


RESEARCH ARTICLE

WILEY

Environmental drivers of benthic fish assemblages and fish-habitat associations in offshore areas of a very large river

Zoltan Szalóky¹ | Vivien Füstös² | Balazs Tóth³ | Tibor Erős¹ 

¹Balaton Limnological Institute, Centre for Ecological Research, Tihany, Hungary

²Department of Hydraulic and Water Resources Engineering, Budapest University of Technology and Economics, Budapest, Hungary

³Danube-Ipoly National Park Directorate, Budapest, Hungary

Correspondence

Tibor Erős, Balaton Limnological Institute, Centre for Ecological Research, Klebelsberg K. u. 3, H-8237, Tihany, Hungary.
Email: eros.tibor@ecolres.hu

Abstract

Fish-habitat associations are poorly known in offshore areas of very large rivers. We examined physical habitat structure and its effect on habitat use and assemblage formation of benthic fishes in the main channel of the Danube River, Hungary. Principal component analysis of physical variables showed that sample unit (i.e., 500 m long reaches) and cross-channel transect-level data of corresponding reaches were highly correlated. We found clear gradients in physical variables from areas with high velocity and higher proportion of hard substratum (pebble and cobble) to areas with low velocity, high mean depth and finer substratum (mainly sand) composition. Variation in velocity was coupled with variation in both mean depth and substratum composition (i.e., Shannon diversity of sediment composition) and higher proportion of silt material. Differences in physical habitat structure (flow, substrate) also manifested among river segments. Classification and regression tree analyses (CART) and fish abundance – occupancy patterns in the PCA template revealed that many species showed clear responses to environmental heterogeneity (barbel, *Barbus barbus*; schraetser, *Gymnocephalus schraetser*; Danube streber, *Zingel streber*; whitefinned gudgeon, *Romanogobio vladikovyi*; round goby, *Neogobius melanostomus*) while others (white bream, *Blicca bjoerkna*) showed very elusive habitat use patterns. Multivariate regression tree analysis confirmed the results of CART and indicated that transect-level substratum composition was the most important determinant in the formation of benthic assemblages. These results on habitat use can contribute to the more effective conservation management of offshore fish assemblages, which is important due to increasing inland navigation in the Danube River.

KEYWORDS

benthic fishes, CART, Danube River, geomorphology, habitat use, MVRT, substrate, trawling

1 | INTRODUCTION

Determination of how physical structure of the habitat influences the distribution, abundance and assembly of species is a prerequisite for effective nature conservation and environmental management

(Allan & Flecker, 1993). However, sampling even basic habitat and assemblage data can be problematic in remote, hardly accessible habitats and/or where extreme habitat conditions preclude effective sampling, such as, for example, deep water oceanic habitats or canyons of high mountains. In river ecology, the offshore areas of very large

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *River Research and Applications* published by John Wiley & Sons Ltd.

(or great) rivers have been long considered as areas that are difficult to sample effectively, due mostly to depth and flow conditions (e.g., Dettmers, Gutreuter, Wahl, & Soluk, 2001; Loisl, Singer, & Keckeis, 2014). Not surprisingly, most habitat assessment studies on fish assemblages are restricted to shoreline analyses only, or to smaller streams and rivers (see e.g., Boys & Thoms, 2006; Erős, Tóth, Sevcsik, & Schmera, 2008; Keckeis et al., 2013). Although offshore areas proved to be important feeding, breeding and wintering habitats for many large river fishes (Galat & Zweimüller, 2001; Wolter & Bischoff, 2001), there is still limited knowledge on specific fish-habitat associations and the assembly of species in a variety of biogeographic regions (e.g., Cao, Parker, Edison, & Epifanio, 2019; Dettmers et al., 2001; Ridenour, Starostka, Doyle, & Hill, 2009), and most studies address only rough spatial scales.

Fish-habitat associations of offshore areas may be also difficult to model effectively. For example, environmental gradients in substrate composition and water velocity are relatively short in the potamon, at least compared with the littoral zone or with wadeable streams of the rhithron, where contrasting changes in physical habitat quality and fish assemblages are more common (Allan & Castillo, 2007; Erős, 2017; Erős et al., 2017). This may result very elusive fish-habitat relationships for offshore areas. Fish abundance may also vary largely offshore (Wolter & Freyhof, 2004; Szalóky et al., 2014), which can be due either to the response of fish to subtle changes in physical habitat quality or stochasticity in abundance. Advanced statistical and machine learning techniques, which can handle nonlinear and complex interaction effects, may be better applicable to explore fish-habitat

relationships in these cases than traditional methods such as linear or multiple regressions (Knudby, Brenning, & LeDrew, 2010; Olden, Lawler, & Poff, 2008). In this regard, regression trees proved to be especially useful for modelling fish-habitat relationships (Knudby et al., 2010; Veza, Parasiewicz, Calles, Spairani, & Comoglio, 2014). Although these tools are still largely underutilized in large river fish ecology, and especially not for offshore areas, they have been already successfully applied in modelling the fish assemblage structure of shoreline habitats (Wilkes, Maddock, Link, & Habit, 2016).

Distribution and abundance of fish may not only be influenced by the physical attributes of the sampling area, but by the surrounding habitat (Erős & Grossman, 2005; Schlosser, 1991). However, how the hierarchical structure of the habitat influences the organization of fish assemblages or fish habitat relationships offshore is largely unknown. Therefore, in this study, we examine the benthic fish assemblages of offshore areas of the Danube River, Hungary. We characterize the physical habitat structure of offshore areas and use multivariate regression trees for defining the scale dependent environmental determinants of benthic fish assemblages, and classification and regression trees for examining species specific habitat relationships. Former samplings using a specifically designed benthic framed trawl showed that offshore areas serve as important habitats for many rare and endangered benthic species of high conservation concern in the Danube (Szalóky et al., 2014), the habitat use of which is still poorly known. We predicted that the joint analysis of focal scale (i.e., sample unit level) physical attributes with physical attributes of the surrounding environment (i.e., transect or higher level environmental

