Quantifying river fragmentation from local to continental scales: a data management and modelling toolbox

Jones, J.[1*], Belletti, B.[2, 3], Börger., L.[1], Segura, G.[4], Bizzi, S.[5], van de Bund, W.[6], Garcia de Leaniz, C.[1]

*Corresponding author. j.a.h.jones@swansea.ac.uk

[1] Department of Biosciences, College of Science, Swansea University, Swansea SA2 8PP, UK

[2] Department of Electronics, Information, and Bioengineering, Politecnico di Milano, Milano, Italy

[3] University of Lyon, UMR 5600 CNRS EVS, École Normale Supérieure de Lyon, Lyon, France

[4] ISE, 2 Chemin du Bois de Saget, 31240 Saint-Jean, France

[5] Department of Geosciences, University of Padova, Padua, Italy

[6] European Commission – Joint Research Centre, 21027 Ispra, VA, Italy

Abstract

Restoring river connectivity is a global conservation priority but quantifying river fragmentation has proved difficult due to the paucity of good barrier records, duplicate entries, and other sources of biases. Here we present some tools to help overcome some of these challenges and illustrate their application with case studies drawn across different spatial scales. We begin by proposing a classification of artificial instream barriers that harmonises disparate barrier types into six functional types, and present a binary classification key for ease of use. We then introduce a method for excluding duplicate barrier records that retains most genuine barriers and illustrate its practical use. Sampling bias is a pervasive problem in barrier inventories and we show how to detect and correct for it via bootstrapping of data obtained from standardised field surveys, ad-hoc records provided by citizens, and modelling. Finally, we show how to assess fragmentation when barriers cannot be aligned with the river network, and how to estimate barrier impacts from barrier height and when information on barrier passability or permeability is not known. Collectively, our toolbox will help generate more realistic estimates of river fragmentation and help inform more efficient restoration of river connectivity.

Keywords: dams, barriers, connectivity, duplicate, bootstrap, georeferenced river stressors

1. Introduction: The challenges of addressing river fragmentation

Improving river connectivity is a pre-requisite for more efficient river restoration (Palmer and Ruhi, 2019) and addressing the causes and consequences of river fragmentation is a key target of national and international policies, such as the Water Framework Directive (Water Framework Directive 2000/60/EC) and the Habitats Directives of the European Union (Council Directive 92/43/EEC). However, there is a striking paucity of guidance on how to assess river fragmentation (Kristensen et al., 2018). In comparison to terrestrial ecosystems, the fragmentation of rivers has received less attention, and as consequence many of the metrics of river fragmentation represent adaptations of methods used to assess habitat patchiness that have met with varying success (Fuller et al., 2015).

Assessing river fragmentation, however, poses particular problems. Compared to terrestrial ecosystems, rivers are more dependent on local scale processes due to the dendritic configuration of river networks, the unidirectional nature of river flow, and the strong seasonality of the hydrological regime (Ver Hoef et al., 2006). Another limitation is the difficulty of defining and identifying barriers, and the pervasive influence of sampling bias (Garcia de Leaniz et al., 2019; Jones et al., 2019) which underestimates the abundance of small barriers.

Whilst recent research has started to address these conceptual and methodological gaps (Grill et al., 2019; Schmitt et al., 2018), the data necessary for estimating river fragmentation – the number, location and characteristics of barriers (type, height) – are often incomplete (Januchowski-Hartley et al., 2019). Global estimates of river fragmentation are based on the location of large dams (typically, above 15-30 m in height) detected through remote sensing

(Lehner et al., 2011; Mulligan et al., 2020; World Commision on Dams, 2000; Zhang and Urpelainen, Johannes Schlenker, 2018) but these constitute a very small proportion of all barriers (Garcia de Leaniz et al., 2019). Even national barrier inventories typically miss 60-90% of instream barriers (Jones et al., 2019; Sun et al., 2020) and cannot be used for accurate estimation of river fragmentation.

Another limitation arises from the scarcity of harmonised barrier records that prevent global comparisons. For example, although high quality data on smaller dams exists at various local scales, the databases are incompatible – for example local and national databases across Europe use >50 different definitions of river barriers (Entec UK Ltd, 2010; "LST Biotopkartering vandringshinder," 2020; ONEMA, 2010; SEPA, 2018).

More generally, the assessment of river fragmentation is challenged by three main sources of errors and biases: (1) biases caused by problems of semantics, definitions and metrics; (2) errors caused by poor quality of barrier data, and (3) errors caused by sampling bias and poor representativeness of existing barrier data (Table 1). To overcome these problems it is necessary to integrate and compare existing barrier databases, and estimate fragmentation in a way that accounts for the limitations of existing barrier data. This will enable improved assessment of barrier impacts and a more effective restoration of river connectivity. To this end, we developed a suite of novel tools, and adapted existing ones, to estimate river fragmentation from barrier records across multiple scales using free open-source software QGIS, R and a free mobile phone app. We illustrate the application of our approach with the help of case studies, from local catchment scales (tens of km) to national and continental scales. It is hoped that addressing these difficulties will help to unlock new data sources, previously deemed unsuitable, and result in more realistic estimates of river fragmentation.

2. Methods: Description of Tools

We illustrate below the development of eight tools that address the challenges outlined in Table 1; these require different sources of data and generate various outputs (Table 2) to enable users to classify barriers into coherent, homogenous functional groups (T1); exclude duplicate barrier records (T2); correct for sampling bias and under-reporting via ground-truthed field surveys (T3), citizen science (T4) and machine learning (T5), generate various metrics of connectivity (T6, T7), and estimate barrier impacts across large scales (T8).

Tool 1. A key for defining and classifying barrier types

The first challenge for addressing river fragmentation is the lack of an agreed definition of instream barrier in relation to river connectivity (Wohl, 2017), and the many meanings and types of barriers (Belletti et al., 2018; Garcia de Leaniz et al., 2018; Jones et al., 2019). We therefore define a barrier as "any built structure that interrupts or modifies the flow of water, the transport of sediments, or the movement of organisms and can result in longitudinal discontinuity" (Belletti et al., 2020). We exclude from our definition natural features (such as waterfalls) or artificial processes (such as thermal or pollution barriers) that are not built but that can also result in longitudinal discontinuity. We also exclude structures that only result in vertical or lateral discontinuity, as these are difficult to identify and are rarely surveyed in barrier inventories (Atkinson et al., 2020; Buddendorf et al., 2019; Jones et al., 2019). Unlike many barrier inventories that only consider large dams (Grill et al., 2019) or consider the needs of few target species (Atkinson et al., 2020; Buddendorf et al., 2019), we place no restriction on the type or height of barriers, or on the taxa or process impacted by barriers.

We propose a functional classification of barriers (Figure 1) that considers six broad barrier types that impact on longitudinal connectivity (Figure 2). Structures that do not fit these criteria such as groynes or debris screens are labelled “Other”.

