Environmental predictors of lake fish diversity across gradients in lake age and spatial scale

Martinsen, Kenneth Thorø; Kristensen, Emil; Baastrup-Spohr, Lars; Søndergaard, Martin; Carl, Henrik; Jeppesen, Erik; Sand-Jensen, Kaj; Kragh, Theis

Published in:
Freshwater Biology

DOI:
10.1111/fwb.14090

Publication date:
2023

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Environmental predictors of lake fish diversity across gradients in lake age and spatial scale

Kenneth Thorø Martinsen1 | Emil Kristensen1 | Lars Bastrup-Spoehr1 | Martin Søndergaard2,3 | Henrik Carl4 | Erik Jeppesen2,3,5,6 | Kaj Sand-Jensen1 | Theis Kragh7

Abstract

1. Freshwater ecosystems have experienced a great loss of biodiversity in the recent century due to eutrophication and loss of habitats, particularly in agricultural lowlands. Some of these ecosystems have recently been re-established, or entirely new lakes have been created to improve nutrient removal, biodiversity, and recreation.

2. Fish have an important structuring role in lake ecosystems, but little is known about the temporal development of fish species richness or composition in new or re-established lakes, especially in comparison with natural lakes. We investigated the influence of environmental variables and landscape features on fish species richness and fish assemblage in drainage basins and in 34 new (between 0 and 99 years) and 193 natural lakes in Denmark, using structural equation modelling.

3. Fish species richness in drainage basins is influenced primarily by basin elevation, lake area, and salinity at the basin outlet; low-salinity coastal regions are important migration pathways for many freshwater fish species, not just for anadromous or catadromous species. Land use does not appear to influence richness at the drainage basin scale, indicating that the influence of anthropogenic activities is minor or that significant effects thereof occurred a long time ago (decades to centuries).

4. The drainage basin richness defines the apparent pool of species that may colonise lakes within the basin, but the actual number of species is particularly influenced by stream network connectivity.Disconnected lakes have fewer species, a deficit that is sustained over long time scales.

5. In addition to the influence of stream connectivity, lake characteristics such as elevation, surface area, and water chemistry (alkalinity and pH) are also important predictors of fish species richness, whereas we did not find any effect of lake nutrient concentrations. The direct effect of land use was not evaluated. Ordination analysis shows that fish assemblage depends on both lake type (new/natural) and stream connectivity (connected/disconnected). Furthermore, species-specific
1 | INTRODUCTION

Lakes, reservoirs, wetlands, and rivers cover about 1.8% of the Earth’s surface (Pekel et al., 2016) but support 9.5% of all known animals and one third of all vertebrate species (Balian et al., 2008). Recently, freshwater ecosystems have experienced a much greater biodiversity loss than their terrestrial and marine counterparts (Reid et al., 2019), with almost 40% of European and North American freshwater fish species now considered endangered (Freyhof & Brooks, 2017; Jelks et al., 2008). Fish are of key importance in lake ecosystems, as they influence the food web structure, nutrient dynamics, and water clarity (Hansson et al., 1998; Jeppesen et al., 2012). Multiple stressors have undermined freshwater biodiversity (Brondizio et al., 2019; Reid et al., 2019), not least the large-scale reclamation of shallow lakes and wetlands for agricultural purposes (Brondizio et al., 2019; Reid et al., 2019), not least the large-scale reclamation of shallow lakes and wetlands for agricultural purposes in both North America and Europe, which resulted in a loss of up to 80% of freshwater areas (Biggs et al., 2017; Verhoeven, 2014; Wine & Laronne, 2020). Another stressor has been the deterioration of water quality, mainly due to eutrophication (Jeppesen et al., 2000; Sand-Jensen et al., 2017; Sas, 1990). Interest in restoring historical freshwater habitats, combined with intentions to reduce nutrient loading to coastal areas and improve carbon sequestration, have motivated initiatives to re-establish or create entirely new lake ecosystems (Audet et al., 2020; Badiou et al., 2011; Hoffmann & Baattrup-Pedersen, 2007). Here, the term new lakes encompass all re-established and entirely new lakes, including artificial man-made lakes such as gravel pit lakes less than 100 years old.

Globally, more than 184,000 km² of new permanent water bodies including reservoirs and gravel pit lakes appeared between 1984 and 2015 (Blanchette & Lund, 2016; Mollema & Antonellini, 2016; Pekel et al., 2016). However, the biotic communities in new lakes often fall short in comparison with reference sites. Limited stream connectivity hinders the ability of freshwater fish species to colonise a lake, and movement between basins can be hindered by the few opportunities for overland dispersal and by the high salinity of ocean corridors (Hirsch et al., 2018). While the transport of fish eggs within waterfowl guts has recently been confirmed (Lovas-Kiss et al., 2020; Silva et al., 2019), the practical relevance of this mechanism continues to be viewed as extremely limited.