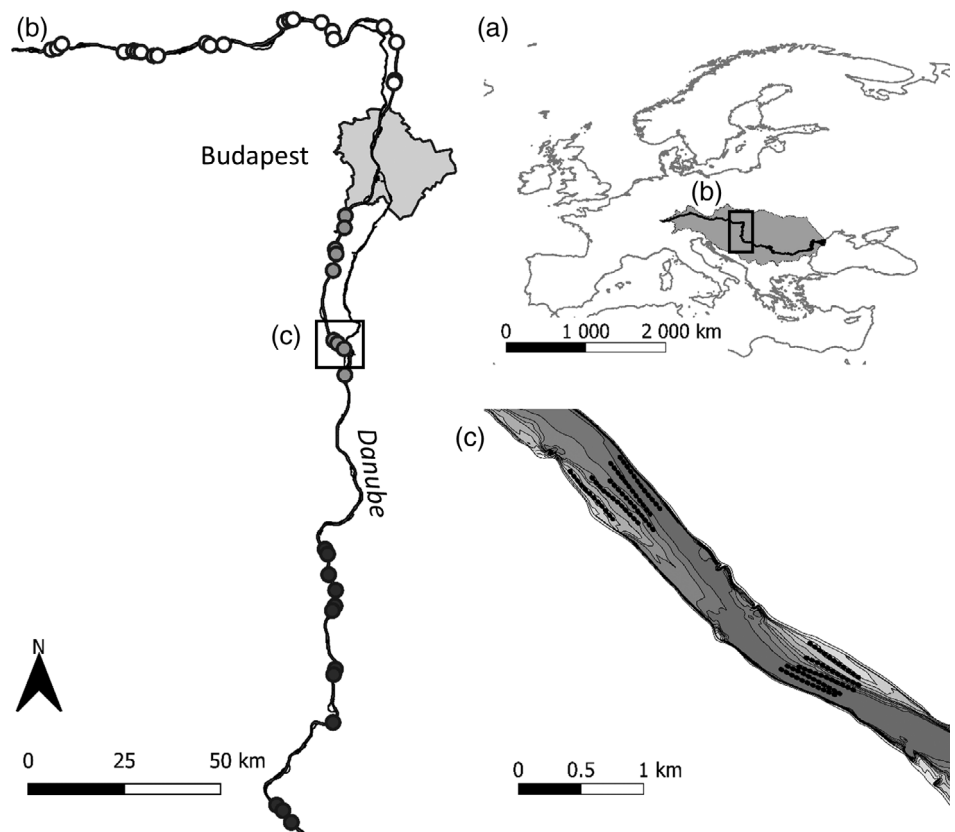


FIGURE 1 Map of the study area and location of sampling sites in the Hungarian portion of the Danube River (a). Three large and separating segments are differentiated using empty (I), grey (II) and black (III) dots (b). Distribution of two exemplar transects and their corresponding sample units (500 m long) are also shown on a bathymetric map (c)


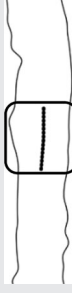

heterogeneity) may better characterize the physical heterogeneity of the relatively homogenous offshore environment and, therefore, may yield stronger fish-habitat relationships than using variable scores of sample level environmental heterogeneity only. Consequently, we were especially interested to explore the relationships between sample unit and transect-level environmental heterogeneity and their influences on fish assemblages and on the habitat use of the assemblage constituting species. In addition, we hypothesized that most benthic species will respond to offshore gradients in habitat structure, even if we experience short environmental gradients. In this regard, we predicted that changes in substrate composition will be the most influential mechanism that govern the habitat use of individual species and the benthic community.

2 | MATERIAL AND METHODS

2.1 | Study area

The Danube has a drainage area of approximately 796,250 km². River regulation, namely the construction of hydroelectric schemes, especially in the Upper Danube (i.e., in Germany and Austria), and channelization have profoundly modified the physical structure of the Danube throughout its course. The Hungary section (Figure 1a), referred to as the “Middle Danube,” runs for 417 km and has a mean annual discharge of 2,000 m³ s⁻¹. The main channel has a substratum dominated by gravel and sand, a mean depth of 4 m and a mean velocity of 0.6 m s⁻¹. The banks are relatively natural (except the

TABLE 1 The code, name, and mean and range values of the measured physical variables in the Middle-Danube, Hungary

Spatial scale (a, b, c)	Code	Name	Mean	Range
(a) 500 m long sampling unit 	a_m_depth	Mean water depth (m)	4.29	1.21–11.1
	a_cv_depth	Coefficient of variation of water depth	0.12	0.01–0.56
	a_min_distance	Minimum distance from shoreline (m)	125.78	5–285
	a_m_velocity	Mean water velocity (cm s ⁻¹)	84.47	15–142
	a_mud	Silt, sediment composition (%)	0.30	0–10
	a_sand	Sand, sediment composition (%)	29.87	0–100
	a_s_gravel	Gravel, sediment composition (%)	12.68	0–75
	a_gravel	Pebble, sediment composition (%)	46.22	0–100
	a_stone	Cobble, sediment composition (%)	7.92	0–90
	a_marl	Marl, sediment composition (%)	3.01	0–100
	a_shan_div_sediment	Shannon diversity index of sediment	0.74	0–1.37
(b) 500 m long cross section 	b_m_depth	Mean water depth (m)	4.23	2.09–8.01
	b_cv_depth	Coefficient of variation of water depth	0.30	0.07–0.7
	b_m_velocity	Mean water velocity (cm s ⁻¹)	84.44	44.65–122.08
	b_cv_velocity	Coefficient of variation of water velocity	0.22	0.05–0.65
	b_mud	Silt, sediment composition (%)	0.35	0–4
	b_sand	Sand, sediment composition (%)	29.79	0–94
	b_s_gravel	Gravel, sediment composition (%)	12.75	0–39
	b_gravel	Pebble, sediment composition (%)	46.18	0–91
	b_stone	Cobble, sediment composition (%)	7.92	0–48
	b_marl	Marl, sediment composition (%)	3.01	0–43
	b_shan_div_sediment	Shannon diversity index of sediment	1.05	0.56–1.51
(c). 2,000 m long section 	c_m_depth	Mean water depth (m)	3.29	2.24–4.44
	c_cv_depth	Coefficient of variation of water depth	0.60	0.38–1.07

Note: (a) sample unit level, where variables were measured within the 500 m long unit; (b) transect level, where values of the variables were calculated using the mean of the sample unit level values; (c) transect neighbourhood level, which embraced 2,000 m long segments, and where we used a grid of 15 × 15 m² systematic, equally distance measurements from a bathymetric map for calculating mean depth and its coefficient of variation for each segment.

section lying within Budapest, the capital of Hungary), interrupted with embanked rip-rap shorelines of ~100–1,000 m long sections.

