PRIMARY RESEARCH PAPER

Longitudinal patterns of fish assemblages in European boreal streams

Tapio Sutela · Teppo Vehanen · Pekka Jounela

Abstract Fish assemblages are known to change from headwaters to river outlets. Still, our knowledge of this change is often approximate or sporadic. In this study, we quantified the average longitudinal change from a large electrofishing data set of boreal streams in Northern Europe. The average species richness increased from headwaters to medium-sized rivers but levelled off when reaching large rivers. Existence of some headwater specialist fish species, e.g. brook trout (*Salvelinus fontinalis*), was interpreted to support the fish zonation concept over the concept of accumulative addition of species downstream. The traditional fish zonation concept developed in Western Europe suggests four zones from headwaters to river outlets, the trout zone, the grayling zone, the barbel zone and the bream zone. Of these, only the trout zone was clearly present with a high dominance in the headwaters of the streams studied. For the North

European boreal streams, we suggest a zonation concept with three dominating fish species from headwaters downstream, brown trout (*Salmo trutta*), bullhead (*Cottus gobio*) and Atlantic salmon (*Salmo salar*). Discovered longitudinal shifts in fish guild compositions offered an opportunity for an ecological interpretation of the data and a promising basis for bioassessment.

Keywords Stream size · Zonation · Species richness · Guild

T. Sutela (✉)
Natural Resources Institute Finland, Oulu University, P.O. Box 413, 90014 Oulu, Finland
e-mail: tapio.sutela@luke.fi

T. Vehanen
Natural Resources Institute Finland, Latokartanonkaari 9, 00790 Helsinki, Finland

P. Jounela
Natural Resources Institute Finland, Itäinen Pitkäkatu 4A, 20520 Turku, Finland

Introduction

Worldwide, fish assemblages typically differ from the tributaries to lower reaches of streams (Hawkes, 1975; Balon & Stewart, 1983; Matthews, 1986). Species richness of fish tends to increase from headwaters to river outlets (Matthews, 1998; Muneepeerakul et al., 2008). Longitudinal changes in the local fish species richness have usually been attributed to biotic zonation (replacement) or accumulative addition of species downstream (Park et al., 2005). Biotic zonation involves discontinuities in river conditions, leading to distinct fish assemblages along the longitudinal gradient (Huet, 1959; Oberdorff et al., 1993; Belliard et al., 1997). In contrast, accumulative addition of species from tributaries to lower reaches is usually

related to environmental gradients containing smooth transition of abiotic factors (Rahel & Hubert, 1991; Park et al., 2005).

The accumulative addition of species downstream and the associated increase in species richness have often been attributed to a downstream increase in habitat diversity (Gorman & Karr, 1978; Schlosser, 1982; Angermeier & Schlosser, 1989) and in environmental stability (Horwitz, 1978; Schlosser, 1982; Park et al., 2005). Small and shallow streams are more variable with greater extremes in the range of conditions experienced by the associated communities (Jackson et al., 2001). Smallest streams are often confronted by temporary droughts (Piniewski et al., 2017).

In some studies, the overall downstream increase in species richness has been judged to result from the converging character of the river network (Fernandes et al., 2004; Muneepeerakul et al., 2008). In the Amazon, tributaries tended to enhance mainstem fish diversity but this effect did not result in overall downstream accumulation of species (Fernandes et al., 2004).

Despite the general trend of longitudinal increase in species richness downstream, some studies suggest that species richness may be highest in midsize streams (Minshall et al., 1985; Oberdorff et al., 1993; Aarts & Nienhuis, 2003). This pattern is consistent with the River Continuum Concept (Vannote et al., 1980) and suits well with the mechanism of species replacement in the zonation concept (Lasne et al., 2007). Species are replaced by other species in the next zone because of the change in environmental conditions. Thereby, species richness and diversity should be highest in the middle reaches, where species from adjacent zones may co-occur (Lasne et al., 2007). However, relatively low species richness measured in the largest rivers may in some cases be explained by a low sampling effort in relation to the size of the habitat (Hughes et al., 2002; Erös 2007, 2017) as well as extensive and intensive hydromorphological and water quality alterations (Rinne et al., 2005; Lintermans, 2007; Herlihy et al., 2020).

Thienemann (1925) proposed six zones for European rivers: spring brook, trout zone, grayling zone, barbel zone, bream zone and brackish-water. Later Huet (1959) put forward a zonation concept with a shortened list consisting of the trout zone, the grayling zone, the barbel zone, and the bream zone. Extension of this concept to Northern Europe evidently does not work because the range of barbel, *Barbus barbus* (Linnaeus, 1758), does not reach Fennoscandia (Britton & Pegg, 2011). No comprehensive fish zonation concept has been proposed for boreal rivers in Northern Europe. There is relatively little knowledge about the longitudinal patterns of stream fish assemblages in boreal rivers compared with European temperate rivers (e.g. Erös, 2007; Lasne et al., 2007). Zonation concepts may still be important for environmental management and typological purposes in bioassessment (Erös et al., 2017).

In ecological studies, fish species are often grouped into ecological guilds that exploit a resource in a similar fashion. The number of guilds usually increases downstream, and longitudinal shifts in guild composition are common (Oberdorff et al., 1993; Aarts & Nienhuis, 2003, McGarvey & Hughes, 2008). Studying the guilds facilitates an ecological interpretation of fish species data yielding new information about riverine habitats and processes. The guild structure in a fish assemblage is often more stable than species composition, because the fish species within a guild can take each other´s functional role. Thereby, fish guild composition may provide a stable basis for bioassessment (Aarts & Nienhuis, 2003).

Our main aims in this study were to 1) discover longitudinal patterns in fish assemblage structure and species richness in European boreal rivers, 2) distinguish traits supporting the concept of biotic zonation or additional accumulation of species downstream, 3) study the longitudinal changes in ecological guild compositions, and 4) outline a new fish zonation concept for boreal rivers in Northern Europe if facilitated by the data.

Materials and methods

Electrofishing data were collected from a national database (Hertta/Koekalastusrekisteri) managed by the Natural Resources Institute Finland (Luke) and hosted by the Finnish Environment Institute (SYKE). Electrofishings from subarctic northern Lapland were ruled out following the national Water Framework Directive (WFD) typology of the rivers. Otherwise, the sampling sites covered the whole area of Finland. Biogeographically, the study area is located in the

boreal region. As lowlands dominate in Finland, the maximum altitude among the sampling sites was only 300 m.

The national electrofishing data had been originally classified into four stream size classes from small to very large streams. Supplementary electrofishing data from small brooks were gathered mainly from Luke and Metsähallitus (a state-owned enterprise responsible for the management of state-owned land and water areas). Extra measurements of the catchment area above the electrofishing sites were made for the small brooks by the VALUE Citrix ArcGIS tool designed in SYKE. The total number of electrofishing samples, finally classified into five stream size classes, was 11,958 (Table 1). Approximate mean discharges (MQ) at stream size boundaries were calculated based on the average annual runoff in Finland, 10 l/s per one km^2 (https://www.syke.fi).