Tool 2. Barrier Duplicate Excluder

Barrier duplicates occur when barrier databases are merged or expanded. These can be ubiquitous (Jones et al, 2019; Belletti et al 2020) and overestimate the extent of fragmentation. To identify and exclude duplicates, a decision tree and consistent protocol must be followed, consisting of various steps, as shown in Figure 3. Barriers are considered to be true records if they meet stringent criteria and represent unique entries as defined by visual checks against satellite imagery. An exclusion radius is then calculated that maximizes the exclusion of false positives (i.e. duplicates) and retains most true records.

First, to identify potential duplicates, pairwise distances between barrier records within 1,000 m of one another (Euclidean distance), regardless of source data, are selected (Jones et al., 2019). This search is performed for each record in the database before proceeding to the next, and is constrained to the same watershed, as it is assumed that barriers that lie in different watershed are not duplicates. Records located within 1,000 m are excluded from further searches as it is assumed that a barrier cannot be the duplicate of more than one point. Secondly, to assess the true status of potential duplicates (duplicate or true record), two or more independent raters assess a sub-sample as well as a common sample of potential duplicates using GIS and satellite imagery. The distance between potential duplicates is used to define the exclusion radius that excludes the 80th percentile of the cumulative distribution, when agreement between raters is >75% on the shared data sample according to Fleiss' kappa (Fleiss, 2003). If the comparison between raters results in <75% agreement, the exclusion distance is increased by the standard error of the distance to the duplicated points, and the visual assessment criteria is repeated until >75% agreement is achieved. In those cases when no true records are found within the 80% exclusion radius, the exclusion radius can be increased to remove a greater number of duplicates. This can be done using the greatest difference between the cumulative distribution of the duplicate records from the cumulative distribution of true records. The exclusion radius is then increased if there is a positive difference between the two cumulative distributions, i.e. if additional duplicates can be removed without removing any true records.

Confidence intervals on the exclusion radii are calculated by bootstrapping (Chao et al., 2013; Efron and Tibshirani, 1994). Bootstrapping is also employed to assess the influence of the number of duplicates that need to be checked visually on the precision of the radius that excludes 80% of potential duplicates, and hence estimate the sample size required to derive robust exclusion values. For this, samples are randomly resampled with replacement (e.g. 10,000 times each). To demonstrate the use of the barrier duplicate excluder tool, we examined records from five barrier databases for France: 62,340 barriers from the French National Agency for Water and Aquatic Environments (http://www.onema.fr/le-roe), 260 barriers from the GlObal geOreferenced Database of Dams (GOODD; Mulligan et al., 2020), 111 barriers from Global Reservoir and Dam Database (GRanD; Lehner et al., 2011), 586 barriers from the European Catchment River Information System (Ecrins; European

Environment Agency, 2012) and 128 barriers from OSM (OpenStreetMap contributors, 2019). We selected all potential duplicates identified by the pairwise distance comparion. Identification of duplicates and bootstrapping were conducted using R 3.6 (R Core Team, 2018), and QGIS 3.10 (QGIS Development Team, 2018) was used for visual assessment using satellite imagery from Google and Bing.

Tool 3. Bootstrapped Field Corrector. Ground truthing barrier records to correct for sampling bias

The existence of sampling bias means that some barriers are more likely to be detected and included in barrier inventories than others. To correct for such bias, and collect new data in cases where no barrier data are available, we developed a simple surveying and monitoring protocol. Previous studies have indicated that surveyed reaches should represent at least 0.1% of the total river network for the area of interest (Jones et al., 2019) using a river network consistent across all areas being studied, e.g. Ecrins (European Environment Agency, 2012) or HydroSHEDS (Lehner et al., 2008). To reduce sampling bias, contiguous 20 km reaches should be surveyed in each test river to avoid missing the more inaccessible barriers, typically located in the headwaters. Surveying should be conducted at low flow conditions (~Q80-Q95) to avoid missing low head barriers that may disappear at high flows. These reaches should be representative of the area being surveyed in terms of altitude, slope, stream order (Strahler, 1957), and land-use cover.

To assess the influence of distance surveyed on barrier discovery rate, the bootstrap approach described in Jones et al. (2019) can be used. Briefly, this involves resampling survey barrier data every km to increase resolution. This also accommodates survey reaches that are shorter than the target 20 km reach length, making the method suitable for a range of spatial scales, from local catchments to broader regional assessments. The ground-truthed bootstrapped density of new barriers/km can then be used to correct barrier density estimates derived from existing inventories and confidence intervals can be calculated.

Tool 4. Barrier Tracker. Using citizen Science to fill data gaps

Citizen science is increasingly being used in conservation and environmental science to fill data gaps and complement existing information (Merenlender et al., 2016). Engaging with citizens has also become a priority for many funding bodies (Warin and Delaney, 2020). Surveying rivers and recording barriers is very time consuming but can be greatly aided with the help of volunteers. For this reason, a smart phone app - the Barrier Tracker (Figure 4), was developed as part of the AMBER project (www.amber.international). The Barrier Tracker harnesses the power of citizen science to provide a more complete picture of barrier abundance. It enables users to locate all types of barriers (classified into 6 main types), take a photograph, and assess their main features, including height, current use and conservation status. The latter information is essential for identifying obsolete barriers and prioritize efforts for mitigation or removal. The information is uploaded it into the cloud where it can used to build a better picture of stream fragmentation. The app has two levels (entry and advanced), depending on skills, that enable users to enter additional information on barrier characteristics. Users can record barriers without any need to register, as well as in remote places without telephone signal (typical of many heavily wooded headwaters). Users can download their data own and barrier records contributed by others. The app is free to use and can be downloaded for Android (Google Play) and IoS (Apple Store), as well as from AMBER website (https://portal.amber.international/). It is currently available in 12 languages:

English, Danish, Dutch, French, German, Italian, Polish, Portuguese, Slovenian, Spanish, Ukrainian, Welsh.

Tool 5. Barrier Modeler. Modelling missing barriers

Modelling through machine learning (ML) can be used to correct for sampling bias and under-reporting of small barriers if high quality datasets are available for use as training datasets. This can complement ground-truthed field estimates of barrier density (Tool 6) and identify where the data gaps may lie. Predictors of barrier abundance and attributes that have been found to be useful include various anthropic and environmental predictors that are typically associated with river discontinuities. These include the type of land-cover, population density, elevation, slope, dendricity, drainage density, road crossings, and average flow, amongst others (Belletti et al., 2020; Januchowski-Hartley et al., 2019). For modelling, Boosted Regression Trees (Januchowski-Hartley et al. 2019) or nonparametric Random Forest Regressor have been used (Belletti et al., 2020).