2.2 | Data collection

Offshore distribution of fishes was examined using a benthic framed trawl. The trawl consisted of a stainless steel frame (2 m wide × 1 m high) to which a drift net was attached (mesh size 5 and 8 mm for the inner and outer mesh bag, respectively) (for details see Szalóky et al., 2014). The frame was electrified with a HansGrassl EL65 IIGI (Hans Grassl GmbH, Schönau am Königssee, Germany) electrofishing device operated with a Vanguard HP20 14.9 kW generator (Briggs & Stratton, Wisconsin, USA). A 6 m long copper cathode cable was connected freely and pulled approx. 2 m before the electrified frame. The fishing team consisted of two people handling the framed net, one handling the electrofishing device and one operating the boat. Trawling was conducted during daytime with a 6.3 m long boat (Alpha Composite Kft., Ebes, Hungary) powered by a 50 hp outboard Mercury four stroke engine (Mercury Marine, Wisconsin, USA). Before starting trawling, the operators lowered the frame to the bottom while the boat was slowly moving downstream with the flow. Trawling route was started to be measured by a GARMIN GPSMAP 60CSx (Garmin Ltd., Kansas City, MO) only after the net reached the bottom, which could be easily felt while holding the central rope, and right after electroshocking started. The direct current (approx. 350 V, 33 A) was applied for 5–8 s. With 3–5 s. Breaks between the operations to minimize fright bias and injury of fish. The applied trawling speed was slightly higher than the current velocity of the river (approx. 0.6 m s^{-1}).

We used a stratified design to select segments (Figure 1b) in order to get a representative coverage of the whole main channel area (excluding the section where the capital, Budapest can be found). In each of these segments several, but a minimum of three transects were selected randomly, perpendicular to the bank. Along each transect, across the width of the main channel, we generally distributed 5–6 trawl paths, 500 m long each, excluding the littoral, less than 2 m deep shoreline zone (Figure 1c). These paths were approximately equispaced and centred over the approximate place of the main channel centreline (Gutreuter, Vallazza, & Knights, 2010). Note that the number of trawl paths (hereafter samples) along the transects varied depending on the river width, and sometimes the trawl was stopped due to interruption by large rocks or logs. This study contains data of 199 samples collected between April 2011–September 2011 period. The collected fish were identified, measured to the nearest mm standard lengths (70% of fish) and then released back to the river. Note, that preliminary exploratory analyses showed that young and adult fish of the same species showed basically the same occupancy patterns in the habitat template, which was characterized by a principal component analysis (see below). Therefore, to save space we did not show size or age group specific analyses in this article, although we could clearly separate young and adult individuals for some species based on length-frequency histograms (see Appendix I).

Physical habitat variables were quantified at three spatial scales, specifically at the sample unit, at the transect, and at the transect neighbourhood scales (Table 1). We measured minimum distance from shoreline, depth, flow velocity and substrate composition data at the sample unit scale. Distance from shoreline was measured using Google Earth. Mean depth of the sampling unit and its coefficient of variation was calculated by reading the records of a Lowrance X50 DS depth finder sonar (Lowrance, Joplin, MO) at an average of 11 points along each 500 m long unit during the fish sampling. Mean water velocity was calculated using three measurements at 60% depth in the upper, middle and lower portion of the sampling unit using a Flow probe (FP211) flow meter (Global Water, College Station, United States). Percentage substrate composition of the units were visually estimated based on three grab samples collected in the upper, middle and lower portion of the sampling unit. The following categories were distinguished (Baranya et al., 2018): silt (<0.25 mm), sand (0.25–2 mm), gravel or fine gravel (2–8 mm), pebble or rough gravel (8–64 mm), cobble or stone (>64 mm), marl and/or clay (rather cemented very fine silty material). Shannon diversity of the substrate material was also calculated based on the percentage data of substrate categories. Transect scale data were calculated using the mean of the sample unit variables of each transect. Finally, we also calculated mean depth and its coefficient of variation at the transect neighbourhood scale, which embraced a 2,000 m long segment (see Table 1). For this we used a grid of $15 \times 15 \text{ m}^2$ systematic, equally distance measurements from a bathymetric map for each segment.

2.3 | Statistical analysis

We used standardized principal component analysis (PCA) to explore the relationships among the physical habitat variables and for the comparison of the sampling sites based on their environmental characteristics. For this, we normalized the variables using either natural logarithmic (velocity, depth, minimum distance from shoreline, Shannon diversity of sediment composition) or square root arcsin transformations (% substratum composition). To visualize habitat use patterns along physical environmental gradients, we plotted the occupancy of fishes on the PCA plots, where the size of the bubbles is proportional to the CPUE-based abundance (i.e., ind 500 m^{-1}) of the species.

We used classification and regression tree analyses (CART) to directly select those key physical variables, which may determine the most the habitat use of the species. CART is a flexible and robust classification and prediction method, and it is ideally suited for modelling non-linear interactions, which often appear in ecological data (Breiman, Friedman, Stone, & Olshen, 1984; De'ath & Fabricius, 2000). Trees explain variation in a single response variable by repeatedly splitting the data (here the CPUE of each species) into more homogenous groups using combinations of the predictor variables (here physical habitat data). Finally, we also used multivariate regression trees (MVRT), a multivariate extension of CART (De'ath, 2002) to model the response of the assemblage to the physical habitat data. We used a cross-validation procedure to find the

optimum tree size and to avoid overfitting the data. Indicator species analysis was then used to find the most characteristic species to each assemblage group (Dufrene & Legendre, 1997).

All statistical analyses were performed in R 3.6.3. (R Development Core team, 2020) using “rpart” (Therneau & Atkinson, 2019), “mvpart” package (De’ath, 2002; De’ath, 2014), and “MVPARTwrap” packages (Ouellette & Legendre, 2013) for CART, MVRT and indval analyses, respectively.

