Ecological Modelling 411 (2019) 108818

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

Scale effects on the performance of niche-based models of freshwater fish distributions: Local vs. upstream area influences

Danijela Markovic[a], Ariane Walz[b], Oskar Kärcher[a,*]

[a] Faculty of Business Management and Social Sciences, Osnabrück University of Applied Sciences, Caprivistr. 30A, 49076 Osnabrück, Germany
[b] Institute of Earth and Environmental Science, University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany

ARTICLE INFO

Keywords:
Catchment order
Conservation planning
Danube
Freshwater fish
Species distribution modelling
Upstream area

ABSTRACT

Niche-based species distribution models (SDMs) play a central role in studying species response to environmental change. Effective management and conservation plans for freshwater ecosystems require SDMs that accommodate hierarchical catchment ordering and provide clarity on the performance of such models across multiple scales. The scale-dependence components considered here are: (a) environment spatial structure, represented by hierarchical catchment ordering following the Strahler system; (b) analysis grain, that included 1st to 5th order catchments; and (c) response grain, the grain at which species respond most, represented by local and upstream catchment area effects. We used fish occurrence data from the Danube River Basin and various factors representing climate, land cover and anthropogenic pressures. Our results indicate that the choice of response grain – local vs. upstream area effects – and the choice of analysis grain, only marginally influence the performance of SDMs. Upstream effects tend to better predict fish distributions than corresponding local effects for anthropogenic and land cover factors, in particular for species sensitive to pollution. Key predictors and their relative importance are scale and species dependent. Consequently, choosing proper species dependent spatial scales and factors is imperative for effective river rehabilitation measures.

1. Introduction

The majority of studies describing distributions of freshwater species rely on niche-based species distribution models (SDMs) (Buisson et al., 2008; Lassalle et al., 2010; Markovic et al., 2012). While the methodological aspects of SDMs towards reducing model uncertainties have been extensively studied (e.g. Thuiller, 2004; Marmion et al., 2009; Grenouillet et al., 2011), the influence of scale components on predictive power of SDMs was less frequently studied. Following Mertes and Jetz (2018), the scale-dependence components of species-environment relationships are: (1) spatial structure of the environment; (2) analysis grain, the grain at which analyses are conducted; and (3) response grain, the grain at which species respond most strongly to their environment. Domisch et al. (2013) and Kärcher et al. (2019) have shown that the second scale component – analysis grain – only marginally influences the performance of SDMs; however, to address the effects of the above-defined three scale components on the performance of SDMs tailored to freshwater species, an integrated modelling framework is needed.

Stream ecologists have long recognized that biodiversity patterns of freshwater species are strongly influenced by the spatial structure of the environment – the dendritic structure of stream and river networks (cf. Fausch et al., 2002). Moreover, Swan and Brown (2017) have shown that restoration activities aiming at increasing biodiversity are more efficient in small isolated streams than in larger, well-connected downstream reaches. Accordingly, the local diversity of streams and rivers is influenced by the watershed structure and land use within the surrounding river and stream valleys at multiple scales (Allan et al., 1997; Allan, 2004; Black et al., 2004; Stephenson and Morin, 2009), and also known as the riverine macrosystem (McCluney et al., 2014). For example, macroinvertebrates and fish are shown to be strongly influenced by riparian land use and instream habitat conditions at the reach scale (Moerke and Lamberti, 2006; Feld and Hering, 2007; Verdonschot, 2009; Sui et al., 2014), and up to a few kilometres upstream of sampling sites (Kail et al., 2009; Kail and Wolter, 2013). In the era of urgent management response to high environmental pressures to freshwater ecosystems, catchment scale hold great promise where fine grain (e.g. reach scale) survey data are unavailable. Given the high vulnerability of freshwater ecosystems to environmental change (Woodward et al., 2010; Markovic et al., 2017) and the fact that

* Corresponding author at: Faculty of Business Management and Social Sciences, Osnabrück University of Applied Sciences, Caprivistr. 30A, 49076 Osnabrück, Germany.
E-mail address: o.kaercher@hs-osnabrueck.de (O. Kärcher).

https://doi.org/10.1016/j.ecolmodel.2019.108818
Received 11 July 2019; Received in revised form 2 September 2019; Accepted 3 September 2019

SDMs play a central role in conservation and restoration planning at multiple scales (e.g. Franklin, 2009), it is of fundamental importance to investigate the scale effects on the performance of SDMs of freshwater species' distributions.

Hierarchical catchment ordering provides a valuable framework to address the effects of the two scale components – analysis grain and spatial structure of the environment – on models of freshwater species' distributions (cf. Allan et al., 1997). In our recent study (Kärcher et al., 2019), we have shown that the choice of the analysis grain, reflected by the hierarchical catchment order, only marginally influences the mean performance of SDMs. The present study incorporates an additional scale dimension into this discourse – the response grain, implemented through the combination of local factors and those of the entire upstream catchment. Namely, while the relative importance of factors affecting species distributions is still debated (Tudesque et al., 2014), there is clear evidence for the influence of the upstream area on the occurrence probability of species in a particular stream reach or a site (Angermeier and Winston, 1988; Allan, 2004). Despite this evidence and the increased awareness that cause–effect relationships should be simultaneously studied at both fine and broad scales (see Soranno et al., 2014), the combination of local factors and those of the entire upstream catchment in shaping distribution patterns of freshwater species had not been sufficiently studied. For instance, the upstream effects were commonly not considered for the whole upstream catchment, but only up to a certain distance from a sampling site (e.g. up to 7.5 km in Kail and Wolter, 2013). The few studies addressing the relationship between a species' occurrence probability and upstream effects suggest strong impacts of intensive agricultural land use (Allan et al., 1997; Allan, 2004) and a higher relative importance of cumulative upstream area conditions than environmental conditions at the reach scale (Esselman and Allan, 2010).

Here we present a comprehensive modelling framework that addresses the scale effects on the performance of SDMs describing freshwater fish species distribution patterns. The study area includes the Danube River Basin. The spatial structure of the environment and analysis grain are based on the most widely used hierarchical catchment ordering, the Strahler system (Strahler, 1964). Strahler order reflects the hierarchical level of each reach in the whole river network, with 1st-order assigned to all reaches with no tributaries, 2nd-order to the confluences of two first-order reaches, and so on. In this study, we focus on the "reach scale" (here, the catchment area of the river reach where an occurrence of the freshwater fish was observed; hereafter called "1st order catchment") and the corresponding 2nd to 5th order catchments. At each of the hierarchical catchment orders, we evaluated the species' response grain using environmental factors summarizing effects across catchments of a particular order, across the corresponding whole upstream area, as well as those summarizing cross scale effects. The aims of the present study are: (1) to identify for each individual environmental factor whether the entire upstream catchment effect, the local reach effect or their interaction are better related to fish distribution patterns and (2) to test whether there is a decrease in the performance of SDMs with coarsening analysis grain (i.e., catchment order). To test the effects on model performance (objective 2 above), we compared the performance of SDMs for each catchment order based (a) solely on local catchment effects, (b) solely on upstream catchment effects and (c) on the combination of local- and upstream catchment effects.

2. Methods

As this study is an extension of our previously published work (Kärcher et al., 2019) and thus partially follows the same methodology, for convenience, the descriptions of the methods are kept the same wherever appropriate.

2.1. Study area

The Danube is Europe's second largest river basin (801,463 km^2) and is shared among more than 80 million people from 19 countries. It extends from Central Europe through the Balkans and drains to the Black Sea (Fig. A1). The Danube River delta is one of the world's largest wetlands, rich in rare fauna and flora and inscribed on UNESCO's World Heritage List in 1991. Due to its large spatial extent and diverse relief, the Danube River Basin also shows great differences in climate. The summed annual precipitation ranges from more than 2300 mm in the high mountains to less than 400 mm in the delta region, while the mean annual discharge reaches 6460 m^{3}s^{-1} at the Danube delta in Romania (ICPDR, 2009). The ecosystems of the Danube River Basin host over 2000 plant species, over 40 mammals and approximately 100 fish species, but are also subject to increasing pressure and serious pollution from agriculture, industry and municipalities (ICPDR, 2009).

2.2. Environmental data

Choice of environmental factors was based on recent studies of fish species distributions (Buisson et al., 2008; Lassalle et al., 2010; Markovic et al., 2012; Isaak et al., 2017), with the predictor selection performed following Markovic et al. (2012). In total, six environmental predictors emerged representing climate, land cover and anthropogenic pressure of the study area (Table 1). Climatic and topographic data were extracted from the Worldclim 30 arc-second (approx. 1 km × 1 km) gridded information (Hijmans et al., 2007). The climatic data set included annual mean temperature (AnnTMean) and annual mean precipitation (AnnPMean). Land use of the study area was described by applying three land cover classes: percentage of area covered by row crops (RowCrops), by forest (Forest) and by artificial surface (BuiltUp). Land cover information was extracted from the CORINE land cover database (EEA, 2011). The number of inhabitants per area (Population) was used as a measure of anthropogenic pressure (Global Rural-Urban Mapping Project version1, available at http://sedac.ciesin.columbia.edu/gpw). Initially, topography of the study area was considered using altitude and slope; however, due to high correlations (above $|0.7|$) with the annual mean temperature both, altitude and slope were omitted from further analyses (Fig. A2). As indicated in Table 1, the study area encompasses a wide range of environmental conditions.