Tool 6. Barrier Free Length (BFL), a taxon-free metric of river connectivity

Two common constraints of existing methods for assessing river connectivity (Table 3) are that they are often (1) taxa specific, requiring information on the “passability” for each barrier and target species, and (2) that barriers must be “snapped” or aligned to the correct river reach, or the resulting metric of connectivity would be in error. While it is possible to meet these constraints at small spatial scales, or when dealing with a limited number of barriers (Cote et al., 2009; Grill et al., 2019, 2015, 2014), this is impractical when a large number of barriers are being considered or - as is often the case, passability is unknown for most species. Two alternatives to the above shortcomings are the calculation of (1) barrier-free length

(Grizzetti et al., 2017; Jones et al., 2019) and (2) barrier density at the sub-catchment or catchment scale (Belletti et al., 2020; Jones et al., 2019). Barrier free length (BFL) can be calculated as the average length between consecutive barriers (Jones et al., 2019). If information is missing for some barrier types, BFL can be calculated using the dominant barrier types only.

Tool 7. Barrier Density (BD), a metric of connectivity that does not require reach alignment

Unlike most other metrics of river connectivity that require accurate knowledge of barrier coordinates, and precise alignment of barrier position to the correct river reach (Grill et al., 2019), barrier density is largely invariant to barrier location, or to the resolution of the underlying river network. All that is required is that barriers are assigned to the correct catchment or basin. Barrier density can be expressed as barrier per unit of stream length (No/km) or per catchment area (No/km2). The former is preferred as rivers can differ widely in dendricity (i.e. river length/No. river segments) and drainage density (i.e. river length/catchment area) that would introduce spurious errors if this variation is ignored (Kristensen and Globevnik, 2014; Vogt et al., 2008). One useful feature of barrier density is that it can be used to predict barrier free length as BFL decreases with the power function 1/density if barriers are assumed to be distributed at random, which is a good starting approximation (Figure 5).

Tool 8. Barrier Impact Modeler

Barrier impacts on fish passage are traditionally estimated using protocols that take into account the topographical and hydraulic features of a barrier and the swimming performance of one or more target species. This information is used to derive estimates of "passability"

(Cote et al. 2009), that can range from 0 (passage impossible) to 1 (unimpeded passage). Examples of such protocols include SNIFFER (SNIFFER, 2010), ICE (Baudoin et al., 2015), and ICF (Solà et al., 2011) in Europe. However, all these approaches require detailed knowledge of every barrier in a catchment to assess fragmentation, which is labour intensive and not always possible. They also rely on having information on swimming performance to derive passabilities, but that is not available for all species (Furniss et al., 2006; Mahlum et al., 2014) and can vary a lot depending on the method being used. Finally, they do not consider barrier impacts on sediment transport or other river processes. As barrier impacts on animal movement (Baudoin et al., 2015; Bourne et al., 2011; Kemp and O'Hanley, 2010) and sediment and water storage (Ramos-Diez et al., 2016; Stephens, 2010) are typically a function of barrier height, we propose a simple impact function that considers that barrier impacts increase with barrier height according to a sigmoid function. The function consists of three distinct phases: an initial exponential phase, an approximately linear phase with an inflexion point where the slope is greatest, and an asymptotic phase (Figure 6). We have set the asymptotic maximum impact (1.0) to barriers of 4.0 m in height, which is close to the maximum head drop that can be overcome by Atlantic salmon (3.7 m, Mills, 1989), the strongest fish swimmer in Europe (Baudoin et al., 2015). This asymptotic value is also consistent with the observation that small dams (<5 m) cause only limited sediment retention (MacBroom, 2005). We therefore consider that any barrier greater than 4 m causes essentially the same (maximum) impact. However, other slopes and asymptotes can be used and a sensitivity analysis carried out to see how this changes the estimates of river fragmentation. The parameterisation of various sigmoidal curves (e.g. logistic, Gompertz, log-logistic, Weibull) can be achieved with the *aomisc* R package (Onofri, 2020).

To model barrier height, various approaches are possible. If, as is often the case, barrier type is known, the easiest approach is to assign to each barrier with missing height the mean (or mode, or median) height for that barrier type. Another potential approach to model barrier height is to use machine learning and take into account not just barrier type, but also their location in the stream, as well as some height covariates. For example, Januchowski-Hartley et al. (2019) used Boosted Regression Trees to model the height of 20,077 barriers with missing heights in France. The most important predictors of barrier height were stream reach length, elevation, gradient, average flow, and agricultural land cover, contributing 68% of the variance explained. A third approach that might be useful is based on the fact that barriers with missing heights are more likely to be small than large due to the nature of sampling bias (Januchowski-Hartley et al. 2019), and that complete barrier inventories are only available for structures (typically dams) beyond a certain height. The distribution of barrier heights is therefore left truncated, or under-reported for small heights. It might be possible to reconstruct the abundance and height of missing barriers from knowledge of the abundance and height of existing ones using the Cullen and Frey approach in the *fitdistrplus* R package (Delignette-Muller et al., 2020).

3. Results & Discussion

We have identified some of the main challenges for quantifying stream fragmentation and developed some tools (T) that can be used for filling data gaps, quantifying uncertainty, and reduce bias in river connectivity estimates. These are illustrated with real case studies and examples below.

Harmonising barrier type definitions (T1)

The lack of an agreed definition of barrier has made it difficult to compare studies that consider different barrier heights, or barrier types. Our barrier identification key does not impose any restriction on barrier height and was able to classify more than 290 different types of longitudinal instream barriers present in Europe into six functional types (Belletti et al., 2020). These differ in use, size, and location within the watershed. They are also easy to recognise and can be used as part of a citizen science programme. Our results suggest that barrier types differ significantly in height, and thus likely also in impact, and that this is a field that is known in most databases. For example, in Spain barrier type is known for 93% of barriers. The application of the barrier classification key to a large database of barrier heights in Spain (N = 18,935) indicates that barrier types differ considerably in height ($F_{7,18927}$= 511.4, P<0.001) and that it is hence possible to estimate missing heights if barrier type is known, albeit with some uncertainty (R^2adj. = 0.156).

Data cleaning and exclusion of duplicate records (T2)

Data cleaning typically takes 30-80% of the time in big data applications before it can be used (Wang and Wang, 2019). Barrier data management is no exception. Excluding barrier duplicates is an essential aspect of data cleaning, and as an example we applied the barrier exclusion tool to the analysis of five barrier databases with information for France. This identified 1,497 duplicated barriers using a radius of 302 m that excluded 80% of duplicates and only 14 true records. The analysis of 190 potential duplicates by three independent observers indicated that 136 barriers (72%) were duplicates and 54 barriers (28%) were true records. Agreement between observers on a common sample of 50 barriers was high (Fleiss's kappa = 0.88, 95CI: 0.83-0.92, P < 0.001), indicating that the procedure we developed was robust and the decision tree used to identify duplicates repeatable. This yielded 61,960 unique barriers in France (Belletti et al., 2020), which is ~19% fewer than the 76,292 structures listed in the French National Agency for Water and Aquatic Environments which includes records listed as destroyed, planned, under construction, invalid, or duplicated (http://www.onema.fr/le-roe).