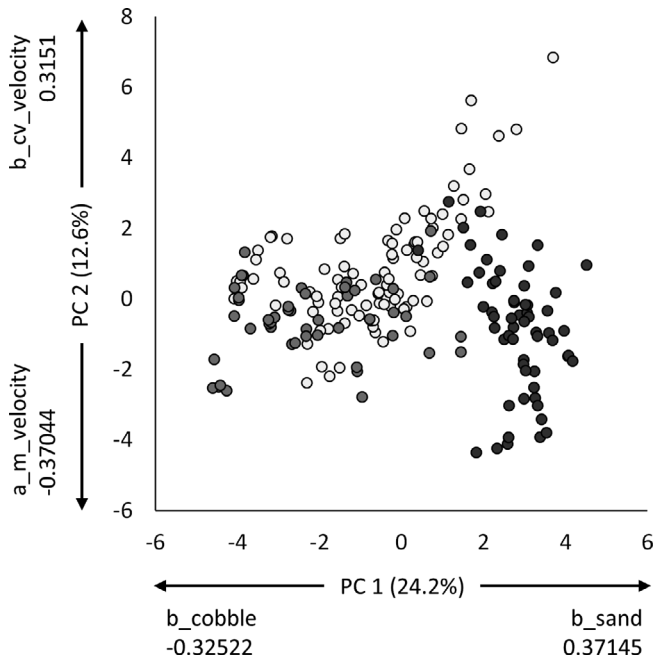


FIGURE 2 Principal component analysis (PCA) ordination of the sample units based on their physical habitat variables and their corresponding transect level physical variables. To show the degree of macroscale-level separation of sample units in the ordination space, sample units of the three large segments are differentiated using empty (I), grey (II) and black (III) dots (see Fig. 1)

3 | RESULTS

3.1 | Habitat characteristics

The first two components of the PCA explained 36.8% of the total variance and indicated high degree of correlation among many physical variables (Figure 2). Component 1 characterized a gradient which was determined mainly by substratum composition. Samples with strong positive loadings had rather fine substrate (i.e., sand and gravel) at both the sample unit and transect levels. Conversely, samples with strong negative loadings had a substratum dominated by pebble and cobble at both the sample unit and transect levels. Component 2 was mainly determined by velocity parameters, and especially variability in velocity at the transect level (positive loadings) and mean velocity at both the sample unit and transect levels (negative loadings). However, in general, several physical variables contributed to the distribution of samples in the ordination place, and consequently their joint consideration is necessary when comparing the environmental features of the samples in the ordination plane (Appendix II).

Segment-level distribution of samples in the ordination plane showed that most environmental heterogeneity occurred at the macroscale (i.e., between large river segments), especially between the lower and the two upper segments, while within segment, mesoscale-level physical heterogeneity was relatively low (Figure 2).

3.2 | Fish-habitat associations

We collected 33 species and 9,274 specimens during the 199 trawling paths. The most dominant benthic species (Table 2) composed of 92.79% of the total abundance and were as follows: the round goby, *Neogobius melanostomus* (57.01%); Danube streber, *Zingel streber* (10.58%); white bream, *Blicca bjoerkna* (8.05%); white-finned gudgeon, *Romanogobio vladkovi* (7.13%); schraetzer, *Gymnocephalus schraetser* (4.21%); zingel, *Zingel zingel* (1.53%); barbel, *Barbus barbus* (1.50%);

TABLE 2 Species list and abbreviation code of the collected species and their relative abundance (RA%), frequency of occurrence (FRO%) and mean CPUE data

Species name	Common name	Code	Total number	RA %	FRO %	Mean CPUE (ind 500 m ⁻¹) ± SD
<i>Abramis brama</i>	Freshwater bream	Abrbra	80	0.86	15.08	0.40 ± 1.69
<i>Babka gymnotrachelus</i>	Racer goby	Babgym	75	0.81	12.06	0.38 ± 1.57
<i>Barbus barbus</i>	Barbel	Barbar	139	1.50	34.17	0.69 ± 1.24
<i>Blicca bjoerkna</i>	White bream	Blibjo	747	8.05	42.71	3.75 ± 17.26
<i>Gymnocephalus schraetser</i>	Schraetzer	Gymsch	390	4.21	31.66	1.96 ± 7.08
<i>Neogobius melanostomus</i>	Round goby	Neomel	5,287	57.01	59.30	26.57 ± 117.99
<i>Ponticola kessleri</i>	Bighead goby	Ponkes	103	1.11	17.59	0.52 ± 1.79
<i>Romanogobio vladkovi</i>	White-finned gudgeon	Romvla	661	7.13	67.34	3.32 ± 6.01
<i>Zingel streber</i>	Danube streber	Zinstr	981	10.58	74.87	4.93 ± 7.75
<i>Zingel zingel</i>	Zingel	Zinzin	142	1.53	22.61	0.71 ± 2.08
Number of individuals			8,605			
Number of samples			199			

bighead goby, *Ponticola kessleri* (1.11%); freshwater bream, *Abramis brama* (0.86%); racer goby, *Babka gymnotrachelus* (0.81%). In general, most species occupied the full habitat gradients characterized by the PCA. However, there were clearly recognizable differences between some species regarding their abundance patterns in the ordination plane (Figure 3). For example, the barbel and the Danube streber preferred areas with higher velocity and dominance of pebbles and cobbles, while the schraetser and zingel were found in areas with finer substratum, lower velocity and higher mean depth. The invasive Ponto-Caspian gobies, such as the round goby, the bighead goby and the racer goby, showed relatively similar habitat affinity, which differed to some degree from the native species. Specifically, their habitat use pattern could be best characterized by considering both PC1 and PC2 axes, where a gradient in velocity and variation in velocity were the main physical determinants, and secondarily the proportion of rarer substratum materials, such as silt.

The CART analysis generally supported the results obtained by the visualization of the abundance – occupancy patterns in the PCA. However, it specifically selected some cut-off values both for the predictor physical variables and for the fish abundance values, which help to more specifically quantifying the habitat use of the species (Figure 4). In this regard, CART also revealed that no valid models could be obtained for some species (racer goby, white bream, bighead goby, zingel), due to the lack of clear response of these species to the examined physical variable gradients.

3.3 | Formation of benthic fish assemblages

The MVRT analysis indicated that transect-level substratum composition was the most important determinant in the formation of benthic fish assemblages (Figure 5). Areas with lower portion of pebble (<37.5%), and consequently, higher portion of finer bed material, had a benthic assemblage which could be mainly characterized by the whitefinned gudgeon, the schraetser and the white bream. The dominance of pebble (>37.5%) was the best associated with the Danube streber, the barbel and the round goby. Although the portion of gravel was also a determinant of the separation of gobiid species (>4%), and especially for the round goby since the relative abundance of the bighead and the racer goby was very low in this assemblage. Overall, the assemblage-level analyses with MVRT corresponded well with the species specific result of CART analyses.

