1002/1099-1646(200101/02)17:1<37::AID-RRR606>3.0.CO;2-Q)
- Walsh, W. A., & Kleiber, P. (2001). Generalized additive model and regression tree analyses of blue shark (*Prionace glauca*) catch rates by the Hawaii-based commercial longline fishery. *Fisheries Research*, 53(2), 115–131. [https://dx.doi.org/10.1016/S0165-7836\(00\)00306-4](https://dx.doi.org/10.1016/S0165-7836(00)00306-4)
- Wisn, M. S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H., Guisan, A. and NCEAS Predicting Species Distributions Working Group (2008). Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, 14(5), 763–773. <https://dx.doi.org/10.1111/j.1472-4642.2008.00482.x>
- Yee, T., & Mitchell, N. (1991). Generalized additive models in plant ecology. *Journal of Vegetation Science*, 2(5), 587–602. <https://dx.doi.org/10.2307/3236170>.
- Yee, T. W., & Mackenzie, M. (2002). Vector generalized additive models in plant ecology. *Ecological Modelling*, 157(2–3), 141–156. [http://dx.doi.org/http://dx.doi.org/10.1016/S0304-3800\(02\)00192-8](http://dx.doi.org/http://dx.doi.org/10.1016/S0304-3800(02)00192-8)
- Young, M. A., Iampietro, P. J., Kvitek, R. G., & Garza, C. D. (2010). Multivariate bathymetry-derived generalized linear model accurately predicts rockfish distribution on Cordell Bank. *Marine Ecology Progress Series*, 415, 247–261. <https://dx.doi.org/10.3354/meps08760>
- Zengin, M., & Aksungur, M. (2008). Effects of environmental factors on the reproduction migration of the Black Sea trout (*Salmo trutta labrax*) in the Eastern Black Sea Region (Turkey). Antalya, Turkey, European Inland Fisheries Advisory Commission.