The electrofishing sites usually represented wadeable riffles with stony bottoms. Escape nets were not used at the sampling sites mostly covering 100–300 m^2. The majority of the sites had been sampled in late July–October and in this century. Calculated fish densities (ind./100 m^2) represent the catch of one electrofishing run. As a rule, European standard EN 14011:2003 (Water quality—sampling of fish with electricity) was followed in sampling. Samples were not pooled for later analyses. The number of fish species recorded in a single standard electrofishing was used as a unit for species richness to guarantee comparability between streams, with a varying number of samples per stream. Total species richness is usually higher than the species richness obtained from a sample (Cao et al., 2001).

Fish species were grouped into guilds (Table 2) basically following Holzer (2008). Published allocations of fish species to feeding guilds are often contradictory. For example, brown trout has been classified in different studies as invertivorous, zoobenthivorous, piscivorous, or some of their combinations (Oberdorff & Hughes, 1992; Aarts & Nienhuis, 2003; Holzer, 2008). Almost all fish species change their diet during ontogeny (Miller, 1979) and flexibility in river fish diets is common (Welcomme et al., 2006). In this study, we chose the alternative that reflects the current diet of the fish in the stream, and thereby invertivory was chosen for brown trout, most often caught by electrofishing with less than 20 cm TL in an ontogenetic phase and size dominated by invertivory.

Statistical analyses

A dendrogram of a hierarchical cluster analysis was used to visualize clusters of fish species possibly reflecting stream size gradient (IBM SPSS Statistics, version 25). Rare species with an average density of less than 0.0047 ind./100 m^2 in the whole data were excluded from the abundance data. Between groups linkage with Pearson correlation was chosen. Z-scores were used to standardize distributions of variables (Zorn et al., 2002).

Densities of fish species (ind./100 m^2) were used to predict five stream size classes with a support vector machine (SVM, Vapnik 1995, 1998), which is a group of supervised, semi-supervised and unsupervised machine learning methods used for classification, regression, clustering, anomaly detection and distribution estimation for complex data. Rare species were excluded likewise in the hierarchical cluster analysis. To balance the number of samples in the five stream size categories, a random sample of streams representing about 500 electrofishing samples was taken in the categories small rivers and medium rivers. Specifically, this study used the Java version of mySVM classification model with a dot (linear) kernel (Rüping, 2000; Mierswa et al., 2006). This model type is based on the optimization algorithm of SVMlight described

Table 1 Number of electrofishings (N) in the five stream size classes with an approximate of annual mean discharge (MQ) at the lower boundary of the stream size range

	Headwaters	Brooks	Small rivers	Medium rivers	Large rivers
Drainage basin area (km^2)	< 10	10–100	100–1000	1000–10,000	> 10,000
MQ (m^3/s)		0.1	1	10	100
N	510	416	4648	2867	517

Table 2 Guilds of the fish species

	Feeding	Feeding habitat	Reproduction
Salmo trutta Linnaeus, 1758	I	WC	LITH
Salmo salar Linnaeus, 1758	I	WC	LITH
Thymallus thymallus (Linnaeus, 1758)	I	WC	LITH
Salvelinus fontinalis (Mitchill, 1814)	I	WC	LITH
Cottus gobio Linnaeus, 1758	I	B	SPEL
Cottus poecilopus Heckel, 1837	I	B	SPEL
Barbatula barbatula (Linnaeus, 1758)	I	B	LITH
Gasterosteus aculeatus Linnaeus, 1758	I	WC	PHYT
Pungitius pungitius (Linnaeus, 1758)	I	WC	ARIAD
Rutilus rutilus (Linnaeus, 1758)	O	WC	PHLI
Leuciscus leuciscus (Linnaeus, 1758)	O	WC	LITH
Alburnus alburnus (Linnaeus, 1758)	PLAN	WC	PHLI
Squalius cephalus (Linnaeus, 1758)	O	WC	LITH
Gobio gobio (Linnaeus, 1758)	I	B	PSAM
Phoxinus phoxinus (Linnaeus, 1758)	I	WC	LITH
Lota lota (Linnaeus, 1758)	P	B	LITH
Esox lucius Linnaeus, 1758	P	P	WC
Perca fluviatilis Linnaeus, 1758	P	WC	PHLI
Gymnocephalus cernua (Linnaeus, 1758)	I	B	PHLI
Anguilla anguilla (Linnaeus, 1758)	I	B	PELA
Lampetra fluviatilis (Linnaeus, 1758)	P	B	LITH

I invertivorous, *O* omnivorous, *P* piscivorous, *PLAN* planktivorous, *WC* water column feeder, *B* benthic feeder, *LITH* lithophilic, *PHYT* phytophilic, *SPEL* speleophilic, *PHLI* phytolithophilic, *ARIAD* ariadnophilic, *PELA* pelagophilic, *PSAM* psammophilic (Oberdorff & Hughes 1992; Holzer 2008)

in Joachims (1999). Error-correcting output coding (ECOC, Dietterich & Bakiri, 1995) was used for decomposing a multiway classification problem into many binary classification tasks, and then combining the results of the subtasks into a hypothesized solution to the original problem. The absolute (plus or minus) importance of each predictor on predicted stream size classes was estimated using sensitivity analysis (Olden & Jackson, 2002) with the best model. The predictors of five stream size classes were normalized using a zeroed mean with a variance of one with the aim of avoiding bias caused by the very high or very low values of some predictors. Information leaks from training to validation were prevented by the separate normalization of the training and validation sets. Emphasis was put on the avoidance of over- and underfitting (Bishop, 2006; Hastie et al., 2008), by carefully optimizing the complexity parameter C (also called "capacity" and "regularization"). Too large C values can lead to overfitting and too small values to overgeneralization. The best model fit (accuracy) was sought by optimizing SVM parameters C and the insensitivity (also called "slack") parameter ε using 10-fold cross-validation (Kohavi, 1995) applied to

sequential grid-search. The electrofishing data for training and validation comprised 60% of the randomly selected data rows (divided by stratified sampling). Having identified the best model parameters, the remaining 40% of the data rows were used as a holdout set to test the best model's performance. The total number of predictors was rather low (22 species) compared to the capabilities of the SVM model, and therefore the feature (predictor) selection methods (Guyon & Elissee, 2003) were not used. The SVM analyses were performed using RapidMiner software (https://rapidminer.com, version Studio Large 9.5.001, Mierswa et al., 2006).

Results

Average species richness of fish in the electrofishing samples increased steadily from headwaters to medium rivers, but levelled off when shifting from medium to large rivers (Fig. 1, *t* test between medium and large rivers, $P = 0.742$). Electrofishing trials with no catch (species richness = 0) were included in these
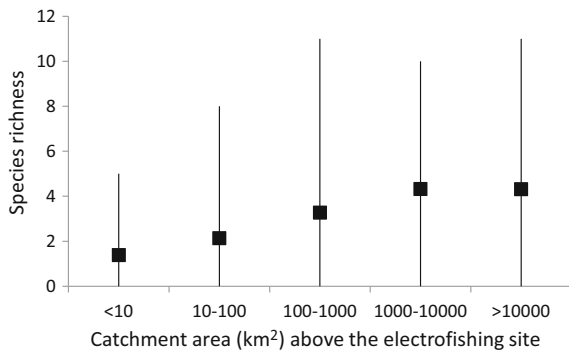
Fig. 1 Species richness (average and range) in a single electrofishing catch in the five stream size groups

calculations, and they were clearly highest (12.2%) in the headwaters.