2.3. Hierarchical catchment orders and predictor data sets

The analyses and modelling were conducted for five different catchment orders (WaterShed Order WSO1 to WSO5) based on the Strahler order of the river reaches from the CCM2 pan-European catchments database (CCM version 2.1, de Jager and Vogt, 2010). The 1st order catchments (WSO1) are defined according to the drainage areas of the individual river reaches while higher order catchments (WSO2-WSO5) result from groupings of the lower-order catchments in a hierarchical way. More specifically, the WSO1 represents here drainage areas of the river reaches with observed species' occurrences. Catchment order increase is reflected in an increasing catchment size, with an average ranging from 12 km^2 for the WSO1 to 2148 km^2 for the WSO5.

For each WSO, environmental factors were calculated by averaging gridded data across the corresponding local catchment areas ("lca") reflecting the river reach scale and across the cumulative upstream catchment areas ("uca") (Fig. A1), resulting in two distinct predictor data sets. The upstream catchment-related factors were derived using the Pfafstetter coding system of the CCM2 database in R environment (R Development Core Team, 2019). Regarding the "lca" and the "uca" predictor data sets, the environmental factors represent different statistics resulting from the processing of the gridded data for the local- and the upstream catchments: mean values for climatic factors, percentage values for the land use classes and the sum of the number of

Table 1
Mean and range of the environmental predictors describing local catchment influences ("lca") and the cumulative upstream influences ("uca") across the studied scales (WSO1-WSO5).

Predictor	Description	Approach	WSO1	WSO2	WSO3	WSO4	WSO5
AnnTMean [°C]	Annual mean temperature	lca	7.6 (-0.7-11.3)	7.3 (-2.1-11.2)	6.9 (-1.5-11.1)	6.5 (-0.9-11.1)	7.1 (0.1-11.0)
		uca	5.8 (-2.1-11.0)	5.6 (-2.1-11.0)	5.5 (-1.5-11.0)	5.4 (-0.9-11.0)	6.0 (0.1-11.0)
AnnPMean [mm]	Annual mean precipitation	lca	870 (476-1486)	888 (483-1522)	924 (494-1537)	957 (511-1538)	908 (523-1479)
		uca	958 (519-1559)	968 (519-1548)	992 (519-1537)	1011 (531-1538)	968 (542-1479)
RowCrops [%]	Row crop percentage	lca	26.6 (0-100)	24.6 (0-94.8)	23.3 (0-86.0)	23.0 (0-86.0)	28.3 (0-86.0)
		uca	17.4 (0-97.4)	17.1 (0-87.0)	17.1 (0-79.2)	17.4 (0-73.6)	20.7 (0-66.4)
Forest [%]	Forest percentage	lca	43.0 (0-100)	46.4 (0-100)	46.3 (0-100)	44.3 (0-94.4)	40.7 (3.1-84.0)
		uca	53.1 (0-100)	51.4 (0-100)	48.4 (0-96.5)	46.2 (0-94.4)	45.0 (3.8-83.9)
BuiltUp [%]	Built up area percentage	lca	9.0 (0-100)	7.8 (0-100)	6.3 (0-78.8)	5.6 (0-78.8)	5.4 (0.1-18.1)
		uca	3.3 (0-46.2)	3.5 (0-46.2)	3.7 (0-26.3)	3.7 (0-16.9)	4.3 (0.1-13.4)
Population [No. inh.]	Inhabitants per area	lca	2042 (0-3.2·10^5)	6198 (0-3.6·10^5)	18,129 (22-5.9·10^5)	60,115 (214-18.8·10^5)	219,284 (528-26.8·10^5)
		uca	300,837 (0-275·10^5)	357,999 (0-275·10^5)	614,163 (63-276·10^5)	934,494 (973-276·10^5)	2,078,784 (5746-279·10^5)

inhabitants per area. Also, it is noteworthy that cumulative upstream environmental factors for small 5th order catchments might be very similar over all lower hierarchical WSO levels. In addition to the "lca" and the "uca" predictor data sets, we also defined their combined cross-scale interactions ("csi"). Specifically, the interactions between the environmental factors were measured using the product of the corresponding "lca" and the "uca" predictors. To ensure interpretability of our outcomes, we used only pairs of identical environmental factors (e.g., "lca" and "uca" related percentage of area under row crops) and not all possible environmental factor combinations (e.g., a possible, but poorly interpretable combination would be the product of "lca" annual mean precipitation and "uca" number of inhabitants).

2.4. Fish data

Species occurrence data for the Danube River Basin were provided by partners of the project Biofresh (www.freshwaterplatform.eu, Schinegger et al., 2016) for 1364 sites (see also Fig. A1). Fisheries data were sampled by either single-pass or double-pass electrofishing (mainly between 1985 and 2002). To ensure an accurate estimate of the species distributions, only species with at least 50 occurrences at the largest analysed scale, the WSO5, were included into the analysis (cf. Coudon and Gégout, 2007). In total, eight species are considered (Table 2). Species ecological characterizations follow Kottelat and Freyhof (2007): the bleak (*Alburnus alburnus*) is a small cyprinid, and prefers open waters of lakes and medium to large rivers. The stone loach (*Barbatula barbatula*) is usually found in medium-sized rivers with gravel to stone bottom. *Barbus barbus* is a fish of the cyprinid family preferably inhabiting faster flowing, summer-warm, medium to large-sized rivers. Gudgeons of the genus *Gobio*, here represented by *Gobio obtusirostris*, are riverine cyprinids, too, which tolerate – in contrast to barbels – lower flow velocities and finer spawning substrates. The bullhead (*Cottus gobio*) inhabits cold, clear and fast-flowing water of small streams to medium-sized rivers. The roach (*Rutilus rutilus*) is a small fish of the cyprinid family mainly found in nutrient-rich large to medium sized lowland rivers and backwaters. The trout (*Salmo trutta*) is a species of salmonid fish preferring cold, well-oxygenated streams in the mountainous areas. The chub (*Squalius cephalus*) is a fish of the cyprinid family found in slow-flowing lowland rivers, very small mountain streams, and in large streams of barbel zone.

Table 2
Freshwater fish species and their prevalence. The total number of catchments for the studied scales (WSO1-WSO5) is indicated in parentheses.

Name	Code	WSO1 (1363)	WSO2 (1059)	WSO3 (681)	WSO4 (350)	WSO5 (126)
Alburnus alburnus	Albual	0.20	0.22	0.29	0.38	0.62
Barbus barbus	Barbba	0.20	0.21	0.23	0.28	0.45
Barbatula barbatula	Barbbr	0.31	0.33	0.34	0.35	0.45
Cottus gobio	Cottgo	0.44	0.42	0.43	0.48	0.51
Gobio obtusirostris	Gobris	0.31	0.33	0.36	0.37	0.51
Rutilus rutilus	Rutiru	0.22	0.25	0.30	0.38	0.58
Salmo trutta	Saltta	0.71	0.70	0.72	0.74	0.69
Squalius cephalus	Squace	0.44	0.46	0.49	0.53	0.72

For each of the five WSOs, fish occurrence data were aggregated to presence/absence information (hereafter called "catchment-scale mapping"). The final number of catchments of each particular order and their spatial arrangement was thus directly constrained by fish data availability. With the WSO increase, the number of catchments under consideration decreases (from 1363 for the WSO1 to 126 for the WSO5), with a corresponding increase in species prevalence (Table 2). We note here that, because of the dendritic structure of river networks, catchment-scale mapping is more appropriate for freshwater species than the point-to-grid mapping, used for mapping terrestrial species' occurrences (see Fagan, 2002). In addition, given that catchments serve as units for freshwater management and conservation (commonly referred to as the Catchment Based Approach – CaBA, see DEFRA, 2013), catchment-scale mapping of freshwater species' occurrences ensures compatibility between the management and the analysis scales, as well as the optimisation of ecological restoration efforts (Lévêque et al., 2008; Markovic et al., 2017; Kuemmerlen et al., 2019).