Fish assemblage is structured by spatiotemporal filtering occurring at multiple scales, ranging from geographic to biotic and abiotic (Jackson et al., 2001; Mehner & Bru cet, 2022). The geographic range of freshwater fish in Europe is highly influenced by migrations during the low-sea-level Quaternary period, when the connectivity between contemporary river basins was at its highest (Carvajal-Quintero et al., 2019). While low-sea periods promoted species distribution between basins, frequent glaciations (the most recent ending 11,700 years ago in Denmark) caused many local extinctions (Moyle & Herbold, 1987).

Fish species richness at the drainage basin level (S basin) is influenced by immigration history, but the extent and diversity of specific lake habitats are also likely to play important roles in subsequent population stability. Drainage basins with abundant fish habitats tend to harbour more species, both because larger populations reduce the risk of extinction and because larger habitat size tends to increase habitat diversity (Eckmann, 1995; Jackson et al., 2001). The types of habitats available can further be increased by landscape-level features such as variation in elevation (Carvajal-Quintero et al., 2015). In addition, anthropogenic activity can affect S basin both positively, through the stocking of new fish species from neighbouring basins, and negatively, through eutrophication and physical habitat deterioration (Bru cet et al., 2013; Matern et al., 2019; Zhao et al., 2016). Assuming dispersal between drainage basins is low or occurs sporadically over longer time scales, S basin represents the imminent species pool that can (re-) colonise habitats within the basin, including new lakes.

While the colonisation of isolated lakes can be limited, and dispersal presumably happens during the infrequent floods that connect water bodies within a basin, transfer of fish egg by birds, either through the gut or attached to their feet, have shown to be rare (Hirsch et al., 2018; Lovas-Kiss et al., 2020). In contrast, the intentional transfer of fish by humans can play an important role in species composition and fish species richness in lakes (S lake), not least in lakes located close to cities (Copp et al., 2007; Ruesink, 2005; Trochine et al., 2018).

For lakes connected to stream networks, fish species may colonise unoccupied areas where surface water facilitates movement and environmental conditions allow persistence (Kristensen et al., 2020; Olden et al., 2001). The impact of connectivity on species diversity depends on the organisation, setting, and permeability of patches and corridors of waterways (Mehner et al., 2014). For some freshwater fish species, such as bream (Abramis brama), the migration and movement seem almost random, while others, such as roach (Rutilus rutilus), can be driven by predator avoidance (Brodersen et al., 2019).

Once a fish species arrives at a lake, its ability to survive depends on both abiotic variables and biotic interactions between and withi...
species, such as predation, reproduction, and resource competition (Leibold et al., 2004; Lucas et al., 2008). Lake size and depth, both of which may be proxies for heterogeneity, are often positively correlated with $S_{\text{lake}}$ (Matuszek & Beggs, 1988; Mehner et al., 2005, 2007; Olin et al., 2002). Among the lake environmental variables, low pH in particular is known to negatively influence $S_{\text{lake}}$ (Matuszek & Beggs, 1988), an impact that has been widespread in the Northern temperate zone due to acid rain (Schindler, 1988). Other common anthropogenic effects include eutrophication and shoreline modifications. While nutrients normally influence $S_{\text{lake}}$ positively (Brucet et al., 2013; Jeppesen et al., 2000; Olin et al., 2002), shoreline modifications usually have a negative effect (Jennings et al., 1999).

The development of freshwater fish communities has been studied in reservoirs (Agostinho et al., 2016; Paller et al., 1992), new lakes (Degani et al., 1998; Kristensen et al., 2020), ponds (Ray et al., 2004), and wetlands (Langston & Kent, 1997). To our knowledge, Emmrich et al. (2014) and Matern et al. (2022) are the only studies investigating $S_{\text{lake}}$ and assemblages in natural and new lakes; however, these studies only include gravel pit lakes as new lakes and did not evaluate the effect of connectivity. We hypothesised that: (1) $S_{\text{basin}}$ is influenced by lake habitat area, landscape history, and anthropogenic pressure; (2) $S_{\text{lake}}$ is structured by a combination of dispersal and environmental filters; and (3) rapid dispersal of most fish species will result in rapid development of fish communities in new lakes connected to a stream network, resulting in minor differences between new and natural lakes, while disconnected new lakes will have fewer species than comparable natural lakes.