4 | DISCUSSION

To our knowledge, this is the first study that directly models the relationships between physical habitat variables and the abundance of several offshore benthic fish species in the very large Danube River. Although former studies addressed patterns in the occurrence and abundance of species offshore to some degree, these studies focused more on the comparison of inshore versus offshore assemblages and the effect of sampling methods on assemblage structure (see

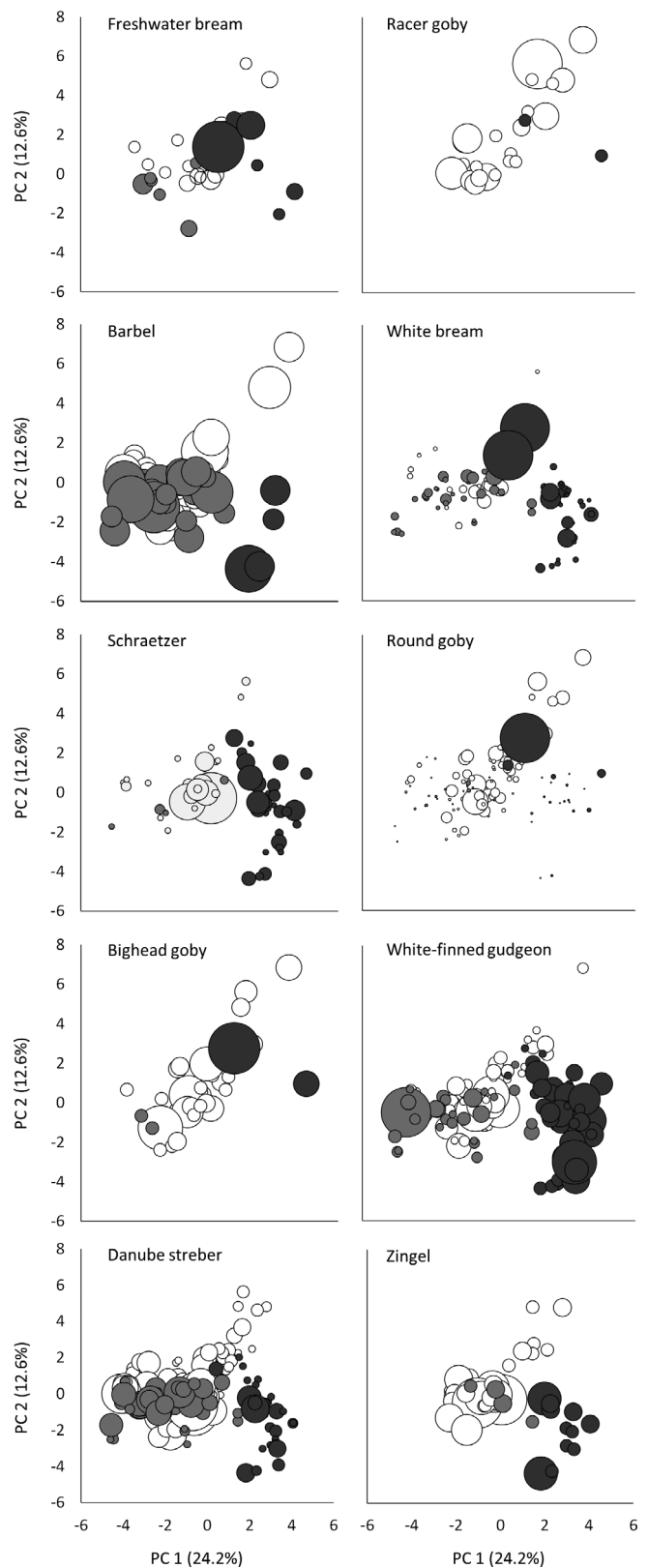
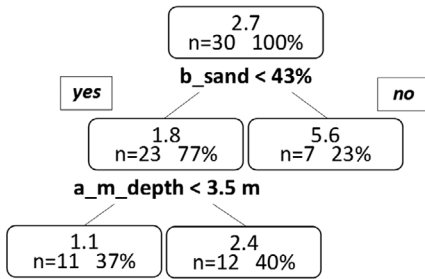


FIGURE 3 Abundance and occupancy pattern of each species in the PCA ordination space. The size of the circles is proportional to the relative abundance of the species. To show the degree of macroscale-level separation of sample units in the ordination space, sample units of the three large segments are differentiated using empty (I), grey (II) and black (III) dots (see Fig. 1)

Freshwater bream



Barbel

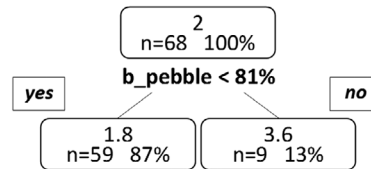
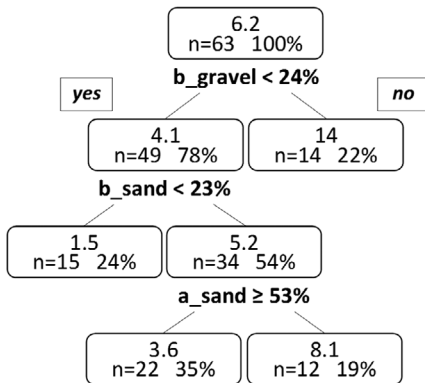
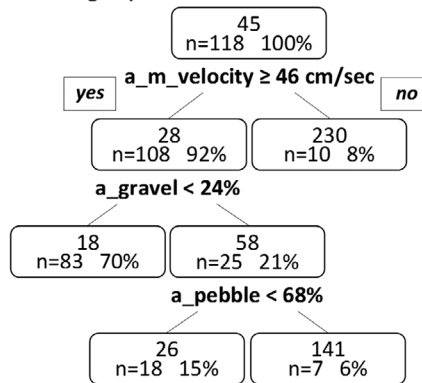


FIGURE 4 Results of the classification and regression tree analyses (CART) for each species. For each tree, the boxes show the mean CPUE of the fish species, the sample number (n) and percentage of the samples (%) to the total number of samples. The key response variable, which defines each split in the tree and its mean value is also shown