2.5. Data analysis and modelling

A schematic view of the data analysis and modelling workflow is provided in Fig. 1. Univariate strength of the environmental predictor variables was quantified using the weight of evidence (WOE) and information value concepts as implemented in the R library "Information"

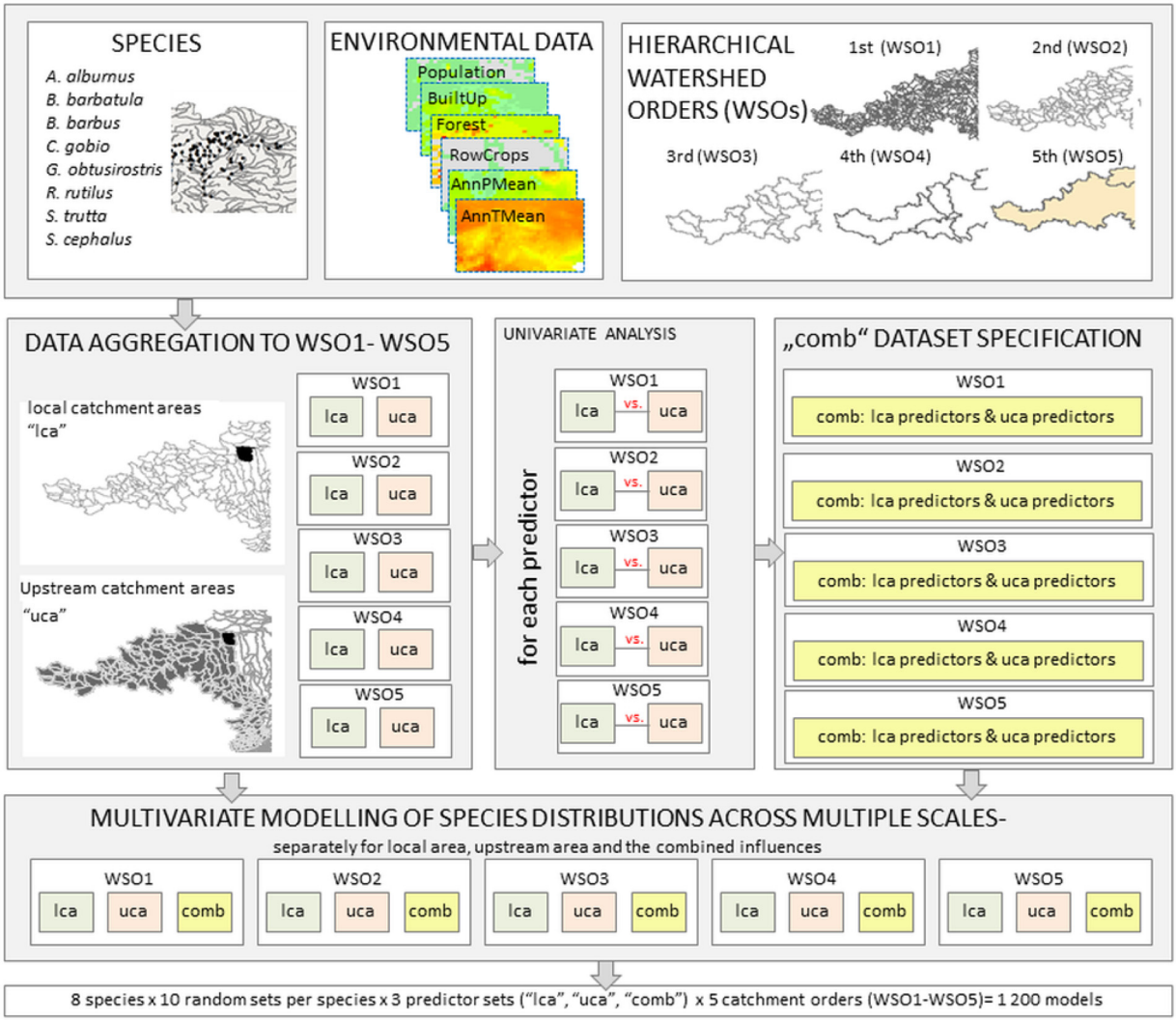

Fig. 1. Conceptual modelling framework. With an increase in the hierarchical watershed order (WSO), there is an increase in the average catchment size. The local effects were calculated by averaging environmental gridded data across the corresponding local catchment areas ("lca", i.e., by averaging environmental data across the polygons representing the catchments of the particular order). The upstream effects summarize the environmental variability across the whole upstream catchment area ("uca") of an individual catchment of a particular order. For example, for the black shaded polygon in the box "Data Aggregation" representing a 1st order catchment (i.e., catchment area of the individual river reaches), the upstream effects are calculated by averaging environmental parameters across the area consisting of the catchment itself and the corresponding "uca" (grey shaded polygons).

(Larsen, 2016). While WOE describes the relationship between a predictor variable X (here environmental predictors listed in Table 2) and a binary dependent variable Y (here species occurrence in a particular catchment), information value measures the strength of the $Y - X$ relationship. Specifically, if b_i, $i = 1, ...,k$, denote k discrete bins for the predictor X, then, the strength of the predictor in describing Y can be quantified as $\sum_i (P(X \in b_i|Y = 1) - P(X \in b_i|Y = 0)) \times WOE_i$ (cf. Larsen, 2016). As such, information value is suitable for comparing the strength between local and upstream area effects on freshwater fish distribution across multiple scales (i.e., for comparing the predictive power of the environmental factors calculated for the local catchment areas of each watershed order and for the corresponding cumulative upstream catchment areas).

Species distribution modelling was performed using Generalized Additive Models (GAM) (R library "gam"; Hastie, 2005). GAM is a non-parametric extension of generalized linear methods, and is widely used for modelling current and future distribution patterns of fish species. Previous investigations using various SDMs have shown that GAM-, GLM- (Generalized Linear Models) and regression tree based SDMs have similar validation performance. The first two kinds of models had also similar calibration performance, while regression tree based SDMs tend to overfit during the calibration phase (Markovic et al., 2012). The improved performance of consensus or ensemble methods in providing more accurate and robust projections of species distribution have been already demonstrated (Marmion et al., 2009; Buisson et al., 2010; Lauzeral et al., 2013); however, the main objective of this study was determining the effect and importance of variation in spatial scale rather than the performance of different SDMs. Therefore, we focus here only on GAM based SDMs, but acknowledge the importance of using multiple SDMs when the study goal is predicting future species distribution patterns (see Markovic et al., 2012). Moreover, GAMs are very flexible models, and have well performance at high collinearity (Dormann et al., 2013).

Species occurrence probabilities resulting from the application of GAM were transformed to presence/absence information using the thresholds, which maximize both sensitivity (the true positive rate) and the specificity (the true negative rate). Thus, the percentage of correctly predicted presences (sensitivity) and the percentage of correctly predicted absences (specificity) are jointly maximized. We used repeated random splitting (10 times) of the fish data into calibration (70%) and

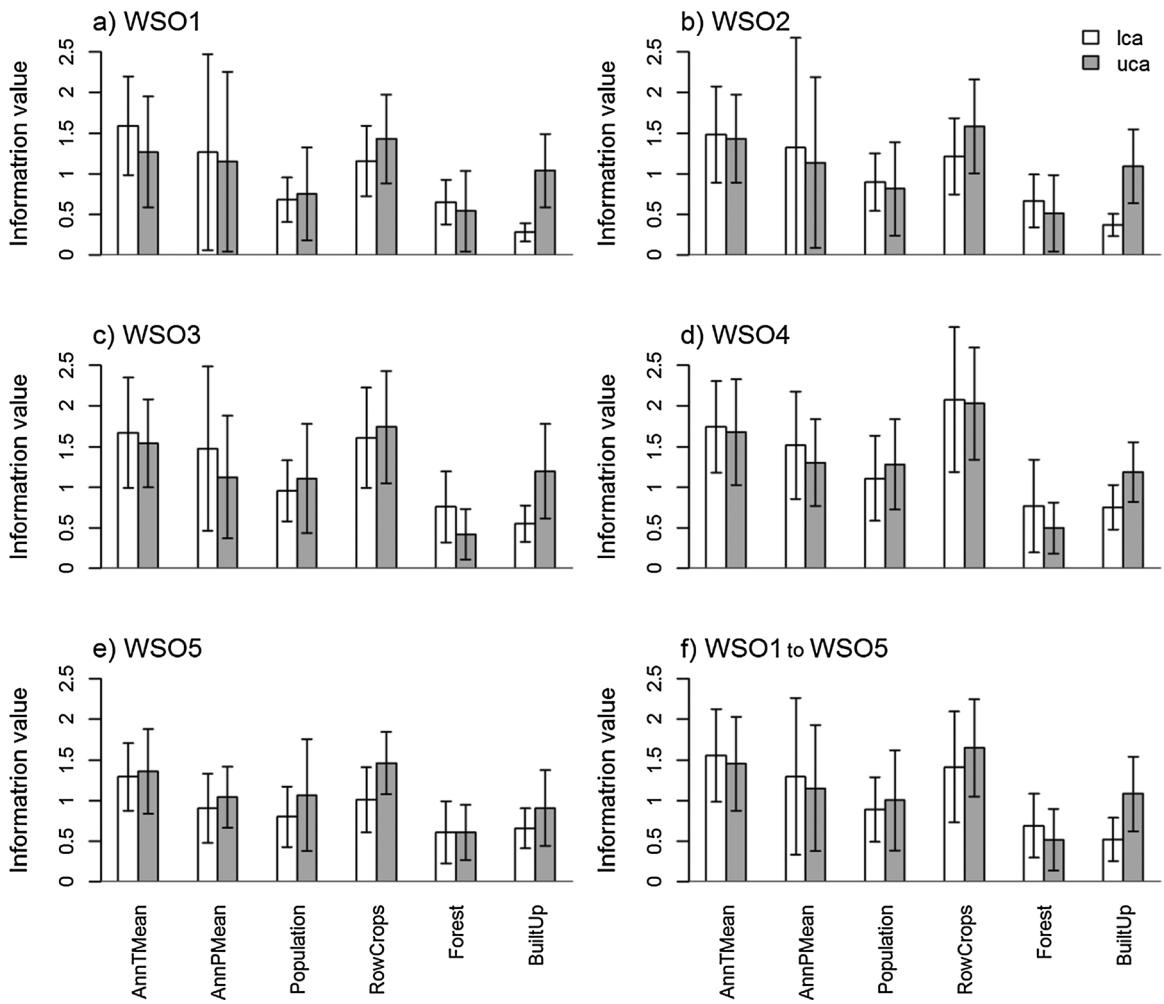

Fig. 2. Information value as a measure of the univariate strength of the environmental predictor variables derived across the corresponding local catchment areas ("lca") and across the whole upstream catchment area ("uca"). a–e) For each of the five spatial scales the information value was calculated as an average predictive strength of the environmental predictor across all studied species; f) The overall information value of an environmental predictor was calculated as an average predictive strength across all scales and all considered species. Error bars represent one standard deviation of the estimates.