The spatial distribution of duplicate and true barrier records can be characterised by plotting the distances to nearest barrier against the cumulative proportion of duplicates (Figure 7A) and true records (Figure 7B). In the case of the French example, this follows a rectangular hyperbola that can be described by the Michaelis-Menten two-parameter equation:

$$Y = \frac{aX}{(b + X)}$$

The proportion of duplicates that are excluded (Y) increases rapidly as the exclusion radius (X) increases and reaches a plateau given by the asymptote a. The second parameter (b) represents the radius that excludes 50% of the duplicate records. These can be easily parametrised with the *nlstools* R package (Baty et al., 2015) and can be used to derive exclusion radii for different geographical areas.

Bootstrapping was used to estimate the sample size required to obtain precise estimates of exclusion radii. In this example, the optimal duplicate exclusion radius is reached after detecting ~100 duplicates, although the variance does not stabilize until ~120 duplicates have been classified (Figure 8).

Detecting and correcting for sampling bias and barrier under-reporting (T3-T5)

Sampling bias occurs when more accessible or more conspicuous items are more likely to be detected (Araujo and Guisan, 2006). It is a pervasive problems in species distribution modelling (Araujo and Guisan, 2006), in conservation (Reddy and Dávalos, 2003), and also in barrier inventories, where small barriers tend to be under-reported (Belletti et al., 2020; Jones et al., 2019). To detect and account for barrier sampling bias, systematic field surveys can be used to ground-truth barrier records. For example, standardised barrier surveys conducted across 2,715 km of rivers in 26 European countries (Belletti et al., 2020) indicate that barrier densities range more than 10-fold (Table 4) and reveal widespread under-reporting (Belletti et al., 2020).

As with the estimation of optimal exclusion radius, bootstrapping can also be used to estimate the minimum length of river that needs to be sampled to derive precise estimates of barrier density (Figure 9). For example, in Great Britain the barrier discovery rate (i.e. those barriers not recorded in existing databases) reaches an asymptote after 68 km of river length, but at least 200–250 km need to be surveyed to obtain a precise estimate of barrier density (Jones et al., 2019). Across Europe, the sample mean tends to overestimate the bootstrapped 50% barrier discovery rate (mean difference: 0.027 barriers/km , SE = 0.01), while the sample median underestimates it (mean difference= -0.11 barriers/km, SE = 0.04; Table 4, Figure 9). The bootstrapped median barrier discovery rate tends to stabilise after ~100 km of surveying in most countries, and 100 km can therefore be taken as an adequate survey stream length for barrier density estimates in most cases.

Results from standardised barrier surveys carried by trained personnel tend to produce the most accurate barrier density estimates (Atkinson et al., 2020; Buddendorf et al., 2019; Jones et al., 2019; Sun et al., 2020), but field work is expensive and time consuming. Barriers can also be located by citizens using a smartphone app, and this can help to fill data gaps on an ad-hoc basis. For example, using the Barrier Tracker app, 5,530 barriers have been uploaded into the AMBER barrier database by +2,000 users from 36 countries (Figure 10). Duplicates and sampling bias can be an issue with records provided by citizens, but the duplicate exclusion tool described above could be used to automate the filtering out of duplicates before they are added to the database. Barrier locations generated by an app are collected in a uniform way and are immediately publicly available. Also, because photographic records of barriers are stored these can be reanalysed at a later stage, and additional data can be obtained using image recognition. As barrier records are date stamped and can be generated by multiple users, these can be used to monitor changes in barrier use, and be used to initiate the decommissioning of abandoned structures.

Modelling through machine learning can also be used to estimate the number, and possibly also the characteristics and location, of missing barriers, although this depends on having good training datasets and suitable predictors (Januchowski-Hartley et al., 2019, 2013). Modelling cannot be a substitute for field work, but it can complement it and direct sampling efforts where the biggest data gaps lie (Belletti et al., 2020).

Estimating river fragmentation when accurate barrier-to-reach alignment is not possible (T6), or when barrier passability is unknown (T7)

Most river connectivity metrics are overly restrictive for use at large spatial scales because they require that both barrier location and stream reach are well aligned, and measured without error (Grill et al., 2019). Barrier density is largely invariant to these constraints and can also be computed for basin area. This is useful when there are errors in barrier coordinates, but also when the underlying river networks are not detailed enough, or differ in resolution and quality, as is often the case across large spatial scales (Kristensen and Globevnik, 2014; Vogt et al., 2008). Our results for Great Britain (Jones et al., 2019) indicate that BFL decreases approximately as a power function of 1/density, as would be predicted if barriers were distributed at random through the catchments (Figure 11). This has two practical implications. First, it shows that BFL can be predicted from linear barrier density, which makes comparisons of river connectivity across catchments possible. Secondly, and more importantly, it shows that connectivity drops rapidly with the first few barriers as a 50% reduction in BFL is predicted to occur with just 5% of barriers. From a barrier management perspective, this means that removing or mitigating a few barriers in heavily fragmented rivers (density > 1.0 barriers/km) is not effective as no significant increase in connectivity is likely to occur until density is decreased to ~0.25 barriers/km (Figure 11). This assumes that barriers are removed or mitigated at random, which will seldom be the case, but it provides a useful baseline against which the connectivity gain predicted by more targeted prioritization approaches (Erős et al., 2018; Kemp and O'Hanley, 2010; King et al., 2017; King and O'Hanley, 2016; O'Hanley, 2015; O'Hanley et al., 2011; Roni et al., 2002) can be gauged.

Estimating barrier impacts (T8)

Barrier height is missing from many barrier records (Atkinson et al., 2020; Januchowski-Hartley et al., 2019; Jones et al., 2019), or has often been estimated via remote sensing which can be inaccurate (Jones et al., 2019). This precludes an analysis of barrier impacts and a more in depth estimate of river fragmentation. However, barrier height can be estimated via modelling through machine learning (Januchowski-Hartley et al., 2019) or, more easily (but only approximately) from knowledge of barrier type. Analysis of a large barrier database from Spain (N= 30,677) indicates that barrier height was missing from 61% of records, but barrier type was known in 88% of cases. This is typical of many barrier databases. As shown in a previous study (Jones et al. 2019), barrier types differed considerably in height ($F_{7,18927}$= 511.4, P<0.001) which makes it possible to estimate missing heights if barrier type is known, albeit with some uncertainty (R^2adj. = 0.156).

Some measure of impact or passability for each barrier is also required for the assessment of longitudinal river connectivity at the catchment scale (Cote et al., 2009), but this information is difficult to obtain for most species and barriers, particularly when barrier height is unknown. We illustrate how the application of the tools outlined above can be used to overcome these difficulties and estimate barrier impacts using the Spanish case study. This consists of six steps (Figure 12): (1) we first harmonised all the barrier types into six coherent functional groups using Tool 1; (2) we then used Tool 2 to clean the data and exclude duplicates when different databases were merged. (3) field surveys were then used to ground-truth the existing barrier inventory (Tool 3) and derive correction factors to estimate the abundance of different barriers while accounting for under-reporting (Figure 12A); (4) the mean height of each barrier (Figure 12B) was used to fill data gaps in those cases where height was missing; (5) we then used the height-impact function (Tool 6), to calculate the barrier-specific (per capita) impact for each barrier type (Figure 12C); (6) finally, we combined information on the corrected abundance of each barrier type with their per capita impact to estimate net barrier impacts (Figure 12D).