Frequency of occurrence of ninespine stickleback (*Pungitius pungitius*), three-spined stickleback (*Gasterosteus aculeatus*) and brook trout (*Salvelinus fontinalis*) was highest in the headwaters indicating clear preference on the smallest stream size class (Fig. 2). The existence of these apparent small stream specialist species was interpreted to support the fish zonation concept. At the other edge of the stream size gradient, especially stone loach (*Barbatula barbatula*), Atlantic salmon (*Salmo salar*), chub (*Squalius cephalus*) and dace (*Leuciscus leuciscus*) showed a preference for the largest streams (Fig. 2b). The four salmonid species in our material displayed clearly different preferences (Fig. 2c). Brown trout (*Salmo trutta*) occurred frequently in all stream sizes (Fig. 2), but its average density decreased from headwaters to large rivers (Table 3).

Average total density of fish was steady from headwaters to small rivers but peaked at medium rivers (Table 3). Altogether eight fish species displayed their highest average density in large rivers (Table 3). Of all the fish species, brown trout showed the highest average density in headwaters and brooks, bullhead (*Cottus gobio*) in small and medium rivers, whereas Atlantic salmon dominated in the large rivers (Table 3). These three fish species also dominated in terms of density proportions in those river classes (Fig. 3). In frequencies of occurrence, brown trout dominated in headwaters (0.62) and brooks (0.62), and bullhead in small (0.62) and medium (0.81) rivers. In the large rivers, stone loach dominated (0.68) but
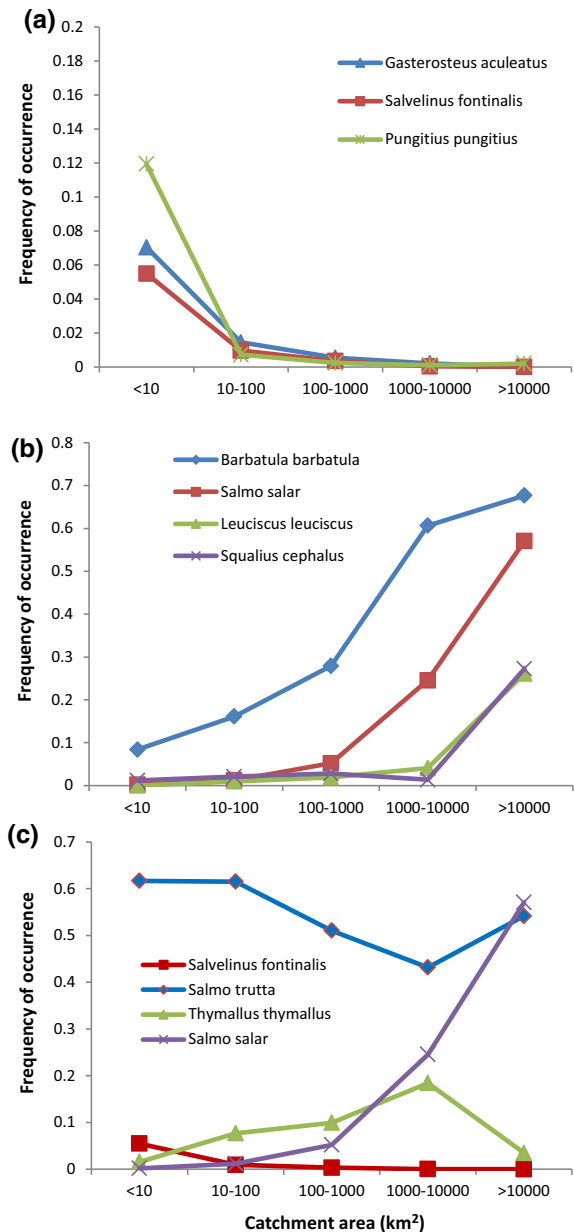

Fig. 2 Frequencies of occurrence of some fish species favouring small streams (**a**), favouring large streams (**b**), and a combination of salmonid species (**c**)

Atlantic salmon showed the second highest (0.57) frequency of occurrence (Fig. 2).

In the guild classification of feeding form, a clear dominance of water column feeders in headwaters was gradually shifted to a slight dominance of benthic feeders in medium rivers (Fig. 4). Invertivory was the prevalent feeding guild in all stream size classes but
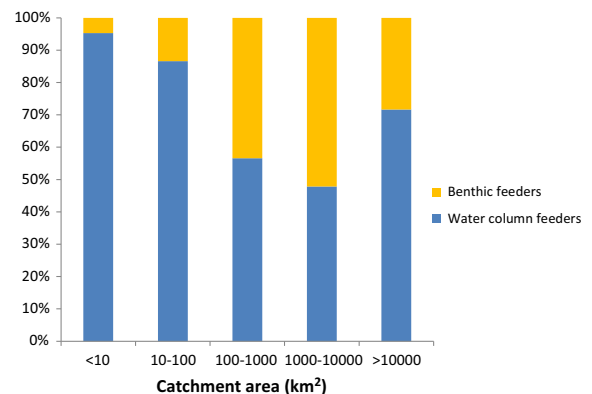
Table 3 Mean densities of the fish species in one electrofishing run in the five river basin size classes (km²)

Catchment size (km²)	< 10	10–100	100–1000	1000–10,000	> 10,000
Salmo trutta	*12.020*	*9.407*	5.181	2.225	1.877
Salmo salar	0.037	0.167	0.257	2.310	**8.426**
Thymallus thymallus	0.045	0.226	0.185	**0.349**	0.055
Salvelinus fontinalis	**0.946**	0.054	0.005	0.000	0.000
Cottus gobio	0.044	0.912	*5.676*	**10.825**	2.280
Cottus poecilopus	0.000	0.082	0.051	0.071	**0.174**
Barbatula barbatula	0.495	1.062	1.792	3.641	**3.852**
Pungitius pungitius	**2.609**	0.051	0.005	0.001	0.001
Gasterosteus aculeatus	**0.515**	0.132	0.034	0.016	0.000
Rutilus rutilus	0.251	1.091	2.327	**2.377**	0.774
Leuciscus leuciscus	0.000	0.024	0.069	0.131	**1.984**
Alburnus alburnus	0.084	0.022	0.621	**1.305**	0.854
Squalius cephalus	0.086	0.020	0.122	0.054	**1.461**
Gobio gobio	0.069	0.044	**0.285**	0.129	0.179
Phoxinus phoxinus	1.530	**4.092**	0.231	3.783	1.217
Lota lota	0.223	0.435	0.658	**1.125**	0.446
Esox lucius	0.270	**0.458**	0.210	0.130	0.086
Perca fluviatilis	0.268	1.718	1.886	1.851	**2.093**
Gymnocephalus cernua	0.001	0.019	0.053	0.050	**0.060**
Anguilla anguilla	0.000	0.000	0.005	0.007	**0.442**
Lampetra planeri	0.075	**0.137**	0.008	0.001	0.005
Lampetra fluviatilis	**0.013**	0.008	0.006	0.006	0.009
Others	0.003	0.015	0.078	0.048	0.112
Sum	19.583	20.176	19.744	30.431	26.387

Numbers in bold represent the highest species-specific value (in lines), whereas the numbers in italics represent the highest value within each size category (in columns)
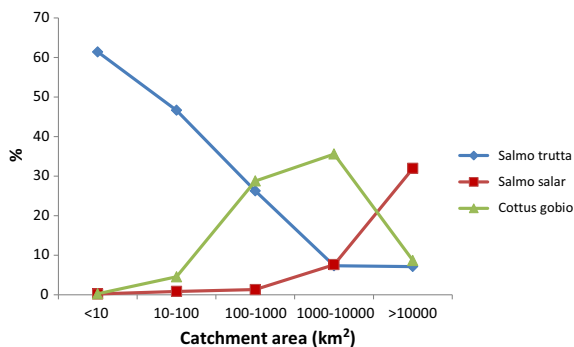
Fig. 3 Percentage proportions of the three dominant fish species of the total fish density in the five stream size categories

Fig. 4 Percentages of fish belonging to the two feeding habitat guilds in the five stream size groups

especially in headwaters, whereas planktivory was met only at the three largest stream size classes (Fig. 5). In reproductive guilds, lithophilic fish dominated as a rule but speleophilic fish took a substantial share in small and medium rivers (Fig. 6). The number of guilds remained the same or slightly increased from headwaters to larger stream size classes.