validation (30%), i.e., each of the 10 models was calibrated using a different 70% data sample and validated using the remaining 30%. The repetitive modelling procedure allowed for quantifying the uncertainty of the estimates. Agreement between the observed and modelled species distribution patterns was quantified by sensitivity and specificity, while the performance of the calibrated models was estimated using the Area under the Receiver Operator Curve (AUC) and the True Skill Statistic (TSS) (Allouche et al., 2006). An AUC of 0.5 and a TSS of 0 indicate that a model has no discriminatory power, while an AUC or TSS of 1 indicate that presences and absences are perfectly discriminated.

Within the multivariate modelling of species distributions three predictor data sets were created: one describing only local reach-scale effects, a second describing only cumulative upstream area effects, and a third combined predictor data set ("comb") merging the local, the upstream and the cross scale interactions predictor sets. In the "comb" predictor data set, the environmental factor with the higher information value of each of the three source data sets was used in the modelling. Consequently, models of distribution patterns of each of the studied species were initiated using three distinct predictor sets resulting in 30 multivariate models per species (10 repetitions × 3 predictor data sets).

For each predictor data set used in modelling of species distribution patterns, the condition of pairwise correlations below |0.70| had to be satisfied at all scales (i.e. no multicollinearity; correlation matrix for the reach-scale is provided in Fig. A2). Thereby, within each of the 30 multivariate model runs per species, the search for a parsimonious model involved analyses of the model improvement based on the Akaike Information Criterion (AIC) through simultaneous forward and backward predictor selection.

Statistical significance between the AUCs for different predictor data sets ("lca", "uca" and "comb") was measured using the two sample *t*-test (e.g. Wilks, 1995). By including statistical testing, data sets leading to significantly better performances can be identified and recommended for explaining the fish distributions across the studied scales. To quantify the relative predictor importance, the variance partitioning method, implemented within the R library "relaimpo" (Grömping, 2006) was used. The advantage of the variance partitioning method by Lindeman et al. (1980) is that it considers sequential sums of squares over all predictor permutations, and thus considers the intercorrelation effects among the individual predictors. We note that high predictor relative importance does not necessarily imply causation.

3. Results

3.1. Local and cumulative upstream effects

Overall, the annual mean temperature and the percentage of area under row crops have the highest information value irrespective of the considered spatial scale ("lca" vs. "uca" and WSO1 to WSO5, Fig. 2). A

Table 3
Mean and standard deviation of the number of statistically significant model parameters per species across the studied scales (WSO1-WSO5) for the "lca", "uca" and the "comb" predictor sets.

Species	Approach	WSO1	WSO2	WSO3	WSO4	WSO5
A. alburnus	uca	3.8 (0.6)	3.1 (0.7)	2.7 (0.5)	2.7 (0.5)	2.3 (0.5)
	comb	4.3 (0.5)	3.3 (0.8)	2.9 (0.9)	2.6 (0.7)	2.1 (0.3)
	lca	3.5 (0.5)	3.1 (0.7)	2.8 (0.9)	2.7 (1.1)	1.5 (0.5)
B. barbatula	uca	4 (0.7)	4.7 (0.7)	4 (0)	2.8 (0.4)	2 (0.7)
	comb	4.5 (0.5)	4.9 (0.6)	3.4 (0.5)	3 (0)	1.6 (0.7)
	lca	4.1 (0.3)	3.8 (0.6)	3.6 (0.8)	2.3 (0.5)	1.5 (0.5)
B. barbus	uca	5 (0)	5 (0)	2.5 (0.5)	2.5 (0.7)	1.5 (0.7)
	comb	5.3 (0.7)	4.6 (0.5)	3.1 (0.6)	2.6 (0.8)	2.6 (0.7)
	lca	4.5 (0.5)	4.7 (0.7)	2.6 (0.5)	3 (0.9)	2.5 (0.8)
C. gobio	uca	4.9 (0.3)	5 (0)	4 (0)	3.8 (0.4)	2.1 (0.3)
	comb	4.7 (0.8)	4.5 (0.7)	3.6 (0.7)	3.3 (0.5)	1.3 (0.7)
	lca	3.7 (0.5)	3.9 (0.6)	3 (0)	2 (0.5)	1.2 (0.4)
G. obtusirostris	uca	4 (0.5)	5 (0)	3.4 (0.5)	2.8 (0.4)	1.6 (0.7)
	comb	4.1 (0.6)	3.9 (0.6)	4.4 (0.5)	3.6 (0.5)	2.8 (0.4)
	lca	3.8 (0.9)	3.4 (0.5)	4.1 (0.9)	3 (0.5)	2.5 (0.5)
R. rutilus	uca	4.8 (0.4)	4.6 (0.5)	3.9 (0.3)	3.8 (0.4)	2.7 (0.8)
	comb	5.1 (0.3)	4.8 (0.4)	3.9 (0.3)	2.5 (0.5)	2.1 (0.9)
	lca	4.5 (0.5)	3.9 (0.7)	3.9 (0.7)	1.8 (0.8)	2.1 (0.7)
S. trutta	uca	5 (0)	5 (0)	3.8 (0.4)	3.3 (0.8)	2.4 (0.5)
	comb	4.2 (0.4)	4.1 (0.7)	2.5 (0.5)	1.7 (0.5)	2 (0.5)
	lca	3.9 (0.3)	3.5 (0.7)	2.8 (0.4)	1.4 (0.7)	1.8 (0.4)
S. cephalus	uca	4 (0)	4.5 (0.5)	3.1 (0.3)	2.4 (0.5)	2.1 (0.3)
	comb	5 (0)	3.4 (0.5)	2.5 (0.5)	2 (0)	1.8 (0.9)
	lca	3.9 (0.3)	3 (0)	2.3 (0.5)	2.1 (0.3)	1.7 (0.8)

Table 4
Mean and standard deviation for validation AUCs of the multivariate models of species distribution patterns across the studied scales (WSO1-WSO5) catchments for the "lca", "uca" and the "comb" predictor data sets.

Species	Approach	WSO1	WSO2	WSO3	WSO4	WSO5
A. alburnus	lca	0.92* (0.01)	0.92* (0.01)	0.93* (0.02)	0.91 (0.03)	0.89 (0.05)
	uca	0.89 (0.01)	0.89 (0.02)	0.89 (0.02)	0.91 (0.03)	0.90 (0.07)
	comb	0.88 (0.01)	0.89 (0.02)	0.91 (0.01)	0.91 (0.03)	0.88 (0.07)
B. barbatula	lca	0.89* (0.01)	0.88* (0.01)	0.87* (0.02)	0.87 (0.03)	0.73 (0.08)
	uca	0.86 (0.01)	0.84 (0.02)	0.83 (0.03)	0.86 (0.03)	0.79 (0.08)
	comb	0.84 (0.02)	0.84 (0.01)	0.83 (0.02)	0.85 (0.04)	0.78 (0.04)
B. barbus	lca	0.77 (0.02)	0.79 (0.01)	0.77 (0.02)	0.74 (0.05)	0.64 (0.08)
	uca	0.82* (0.01)	0.85* (0.02)	0.87* (0.03)	0.8* (0.03)	0.77* (0.06)
	comb	0.82 (0.02)	0.84 (0.03)	0.87 (0.02)	0.8 (0.04)	0.74 (0.08)
C. gobio	lca	0.78 (0.02)	0.8 (0.02)	0.77 (0.03)	0.81 (0.03)	0.79 (0.06)
	uca	0.81 (0.03)	0.84* (0.02)	0.85* (0.03)	0.81 (0.03)	0.81 (0.08)
	comb	0.82 (0.02)	0.84 (0.02)	0.84 (0.02)	0.78 (0.04)	0.82 (0.07)
G. obtusirostris	lca	0.86 (0.02)	0.87 (0.01)	0.83 (0.03)	0.81 (0.03)	0.70 (0.05)
	uca	0.87 (0.01)	0.88 (0.01)	0.88 (0.01)	0.84 (0.04)	0.78 (0.04)
	comb	0.88 (0.01)	0.89 (0.02)	0.88 (0.01)	0.86 (0.03)	0.86* (0.06)
R. rutilus	lca	0.89 (0.02)	0.88 (0.01)	0.90 (0.02)	0.89 (0.03)	0.82 (0.06)
	uca	0.89 (0.02)	0.89 (0.02)	0.89 (0.02)	0.88 (0.04)	0.80 (0.06)
	comb	0.87 (0.02)	0.89 (0.02)	0.89 (0.02)	0.88 (0.03)	0.79 (0.07)
S. trutta	lca	0.91 (0.02)	0.92 (0.01)	0.94 (0.01)	0.97 (0.01)	0.96 (0.03)
	uca	0.93 (0.02)	0.94* (0.01)	0.97* (0.02)	0.97 (0.02)	0.97 (0.03)
	comb	0.93 (0.01)	0.94 (0.01)	0.97 (0.01)	0.97 (0.03)	0.95 (0.05)
S. cephalus	lca	0.86 (0.01)	0.85 (0.02)	0.84 (0.02)	0.86 (0.04)	0.78 (0.07)
	uca	0.88 (0.01)	0.87 (0.01)	0.85 (0.01)	0.86 (0.03)	0.81 (0.07)
	comb	0.87 (0.02)	0.87 (0.02)	0.86 (0.02)	0.86 (0.03)	0.81 (0.09)