### 2 | MATERIALS AND METHODS

#### 2.1 | Study region and data sources

Denmark has approximately 180,000 lentic water bodies of mostly small size (mean area 4,051 m² and median area 490 m²; SDFE, 2021). The country has a temperate climate (annual mean air temperature 8.1°C and annual precipitation 704 mm), a low-relief landscape, and land use is dominated by agriculture (60%). Historically, numerous shallow lakes and wetlands had been reclaimed for agriculture, of which some are now re-established (Hansen, 2014). To investigate the predictors of $S_{\text{lake}}$, we collected data from 25 new lakes and drew on data from 193 natural lakes and nine additional new lakes that had been collected by the Danish National Environmental Monitoring Program (NOVANA; MFVM & DCE, 2021). To further improve coverage at the drainage basin level, we also included data from a national atlas survey of freshwater fish (Fish Atlas; Carl & Møller, 2012). We determined a range of geospatial and environmental variables believed to be relevant for fish species diversity in drainage basins and lakes.

#### 2.2 | Drainage basin delineation

We delineated 1,723 (mean area 2,283 ha and median area 307 ha) topographical drainage basins, the area where water flows downslope along an elevation gradient before approaching the coastline, using a digital elevation model (DEM; SDFE, 2021). The DEM is created from country-wide high resolution (1.6 m) LIDAR surveys resampled to a resolution of 10 m. We determined flow directions and delineated drainage basins using GRASS GIS (v. 7.6; Jasiwicz and Metz, 2011; Metz et al., 2011; Neteler and Mitasova, 2013).

#### 2.3 | Drainage basin richness

$S_{\text{basin}}$ was assessed using data from our own surveys (described in the next section), fish species data collected in lakes and streams between 1990 and 2020 by the national environmental monitoring programme, and the Fish Atlas, resulting in the most comprehensive collection of distribution records of freshwater fish in Denmark. The study region and data sources are shown in the Table 1 and Figure 1.
Denmark. The Fish Atlas is a country-wide survey containing georeferenced presence records of Danish freshwater fish compiled from published sources and further supplemented by sampling using multiple methods (Carl & Møller, 2012). Species records from all sources were combined and fish species richness was summed for each drainage basin. Species that are not found in lake habitats, hybrids, and infertile species (e.g., species stocked too far outside their natural range) were not included in the analysis (Table S1).

2.4 | Drainage basin predictors

We gathered several variables believed to influence $S_{basin}$ (Table 1). We calculated drainage basin morphometry and other characteristics using the DEM, maps of land use (Corine Land Cover; Bossard et al., 2000), and national maps of lakes and streams (SDFE, 2021). The potential predictors include the sum of stream length, the total lake area, mean elevation, elevation range, the proportion of agriculture, the proportion of artificial areas (urban, industrial, construction sites, and parks), basin area and circumference, and a binary variable (ice cover) indicating whether the basin was ice-covered during the last glacial maximum (Weichsel glaciation 115,000–11,700 years ago; GEUS, 2021). As some species use the coastal environment as a migration pathway between drainage basins, we also determined salinity at the outlet of each drainage basin (E.U. Copernicus, 2021).

2.5 | Lake richness

To determine $S_{lake}$, we compiled data from 193 natural lakes and 34 newly established lakes of varying ages (0–99 years, median = 13), all surveyed between 2006 and 2020. The lake survey protocol was identical to that of the national monitoring programme. In short, a varying number (depending on the lake area and depth) of multi-mesh gillnets with EU-standardised dimensions (30 by 1.5 m with 2.5 m sections of different mesh sizes; CEN, 2015) were set in the evening and retrieved in the following morning during between 15 August and 15 September. In addition to gill nets, each lake was sampled in a standardised manner with either electrofishing or fyke nets, which aided in catching species having a lower tendency of being caught in the gill nets. For lakes that were sampled more than once between 2006 and 2020, we used the sample with the highest number of species.

2.6 | Lake predictors

To model influences on $S_{lake}$, we determined a range of environmental and geospatial predictor variables (Table 2). Data on environmental variables were acquired for all selected lakes from either the NOVANA programme (MFVM & DCE, 2021) or as part of our sampling following the same protocol. We used the average of monthly surface water measurements, collected at a mid-lake station during summer (May–September), of alkalinity, pH, total phosphorus, total nitrogen, chlorophyll $a$, and Secchi depth. For each lake, we also determined multiple geospatial variables: surface area, shoreline length, mean depth, maximum depth, volume, elevation, and stream network connectivity (binary, disconnected = 0, connected = 1). We also included a lake type variable (binary, new = 0, natural = 1).

2.7 | Statistical analysis

We analysed fish species richness at both the drainage basin and lake level in a piecewise structural equation model (PSEM) framework, which allows hypothesised relationships between multiple predictor variables and models to be specified as paths (Lefcheck, 2016). This approach was motivated by the hierarchical structure of the data, i.e., $S_{basin}$ is expected to influence $S_{lake}$ and is thus a response variable in the drainage basin model and a predictor variable in the lake model. We specified a PSEM consisting of two sub-models of fish species richness at the drainage basin and lake levels.