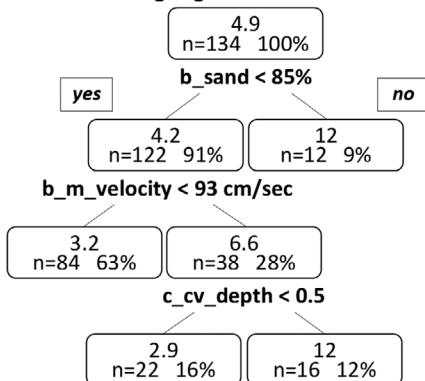
Schraetzer



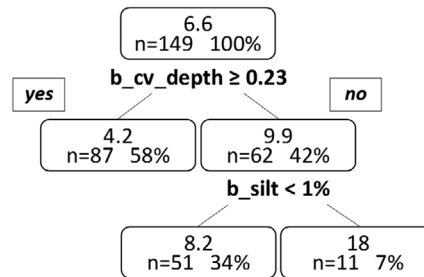
Round goby



White-finned gudgeon



Danube streber



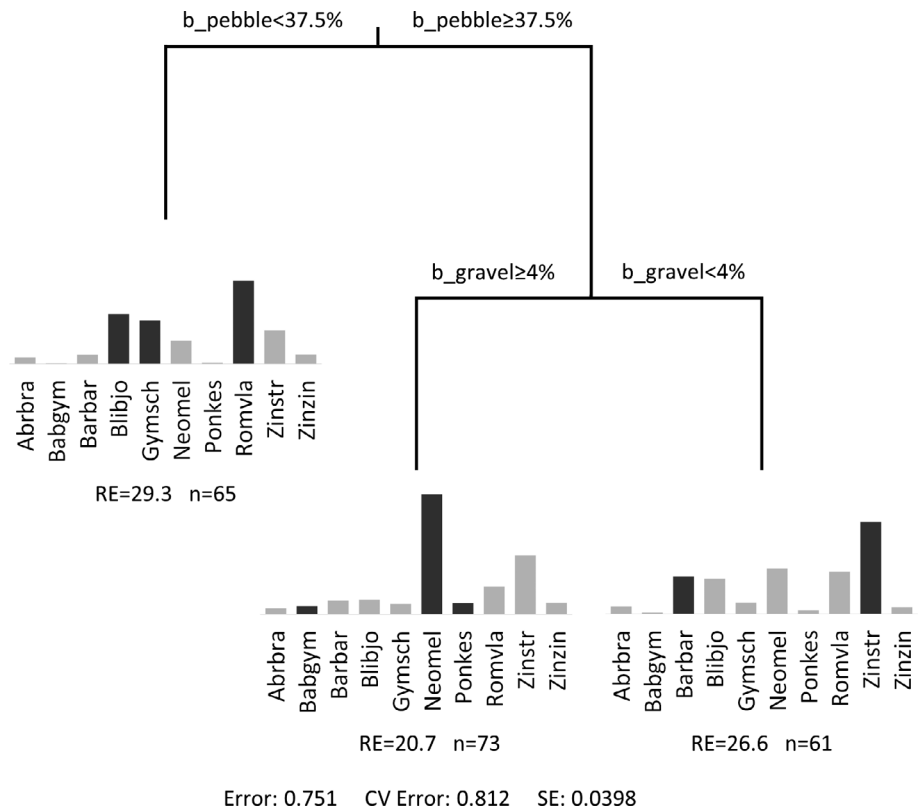
e.g., Erős et al., 2017; Loisl et al., 2014; Szalóky et al., 2014). Here, we found that while all species occurred along the full extent of physical habitat variability gradients some species showed clear affinities, while others showed only very elusive habitat affinity to offshore environmental heterogeneity.

We found strong associations in the values of sample unit and transect-level physical variables (Figure 2), which indicates relatively small mesoscale level (10^1 – 10^3 m) environmental heterogeneity offshore in this very large river, as expected. Although the ranges in the values of physical variables were relatively small, clear and easily interpretable relationships were explored between several physical variables. For example, areas with relatively high velocity had higher proportion of hard substratum (especially pebble and cobble), and areas with low velocity had higher mean depth and finer substratum composition. Further, variation in velocity was coupled with variation in mean depth and substratum composition (i.e., Shannon diversity of

sediment composition) and higher proportion of silt material. These results thus show that offshore areas which are mostly considered as homogenous mesoscale-level units in fish habitat evaluations (Baranya et al., 2018; Habersack, Tritthart, Liedermann, & Hauer, 2014; Wegscheider, Linnansaari, & Curry, 2020) do show some clearly recognizable environmental heterogeneity. However, the hierarchical evaluations and visualization of physical data also show that most offshore environmental heterogeneity occurs at the macroscale (10^4 – 10^5 m) that is between larger river segments in the Middle-Danube, and that within mesoscale level (i.e., between sample units within transects) environmental heterogeneity is relatively low offshore.

Of the examined species the barbel and the Danube streber showed the highest affinity to pebble and boulder covered, high velocity areas. This finding is consistent with former results about the habitat preference of these species from smaller rivers (Brinker et al., 2018; Erős, 2007; Gutmann Roberts, Bašić, Britton, Rice, &

FIGURE 5 Result of the multivariate regression tree analyses (MVRT) for the benthic fish assemblage. For each tree, bars show the relative abundance of fishes, and the significant indicator species. The key response variable, which defines each split in the tree and its mean value is also shown



Pledger, 2020) and from the Austrian portion of the Danube (Loisl et al., 2014). Interestingly, the Danube streber also occupied areas with relatively high portion of sand or mixed substratum composition, although its abundance was low in these habitats. In fact, these deep, slower flowing areas with higher variation in substratum composition and flow were more preferred by the zingel, albeit this species was the most abundant among mean habitat conditions (note, that this is indicated by its high occurrence and abundance values around the origo in the PCA plot). The mesoscale-level abundance patterns of these two endemic zingel species and their distribution along the examined physical habitat gradients thus explain well why the zingel is more bounded to deep very large rivers, and why the Danube streber is more abundant in fast flowing highland rivers (Bănăduc & Curtean-Bănăduc, 2014; Erős, 2007; Harka & Sallai, 2004).

Of the invasive Ponto-Caspian gobies, the round goby was the most dominant species offshore. It occurred along a variety of habitat conditions, although its abundance was lower in the most fast flowing and deepest habitats. The CART analysis further highlighted that its abundance was the highest in relatively slow flowing areas with higher portion of rough substratum (gravel and pebble). This result is consistent with our former, more local scale study, which directly quantified offshore habitat preference curves for the species (Baranya et al., 2018). The habitat use of the racer goby differed to some degree from the round goby. Specifically, the racer goby was more abundant in slow flowing areas with silt substratum and higher variation in depth. Finally, the bighead goby was very rare offshore. It occurred among mean habitat conditions and, not surprisingly, the CART analysis did not select physical variables which would

significantly influence its abundance pattern. Overall, our study prove that several goby species occur offshore, although of these, only the round goby is the only species which have abundant populations offshore in the Middle-Danube (see also Szalóky et al., 2015). Several studies justify that riprap covered shorelines are preferred habitats for most goby species inshore, and especially for the round and bighead goby (e.g., Erős et al., 2008; Ramler & Keckeis, 2019, 2020), which habitat type is lacking offshore. Regarding the offshore distribution and abundance of gobies, our results in general correspond with results on habitat use patterns along natural habitat gradients inshore (Erős, Sevcsik, & Tóth, 2005; Płachocki, Kobak, Poznańska-Kakareko, & Kakareko, 2020).