* Marks the preferred approach, i.e. the approach that led to the rejection of the hypothesis implying no significant differences between AUCs for all predictor set combinations ("lca"-"uca"; "uca"-"comb" and "lca"- "comb") at the $\alpha = 1\%$ significance level.

comparison between the information values of the factors derived for the local catchment areas and for the corresponding upstream catchment areas indicated higher predictive strength of the local effects for the climatic predictors (AnnTMean and AnnPMean) up to the fourth order catchments. For anthropogenic and land cover factors the information value results suggest that cumulative upstream effects of population (Population), percentage of area covered by row crops (RowCrops) and artificial surface (BuiltUp) are likely to be better predictors of fish distribution patterns than the corresponding local catchment effects (Fig. 2f). The information values of the cross scale interactions ("csi"), here measured as the product of the "lca" and "uca" related environmental factors, were significantly lower than the information value of the "lca" and "uca" factors themselves. Consequently, within the "comb" predictor data set, the factors Population, RowCrops and BuiltUp were selected from the "uca" predictor data set, while the remaining predictors were selected from the "lca" predictor data set (Fig. 2f). As such, none of the "csi" predictors qualified to be included in the modelling process.

3.2. Catchment order specific predictor set and model performance

Due to multicollinearity condition, the number of statistically

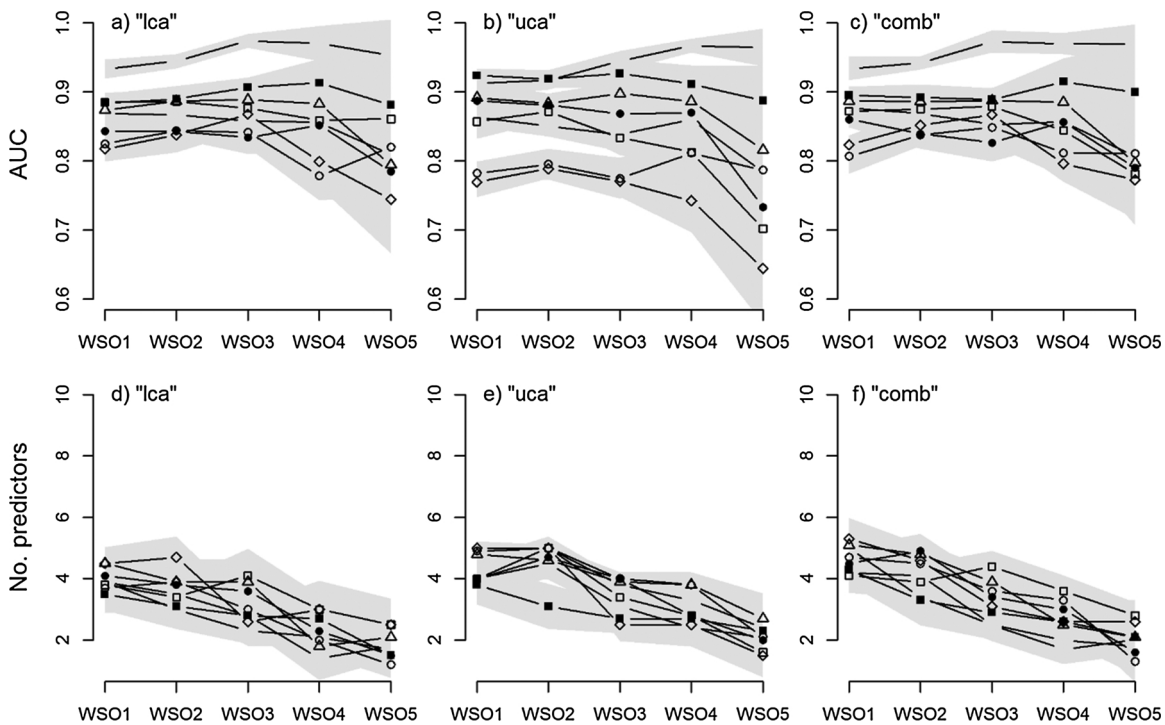

Fig. 3. Mean validation AUCs (a–c) and the mean number of statistically significant predictors (d–f) of species distribution patterns across the studied spatial scales (WSO1 to WSO5) for the "lca", "uca" and the "comb" predictor data sets. The averaging was done across corresponding model output statistics from 10 model repetitions per species. Each line corresponds to a single species, with the shaded area denoting the standard error of the estimates.

applicable model predictors varied with catchment order. Of the initially considered eight factors (Figs. A2-A3), the predictor set used at WSO1 included six environmental factors (Table 1) for the "lca" and "comb" data sets and five factors for the "uca" data set. Although GAMs have well performance at high collinearity (Dormann et al., 2013), since the majority of SDMs are vulnerable to collinearity, at higher catchment orders a subset of that at WSO1 was used satisfying the condition of pairwise correlations < |0.70|. Furthermore, keeping highly correlated predictors in the model would limit both, interpretability of the results and drawing inferences on the performance effects of the individual factors.

Within the predictor selection process, the one of two highly correlated predictors was used having higher information value (Fig. 2). For instance, within the "lca" predictor set at WSO5, annual mean temperature and annual mean precipitation could not be both used in the modelling due to correlation > |0.7|, wherefore only annual mean temperature was left in the predictor set owing to its higher information value (Fig. 2). As a result, at WSO5 of the six "lca" environmental factors used at WSO1, only four were applicable (annual mean temperature, population and percentage of area covered by forest and artificial surface).

The number of statistically significant model predictors is shown to vary across species and catchment orders (Table 3). Regardless of the species studied, the mean number of required parameters decreased with catchment order increase. At the WSO1, the mean number of statistically significant predictors from the 10 independent model runs per species is between 4 and 5, whereas for describing fish distribution patterns across WSO5, only 1–3 predictors were required (Table 3). Overall, species distribution models calibrated with the locally estimated predictors required the least number of predictors compared to those calibrated using cumulative upstream area related effects ("uca" and "comb"). The average of the validation AUCs across all repetition models of species distribution patterns was above 0.85 for all combinations of catchment order and predictor sets, indicating highly accurate models (Table 4, Fig. 3). The results above were confirmed by the mean sensitivity and mean specificity, which both were above 80% for the majority of the models resulting in mean TSS of 0.58. When looking at the model performance per individual species, differences in the validation AUCs between the models calibrated using the three different predictor data sets ("lca", "uca" and "comb") were up to 0.09, but statistically significant only for *B. barbus* (WSO2 to WSO5), *C. gobio* and *S. trutta* (WSO2 to WSO3) in favour of the "uca" predictor set and for *A. alburnus* and *B. barbatula* (WSO1 to WSO3) in favour of the "lca" set (Table 3). For other species/WSOs, consideration of the upstream area effects (i.e. "uca" and "comb" predictor set) did not contribute to a statistically significant improvement of the model performance when compared to the models using "lca" predictor set. Overall, when coarsening the grain size (i.e. increasing catchment order), both a slight model degradation and a slight model improvement was observed (Fig. 3) accompanied by an increase in the standard deviation of the validation AUCs.