In the dendrogram of the hierarchical cluster analysis, the six lowermost fish species in the figure, clustered in the first split, are all small stream species (Fig. 7). This cluster includes all of the five species having their maximum average density in headwaters,
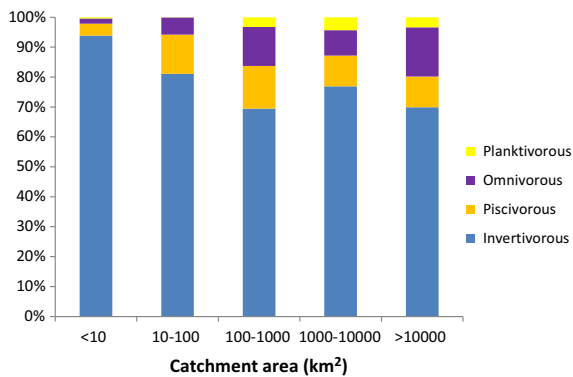
Fig. 5 Percentages of fish belonging to the four feeding guilds in the five stream size groups
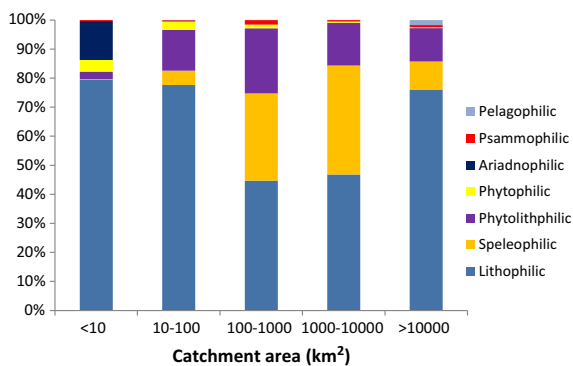

Fig. 6 Percentages of fish belonging to the seven reproductive guilds
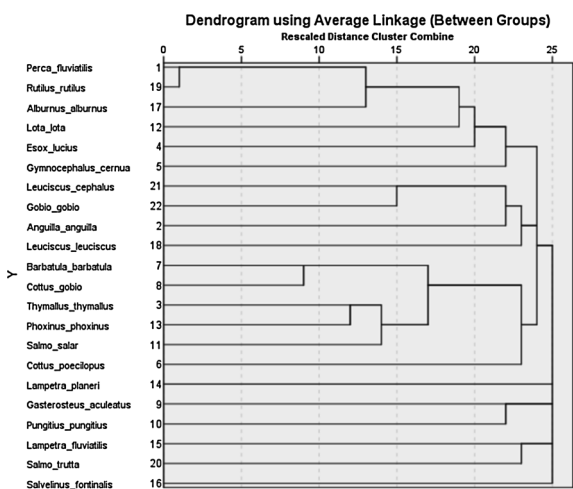

Fig. 7 Dendrogram of the hierarchical cluster analysis from fish abundance data

and brook lamprey (*Lampetra planeri*), with the second highest density in headwaters (Table 3).

The confusion matrix of the SVM classification model suggested that species composition in large rivers differed somewhat from that in medium-sized and small rivers (Table 4). This is because the class precision of large stream types (71.53%) is much higher than that in medium and small stream types, which can be confused (36.68% and 45.65%, respectively; Table 4). In addition, brooks are confused with headwaters, suggesting that species composition in brooks is similar to that in headwaters. Class recall was highest with headwaters, which indicates the most unique species composition in this stream size class (Table 4). In absolute terms, eel (*Anguilla anguilla*) and ninespine stickleback (*Pungitius pungitius*) are the two most important predictors for the classification of river size classes (Fig. 8). The cross-validation performance of the SVM model was 49.19%; the holdout set performance was 47.77%.

Discussion

In this study, distinct longitudinal patterns in fish species composition, species richness and guild composition in boreal European streams were found. In rivers, the uni-directional character of water tends to lead to a linear structure along a gradient of environmental conditions, and thus, biological assemblages are organized longitudinally (Park et al., 2005). Hydrological characteristics may be viewed as an environmental filter for stream assemblages by ultimately selecting the pool of species that can meet their biological requirements (Costa et al., 2018).

Species richness

Our results indicated an overall increasing trend in species richness along stream size, concordant with the mainstream in the literature (e.g. Matthews 1998, Esselman et al., 2006, Reyjol et al., 2007). However, species richness levelled off when shifting from medium to large rivers. In the Seine basin in northern France, fish species richness increased with river size, reached a maximum in midsized rivers, and again decreased in large rivers (Oberdorff et al., 1993). Also, in two other Central European river systems, species richness was highest in midsized streams (Aarts &

Table 4 Confusion matrix of the holdout set using a SVM classification model (predicted vs observed, accuracy 47.77%)

Stream size	Large rivers (obs.)	Medium rivers (obs.)	Small rivers (obs.)	Brooks (obs.)	Headwaters (obs.)	Precision (%)
Large rivers (pred.)	98	27	5	5	2	71.53
Medium rivers (pred.)	82	157	108	56	25	36.68
Small rivers (pred.)	13	9	42	17	11	45.65
Brooks (pred.)	2	5	4	13	2	50.00
Headwaters (pred.)	16	9	42	77	163	53.09
Class recall (%)	46.44	75.84	20.89	7.73	80.29	
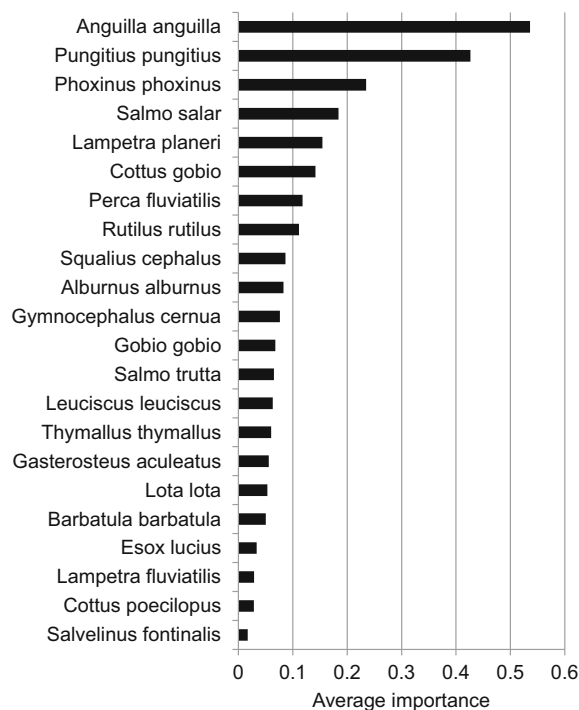
Fig. 8 Predictor importances in classifying five river size classes with a support vector machine

Nienhuis, 2003). These results from the two studies are in accordance with the River Continuum Concept (Vannote et al., 1980) which predicts the maximum biotic diversity for midsized streams. Our results of the average species richness in electrofishing samples settled in between these two dominant concepts, the general longitudinal increase in fish species richness downstream (Matthews, 1998) and the midsize stream maximum in species richness predicted by the River Continuum Concept.