The results show that, by far, the biggest impact on connectivity in Spain is caused by a large number of relatively small weirs (mean height = 1.91 m), and that the impact of large dams is essentially the same as that of much smaller (but much more abundant) ramps built to control channel scouring and bank erosion. Although we illustrate this approach with an example at a relatively large (country) scale, the same approach can be used at the catchment scale.

Concluding remarks

Reconnecting rivers is a priority in river restoration (European Union, 2020; Grizzetti et al., 2017; Hering et al., 2010) but efforts have been hampered by the paucity of good barrier records (Kristensen et al., 2018). Other than at small spatial scales, researches have typically been unable to fully assess river fragmentation because they cannot take advantage of a large number of disparate, duplicate and incomplete records. As a consequence, global estimates of river fragmentation have tended to be based on large dams only (Barbarossa et al., 2020; Grill et al., 2019, 2015, 2014; Zarfl et al., 2019), despite the fact that these are the least common type of barriers (Garcia de Leaniz, 2008; Garcia de Leaniz et al., 2019; Jones et al., 2019). This presents a mismatch between the needs of river managers, who require estimates of connectivity based on complete barrier inventories, and what science has so far been able to offer.

The data management and modelling tools presented here bridge this information gap. We show how to clean and ground truth existing barrier records and estimate barrier impacts more efficiently. For example, we derive robust rules to exclude duplicated barrier records using only a small sample of records (<1%). This means that large scale analysis of thousands, or even tens of thousands of barrier records, can remove 80% of duplicates and retain most genuine records with minimal effort. Likewise, the completeness of barrier records can be assessed and accounted for by a combination of targeted surveying, citizen science and modelling. Collectively, our toolbox will help generate more realistic estimates of river fragmentation and result in more efficient restoration of river connectivity.

Author contributions

J.J., B.B., L.B., G.S., W.v.B. and C.G.L. developed the tools and carried out the analysis. J.J. and C.G.L. wrote the MS with input from L.B. and B.B. All co-authors critically revised the edited manuscript.

Acknowledgements

This study was funded by the EC Horizon 2020 Research & Innovation Programme, AMBER - Adaptive Management of Barriers in European Rivers Project, grant agreement No. 689682 led by C.G.L. Map data copyrighted OpenStreetMap contributors and available from https://www.openstreetmap.org

Table 1. Main challenges for estimating river fragmentation, consequences, and potential solutions illustrated with the application of eight barrier tools developed in this work.

Problem or Challenge	Consequences	Reference examples	Potential Solutions	Illustrated Tool
Definitions				
1. Lack of an agreed definition of barrier	• Difficult to compare connectivity estimates that consider different types of barriers	Wohl, 2017	• Harmonise barrier definitions and barrier type classification	Tool 1 Barrier Classification Key
Data representativeness				
2. Existing barrier records may contain duplicates	• Overestimates fragmentation	Grill et al., 2019; Martin, 2019; Zhang and Urpelainen, Johannes Schlenker, 2018	• Develop algorithms that exclude duplicates	Tool 2 Barrier Duplicate Excluder
3. Barrier sampling bias & under-reporting of small barriers	• Underestimates fragmentation • Not all barrier types are equally represented in inventories • Difficult to compare data from different countries & catchments if they are collected using different criteria	Jones et al., 2019; Sun et al., 2020	• Ground-truth barrier records with field work • Use citizen science • Model barrier abundance	Tool 3 Bootstrapped Field Corrector Tool 4 Barrier Tracker Tool 5 Barrier Modeler
Data quality & connectivity metrics				
4. Errors in mapping barrier coordinates, barriers are not aligned with topologically consistent river networks	• Introduces errors in dendritic connectivity indices and other metrics of connectivity	Hoenke et al., 2014; Martin, 2019	• Use basin-based methods that do not depend on having exact barrier coordinates	Tool 6 Barrier Density Estimator
5. Restrictive nature of existing connectivity metrics	• May apply only to some taxa, typically large sport fish • Require data on 'passabillity' for each species and barrier • Difficult to generalise across contexts	Cooper et al., 2017; Díaz et al., 2019; Mantel et al., 2010; Van Looy et al., 2014	• Use taxon-free metrics	Tool 7 Barrier Free Length
6. Height and passability data are missing for many barriers	• Precludes calculation of barrier impacts and application of some dendritic connectivity indices	Jones et al., 2019; Rincón et al., 2017; Sun et al., 2020	• Model height • Model barrier impacts	Tool 8 Barrier Impact Modeler

Table 2. Data inputs and outputs generated by the application of the eight barrier tools described in the text.

Barrier Tool	Data input	Data output
T1. Barrier Classification Key	Barrier characteristics	Type of barrier
T2. Barrier Duplicate Excluder	Location of barriers Ground-truthed duplicate Ground-truthed unique records	Duplicate exclusion radius Number of duplicates Duplicate-free barrier abundance
T3. Bootstrapped Field Corrector	Number of barriers Location of barriers Ground-truthed field data River length	Barrier discovery rate Corrected barrier density
T4. Barrier Tracker	Barrier photograph Barrier characteristics	Barrier location Barrier abundance
T5. Barrier Modeler	Number of barriers Location of barriers Basin covariates	Predicted barrier abundance
T6. Barrier Density Estimator	Number of barriers Location of barriers River length or basin area	Connectivity metric
T7. Barrier Free Length (BFL)	Number of barriers Location of barriers River length	Connectivity metric
T8. Barrier Impact Modeler	Height-Impact function Barrier height Barrier abundance	Per capita barrier impact Net barrier impact Relative impact share

Table 3. Some common connectivity metrics used to assess the extent of river fragmentation.

Metric	Barrier type	Constraints	Spatial Scale	Reference
Conservation Connectivity Index (CCI)	Dam	Barriers must be snapped to the correct river reach	Reach	Rodeles et al., 2020
Local, upstream, and downstream dam density	Dam and weir	Barriers must be snapped to the correct river reach	Reach and sub-catchment	Cooper et al., 2017; Mantel et al., 2010; Van Looy et al., 2014
Fragmentation Index (FI)	Dam	Barriers must be snapped to the correct river reach	Reach	Díaz et al., 2019
Dendritic Connectivity Index (DCI)	Any	Barriers must be snapped to the correct river reach Passability data needed	Catchment	Cote et al., 2009
Free Flowing River (FFR)	Dam	Barriers must be snapped to the correct river reach	Reach	Grill et al., 2019
Barrier Free Length (BFL)	Any	Barriers must be snapped to the correct river reach	Reach	Jones et al., 2019
Barrier density	Any	Barriers must be assigned to correct catchment	Sub-catchment and catchment	Jones et al., 2019

References

Araujo, M.B., Guisan, A., 2006. Five (or so) challenges for species distribution modelling. J. Biogeogr. 33, 1677–1688.