3.3. Predictor importance per species and catchment order

Predictor importance patterns varied across species and catchment orders (Figs. 4, 5, Fig. A4-A5, Table A1). Overall, irrespective of the type of consideration of the factor influence, that is, irrespective of whether local, cumulative upstream or the combined factor influence was considered, climatic factors are shown to be the main drivers of the distribution patterns of the fish species at all catchment orders (Fig. 4). Also, the relative importance of annual mean temperature generally increased with catchment order (Fig. 4). The relative importance of the annual mean temperature was highest for *S. trutta* (97.2%, "lca", WSO4, Table A1). For *B. barbatula*, *C. gobio*, *G. obtusirostris* and the widely distributed *S. trutta* the mean relative importance of the locally derived annual mean precipitation was higher than that of the annual mean temperature for the first three catchment orders (Fig. A4, Table A1). With regard to the land cover and anthropogenic factors, up to WSO2, the relative importance of their local effects was below 20%, except for the row crops (Table A1). We remark that the factor row crops could

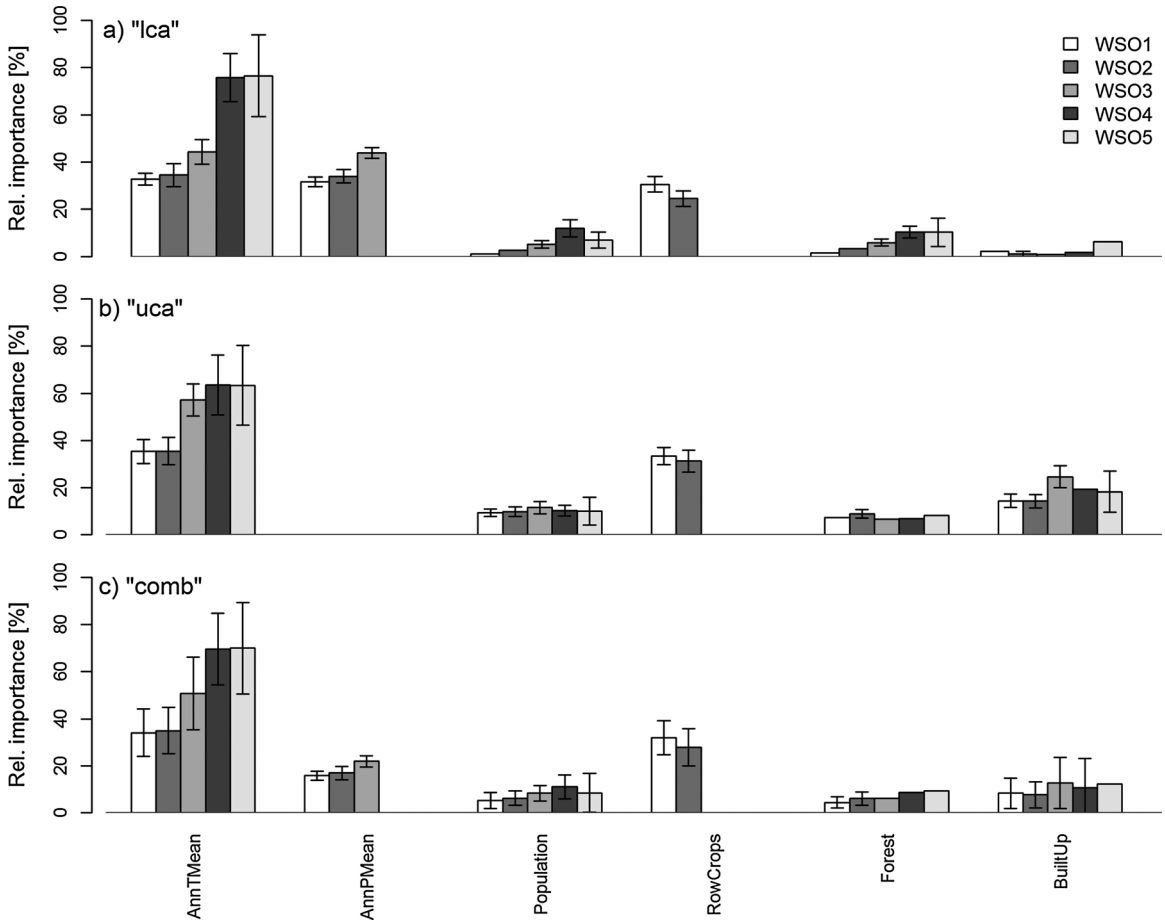

Fig. 4. Mean relative predictor importance resulting from the multivariate SDMs of all studied fish species for the studied scales (WSO1-WSO5) based (a) solely on local catchment influences, (b) solely on upstream catchment influences, and (c) on the combination of local- and upstream catchment influences. Error bars represent one standard deviation of the estimates.

only be used up to WSO2 to avoid multicollinearity. With the catchment order increase, both, a decrease and an increase in relative importance of the land cover and anthropogenic factors were observed (Fig. A5).

The relative importance of the upstream and the combined effects were considered only for those species and catchment orders where their use resulted in the statistically significant improvement of the SDMs, compared to the SDMs based on local influences. In particular, we remind that for *B. barbus* the upstream area consideration of the environmental predictors resulted in statistically significant improvement of the SDMs for all catchment orders (for *C. gobio* and *S. trutta* for WSO2-WSO3), while for *G. obtusirostris*, only for the 5th order catchments there was a statistically significant improvement in the models performance when using the combined local and upstream area influences. For *B. barbus*, the upstream related anthropogenic and land use related influences had higher relative importance than the climatic effects up to WSO4 (Fig. A4b, A5, Table A1). *C. gobio* and *S. trutta* distributions across the 2nd order catchments were strongly influenced by the upstream related area under row crops is (the mean relative importance was 43.3% for *C. gobio* and 30.3% for *S. trutta*). *G. obtusirostris* distributions across the 5th order catchments were mainly influenced by the locally related climatic effects (the mean relative importance of the annual mean temperature was 57.7%).

4. Discussion

The species distribution models of all included fish species were highly accurate across the studied spatial scales, independently of whether or not the cumulative upstream effects were considered. The negligible performance differences across the scales and model setups, affirm the selection of the environmental factors used for explaining fish distribution patterns and the chosen applied statistical methodology GAM.

The upstream catchment related percentage of area covered by row crops and artificial surface, as well as the number of inhabitants per area, tend to be better predictors of fish distribution patterns than the corresponding local catchment effects. These findings correspond to results from previous studies, which highlighted the essential influence of the surrounding landscape on the in-stream ecosystem structure and function (Fausch et al., 2002; Linke et al., 2008). With high nitrate flux positively influencing aquatic plant cover as well as water quality and thus negatively fish species richness, the surrounding area covered by row crops is essentially connected to fish distribution patterns (Strayer et al., 2003). The Danube River Basin has already experienced significant changes in water quality including chemical alterations due to nitrate and other nutrient pollutants from agriculture and various land use factors (Chapman et al., 2016), which are known to favour algal blooms. Further decreases of oxygen concentrations resulting from water quality degradation are expected, which can drastically influence fish species and their ability to cope with warming (Chapman et al., 2016; Verberk et al., 2016). Especially gravel-spawning riverine fish species are affected by intensely used cropland. Specifically, intensely used croplands provide a continuous source of fine sediments and siltation of coarse substrates, which is followed by the decline of gravel-spawning riverine fish species (Soulsby et al., 2001; Lapointe et al., 2004; Greig et al., 2005; Jensen et al., 2009). In contrast to the above discussed land use related factors, the climatic factors are shown to