Before modelling, we examined the density distribution of the predictor variables. Some variables were log$_{10}(x)$, or log$_{10}(x+1)$ transformed to reduce skewness. To avoid intercorrelation between predictor variables, we calculated the variance inflation factor and iteratively discarded predictor variables with high variance inflation factors (>3) for each sub-model (Tables 1 and 2; Zuur et al., 2010).

As part of the initial data exploration, we examined the relationship between $S_{lake}$, lake age, and connectivity for the new lakes using a Poisson-type generalised linear model. For the PSEM analysis, the sub-models are Poisson-type generalised linear mixed models with basin included as a random effect (Bates et al., 2015), and the combined PSEM model was fitted using the `piecewiseSEM` R-package (Lefcheck, 2016). We report the marginal ($R^2_{marg}$, fixed effects only) and conditional ($R^2_{cond}$, both fixed and random effects) trigamma pseudo-$R^2$ values (Nakagawa et al., 2017).

To assess potential differences in fish assemblages in new and natural lakes, we performed nonmetric multidimensional scaling (NMDS) using the vegan R-package (Oksanen et al., 2020). The NMDS was performed on lake-level incidence data (presence/absence) using Sørensen dissimilarity. Differences were compared using permutation test followed by pairwise comparison of four groups resulting from the combination of the lake type and stream connectivity variable with Holm’s correction for multiple comparisons. We determined the frequency of fish species occurrences in natural and new lakes by quantifying the frequency with which a species present in the drainage basin also exists in that basin’s lake(s). This could highlight specific species that are present in natural but not in new lakes, or vice-versa. All data analyses were conducted using R (version 4.1; R Core Team, 2021).
TABLE 2 Mean and range of fish species richness and predictor variables of the lakes included in the analysis (n = 227).

<table>
<thead>
<tr>
<th>Lake type</th>
<th>Variable</th>
<th>Unit</th>
<th>Role</th>
<th>Min.</th>
<th>Mean</th>
<th>Max.</th>
<th>Min.</th>
<th>Mean</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Richness</td>
<td>N</td>
<td>Response</td>
<td>0</td>
<td>6.1</td>
<td>13</td>
<td>0</td>
<td>4.9</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Alkalinity</td>
<td>meq./L</td>
<td>Predictor</td>
<td>-0.058</td>
<td>2</td>
<td>5.1</td>
<td>0.44</td>
<td>2.2</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td>Surface area</td>
<td>km²</td>
<td>Predictor</td>
<td>0.04</td>
<td>1.1</td>
<td>40.0</td>
<td>0.02</td>
<td>1.0</td>
<td>8.9</td>
</tr>
<tr>
<td></td>
<td>Maximum depth</td>
<td>m</td>
<td>Predictor</td>
<td>0.36</td>
<td>6</td>
<td>38</td>
<td>1</td>
<td>3.4</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Elevation</td>
<td>m</td>
<td>Predictor</td>
<td>-1</td>
<td>24</td>
<td>100</td>
<td>-1</td>
<td>18</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Stream connection</td>
<td>0/1</td>
<td>Predictor</td>
<td>4.3</td>
<td>8.1</td>
<td>9.6</td>
<td>7.1</td>
<td>8.3</td>
<td>9.3</td>
</tr>
<tr>
<td></td>
<td>Secchi depth</td>
<td>m</td>
<td>Predictor</td>
<td>0.2</td>
<td>1.5</td>
<td>6.8</td>
<td>0.26</td>
<td>1.2</td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td>Total nitrogen</td>
<td>mg/L</td>
<td>Predictor</td>
<td>0.28</td>
<td>1.3</td>
<td>5.4</td>
<td>0.38</td>
<td>1.8</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Total phosphorus</td>
<td>mg/L</td>
<td>Predictor</td>
<td>0.01</td>
<td>0.13</td>
<td>0.99</td>
<td>0.011</td>
<td>0.19</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>Volume</td>
<td>m³×10⁶</td>
<td>Discarded</td>
<td>0.01</td>
<td>5.4</td>
<td>230</td>
<td>0.025</td>
<td>1.3</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Mean depth</td>
<td>m</td>
<td>Discarded</td>
<td>0.12</td>
<td>2.7</td>
<td>16</td>
<td>0.37</td>
<td>1.4</td>
<td>4.3</td>
</tr>
<tr>
<td></td>
<td>Chlorophyll a</td>
<td>μg/L</td>
<td>Discarded</td>
<td>0.5</td>
<td>52</td>
<td>420</td>
<td>2</td>
<td>59</td>
<td>240</td>
</tr>
<tr>
<td></td>
<td>Shoreline length</td>
<td>km</td>
<td>Discarded</td>
<td>0.63</td>
<td>4.9</td>
<td>44</td>
<td>0.58</td>
<td>6.5</td>
<td>32</td>
</tr>
</tbody>
</table>

Note: The role of each variable in the modelling is also denoted, as some have been discarded due to high intercorrelation with other predictor variables.