The species–area relationship (Preston, 1962) explains the species richness as a power function of surface area. Species–volume relationship stresses the linkage between habitat volume and fish species richness in rivers (Angermeier & Schlosser, 1989). Species–discharge relationships (SDR) are analogues of these relationships (McGarvey & Ward, 2008). The low species richness discovered in the narrow and shallow headwater streams with low discharge suits these hypotheses well.

Species richness observed in a field study usually results in an underestimate, because some species are not included in the samples (Cao et al., 2001). Some studies have addressed the need for much larger sample areas in large rivers to balance the representativeness of the samples (Hughes et al., 2002; Erös, 2017). In this perspective, our material gained from a national database was not optimal.

In this study, species richness was studied based on sampling in wadeable riffles. Deep river sections, which often have a low current, and which are most frequently encountered in large rivers, were not sampled. The proportion of wadeable riffles of the streambed area usually decreases downstream. Habitat diversity tends to increase downstream (Gorman & Karr, 1978; Angermeier & Schlosser, 1989). Supplementary sampling of the streams studied with different equipment in different habitats could give a different perspective to species richness examination (Galat et al., 2005; Loisl et al., 2014).

Longitudinal changes in fish species composition

To sum up the results about the frequency of observations of fish species, their average densities,

the dendrogram from the hierarchical cluster analysis and predictions of SVM classification model, longitudinal changes in fish species composition were evident, being most highlighted at the headwater edge of the stream size gradient. Fish assemblages are structured by environmental filtering, especially in headwater streams (Zbinden & Matthews, 2017). The relatively high accuracy of the SVM model also suggests longitudinal changes in species composition from large rivers to headwaters.

Fishes in small streams are often habitat specialists (Gorman & Karr, 1978). Within small streams, the impact of predation may exclude other species, thereby leading to mutually exclusive distributions and strong differences in community composition (Jackson et al., 2001). A study of boreal rivers by Eklöv et al. (1998) suggested that brown trout may exclude ninespine stickleback by predation or competition in small streams. In our study, no clear indications of such an interaction was noticed, as brown trout was discovered at about 73% of the sites occupied by ninespine stickleback.

Distribution of brook trout, an alien species in Europe, was closely restricted to headwaters. Brook trout also favours small streams in its home district in North America where brown trout as an alien species has expelled brook trout to still smaller brooks (Rahel & Nibbelink, 1999; Kanno et al., 2015). In Finland, brook trout was found to exclude brown trout in tributary streams (Korsu et al., 2007). This result accords with brook trout's high preference for headwaters (Fig. 2; Table 3) and the dendrogram from the cluster analysis indicating brook trout's relatively distinct position (Fig. 7). In Sweden, the frequency of brook trout one-species systems increased in the smallest streams (Öhlund et al., 2008). According to Carlsson et al. (2004), the upper reaches of a Swedish stream supported brook trout. North European streams are species-poor, mirroring the legacy of the last ice age (Reyjol et al., 2007). Korsu et al. (2012) suggested that streams in Northern Europe lack a headwater specialist fish. This offers opportunities for alien invaders tolerant of headwater conditions, such as brook trout, to establish breeding populations. Our results about the high preference of brook trout on small streams often in single-species populations fits well with the concept put forward by Korsu et al. (2012).

In our study, salmonids (brown trout and brook trout) dominated in the majority of the smallest tributaries. High swimming skills of salmonids may be valuable in the smallest brooks where drought periods may force the fish to move downstream occasionally for refuge. Juvenile brown trout escaped downstream to stream outlets during periods of drought in the island of Bornholm in the Baltic Sea (Østergaard et al., 2003). On their return migration back to tributaries, salmonids can negotiate natural barriers such as waterfalls and high-gradient stream sections better than cottids (Maret et al., 1997). In rivers, cottids migrate less than salmonids (Radinger & Wolter, 2014). Dominance of salmonids and low frequency of bottom-dwelling fish species (cottids and stone loach) in the headwaters studied is suggested to be linked to their swimming skills.

In French rivers, longitudinal occurrence patterns with high resemblance to our material were found for many species, e.g. brown trout and bullhead (Oberdorff et al., 2001). Also in the Seine River basin studied by Belliard et al. (1997) most fish species displayed quite similar patterns. The clearest difference to our study was in the two stickleback species which occupied headwaters as well as the lower reaches in the Seine River basin. Also the occurrence of stone loach and bullhead was more uniform across stream size classes compared to our study. Contrary to our results, in southern France bullhead inhabited chiefly small and shallow streams close to the source (Legalle et al., 2005). In a Swedish boreal stream Färsån, brown trout and brook trout occupied the headwaters while bullhead, grayling *(Thymallus thymallus)* and European minnow (*Phoxinus phoxinus*) accompanied them in the lower reaches as in our study (Carlsson et al., 2004).

Fish zonation

Two basic concepts have been put forward to describe the longitudinal change in the fish assemblages, the zonation concept with species replacement downstream (e.g. Huet, 1959; Schlosser, 1982), and the accumulative addition of species downstream (Sheldon, 1968; Rahel & Hubert, 1991, Park et al., 2005). In our material, the existence of the three headwater specialists (brook trout, two stickleback species) supports the zonation concept. However, gradual accumulative addition of species from headwaters to

large rivers could also be recognized. Our results are concordant with Rahel & Hubert (1991) and Lasne et al. (2007) who suggest that these two mechanisms may coexist.

A comparison of our results of the fish species compositions to European zonation concept popularized by Huet (1959) with trout, grayling, barbel and bream zones from headwaters downstream reveals high correspondence to our material in the case of the trout zone. Frequency and average density of grayling was highest in medium rivers expressing some correspondence with Huet (1959) zonation. However, no signs of barbel zone or bream zone was observed in our material. Our study area is out the geographic range of barbel, whereas bream (*Abramis brama*), occasionally recorded in our material, was totally missing from the large rivers sampled.

The longitudinal profile of large rivers in Central Europe (e.g. the Rhine and Danube) is often concave, and thereby lower reaches are characterized by low slope and low currents (Welcomme, 1983). In contrast, the longitudinal profile of the large rivers in our study is rather linear facilitating riffles also in the lower reaches if not harnessed by power plants. In our view, this difference at least partly explains the dominance of the rheophilic Atlantic salmon in the boreal large rivers studied and the dominance of the limnophilic bream in the large rivers of Central Europe.