Atkinson, S., Bruen, M., O' Sullivan, J.J., Turner, J.N., Ball, B., Carlsson, J., Bullock, C., Casserly, C.M., Kelly-Quinn, M., 2020. An inspection-based assessment of obstacles to salmon, trout, eel and lamprey migration and river channel connectivity in Ireland. Sci. Total Environ. 719, 137215. https://doi.org/https://doi.org/10.1016/j.scitotenv.2020.137215

Barbarossa, V., Schmitt, R.J.P., Huijbregts, M.A.J., Zarfl, C., King, H., Schipper, A.M., 2020. Impacts of current and future large dams on the geographic range connectivity of freshwater fish worldwide. Proc. Natl. Acad. Sci. 117, 3648–3655.

Baty, F., Ritz, C., Charles, S., Brutsche, M., Flandrois, J.-P., Delignette-Muller, M.-L., 2015. A toolbox for nonlinear regression in R: the package nlstools. J. Stat. Softw. 66, 1–21.

Baudoin, J.-M., Burgun, V., Chanseau, M., Larinier, M., Ovidio, M., Sremski, W., Steinbach, P., Voegtle, B., 2015. Assessing the passage of obstacles by fish. Concepts, design and application.

Belletti, B., Bizzi, S., Castelletti, A., Garcia De Leaniz, C., Börger, L., Jones, J., Olivo del Amo, R., Segura, G., Tummers, J.S., van der Bund, W., AMBER, 2018. Small isn't beautiful: the impact of small barriers on longitudinal connectivity of European Rivers, in: European Geosciences Union.

Belletti, B., Garcia de Leaniz, C., Jones, J., Bizzi, S., Börger, L., Segura, G., Castelletti, A., Bund, W. Van de, Aarestrup, K., Barry, J., Belka, K., Berkhuysen, A., Birnie-Gauvin, K., Bussettini, M., Carolli, M., Consuegra, S., Dopico, E., Feierfeil, T., Fernández, S., Garrido, P.F., Garcia-Vazquez, E., Garrido, S., Giannico, G., Gough, P., Jepsen, N., Jones, P.E., Kemp, P., Kerr, J., King, J., Łapińska, M., Lázaro, G., Lucas, M.C., Marcello, L., Martin, P., McGinnity, P., O'Hanley, J., Amo, R.O. del, Parasiewicz, P., Rincon, G., Rodriguez, C., Royte, J., Schneider, C.T., Tummers, J.S., Vallesi, S., Vowles, A.S., Verspoor, E., Wanningen, H., Wantzen, K.M., Wildman, L., Zalewski, M., 2020. Broken rivers: ground-truthing the world's most fragmented rivers. Authorea. https://doi.org/10.22541/au.159355955.53596231

Bourne, C.M., Kehler, D.G., Wiersma, Y.F., Cote, D., 2011. Barriers to fish passage and barriers to fish passage assessments: the impact of assessment methods and assumptions on barrier identification and quantification of watershed connectivity. Aquat. Ecol. 45, 389–403. https://doi.org/10.1007/s10452-011-9362-z

Buddendorf, W.B., Jackson, F.L., Malcolm, I.A., Millidine, K.J., Geris, J., Wilkinson, M.E., Soulsby, C., 2019. Integration of juvenile habitat quality and river connectivity models to understand and prioritise the management of barriers for Atlantic salmon populations across spatial scales. Sci. Total Environ. 655, 557–566. https://doi.org/https://doi.org/10.1016/j.scitotenv.2018.11.263

Chao, A., Wang, Y.T., Jost, L., 2013. Entropy and the species accumulation curve: a novel entropy estimator via discovery rates of new species. Methods Ecol. Evol. 4, 1091–1100.

Cooper, A.R., Infante, D.M., Daniel, W.M., Wehrly, K.E., Wang, L., Brenden, T.O., 2017. Assessment of dam effects on streams and fish assemblages of the conterminous USA. Sci. Total Environ. 586, 879–889. https://doi.org/10.1016/J.SCITOTENV.2017.02.067

Cote, D., Kehler, D.G., Bourne, C., Wiersma, Y.F., 2009. A new measure of longitudinal connectivity for stream networks. Landsc. Ecol. 24, 101–113. https://doi.org/10.1007/s10980-008-9283-y

Delignette-Muller, M.-L., Dutan, C., Pouillot, R., Denis, J.-B., Siberchicot, A., 2020. fitdistrplus: Help to Fit of a Parametric Distribution to Non-Censored or Censored Data.

Díaz, G., Arriagada, P., Górski, K., Link, O., Karelovic, B., Gonzalez, J., Habit, E., 2019. Fragmentation of Chilean Andean rivers: expected effects of hydropower development. Rev. Chil. Hist. Nat. 92, 1. https://doi.org/10.1186/s40693-019-0081-5

Efron, B., Tibshirani, R.J., 1994. An introduction to the bootstrap. CRC press.

Entec UK Ltd, 2010. Mapping Hydropower Opportunities and Sensitivities in England and Wales. Bristol.

Erős, T., O'Hanley, J.R., Czeglédi, I., 2018. A unified model for optimizing riverscape conservation. J. Appl. Ecol. 55, 1871–1883.

European Environment Agency, 2012. EEA Catchments and Rivers Network System ECRINS v1.1. Copenhagen.

European Union, 2020. Bringing nature back into our lives. EU 2030 Biodiversity Strategy. Brussels.

Fuller, M.R., Doyle, M.W., Strayer, D.L., 2015. Causes and consequences of habitat fragmentation in river networks. Ann. N. Y. Acad. Sci. 1355, 31–51.

Furniss, M., Love, M., Firor, S., Moynan, K., Llanos, A., Guntle, J., Gubernick, R., 2006. FishXing, version 3.0. US For. Sevice, San Dimas Technol. Dev. Center, San Dimas, Calif.

Garcia de Leaniz, C., 2008. Weir removal in salmonid streams: Implications, challenges and practicalities. Hydrobiologia 609, 83–96. https://doi.org/10.1007/s10750-008-9397-x

Garcia de Leaniz, C., Belletti, B., Bizzi, S., Segura, G., Borger, L., Jones, J., Olivo del Amo, R., Wanningen, H., Tummers, J., Kerr, J., Kemp, P., van de Bund, W., the AMBER consortium, 2018. The importance of having a good database for restoring river connectivity: the AMBER Barrier Atlas in Europe, in: Brink, K., Gough, P., Royte, J., Schollema, P.P., Wanningen, H. (Eds.), From Sea to Source 2.0. Protection and Restoration of Fish Migration in Rivers Worldwide. World Fish Migration Foundation, Groningen, pp. 142–145.