3 | RESULTS

3.1 | Drainage basin richness

Of the 27 identified lake fish species (Table S1), one or more species were found in 894 drainage basins (Figure 1) that, together, cover 89.6% of Denmark’s area (c. 43,000 km²). A total of 829 basins contained no fish species and were generally small, coastal, and without lakes present. Mean \( S_{basin} \) was 3 (5.9 when excluding drainage basins with zero species) and the highest \( S_{basin} \) (25 species) were found in the two largest drainage basins of River Guden and River Skjern, and also in the smaller drainage basin of River Von and Lake Stadil.

The 227 lakes included in this analysis are located in 98 drainage basins (Figure S1), one or more species was also found in 226 S basins (Figure 2a) that, together, cover 89.6% of Denmark’s area (c. 43,000 km²). A total of 829 basins contained no fish species and were generally small, coastal, and without lakes present. Mean \( S_{basin} \) was 3 (5.9 when excluding drainage basins with zero species) and the highest \( S_{basin} \) (25 species) were found in the two largest drainage basins of River Guden and River Skjern, and also in the smaller drainage basin of River Von and Lake Stadil.

The 227 lakes included in this analysis are located in 98 drainage basins that span almost the entire range of \( S_{basin} \) (Table 1). Furthermore, the candidate predictor variables span large differences in land use, geology, outlet salinity, and geomorphology, making the analysis generalisable to other temperate, lowland regions.

3.2 | Fish species richness in new and natural lakes

The natural (\( n = 193 \)) and new lakes (\( n = 34 \)) included in the analysis span wide gradients of water clarity, alkalinity, bathymetry, and nutrient state, and are representative of lakes in Denmark (Table 2). Compared to natural lakes, new lakes were more often disconnected from their basin’s stream network, and also shallower, more nutrient-rich, and situated at lower elevations. Figure 2a shows that the mean \( S_{lake} \) in natural lakes is significantly higher than that in new lakes (Mann–Whitney U-test, \( U = 2.371.5, p < 0.0096 \); Table 2). When including stream network connectivity, \( S_{lake} \) in new and natural lakes with similar connectivity does not differ significantly while all other pairwise comparisons do (Figure 2b; \( p = 0.004 \) for new connected vs. natural disconnected and \( p < 0.001 \) for all other combinations; Kruskal–Wallis test followed by Dunn’s test of multiple comparisons with Bonferroni correction). The new lakes vary in age from 0 to 99 years, with the majority below 30 years (mean = 17 years). The apparent increase in \( S_{lake} \) during the first 30 years (Figure 2c) is not an effect of age (Poisson-type GLM reported as an estimate (±SE), 0.04 ±0.028, \( p = 0.143 \) when including the effect of stream network connectivity which contributed significantly (1.45 ±0.37, \( p < 0.001 \); interaction not significant).

3.3 | Predictors of richness in drainage basins and lakes

Acknowledging the hierarchical nature of the data, i.e., the influence of \( S_{basin} \) on \( S_{lake} \), we applied a piecewise structural equation model framework. Following the removal of highly intercorrelated predictor variables, the PSEM could successfully be fitted (non-significant global goodness-of-fit test; Fisher’s C = 43.245, \( p = 0.133 \), df = 34; Figure 3). We found that \( S_{basin} \) could be reasonably explained (\( R^2_{marg} = 0.57 \) and \( R^2_{cond} = 0.59 \)) by three of the candidate predictor variables (direction of effect in parentheses): total lake area (positive), elevation (positive), and salinity at the drainage basin outlet (negative). Land use (agriculture and artificial surfaces) and landscape slope were not significant, and neither was ice cover during the most recent Weichselian glaciation despite being close to the traditional 0.05 significance level (\( p = 0.053 \)). \( S_{lake} \) was also reasonably well explained (\( R^2_{marg} = 0.52 \) and \( R^2_{cond} = 0.52 \)) using multiple predictors (direction of effect in parentheses): pH (positive), alkalinity (positive), surface area (positive), elevation...
(negative), lake type (positive), stream connection (positive), and the interaction of the latter two (negative). Furthermore, as expected, $S_{\text{basin}}$ positively affected $S_{\text{lake}}$. Lake type, stream connection, and their interaction, were highly significant showing that connectivity, modulated by lake type, had a pronounced influence on $S_{\text{lake}}$ in new lakes.

3.4 | Fish assemblage and species occurrences in natural and new lakes

To assess potential differences in fish communities between natural and new lakes, we used NMDS based on species incidence data. The NMDS analysis (stress = 0.14) indicated differences in fish assemblages between different lake types and stream network connectivity (Figure 4). A permutation test showed significant pairwise differences for (adjusted $p$-values in parentheses): natural connected and natural disconnected ($p = 0.006$), natural connected and new disconnected ($p = 0.006$), and natural disconnected and new disconnected lakes ($p = 0.008$). The results, which are very similar to those for $S_{\text{lake}}$, highlight that both richness and fish assemblage depend on both lake type and stream network connectivity. However, lakes connected to the stream network, whether new and natural, have similar $S_{\text{lake}}$ and assemblages.