Schraetser, a Danubian endemic species, showed clear affinity towards areas with high proportion of fine gravel and sand substratum. The preference towards sandy substrate was also observed for the Danube whitefinned gudgeon, although the species also occurred in high abundance in fast flowing pebble covered habitats, and generally showed wider habitat occupancy pattern. The freshwater bream also reached the highest abundance in sand and silt dominated, deep areas, which confirms knowledge about the habitat of the species in large rivers (Wolter & Bischoff, 2001). Of the examined species the white bream proved to be the most ubiquitous. The species occurred along the whole physical variability gradients, although it generally reached the highest abundance among mean habitat conditions confirming former findings (Wolter & Bischoff, 2001; Wolter & Freyhof, 2004). Not surprisingly the CART analysis did not select specific variables, which could be related the most with the abundance of the species. However, it should be noted that results of the CART

analysis should be considered with caution based on results of the PCA, which showed that many physical variables correlate with each other and it is hard to precisely select the one and only variable which is responsible the most for the habitat use of these benthic species.

Mesoscale-level formation of fish assemblages was elusive offshore, which was indicated by the relatively high cross-validation error of MVRT. Nevertheless, the assemblage constituting species differed to some degree in relative abundance among the differentiated assemblage types. Differences in substratum composition were the most influential separating variables in assemblage patterns, which is not surprising for this benthic assemblage (Greenberg, 1991). Although we found that the relative proportion of pebble was one of the key variables in the formation of this offshore assemblage, again, it must be emphasized that substratum composition in general showed correlation with other physical variables.

In conclusion, our findings demonstrate that many benthic species do respond to offshore physical habitat heterogeneity in the Danube River. Although most environmental heterogeneity was related to the macroscale (10^4 – 10^5 m), we found that mesoscale-level (10^1 – 10^3 m) differences in physical habitat quality clearly influence the formation of fish assemblages in this very large river. Consequently, we encourage researchers and managers to pay attention not only to inshore but also to offshore physical and biological heterogeneities of large rivers, even in seemingly homogenous habitats, where geomorphological breaks in flow, depth and substratum patterns cannot be easily recognized. This may contribute to the more effective conservation and management of offshore fish assemblages, which is critically important in the era of increasing inland navigation.

ACKNOWLEDGEMENTS

We would like to express the numerous people who helped during the years of the Danubian field works, and especially for András Sevcsik, Béla Csányi, Ágnes György-Irma, József Szekeres, István Czeglédi, András Specziár, Péter Sály, András Weiperth. This study was supported from an Ecology and Society (MTA KEP) grant, and the ANN-OTKA 135082 project. We acknowledge the support of KDP-2020 Cooperative Doctoral Programme of the National Research, Development and Innovation Office to Vivien Füstös.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author, [TE], upon reasonable request.

ORCID

Tibor Erős  <https://orcid.org/0000-0002-2252-3115>

REFERENCES

- Allan, J. D., & Castillo, M. M. (2007). *Stream ecology: Structure and function of running waters* (2nd ed.). New York: Chapman and Hall. <https://doi.org/10.1007/978-1-4020-5583-6>.
- Allan, J. D., & Flecker, A. S. (1993). Biodiversity conservation in running waters. *Bioscience*, 43(1), 32–43.
- Bănăduc, D., & Curtean-Bănăduc, A. (2014). Monitoring elements for *Zingel streber* (Siebold, 1863) in the context of Natura 2000 in Croatia. *Romanian Journal of Biology*, 59, 59–74.
- Baranya, S., Fleit, G., Józsa, J., Szalóky, Z., Tóth, B., Czeglédi, I., & Erős, T. (2018). Habitat mapping of riverine fish by means of hydro-morphological tools. *Ecohydrology*, 11(7), e2009.
- Boys, C. A., & Thoms, M. C. (2006). A large-scale, hierarchical approach for assessing habitat associations of fish assemblages in large dryland rivers. *Hydrobiologia*, 572(1), 11–31.
- Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. Taylor & Francis.
- Brinker, A., Chucholl, C., Behrmann-Godel, J., Matzinger, M., Basen, T., & Baer, J. (2018). River damming drives population fragmentation and habitat loss of the threatened Danube streber (*Zingel streber*): Implications for conservation. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 28(3), 587–599.
- Cao, Y., Parker, J., Edison, T., & Epifanio, J. (2019). Sampling the central channel provides additional information on fish assemblages in a large Boatable River in the US Midwest. *Journal of Fish and Wildlife Management*, 10(2), 432–441.
- De'Ath, G. (2002). Multivariate regression trees: A new technique for modeling species–environment relationships. *Ecology*, 83(4), 1105–1117.
- De'ath, G. (2014). mvpart: Multivariate partitioning. R Package Version 1.6.2. Retrieved from <https://cran.r-project.org/src/contrib/Archive/mvpart/>.
- De'ath, G., & Fabricius, K. E. (2000). Classification and regression trees: A powerful yet simple technique for ecological data analysis. *Ecology*, 81(11), 3178–3192.
- Dettmers, J. M., Gutreuter, S., Wahl, D. H., & Soluk, D. A. (2001). Patterns in abundance of fishes in main channels of the upper Mississippi River system. *Canadian Journal of Fisheries and Aquatic Sciences*, 58(5), 933–942.
- Dufrêne, M., & Legendre, P. (1997). Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecological Monographs*, 67(3), 345–366.
- Erős, T. (2007). Partitioning the diversity of riverine fish: The roles of habitat types and non-native species. *Freshwater Biology*, 52(7), 1400–1415.
- Erős, T. (2017). Scaling fish metacommunities in stream networks: Synthesis and future research avenues. *Community Ecology*, 18(1), 72–86.
- Erős, T., Bammer, V., György, Á. I., Pehlivanov, L., Schabuss, M., Zornig, H., ... Szalóky, Z. (2017). Typology of a great river using fish assemblages: Implications for the bioassessment of the Danube river. *River Research and Applications*, 33(1), 37–49.
- Erős, T., & Grossman, G. D. (2005). Fish biodiversity in two Hungarian streams: A landscape-based approach. *Archiv für Hydrobiologie*, 162(1), 53–71.
- Erős, T., Sevcsik, A., & Tóth, B. (2005). Abundance and night-time habitat use patterns of Ponto-Caspian gobiid species (Pisces, Gobiidae) in the littoral zone of the river Danube, Hungary. *Journal of Applied Ichthyology*, 21(4), 350–357.
- Erős, T., Tóth, B., Sevcsik, A., & Schmera, D. (2008). Comparison of fish assemblage diversity in natural and artificial rip-rap habitats in the littoral zone of a large river (river Danube, Hungary). *International Review of Hydrobiology*, 93, 88–105.
- Galat, D. L., & Zweimüller, I. (2001). Conserving large-river fishes: Is the highway analogy an appropriate paradigm? *Journal of the North American Benthological Society*, 20(2), 266–279.
- Greenberg, L. A. (1991). Habitat use and feeding behavior of thirteen species of benthic stream fishes. *Environmental Biology of Fishes*, 31(4), 389–401.
- Gutmann Roberts, C., Bašić, T., Britton, J. R., Rice, S., & Pledger, A. G. (2020). Quantifying the habitat and zoogeomorphic capabilities of spawning European barbel *Barbus barbus*, a lithophilous cyprinid. *River Research and Applications*, 36(2), 259–279.
- Gutreuter, S., Vallazza, J. M., & Knights, B. C. (2010). Lateral distribution of fishes in the main-channel trough of a large floodplain river: Implications for restoration. *River Research and Applications*, 26(5), 619–635.