Garcia de Leaniz, C., Berkhuysen, A., Belletti, B., 2019. Beware small dams, they can do damage, too. Nature 570. https://doi.org/10.1038/d41586-019-01826-y

Grill, G., Dallaire, C.O., Chouinard, E.F., Sindorf, N., Lehner, B., 2014. Development of new indicators to evaluate river fragmentation and flow regulation at large scales: A case study for the Mekong River Basin. Ecol. Indic. 45, 148–159.

Grill, G., Lehner, B., Lumsdon, A.E., MacDonald, G.K., Zarfl, C., Liermann, C.R., 2015. An index-based framework for assessing patterns and trends in river fragmentation and flow regulation by global dams at multiple scales. Environ. Res. Lett. 10, 15001.

Grill, G., Lehner, B., Thieme, M., Geenen, B., Tickner, D., Antonelli, F., Babu, S., Borrelli, P., Cheng, L., Crochetiere, H., Ehalt Macedo, H., Filgueiras, R., Goichot, M., Higgins, J., Hogan, Z., Lip, B., McClain, M.E., Meng, J., Mulligan, M., Nilsson, C., Olden, J.D., Opperman, J.J., Petry, P., Reidy Liermann, C., Sáenz, L., Salinas-Rodríguez, S., Schelle, P., Schmitt, R.J.P., Snider, J., Tan, F., Tockner, K., Valdujo, P.H., van Soesbergen, A., Zarfl, C., 2019. Mapping the world's free-flowing rivers. Nature 569, 215–221. https://doi.org/10.1038/s41586-019-1111-9

Grizzetti, B., Pistocchi, A., Liquete, C., Udias, A., Bouraoui, F., van de Bund, W., 2017. Human pressures and ecological status of European rivers. Sci. Rep. 7, 205. https://doi.org/10.1038/s41598-017-00324-3

Hering, D., Borja, A., Carstensen, J., Carvalho, L., Elliott, M., Feld, C.K., Heiskanen, A.-S., Johnson, R.K., Moe, J., Pont, D., 2010. The European Water Framework Directive at the age of 10: a critical review of the achievements with recommendations for the future.

Sci. Total Environ. 408, 4007–4019.

Hoenke, K.M., Kumar, M., Batt, L., 2014. A GIS based approach for prioritizing dams for potential removal. Ecol. Eng. 64, 27–36. https://doi.org/https://doi.org/10.1016/j.ecoleng.2013.12.009

Januchowski-Hartley, S.R., Jézéquel, C., Tedesco, P.A., 2019. Modelling built infrastructure heights to evaluate common assumptions in aquatic conservation. J. Environ. Manage. 232, 131–137. https://doi.org/10.1016/J.JENVMAN.2018.11.040

Januchowski-Hartley, S.R., McIntyre, P.B., Diebel, M., Doran, P.J., Infante, D.M., Joseph, C., Allan, J.D., 2013. Restoring aquatic ecosystem connectivity requires expanding inventories of both dams and road crossings. Front. Ecol. Environ. 11, 211–217.

Jones, J., Börger, L., Tummers, J., Jones, P., Lucas, M., Kerr, J., Kemp, P., Bizzi, S., Consuegra, S., Marcello, L., Vowles, A., Belletti, B., Verspoor, E., Van de Bund, W., Gough, P., Garcia de Leaniz, C., 2019. A comprehensive assessment of stream fragmentation in Great Britain. Sci. Total Environ. https://doi.org/10.1016/J.SCITOTENV.2019.04.125

Kemp, P.S., O’Hanley, J.R., 2010. Procedures for evaluating and prioritising the removal of fish passage barriers: a synthesis. Fish. Manag. Ecol. 17, 297–322.

King, S., O’Hanley, J.R., 2016. Optimal fish passage barrier removal—revisited. River Res. Appl. 32, 418–428.

King, S., O’Hanley, J.R., Newbold, L.R., Kemp, P.S., Diebel, M.W., 2017. A toolkit for optimizing fish passage barrier mitigation actions. J. Appl. Ecol. 54, 599–611.

Kristensen, P., Globevnik, L., 2014. European small water bodies. Biol. Environ. Proc. R. Irish Acad. 114B, 281–287. https://doi.org/10.3318/bioe.2014.13

Kristensen, P., Whalley, C., Néry, F., Zal, N., Christiansen, T., 2018. European waters -- Assessment of status and pressures 2018.

Lehner, B., Liermann, C.R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J.C., Rödel, R., Sindorf, N., Wisser, D., 2011. High-resolution mapping of the world’s reservoirs and dams for sustainable river-flow management. Front. Ecol. Environ. 9, 494–502. https://doi.org/10.1890/100125

Lehner, B., Verdin, K., Jarvis, A., 2008. New Global Hydrography Derived From Spaceborne Elevation Data. Eos, Trans. Am. Geophys. Union 89, 93. https://doi.org/10.1029/2008EO100001

LST Biotopkartering vandringshinder [WWW Document], 2020. URL https://ext-geodatakatalog.lansstyrelsen.se/GeodataKatalogen/GetMetaDataById?id=33e74ce9-d0ba-4414-991f-bb4fdeb8bf58 (accessed 3.16.20).

MacBroom, J.G., 2005. Evolution of channels upstream of dam removal sites, in: Managing Watersheds for Human and Natural Impacts: Engineering, Ecological, and Economic Challenges. pp. 1–12.

Mahlum, S., Cote, D., Wiersma, Y.F., Kehler, D., Clarke, K.D., 2014. Evaluating the barrier assessment technique derived from FishXing software and the upstream movement of brook trout through road culverts. Trans. Am. Fish. Soc. 143, 39–48.

Mantel, S.K., Hughes, D.A., Muller, N.W.J., 2010. Ecological impacts of small dams on South African rivers Part 1: drivers of change-water quantity and quality. Water Sa 36, 351–360.

Martin, E.H., 2019. Assessing and Prioritizing Barriers to Aquatic Connectivity in the Eastern United States. J. Am. Water Resour. Assoc. 55, 401–412. https://doi.org/10.1111/1752-1688.12694

Merenlender, A.M., Crall, A.W., Drill, S., Prysby, M., Ballard, H., 2016. Evaluating environmental education, citizen science, and stewardship through naturalist programs. Conserv. Biol. 30, 1255–1265.

Mills, D., 1989. Ecology and management of Atlantic salmon. Springer Science & Business Media.

Mulligan, M., van Soesbergen, A., Sáenz, L., 2020. GOODD, a global dataset of more than 38,000 georeferenced dams. Sci. Data 7, 31. https://doi.org/10.1038/s41597-020-0362-5

O'Hanley, J.R., 2015. Let a river run through it: Optimising river connectivity restoration.

O'Hanley, J.R., Diebel, M., Fedora, M.A., Wright, M., 2011. An optimization planning model for maximizing habitat connectivity for stream resident fish.

ONEMA, 2010. La restauration des cours d'eau. Recueil d'expériences sur l'hydromorphologie (2010) [WWW Document]. URL http://www.onema.fr/recueil_restauration_hydromorphologie

Onofri, A., 2020. aomisc: Statistical methods for the agricultural sciences.