The use of zonation concepts in ichthyology has been limited for a long time because of some serious shortcomings. The concept suggests discrete zones, whereas in real rivers gradual transition of conditions and fish assemblages prevail (Aarts & Nienhuis, 2003). Zonation concepts can describe, but cannot explain the longitudinal changes. European fish zonation concepts (e.g. Huet, 1959) have been deemed too simplistic in order to acquire any ecological significance (Balon & Stewart, 1983). As confronted in this study, the zonation concepts usually cannot be extended to other ecoregions. Fish assemblages are too complex to be identified with a single species (Ibarra et al., 2005). The focus in zonation concepts should be broadened from a single species giving the name for the zone to the entire fish assemblage in the more or less discrete zone of the stream. As Lasne et al. (2007) suggests, zonation concepts can be useful in conceptual and conservation perspectives, if refined

by additional information on individual species patterns.

In our material, brown trout dominated in headwaters and brooks, bullhead in small and median rivers, and Atlantic salmon in large rivers. With the preconditions described in the previous paragraph, we suggest a three-step zonation named after these fish species for boreal European streams. These three fish species have been classified as intolerant fish species (Vehanen et al., 2010), and their role in the boreal riverine fish assemblages could be even higher without human-induced impairment in water quality and river connectivity. Disturbed fish zonation in Warta River, Poland, was interpreted to reflect pollution and hydromorphological degradation (Kruk et al., 2017). Disturbed zonation in boreal rivers could also be used as an indicator of human-induced riverine impairment.

Our data combined from thousands of streams give no direct indication of how distinct this brown trout-bullhead-salmon zonation is in a single stream or catchment area scale. The zones will probably also appear as more or less indiscrete in single streams, as they naturally appear in our average data drawn from extensive material. As cited earlier, gradual downstream transitions of fish assemblages usually prevail (Aarts & Nienhuis, 2003).

Ecological guilds and fish-based bioassessment

The proportion of invertivorous fish has been documented to decrease from headwaters to mainstream in several studies (Matthews, 2008; Ibanez et al., 2009; Wolff et al., 2013), which was also the case in our study. Emergence of planktivorous fish only at the three largest size classes was quite predictable when considering the food resources available in the longitudinal river gradient downstream. This pattern is in concordance with several other studies (e.g. Matthews 2008; Aarts & Nienhuis, 2003). The high percentage of cottids in the small and medium rivers was reflected as high incidence of benthic and speleophilic guilds. Any close resemblance to this sequence in other river systems was not encountered in the literature.

Since the pioneering work of Karr (1981), numerous multimetric indices (MMIs) of biological condition have been developed and applied globally (Ruaro et al., 2020). Recently, bioassessment has been increasingly conducted at regional and continental scales using extensive fish datasets (Esselman et al.,

2013; Schinegger et al., 2016). To do so, many MMI metrics have been calibrated for stream size (Pont et al., 2009). Bioassessment at large geographical extents involves the incorporation of ecoregions and fish assemblage types (Whittier et al., 2007; Schinegger et al., 2016), which vary with stream size. Fish guild composition has been suggested to offer a stable basis for bioassessment (Karr, 1981; Schmutz et al., 2000). Temporal changes in guild composition often reflect human disturbances (Aarts & Nienhuis, 2003). Guilds may function as ´super-species´ in their response to flow and morphological changes (Austen et al., 1994). One advantage of the guild concept is its suitability for different ecoregions (Schmutz et al., 2000).

In fish-based bioassessment, reference sites should be of the same stream size class (or zone) as the affected sites. Alternatively, a reference condition modelling approach based on a continuous stream size gradient can be applied (McCormick et al., 2001; Pont et al., 2006, 2009; Whittier et al., 2007). Specific patterns discovered in fish assemblages of this study from headwaters to river outlets stress the need to incorporate longitudinal gradient in bioassessment. Vile & Henning (2018) developed a special method for bioassessment in headwater streams because of distinct fish fauna therein. Low fish species richness with associated difficulty in developing fish indices has also been encountered in Mediterranean rivers (Miller et al., 1988, Hermoso et al., 2010). North and West European rivers maintain low numbers of fish species due to the most recent ice age (Reyjol et al., 2007). This is a challenge for fish-based bioassessment especially with the smallest streams expressing very low species richness and high probability for zero catches as found in this study.

References

Aarts, B. G. W. & P. H. Nienhuis, 2003. Fish zonations and guilds as the basis for assessment of ecological integrity of large rivers: aquatic Biodiversity. Hydrobiologia 500: 157–178.

Angermeier, P. L. & I. J. Schlosser, 1989. Species-area relationships for stream fishes. Ecology 70: 1450–1462.

Austen, D. J., P. B. Bayley & B. W. Menzel, 1994. Importance of the guild concept to fisheries research and management. Fisheries 19: 12–20.

Balon, E. K. & D. J. Stewart, 1983. Fish assemblages in a river with unusual gradient (Luongo, Africa-Zaire system), reflections on river zonation, and description of another new species. Environmental Biology of Fishes 9: 225–252.

Belliard, J., P. Boët & E. Tales, 1997. Regional and longitudinal patterns of fish community structure in the Seine River basin, France. Environmental Biology of Fishes 50: 133–147.

Bishop, C. M., 2006. Pattern Recognition and Machine Learning. Springer, New York.

Britton, J. R. & J. Pegg, 2011. Ecology of European Barbel *Barbus Barbus*: implications for river, fishery, and conservation management. Reviews in Fisheries Science 19: 321–330.

Cao, Y., D. P. Larsen & R. M. Hughes, 2001. Evaluating sampling sufficiency in fish assemblage surveys: a similarity-based approach. Canadian Journal of Aquatic Sciences 58: 1782–1793.

Carlsson, J., K. Aarestrup, F. Nordwall, I. Näslund, T. Eriksson & J. E. L. Carlsson, 2004. Migration of landlocked brown trout in two Scandinavian streams as revealed from trap data. Ecology of Freshwater Fish 13: 161–167.

Costa, I. D., A. C. Petry & R. Mazzoni, 2018. Responses of fish assemblages to subtle elevations in headwater streams in southwestern Amazonia. Hydrobiologia 809: 175–184.

Dietterich, T. G. & G. Bakiri, 1995. Solving multiclass learning problems via error-correcting output codes. Journal of Artificial Intelligence Research 2: 263–286.

Eklöv, A. G., A. Greenberg, P. Brönmark & O. Berglund, 1998. Response of stream fish to improved water quality: a comparison between the 1960s and 1990s. Freshwater Biology 40: 771–782.

Erös, T., 2007. Partitioning the diversity of riverine fish: the roles of habitat types and non-native species. Freshwater Biology 52: 1400–1415.

Erős, T., 2017. Scaling metacommunities in stream networks: synthesis and future research avenues. Community Ecology 18: 72–86.

Erős, T., V. Bammer, A. I. György, L. Pehlivanov, M. Schabuss, H. Zornig, A. Weiperth & Z. Szaloky, 2017. Typology of a great river using fish assemblages: implications for the bioassessment of the Danube river. River Research and Applications 33: 37–49.

Esselman, P. C., M. C. Freeman & C. M. Pringle, 2006. Fish-assemblage variation between geologically defined regions and across a longitudinal gradient in the Monkey River Basin, Belize. Journal of the North American Benthological Society 25: 142–156.