The average frequency of occurrences is similar to or slightly higher in natural lakes for most species (Figure 5). Some exceptions to this are the eel (*Anguilla anguilla*), goldfish (*Carassius auratus*), crucian carp (*Carassius carassius*), and trout (*Salmo trutta*), which are more frequent in new lakes compared to natural lakes.
Four species, spined loach (**Cobitis taenia**), burbot (**Lota lota**), eastern mudminnow (**Umbra pygmaea**), and whitefish (**Coregonus maraena**), are completely absent in new lakes and present in only a few drainage basins. The most common fish species in drainage basins, perch (**Perca fluviatilis**), roach, and pike (**Esox lucius**), are also very likely to be present in individual lakes. Species such as the eel, nine-spined stickleback (**Pungitius pungitius**), and trout are widely present in drainage basins but are not commonly observed in the investigated lakes. Of the 27 species included in the analysis, 3 are considered non-indigenous (goldfish, carp (**Cyprinus carpio**), and eastern mudminnow) with goldfish and carp being particularly common in many drainage basins.

**FIGURE 3** Diagram showing the fitted piecewise structural equation model consisting of two Poisson type generalised linear mixed models of fish species richness at the drainage basin and lake level. Solid lines are significant paths with negative relationships shown in red. Parameter estimates (±SE) are reported in italics and level of significance with asterisks (* for p < 0.5, ** for p < 0.01, and *** for p < 0.001) for significant paths (estimates are not standardised).

We found one or more fish species in the drainage basins that cover most (89.6%) of Denmark’s area, the rest being small coastal basins with little or no lake area. $S_{\text{basin}}$ generally followed our a priori expectation that larger basins harboured a higher number of species. This was confirmed in our PSEM analysis of 98 drainage basins. Lake surface area at the basin (total lake area) and lake level were found to have a positive influence on richness. Using the total area of lakes in a drainage basin as a predictor variable as opposed to the drainage basin area itself is a natural choice, as we only considered fish species associated with lake habitats. The effect of total lake area on richness is in line with the *island biogeography theory* (MacArthur & Wilson, 1967) where each basin can be considered an island surrounded by inhospitable terrestrial or marine habitat. Larger lakes not only have more potential for habitat diversity, but also have a lower risk of species disappearing from the system, due to their greater populations and higher probability of recolonisation following a local extinction (Baber et al., 2002; Kragh et al., 2020; Mehner & Brucet, 2022). Recolonisation of surviving populations is likely to take place throughout the stream network (Kristensen et al., 2020).

Therefore, the area of lakes within a drainage basin lake might influence dispersal throughout the basin lake area, as larger lakes might provide better refuges and subsequently act as a source of species, while species in small lakes or ponds might be more prone to extinction due to, e.g., episodic anoxia (Sand-Jensen et al., 2019).

We also explored the ability of species to move between basins through coastal-based pathways. Here, we found a significant negative effect of salinity at the drainage basin outlet on $S_{\text{basin}}$, i.e., $S_{\text{basin}}$ is higher in drainage basins with outlets near low-salinity coastal waters. Salinity in coastal waters in Denmark varies from
Many years ago, as 60%–70% of the Danish land area has been under intensive cultivation for the last 120 years. Another possibility is that we studied richness and not the abundance of individual species. Although $S_{\text{basin}}$ may remain the same, land-use-induced eutrophication may well increase the abundance of generalist species at the expense of rare species (Jeppesen et al., 2000; Nielsen et al., 2012; Søndergaard et al., 2005). The second land use variable, artificial areas, did not have a significant influence either, probably reflecting the dominance of mainly urban areas across Denmark. The density of developed areas or cities was expected to influence the presence of species such as carp and goldfish (Copp et al., 2005), where stocking originates from garden ponds or, especially for carp, motivated by improving local opportunities for recreational fishing. However, species such as the carp and goldfish are already widely distributed, and while these species are non-indigenous, stocking in Danish lakes may both be authorised and unauthorised. Other studies have found positive effects of anthropogenic activity on fish species richness due to the introduction of exotic species in populated and easily accessible areas (Copp et al., 2007; Mor et al., 2012; Ruesink, 2005) or illegal introductions of game fish (Johannsen et al., 2009). In contrast, negative influences on richness of artificial shoreline modifications (Jennings et al., 1999) and constructed river dams (Kirsch & Peterson, 2014) have been observed in North America, but neither of these are common pressures in Denmark.