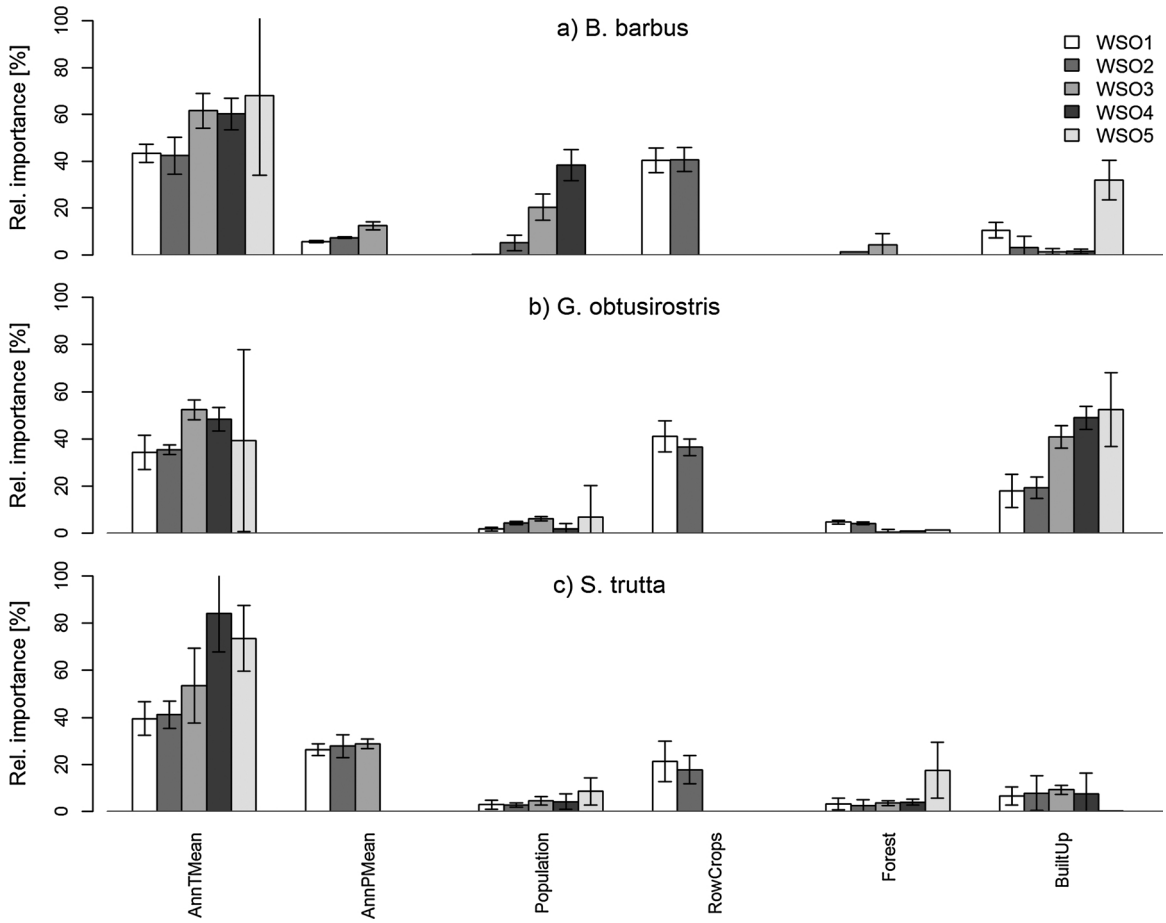

Fig. 5. Relative predictor importance in describing distribution patterns of *B. barbus* (a), *G. obtusirostris* (b) and *S. trutta* (c) across the studied scales (WSO1-WSO5) using predictors describing local catchment influences. Error bars represent one standard deviation of the estimates.

display higher information value when calculated for the local catchment areas, rather than for the entire upstream catchments. As such, our finding supports the selection of local climatic conditions in fish species distribution models across all studied scales (see also Kärcher et al., 2019). The observation of climatic factors dominating non-climatic factors for the local approach could also be made for all studied species. The interactions between the environmental factors manifested low information value; however, this might be only a reflection of the studied species-drivers relationships that do not constitute any proof of the negligible impact of the cross scale interactions. Overall, our results suggest the consideration of scale-dependence components when studying species-environment relationships and thus *factor and species dependent selection* between local and upstream area effects, paired with appropriate monitoring of effects and management of mitigation activities tailored to ensure species' long-term persistence.

The decrease in the number of environmental factors used by the multivariate models with increasing catchment order is generally a consequence of the decrease in the environmental information granularity, manifested through a reduction in the environmental variability (Austin, 2007). For instance, the mean area size at WSO5 (2148 km^2) is almost 180 times larger than the mean area size at WSO1 (12 km^2); thus, the multitude of environmental conditions that exist when considering drivers at the WSO1 scale is significantly "flattened" at the WSO5 scale. We underline that the number and combination of predictors was species- and catchment-order specific, implying that the use of arbitrary parameter sets for the calibration of SDMs is inappropriate and should be replaced by a well-founded parameter selection methodology. Specifically, similarly to our past study (Kärcher et al., 2019), our results indicate that for *B. barbus* the relative importance of anthropogenic pressure – manifested by either population density or percentage of area covered by artificial surface and row crops – is higher than the importance of the climatic factors. Damming, river regulation and fine sediment input are known to influence *B. barbus* (Kottelat and Freyhof, 2007). Thus, the results obtained from the SDMs suit well to the ecological classification of this species. Compared to *S. trutta*, which has similar environmental requirements, *B. Barbus* seems also more susceptible to the effects of land-use on water quality and spawning habitat. With barbel preferably colonizing lower elevated river reaches, this identified difference may simply result from higher cumulative anthropogenic pressure at such lower elevated reaches. The influence of anthropogenic pressure is additionally outlined by the ability of cumulative upstream area effects, in particular the influence of row crops, to describe the distribution of trout and gudgeons of the genus *Gobio*, all preferring cold and clean water, better than the locally derived effects.

Besides those limitations that we discussed in Kärcher et al. (2019), the present modelling framework has a number of aspects that can be addressed in future research. One of the key aspects is the consideration of hydrological connectivity among catchments assuming non-disturbed flow. As dams can separate sites form their upstream catchments, the inclusion of this complex disturbance factor can alter the impact that upstream factors have on local fish distributions. Consequently, the conclusions resulting from the upstream area approach should be viewed in the context of the difficulty to quantify accurate dam influences. Also, for large transboundary catchments such as the Danube River catchment, the non-uniform level of available species data needs to be addressed. The latter includes spatial autocorrelation effects as well as non-availability of species data for large catchment

parts. Moreover, including point and non-point pollution sources has also the potential to refine our results. However, with regard to the good performance of the calibrated models and the species-specific outcomes we argue that the key outcomes would not substantially change even if we would have had better input data, or have included additional parameters to our models. Specifically, given the overall good performance of our SDMs, adding new parameters to our models would rather result in an overfitting than in a meaningful improvement in the model accuracy.

5. Conclusion

Negligible effects of the choice between local vs. upstream area effects means, that a significant portion of freshwater fish patterns can be explained by local, reach scale environmental factors. This finding implicates that especially for improving fish assemblages, river rehabilitation measures at the spatial scale of sites and reaches are highly relevant. However, we highlight the importance of using upstream effects for population and land cover related predictors when addressing *species sensitive to pollution* such as *B. barbus* and consequently taking into account the different scale components for investigating species-environment relationships. In particular, our results imply both *catchment order dependent* and *species-dependent* model complexity and importance of environmental drivers, which underlines the importance of compatibility between the scale of factor importance and the species conservation management scale.

Data accessibility statement

The environmental datasets analysed during the current study are available at http://www.worldclim.org/ (Climatic and topographic data), http://sedac.ciesin.columbia.edu/gpw (Global Rural-Urban Mapping Project version1) and https://land.copernicus.eu/pan-european/corine-land-cover (CORINE land cover database). Species occurrence data of the project Biofresh are available at www.freshwaterplatform.eu.

Acknowledgements

This work was supported by the DFG under Grant MA 6593/2-1 and by the Horizon 2020 ECOPOTENTIAL project (ref. 641762).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ecolmodel.2019.108818.