Esselman, P. C., D. M. Infante, L. Wang, A. R. Cooper, D. Wieferich, Y.-P. Tsang, D. J. Thornbrugh & W. W. Taylor, 2013. Regional fish community indicators of landscape disturbance to catchments of the conterminous United States. Ecological Indicators 26: 163–173.

Fernandes, C. C., J. Podos & J. G. Lundberg, 2004. Amazonian Ecology: tributaries enhance the diversity of electric fishes. Science 305: 1960–1962.

Galat, D. L., C. R. Berry, W. M. Gardner, J. C. Hendricson, G. E. Mestl, G. J. Power, C. Stone & M. R. Winston, 2005. Spatiotemporal patterns and changes in Missouri River fishes. American Fisheries Society Symposium 45: 249–291.

Gorman, O. T. & J. R. Karr, 1978. Habitat structure and stream fish communities. Ecology 59: 507–515.

Guyon, I. & A. Elissee, 2003. An introduction to variable and feature selection. Journal of Machine Learning Research 3: 1157–1182.

Hastie, T., R. Tibshirani & J. Friedman, 2008. The elements of statistical learning: data mining, inference and prediction. Springer, New York.

Hawkes, H. A., 1975. River zonation and classification. In Whitton, B. A. (ed.), River Ecology. University of California Press, Berkeley: 312–374.

Herlihy, A. T., J. C. Sifneos, R. M. Hughes, D. V. Peck & R. M. Mitchell, 2020. Relation of lotic fish and benthic macroinvertebrate condition indices to environmental factors across the conterminous USA. Ecological Indicators 112(May): 105958.

Hermoso, V., M. Clavero, F. Blanco-Garrido & J. Prenda, 2010. Assessing the ecological status in species-poor systems: a fish-based index for Mediterranean Rivers (Guadiana River, SW Spain). Ecological Indicators 10: 1152–1161.

Holzer, S., 2008. European Fish Species: Taxa and Guilds Classification Regarding Fish-Based Assessment Methods. Universität Wien, Diplomarbeit: 196.

Huet, M., 1959. Profiles and biology of Western European streams as related to fish management. Transaction of the American Fisheries Society 88: 155–163.

Hughes, R. M., P. R. Kaufmann, A. T. Herlihy, S. S. Intelmann, S. C. Corbett, M. C. Arbogast & R. C. Hjort, 2002. Electrofishing distance needed to estimate fish species richness in raftable Oregon rivers. North American Journal of Fisheries Management 22: 1229–1240.

Ibanez, C., J. Belliard, R. M. Hughes, P. Irz, A. Kamdem-Toham, N. Lamouroux, P. A. Tedesco & T. Oberdorff, 2009. Convergence of temperate and tropical stream fish assemblages. Ecography 32: 658–670.

Ibarra, A. A., Y.-S. Park, S. Brosse, Y. Reyjol, P. Lim & S. Lek, 2005. Nested patterns of spatial diversity revealed for fish assemblages in a west European river. Ecology of Freshwater Fish 14: 233–242.

Jackson, D. A., P. R. Peres-Neto & J. D. Olden, 2001. What controls who is where in freshwater fish communities: the roles of biotic, abiotic, and spatial factors. Canadian Journal of Fisheries and Aquatic Sciences 58: 157–170.

Joachims, T., 1999. Making large-scale SVM learning practical. In: Advances in Kernel Methods-Support Vector Learning, chapter 11. MIT Press.

Kanno, Y., B. H. Letcher, A. L. Rosner & K. P. O´Neil, 2015. Environmental factors affecting brook trout occurrence in headwater stream segments. Transactions of the American Fisheries Society 144: 373–382.

Karr, J. R., 1981. Assessment of biotic integrity using fish communities. Fisheries 6: 21–27.

Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: IJCAI-95 Proceedings of the 14th international joint conference on artificial intelligence. Morgan Kaufmann Publishers, San Francisco, CA, 1137–1143.

Korsu, K., A. Huusko & T. Muotka, 2007. Niche characteristics explain the reciprocal invasion success of stream salmonids in different continents. Proceedings of the National Academy of Sciences of the United States of America 104: 9725–9729.

Korsu, K., J. Heino, A. Huusko & T. Muotka, 2012. Specific niche characteristics facilitate the invasion of an alien fish invader in boreal streams. International Journal of Ecology. Article ID 813016, 10 pp.

Kruk, A., M. Ciepłucha, D. Zięba, G. Błońska, L. Marszał, S. Tybulczuk, M. Tszydel & T. Penczak, 2017. Disturbed fish fauna zonation as an indicator of large-scale human impact: a case study (2011–2012) of the large, lowland Warta River, Poland. Journal of Applied Ichthyology 33: 174–188.

Lasne, E., B. Bergerot, S. Lek & P. Laffaille, 2007. Fish zonation and indicator species for the evaluation of the ecological status of rivers: example of the Loire basin (France). River Research and Applications 23: 877–890.

Legalle, M., F. Santoul, J. Figuerola, S. Mastrorillo & R. Céréghino, 2005. Factors influencing the spatial distribution patterns of the bullhead (*Cottus gobio* L., Teleostei Cottidae): a multi-scale study. Biodiversity and Conservation 14: 1319–1334.

Lintermans, M., 2007. Fishes of the Murray-Darling Basin: An Introductory Guide. Murray-Darling Basin Commission, Canberra.

Loisl, F., G. Singer & H. Keckeis, 2014. Method-integrated fish assemblage structure at two spatial scales along a free-flowing stretch of the Austrian Danube. Hydrobiologia 729: 77–94.

Maret, T. R., C. T. Robinson & G. W. Minshall, 1997. Fish assemblages and environmental correlates in least-disturbed streams of the upper Snake River basin. Transactions of the American Fisheries Society 126: 200–216.

Matthews, W. J., 1986. Fish faunal 'breaks' and stream order in the eastern and central United States. Environmental Biology of Fishes 17: 81–92.

Matthews, W. J., 1998. Patterns in Freshwater Fish Ecology. Chapman & Hall, New York.

McCormick, F. H., R. M. Hughes, P. R. Kaufmann, A. T. Herlihy & D. V. Peck, 2001. Development of an index of biotic integrity for the Mid-Atlantic Highlands Region. Transactions of the American Fisheries Society 130: 857–877.

McGarvey, D. J. & R. M. Hughes, 2008. Longitudinal zonation of Pacific Northwest (USA) fish assemblages and the species-discharge relationship. Copeia 2: 311–312.

McGarvey, D. J. & G. M. Ward, 2008. Scale dependence in the species-discharge relationship for fishes of the southeastern USA. Freshwater Biology 53: 2206–2219.

Mierswa, I., M. Wurst, R. Klinkenberg, M. Scholz & T. Euler, 2006. Yale: rapid prototyping for complex data mining tasks. In: Proceedings of the 12th ACM SIGKDD international conference on knowledge discovery and data mining (KDD '06). ACM, New York, NY, pp. 935–940. URL: https://rapidminer.com.

Miller, R. J., 1979. Relationship between habitat and feeding mechanisms in fishes. In Clepper, H. (ed.), Predatory-Prey Systems in Fisheries Management. Sport Fishing Institute, Washington, DC: 269–280.

Miller, D. L., P. M. Leonard, R. M. Hughes, J. R. Karr, P. B. Moyle, L. H. Schrader, B. A. Thompson, R. A. Daniel, K. D. Fausch, G. A. Fitzhugh, J. R. Gammon, D. B. Halliwell, P. L. Angermeier & D. J. Orth, 1988. Regional applications of an index of biotic integrity for use in water resource management. Fisheries 13: 12–20.