### 4.2 Lake richness

Species richness in the 227 lakes under study was significantly influenced by variables associated with lake chemistry (alkalinity and pH), habitat characteristics (elevation, surface area, new or
natural), connectivity (stream connection), and $S_{\text{basin}}$ (hypothesis 2). Unexpectedly, nutrients (total nitrogen and total phosphorus) and other habitat descriptors such as maximum depth and Secchi depth do not influence $S_{\text{lake}}$, but rather relative species abundance (Jeppesen et al., 2000; Mehnert et al., 2005; Persson et al., 1991). However, the finding is consistent with the non-significant land use effects on $S_{\text{lake}}$, particularly for agriculture, which is correlated to lake nutrient concentrations (Nielsen et al., 2012). Other studies have found lake pH to affect $S_{\text{lake}}$. Matuszek and Beggs (1988) found that $S_{\text{lake}}$ decreased with pH below 6, and pH alone explained 21% of the variation in $S_{\text{lake}}$ among lakes. However, our dataset includes only a few lakes with low pH. Lake area has a positive effect on $S_{\text{lake}}$, which is probably because both habitat area and habitat diversity typically increase with lake size (Rosenzweig, 1995), an effect that has been observed in numerous studies (Matuszek & Beggs, 1988; Minns, 1989). Lake elevation has a negative influence on $S_{\text{lake}}$, in contrast to the positive influence at the drainage basin level. While the range in lake elevation in our study was low (~1–100 meters), excluding a role of climatic differences, lake elevation is a proxy of position within the drainage basin, and lakes near the basin outlet tend to support a higher $S_{\text{lake}}$ (Olden et al., 2001), possibly driven by migratory species (Laske et al., 2016). Additionally, lakes at lower elevations might be subject to more frequent instances of flooding, whereby species can migrate through temporary waterways (Poizat & Corivelli, 1997), while lakes at higher elevations might be affected by discontinuities in the stream network, reducing habitat connectivity.

We found stream network connectivity to be important, supporting the findings of other studies (Carvajal-Quintero et al., 2019; Laske et al., 2016; Olden et al., 2001). The importance of connectivity on both $S_{\text{lake}}$ and fish assemblage suggests that stream network connectivity enables efficient migration within a drainage basin. The interaction between stream network connectivity and lake type and the lack of an effect of age suggests that this is especially important in new lakes (hypothesis 3). Furthermore, it shows that even natural lakes that are disconnected from the stream network do not attain high $S_{\text{lake}}$, indicating that dispersal events to such lakes are extremely limited, even in timescales of decades to centuries. One suggested mechanism for overland dispersal to isolated lakes is endozoochorous transport of fish eggs by waterfowl (Lovas-Kiss et al., 2020). Furthermore, flooding events during extreme precipitation events could promote dispersal to seemingly isolated lakes. However, present-day human control over nature, especially by preventing flooding events to secure both residential areas and farmland but also by modifying the connectivity within stream networks (Carl & Møller, 2012), has reduced the connectivity of lakes within basins. As expected, we found that $S_{\text{basin}}$ influences $S_{\text{lake}}$, suggesting that the species pool does play a role, in addition to the significant influence of multiple lake-level ecosystem characteristics. Species such as the eel, nine-spined stickleback, and trout were common in all drainage basins but generally not observed in the lakes. The nine-spined stickleback is a pioneer species that quickly diminishes in number as other species colonise a lake (Englund et al., 2009), while cold-water-adapted trout thrive only in temperature conditions found in deep, stratified, clear-water lakes with oxygenated conditions below the thermocline (Hari et al., 2006). In the case of eel, this might be a sampling artefact because they are difficult to catch using gill nets (Menezes et al., 2013). Due to the inherent selectivity of some methods for certain species, we supplemented gill nets with either electrofishing or fyke nets to obtain better species coverage, especially in the littoral zone, which generally makes up a large proportion of shallow lakes in Denmark. While electrofishing is a very effective method for determining $S_{\text{lake}}$ (Menezes et al., 2013), fyke nets have been shown to have comparable performance (Eggleton et al., 2010). The use of multiple methods for investigating fish assemblages in lakes is recommended by multiple studies (Diekmann et al., 2005; Fago, 1998), and because we only examine presence/absence records, we do not believe that this introduces significant bias in the analysis.

4.3 | Fish diversity in new lakes

When considering only the new lakes, stream network connectivity was found to have a much greater influence on $S_{\text{lake}}$ than lake age. Many studies have found that species numbers in new lakes increase to a saturation point (Degani et al., 1998; Paller et al., 1992) in a period that ranges from weeks (Baber et al., 2002) to a few years (Kristensen et al., 2020). Our study thus supports the idea that fish communities can develop rapidly in new lakes that are well-connected to stream networks. The species-specific frequency of occurrence shows that fish are generally found at the same frequencies in new and natural lakes. The four generalist species, perch, roach, pike, and rudd (Scardinius erythrophthalmus), are the most common in the drainage basins and occur with high frequency in both natural and new lakes. This finding is in line with results from investigations of gravel pit lakes in Denmark (Søndergaard et al., 2018) and Germany (Emmrich et al., 2014).