- Habersack, H., Tritthart, M., Liedermann, M., & Hauer, C. (2014). Efficiency and uncertainties in micro-and mesoscale habitat modelling in large rivers. *Hydrobiologia*, 729(1), 33–48.
- Harka, Á., & Sallai, Z. (2004). *Magyarország halfaunája* (p. 269). Szarvas, Hungary: Nimfea Természetvédelmi Egyesület.
- Keckeis, H., Schludermann, E., Tritthart, M., Hauer, C., Liedermann, M., & Habersack, H. (2013). Habitat preference of the sublittoral fish assemblage in a free-flowing section of the Danube River, Austria. Paper presented at: 5th Symposium for Research in Protected Areas pp. 353–362.
- Knudby, A., Brenning, A., & LeDrew, E. (2010). New approaches to modelling fish–habitat relationships. *Ecological Modelling*, 221(3), 503–511.
- Loisl, F., Singer, G., & Keckeis, H. (2014). Method-integrated fish assemblage structure at two spatial scales along a free-flowing stretch of the Austrian Danube. *Hydrobiologia*, 729(1), 77–94.
- Olden, J. D., Lawler, J. J., & Poff, N. L. (2008). Machine learning methods without tears: A primer for ecologists. *The Quarterly Review of Biology*, 83(2), 171–193.
- Ouellette, M. H., & Legendre, P. (2013). MVPARTwrap: Additional features for package mvpart. R package version 0.1-9.2. Retrieved from <https://cran.r-project.org/src/contrib/Archive/MVPARTwrap/>.
- Piłchocki, D., Kobak, J., Poznańska-Kakareko, M., & Kakareko, T. (2020). Environmental factors associated with the occurrence of the Ponto-Caspian gobies in a lowland river belonging to the central European invasion corridor. *River Research and Applications*, 36(1), 25–35.
- Ramler, D., & Keckeis, H. (2019). Occurrence of non-native fishes in the Danube east of Vienna (Austria) and potential interactions of invasive gobiids with native fishes. *Journal of Applied Ichthyology*, 35(4), 850–862.
- Ramler, D., & Keckeis, H. (2020). Effects of hydraulic engineering restoration measures on invasive gobies in a large river (Danube, Austria). *Biological Invasions*, 22(2), 437–453.
- Ridenour, C. J., Starostka, A. B., Doyle, W. J., & Hill, T. D. (2009). Habitat used by *Macrhybopsis* chubs associated with channel modifying structures in a large regulated river: Implications for river modification. *River Research and Applications*, 25(4), 472–485.
- R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>.
- Schlosser, I. J. (1991). Stream fish ecology: A landscape perspective. *Bioscience*, 41(10), 704–712.
- Szalóky, Z., Bammer, V., György, Á. I., Pehlivanov, L., Schabuss, M., Zornig, H., ... Erős, T. (2015). Offshore distribution of invasive gobies (Pisces: Gobiidae) along the longitudinal profile of the Danube River. *Fundamental and Applied Limnology/Archiv für Hydrobiologie*, 187(2), 127–133.
- Szalóky, Z., György, Á. I., Tóth, B., Sevcsik, A., Specziár, A., Csányi, B., ... Erős, T. (2014). Application of an electrified benthic frame trawl for sampling fish in a very large European river (the Danube River)—is offshore monitoring necessary? *Fisheries Research*, 151, 12–19.
- Therneau, T., & Atkinson, B. (2019). Rpart: Recursive Partitioning and Regression Trees; R package, version 4.1-15. Retrieved from <https://github.com/bethatkinson/rpart>.
- Veza, P., Parasiewicz, P., Calles, O., Spairani, M., & Comoglio, C. (2014). Modelling habitat requirements of bullhead (*Cottus gobio*) in alpine streams. *Aquatic Sciences*, 76(1), 1–15.
- Wegscheider, B., Linnansaari, T., & Curry, R. A. (2020). Mesohabitat modelling in fish ecology: A global synthesis. *Fish and Fisheries*, 21(5), 927–939.
- Wilkes, M. A., Maddock, I., Link, O., & Habit, E. (2016). A community-level, mesoscale analysis of fish assemblage structure in shoreline habitats of a large river using multivariate regression trees. *River Research and Applications*, 32(4), 652–665.
- Wolter, C., & Bischoff, A. (2001). Seasonal changes of fish diversity in the main channel of the large lowland river Oder. *Regulated Rivers: Research & Management: An International Journal Devoted to River Research and Management*, 17(4–5), 595–608.
- Wolter, C., & Freyhof, J. (2004). Diel distribution patterns of fishes in a temperate large lowland river. *Journal of Fish Biology*, 64(3), 632–642.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Szalóky Z, Füstös V, Tóth B, Erős T. Environmental drivers of benthic fish assemblages and fish-habitat associations in offshore areas of a very large river. *River Res Applic*. 2021;37:712–721. <https://doi.org/10.1002/rra.3793>