References

Allan, J.D., 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. Annu. Rev. Ecol. Evol. Syst. 35, 257–284.
Allan, J.D., Erickson, D.L., Fay, J., 1997. The influence of catchment land use on stream integrity across multiple spatial scales. Freshw. Biol. 37, 149–161.
Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J. Appl. Ecol. 43, 1223–1232.
Angermeier, P.L., Winston, M.R., 1988. Local vs. Regional influences on local diversity in stream fish communities of Virginia. Ecology 79, 911–927.
Austin, M., 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. Ecol. Modell. 200, 1–19.
Black, R.W., Munn, M.D., Plotnikoff, R.W., 2004. Using macroinvertebrates to identify biota-land cover optima at multiple scales in the Pacific Northwest, USA. J. North Am. Benthol. Soc. 23, 340–362.
Buisson, L., Blanc, L., Grenouillet, G., 2008. Modelling stream fish species distribution in a river network: the relative effects of temperature versus physical factors. Ecol. Freshw. Fish 17, 244–257.
Buisson, L., Thuiller, W., Casajus, N., Lek, S., Grenouillet, G., 2010. Uncertainty in ensemble forecasting of species distribution. Glob. Chang. Biol. 16, 1145–1157. https://doi.org/10.1111/j.1365-2486.2009.02000.x.
Chapman, D.V., Bradley, C., Gettel, G.M., Hatvani, I.G., Hein, T., Kovacs, J., ... Varbiro, G., 2016. Developments in water quality monitoring and management in large river catchments using the Danube River as an example. Environ. Sci. Policy 64, 141–154.
Coudon, C., Gégout, J.C., 2007. Quantitative prediction of the distribution and abundance of *Vaccinium myrtillus* with climatic and edaphic factors. J. Veg. Sci. 18, 517–524.
DEFRA, 2013. Catchment Based Approach: Improving the Quality of Our Water Environment. available at:. Department for Environment, Food and Rural Affairs. Policy paper. https://www.gov.uk/government/publications/catchment-based-approach-improving-the-quality-of-our-water-environment.
de Jager, A.L., Vogt, J.V., 2010. Development and demonstration of a structured hydrological feature coding system for Europe. Hydrol. Sci. J. 55, 661–675.
Domisch, S., et al., 2013. Choice of study area and predictors affect habitat suitability projections, but not the performance of species distribution models of stream biota. Ecol. Modell. 257, 1–10.
Dormann, C.F., et al., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36, 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x.
EEA, 2011. CORINE Land Cover. Version 13, Copenhagen: European Environmental Agency (EEA). available at. www.eea.europa.eu/publications/COR0-landcover.
Esselman, P.C., Allan, J.D., 2010. Relative influences of catchment- and reach-scale abiotic factors on freshwater fish communities in rivers of northeastern Mesoamerica. Ecol. Freshw. Fish 19, 439–454.
Fagan, W.F., 2002. Connectivity, fragmentation, and extinction risk in dendritic metapopulations. Ecology 83, 3243–3249.
Fausch, K.D., Torgersen, C.E., Baxter, C.V., Li, H.W., 2002. Landscapes to riverscapes: bridging the gap between research and conservation of stream fishes. BioScience 52, 483–498.
Feld, C., Hering, D., 2007. Community structure of function: effects of environmental stress on benthic macroinvertebrates at different spatial scales. Freshw. Biol. 52, 1380–1399.
Franklin, J., 2009. Mapping Species Distributions: Spatial Inference and Prediction. Cambridge University Press, New York 336 p.
Greig, S.M., Sear, D.A., Carling, P.A., 2005. Fine sediment accumulation in salmon spawning gravels and the survival of incubating salmon progeny: implications for spawning habitat management. Sci. Total Environ. 344, 241–258.
Grenouillet, G., Buisson, L., Casajus, N., Lek, S., 2011. Ensemble modelling of species distribution: the effects of geographical and environmental ranges. Ecography 34, 9–17.
Grömping, U., 2006. Relative Importance for Linear Regression in R: The Package relaimpo. J. Stat. Softw. 17, 1–27.
Hastie, T.J., 2005. Generalized Additive Models, R Package Version 0.94.
Hijmans, R.J., Cameron, S.E., Parra, J.L., 2007. WorldClim Version 1.4. Museum of Vertebrate Zoology. Available at:. University of California. http://www.worldclim.org/.
ICPDR, 2009. The Danube River Basin: Facts and Figures. International Commission for the Protection of the Danube River (ICPDR), Vienna.
Isaak, D.J., Wenger, S.J., Young, M.K., 2017. Big biology meets microclimatology: defining thermal niches of ectotherms at landscape scales for conservation planning. Ecol. Appl. 27, 977–990.
Jensen, D.W., Steel, E.A., Fullerton, A.H., Pess, G.R., 2009. Impact of fine sediment on egg-to-fry survival of Pacific salmon: a meta-analysis of published studies. Rev. Fish. Sci. 17, 348–359.
Kail, J., Jähnig, S.C., Hering, D., 2009. Relation between floodplain land use and river hydromorphology on different spatial scales – a case study from two lower-mountain catchments in Germany. Fundam. Appl. Limnol. 174, 63–73.
Kail, J., Wolter, C., 2013. Pressures at larger spatial scales strongly influence the ecological status of heavily modified river water bodies in Germany. Sci. Total Environ. 454–455, 40–50.
Kärcher, O., Frank, K., Walz, A., Markovic, D., 2019. Scale effects on the performance of niche-based models of freshwater fish distributions. Ecol. Modell. 405, 33–42.
Kottelat, M., Freyhof, J., 2007. Handbook of European Freshwater Fishes. Freyhof, Cornol; Kottelat, Berlin 646 p.
Kuemmerlen, M., Reichert, P., Siber, R., Schuwirth, N., 2019. Ecological assessment of river networks: from reach to catchment scale. Sci. Total Environ. 650, 1613–1627.
Larsen, K., 2016. R Package 'Information', Version 0.0.9.
Lapointe, M.F., Bergeron, N.E., Bérubé, F., Pouliot, M.-A., Johnston, P., 2004. Interactive effects of substrate sand and silt contents, redd-scale hydraulic gradients, and interstitial velocities on egg-to-emergence survival of Atlantic salmon (Salmo salar). Can. J. Fish. Aquat. Sci. 61, 2271–2277.
Lassalle, G., Crouzet, P., Gessner, J., Rochard, E., 2010. Global warming impacts and conservation responses for the critically endangered European Atlantic sturgeon. Biol. Conserv. 143, 2441–2452.
Lauzeral, C., Grenouillet, G., Brosse, S., 2013. Spatial range shape drives the grain size effects in species distribution models. Ecography 36, 778–787.
Lévêque, C., Oberdorff, T., Paugy, D., Stiassny, M.L.J., Tedesco, P.A., 2008. Global diversity of fish (Pisces) in freshwater. Hydrobiologia 595, 545–567.
Lindeman, R.H., Merenda, P.F., Gold, R.Z., 1980. Introduction to Bivariate and Multivariate Analysis. Scott, Foresman, Glenview, IL 444 p.
Linke, S., Norris, R.H., Pressey, R.L., 2008. Irreplaceability of river networks: towards catchment-based conservation planning. J. Appl. Ecol. 45, 1486–1495.
Markovic, D., Freyhof, J., Wolter, C., 2012. Where are all the fish: potential of biogeographical maps to project current and future distribution patterns of freshwater species. PLoS One 7, e40530. https://doi.org/10.1371/journal.pone.0040530.
Markovic, D., Carrizo, S.F., Kärcher, O., Walz, A., David, J.N.W., 2017. Vulnerability of European freshwater catchments to climate change. Glob. Chang. Biol. 23, 3567–3580. https://doi.org/10.1111/gcb.13657.

Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K., Thuiller, W., 2009. Evaluation of consensus methods in predictive species distribution modelling. Divers. Distrib. 15, 59–69.

McCluney, K.E., Poff, N.L., Palmer, M.A., Thorp, J.H., Poole, G.C., Williams, B.S., ... Baron, J.S., 2014. Riverine macrosystems ecology: sensitivity, resistance, and resilience of whole river basins with human alterations. Front. Ecol. Environ. 12, 48–58. https://doi.org/10.1890/120367.

Mertes, K., Jetz, W., 2018. Disentangling scale dependencies in species environmental niches and distributions. Ecography 41, 1604–1615. https://doi.org/10.1111/ecog.02871.

Moerke, A.H., Lamberti, G.A., 2006. Scale-dependent influences on water quality, habitat, and fish communities in streams of the Kalamoazoo River Basin, Michigan (USA). Aquat. Sci. 68, 193–205.

R Development Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

Schinegger, R., Pletterbauer, F., Melcher, A., Schmutz, S., 2016. Metadata describing the European Fish Index Plus (EFI+) database. Freshw. Metadata J. 17, 1–12. https://doi.org/10.15504/fmj.2016.17.

Soranno, P.A., Cheruvelil, K.S., Bissell, E.G., Bremigan, M.T., Downing, J.A., Fergus, C.E., ... Webster, E., 2014. Cross-scale interactions: Quantifying multi-scaled cause-effect relationships in macrosystems. Front. Ecol. Environ. 12, 65–73.

Soulsby, C., Youngson, A.F., Moir, H.J., Malcolm, I.A., 2001. Fine sediment influence on salmonid spawning habitat in a lowland agricultural stream: a preliminary assessment. Sci. Total Environ. 265, 295–307.

Stephenson, J.M., Morin, A., 2009. Covariation of stream community structure and biomass of algae, invertebrates and fish with forest cover at multiple spatial scales. Freshw. Biol. 54, 2139–2154.

Strahler, A.N., 1964. Quantitative geomorphology of drainage basins and channel networks. In: Chow, V.T. (Ed.), Handbook of Applied Hydrology. McGraw-Hill, New York, pp. 4–74.

Strayer, D.L., Beighley, R.E., Thompson, L.C., Brooks, S., Nilsson, C., Pinay, G., Naiman, R.J., 2003. Effects of land cover on stream ecosystems: roles of empirical models and scaling issues. Ecosystems 6, 407–423.

Sui, P., Iwasaki, A., Saavedra, V., O. C, Yoshimura, C., 2014. Modelling basin-scale distribution of fish occurrence probability for assessment of flow and habitat conditions in rivers. Hydrol. Sci. J. 59, 618–628.

Swan, M.C., Brown, B.L., 2017. Metacommunity theory meets restoration: isolation may mediate how ecological communities respond to stream restoration. Ecol. Appl. 27 (7), 2209–2219.

Thuiller, W., 2004. Patterns and uncertainties of species range shifts under climate change. Glob. Change Biol. 10, 2020–2027.

Tudesque, L., Tisseuil, C., Lek, S., 2014. Scale-dependent effects of land cover on water physico-chemistry and diatom-based metrics in a major river system, the Adour-Garonne basin (South Western France). Sci. Total Environ. 466–467, 47–55.

Verberk, W.C.E.P., Durance, I., Vaughan, I.P., Ormerod, S.J., 2016. Field and laboratory studies reveal interacting effects of stream oxy-genation and warming on aquatic ectotherms. Glob. Change Biol. 22, 1769–1778.

Verdonschot, P.F.M., 2009. Impact of hydromorphology and spatial scale on macroinvertebrate assemblage composition in streams. Integr. Environ. Assess. Manag. 5, 97–109.

Wilks, D., 1995. Statistical Methods in the Atmospheric Sciences. Academic Press Inc., San Diego 648 p.

Woodward, G., Perkins, D.M., Brown, L.E., 2010. Climate change and freshwater ecosystems: impacts across multiple levels of organization. Philos. Trans. R. Soc. B 365, 2093–2106.