Recreational fishing is known to influence non-native fish species richness patterns in Europe (Zhao et al., 2016) and North America (Davis & Darling, 2017), where several catchments experienced an increase of almost 50% in richness due to the transferring of game fish to areas where the species were not found previously (Peoples et al., 2020). Studies in Europe have found different patterns in lakes managed for recreational fishing, showing that fish communities are subject to homogenisation (Matern et al., 2019, 2022) with cases of both elevated (Zhao et al., 2016) and unchanged non-native richness (Matern et al., 2019). While most of the investigated lakes are used for recreational fishing, they are not managed for this purpose, in a way similar to fisheries in other countries (Arlinghaus et al., 2021). However, our study shows that game species such as carp and goldfish are nevertheless frequent in the new lakes, probably as a consequence of illegal release of these species for recreational fishing. These different patterns call for larger scale studies, beyond a single country or region, of fish species richness and the role of management (Matern et al., 2022).
The carp is often released by humans due to its recreational value as a game fish (Arlinghaus & Mehner, 2003; Vilizzi, 2012) and has also been observed in Danish gravel pit lakes (Søndergaard et al., 2018). On the contrary, species such as bream and ruffe (Gymnocephalus cernua), which are less targeted for recreational fishing, were less frequently in new compared to natural lakes. Absence in disconnected gravel pit lakes in Denmark (Søndergaard et al., 2018) indicates an overall limited natural dispersal to disconnected waters for many fish species.

Of the 34 new lakes included in this analysis, only a few (five) were gravel pit lakes. These ecosystems are generally disconnected, deeper, and nutrient poor (Seelen et al., 2021) and thus habitats that diverge markedly from their natural Danish counterparts that tend to be connected to stream networks, eutrophic, and shallow (Søndergaard et al., 2018). Because of dispersal limitations, these potentially highly suitable habitats cannot be used by a large fraction of the natural species pool. Managed relocation (also termed assisted migration) of missing species might be particularly useful in such isolated systems (Olden et al., 2011). This should focus on species that are rare within a given drainage basin and have habitat preferences similar to those presented by the new disconnected lakes. Introducing native species with greater biotic resistance might bring the additional advantage of limiting the success of unauthorised introductions of game and ornamental species (Henriksson et al., 2015).

5 | CONCLUSIONS

Unexpectedly, we did not find a significant effect of anthropogenic pressures or landscape history on $S_{basin}$, which is primarily influenced by habitat area and salinity at the outlet, i.e., the accessibility to a natural pathway for migration between drainage basins. Fish species quickly find new lakes that are connected to a stream network, and these lakes do not differ markedly from their natural counterparts. Disconnected new lakes, however, have lower $S_{lake}$ and different species composition compared to connected new or natural lakes. Non-indigenous species such as carp and goldfish were frequent in new lakes, probably due to stocking. A promising avenue for the conservation of fish diversity would be to assist the migration of rare and poorly dispersed species to the new isolated lakes with promising environmental conditions.

AUTHOR CONTRIBUTIONS


ACKNOWLEDGMENTS

We thank Aage V. Jensen Foundation for grants to K.S.J., T.K., M.S., E.J., L.B.S., K.T.M., and E.K. by funding the monitoring programme and investigations of the new lakes. E.J. was also supported by the TÜBITAK BİDEB2232 programme (project 118C250). We thank David Stuligross for constructive comments and careful proofreading. We thank the two anonymous reviewers for their constructive comments which helped improve the manuscript and Anne Mette Poulsen for valuable English editions.

DATA AVAILABILITY STATEMENT

Data from the national environmental monitoring programme and other geospatial data are publicly available from sources cited in the main text. The Fish Atlas data are not publicly available. Scripts used for data processing, analysis, and figures are available at https://doi.org/10.17894/ucph.b637d47a-80d6-42ad-b3ae-a1ad7a447586.

ORCID

Kenneth Thøra Martensen https://orcid.org/0000-0001-8064-513X
Emil Kristensen https://orcid.org/0000-0001-6717-2921
Lars Bastrup-Spohr https://orcid.org/0000-0001-8382-984X
Martin Søndergaard https://orcid.org/0000-0002-1514-0482
Henrik Carl https://orcid.org/0000-0002-1939-1094
Erik Jeppesen https://orcid.org/0000-0002-0542-369X
Kaj Sand-Jensen https://orcid.org/0000-0003-2534-4638
Theis Kragh https://orcid.org/0000-0002-9760-2571

REFERENCES


SUPPORTING INFORMATION

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