Minshall, G. W., R. C. Petersen Jr. & C. F. Nimz, 1985. Species richness in streams of different size from the same drainage basin. American Naturalist 125: 16–38.

Muneepeerakul, R., E. Bertuzzo, H. L. Lynch, W. F. Fagan, A. Rinaldo & I. Rodriguez-Iturbe, 2008. Neutral metacommunity models predict fish diversity patterns in Mississippi-Missouri basin. Nature 453: 220–223.

Oberdorff, T. & R. M. Hughes, 1992. Modification of an index of biotic integrity based on fish assemblages to characterize rivers of the Seine Basin, France. Hydrobiologia 228: 117–130.

Oberdorff, T., E. Guilbert & J.-C. Lucchetta, 1993. Patterns of fish species richness in the Seine River basin, France. Hydrobiologia 259: 157–167.

Oberdorff, T., D. Pont, B. Hugueny & D. Chessel, 2001. A probabilistic model characterizing fish assemblages of French rivers: a framework for environmental assessment. Freshwater Biology 46: 399–415.

Öhlund, G., F. Nordwall, E. Degerman & T. Eriksson, 2008. Life history and large-scale habitat use of brown trout (Salmo trutta) and brook trout (Salvelinus fontinalis): implications for species replacement patterns. Canadian Journal of Fisheries and Aquatic Sciences 65: 633–644.

Olden, J. D. & D. A. Jackson, 2002. Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks. Ecological Modelling 154: 135–150.

Østergaard, S., M. Hansen, V. Loeschke & E. E. Nielsen, 2003. Long-term temporal changes of genetic composition in brown trout (Salmo trutta L.) populations inhabiting an unstable environment. Molecular Ecology 12: 3123–3135.

Park, Y.-S., T. Oberdorff & S. Lek, 2005. Patterning riverine fish assemblages using an unsupervised neural network. In Lek, S., M. Scardi, P. F. Verdonschot, J. P. Descy & Y.-S. Park (eds.), Modelling Community Structure in Freshwater Ecosystems. Springer, Berlin: 43–53.

Piniewski, M., C. Prudhomme, M. C. Acreman, L. Tylec, P. Oglecki & T. Okruszko, 2017. Responses of fish and invertebrates to floods and droughts in Europe. Ecohydrology 10: e1793.

Pont, D., B. Hugueny, U. Beier, D. Goffaux, A. Melcher, R. Noble, C. Rogers, N. Roset & S. Schmutz, 2006. Assessing river biotic condition at a continental scale: a European approach using functional metrics and fish assemblages. Journal of Applied Ecology 43: 70–80.

Pont, D., R. M. Hughes, T. R. Whittier & S. Schmutz, 2009. A predictive index of biotic integrity model for aquatic-vertebrate assemblages of western U.S. streams. Transactions of the American Fisheries Society 138: 292–305.

Preston, F. W., 1962. The canonical distribution of commonness and rarity: Part I. Ecology 43: 185–215.

Radinger, J. & C. Wolter, 2014. Patterns and predictors of fish dispersal in rivers. Fish and Fisheries 15: 456–473.

Rahel, F. J. & W. A. Hubert, 1991. Fish assemblages and habitat gradients in a Rocky Mountain–Great Plains stream: biotic zonation and additive patterns of community change. Transactions of the American Fisheries Society 120: 319–332.

Rahel, F. J. & N. P. Nibbelink, 1999. Spatial patterns in relations among brown trout (Salmo trutta) distribution, summer air temperature, and stream size in Rocky Mountain streams. Canadian Journal of Fisheries and Aquatic Sciences 56: 43–51.

Reyjol, Y., B. Hugyeny, D. Pont, P. G. Bianco, U. Beier, N. Caiola, F. Casals, I. Cowx, A. Economou, T. Ferreira, G. Haidvogl, R. Noble, A. de Sostoa, T. Vigneron & T. Virbickas, 2007. Patterns in species richness and endemism of European freshwater fish. Global Ecology and Biogeography 16: 65–75.

Rinne, J. N., R. M. Hughes & B. Calamusso (eds.). 2005. Historical changes in large river fish assemblages of the Americas. American Fisheries Society, Symposium 45.

Ruaro, R., A. A. Gubiani, R. M. Hughes & R. P. Mormul, 2020. Global trends and challenges in multimetric indices of biological condition. Ecological Indicators. https://doi.org/10.1016/j.ecolind.2019.105862.

Rüping, S., 2000. mySVM-Manual. Universität Dortmund, Lehrstuhl Informatik VIII, 2000. http://wwwai.cs.uni-dortmund.de/SOFTWARE/MYSVM/.

Schlosser, I. J., 1982. Fish community structure and function along two habitat gradients in a headwater stream. Ecological Monographs 52: 395–414.

Schmutz, S., M. Kaufmann, A. Vogel, M. Jungwirth & S. Muhar, 2000. A multi-level concept for fish-based, river-type-specific assessment of ecological integrity. Hydrobiologia 422(423): 279–289.

Sheldon, A. L., 1968. Species diversity and longitudinal succession in stream fishes. Ecology 49: 193–198.

Thienemann, A., 1925. Die Binnengewässer Mitteleuropas. Die Binnengewässer 1: 54–83.

Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell & C. E. Cushing, 1980. Canadian Journal of Fisheries and Aquatic Sciences 37: 130–137.

Vapnik, V., 1995. The Nature of Statistical Learning Theory. Springer, New York.

Vapnik, V., 1998. Statistical Learning Theory. Wiley, New York.

Vehanen, T., T. Sutela & H. Korhonen, 2010. Environmental assessment of boreal rivers using fish data: a contribution to the Water Framework Directive. Fisheries Management and Ecology 17: 165–175.

Vile, J. S. & B. F. Henning, 2018. Development of indices of biotic integrity for high-gradient wadeable rivers and headwater streams in New Jersey. Ecological Indicators 90: 469–484.

Welcomme, R. L., 1983. River basins. FAO Fisheries Technical Paper 202 (60 pp.).

Welcomme, R. L., K. O. Winemiller & I. G. Cowx, 2006. Fish environmental guilds as a tool for assessment of ecological condition of rivers. River Research and Applications 22: 377–396.

Whittier, T. R., R. M. Hughes, J. L. Stoddard, G. A. Lomnicky, D. V. Peck & A. T. Herlihy, 2007. A structured approach for developing indices of biotic integrity: three examples from streams and rivers in the Western USA. Transactions of the American Fisheries Society 136: 718–735.

Wolff, L. L., N. Carniatto & N. S. Hahn, 2013. Longitudinal use of feeding resources and distribution of fish trophic guilds in a coastal Atlantic stream, southern Brazil. Neotropical Ichthyology 11: 375–386.

Zbinden, Z. D. & W. J. Matthews, 2017. Beta diversity of stream fish assemblages: partitioning variation between spatial and environmental factors. Freshwater Biology 62: 1460–1471.

Zorn, T. G., P. W. Seebach & M. J. Wiley, 2002. Distribution of stream fishes and their relationship to stream size and hydrology in Michigan´s lower peninsula. Transaction of the American Fisheries Society 131: 70–85.