OpenStreetMap contributors, 2019. Planet dump retrieved from https://planet.osm.org.

Palmer, M., Ruhi, A., 2019. Linkages between flow regime, biota, and ecosystem processes: Implications for river restoration. Science (80-.). 365, eaaw2087. https://doi.org/10.1126/science.aaw2087

QGIS Development Team, 2018. QGIS Geographic Information System. Open Source Geospatial Foundation Project.

R Core Team, 2018. R: A language and environment for statistical computing.

Ramos-Diez, I., Navarro-Hevia, J., Fernández, R.S.M., Díaz-Gutiérrez, V., Mongil-Manso, J., 2016. Analysis of methods to determine the sediment retained by check dams and to estimate erosion rates in badlands. Environ. Monit. Assess. 188, 405.

Reddy, S., Dávalos, L.M., 2003. Geographical sampling bias and its implications for conservation priorities in Africa. J. Biogeogr. 30, 1719–1727.

Rincón, G., Solana-gutiérrez, J., Alonso, C., Santiago Saura, ·, García De Jalón, D., Solana-Gutiérrez, J., Alonso, C., Saura, S., de Jalón, D.G., 2017. Longitudinal connectivity loss in a riverine network: accounting for the likelihood of upstream and downstream movement across dams. Aquat. Sci. 79, 573–585. https://doi.org/10.1007/s00027-017-0518-3

Rodeles, A.A., Galicia, D., Miranda, R., 2020. A new method to include fish biodiversity in river connectivity indices with applications in dam impact assessments. Ecol. Indic. 117, 106605. https://doi.org/https://doi.org/10.1016/j.ecolind.2020.106605

Roni, P., Beechie, T.J., Bilby, R.E., Leonetti, F.E., Pollock, M.M., Pess, G.R., 2002. A review of stream restoration techniques and a hierarchical strategy for prioritizing restoration in Pacific Northwest watersheds. North Am. J. Fish. Manag. 22, 1–20.

Schmitt, R.J.P., Bizzi, S., Castelletti, A.F., Kondolf, G.M., 2018. Stochastic modeling of sediment connectivity for reconstructing sand fluxes and origins in the unmonitored Se Kong, Se San, and Sre Pok tributaries of the Mekong River. J. Geophys. Res. Earth Surf. 123, 2–25. https://doi.org/10.1002/2016JF004105

SEPA, 2018. Scottish Obstacles to Fish Migration data set [WWW Document]. URL https://www.sepa.org.uk/environment/environmental-data/ (accessed 9.10.18).

SNIFFER, 2010. Trialling of the methodology for quantifying the impacts of obstacles to fish passage.

Solà, C., Ordeix, M., Pou-Rovira, Q., Sellares, N., Queralt, A., Bardina, M., Casamitjana, A., Munné, A., 2011. Longitudinal connectivity in hydromorphological quality assessments of rivers. The ICF index: A river connectivity index and its application to Catalan rivers.

Limnetica 30, 273–292.

Stephens, T., 2010. Manual on small earth dams: a guide to siting, design and construction. Food and Agriculture Organization of the United Nations (FAO).

Strahler, A.N., 1957. Quantitative analysis of watershed geomorphology. Eos, Trans. Am. Geophys. Union 38, 913–920.

Sun, J., Galib, S.M., Lucas, M.C., 2020. Are national barrier inventories fit for stream connectivity restoration needs? A test of two catchments. Water Environ. J.

Van Looy, K., Tormos, T., Souchon, Y., 2014. Disentangling dam impacts in river networks. Ecol. Indic. 37, 10–20. https://doi.org/10.1016/J.ECOLIND.2013.10.006

Ver Hoef, J.M., Peterson, E., Theobald, D., 2006. Spatial statistical models that use flow and stream distance. Environ. Ecol. Stat. 13, 449–464.

Vogt, J. V, Rimaviciute, E., de Jager, A., 2008. CCM2 River and Catchment Database for Europe: Version 2.1 Release Notes.

Wang, X., Wang, C., 2019. Time Series Data Cleaning: A Survey. IEEE Access 8, 1866–1881.

Warin, C., Delaney, N., 2020. Citizen Science and Citizen Engagement. Achievements in Horizon 2020 and recommendations on the way forward. Brussels.

Wohl, E., 2017. Connectivity in rivers. Prog. Phys. Geogr. 41, 345–362. https://doi.org/10.1177/0309133317714972

World Commision on Dams, 2000. Dams and Development: A New Framework for Decision-making: the Report of the World Commission on Dams. Earthscan.

Zarfl, C., Berlekamp, J., He, F., Jähnig, S.C., Darwall, W., Tockner, K., 2019. Future large hydropower dams impact global freshwater megafauna. Sci. Rep. 9, 1–10.

Zhang, A.T., Urpelainen, Johannes Schlenker, W., 2018. Power of the River: Introducing the Global Dam Tracker (GDAT).

Figure legends

Figure 1. Decision tree used for classifying longitudinal stream barriers into six main functional types. Structures that do not meet these criteria are classified as "other".

Figure 2. Definitions and examples of longitudinal instream barriers.

Figure 3. Decision tree used to identify duplicate barrier records.

Figure 4. The *Barrier Tracker* app for recording barriers and filling data gaps developed as part of the AMBER citizen science programme in Europe (www.amber.internantional).

Figure 5. Expected relationship between barrier density (No/km) and Barrier Free Length (%) when barriers are distributed at random (L = 100 km in the example).

Figure 6. Proposed sigmoid impact function relating barrier height to barrier impact. The asymptotic value is set to near the maximum head drop (3.7 m) that migratory Atlantic salmon can overcome (Mills, 1987)

Figure 7. Cumulative distribution of (A) distance to nearest duplicate and non-duplicate records following a visual assessment of 190 potential duplicate barriers in France, and (B) distance to the nearest unique barrier along the river network, where 80% of unique barriers were found to be within 750 m from one another.

Figure 8. Effect of sample size on estimated bootstrapped exclusion radius in France. The radius that removes 80% of duplicated barriers is 302 m (95% CI: 293-304m) and this is reached after detecting ~100 duplicates, although the variance does not stabilize until ~120 duplicates have been classified.

Figure 9. Barrier discovery rate based on bootstrapped estimates of barrier density (median ± 95% CI) based on field surveys in 26 countries in Europe (data from Belletti et al 2020).

Figure 10. Use of the *Barrier Tracker* by country as of March 2020. The app has been used to record 5,530 barriers by +2,000 users in 36 countries.

Figure 11. Relationship between barrier density and (A) median Barrier Free Length, and (B) mean Barrier Free Length in rivers or Great Britain showing the random expectation (blue line) under BFL = 1/density (data from Jones et al, 2019).

Figure 12. Barrier-specific impacts on river connectivity based on (A) corrected abundance (%), (B) mean height (m), (C) estimated per capita impact, and (D) share of net impact (%). Error bars represent 95 CI. Shown are data for Spain described in Belletti et al 2020.

Figure 1

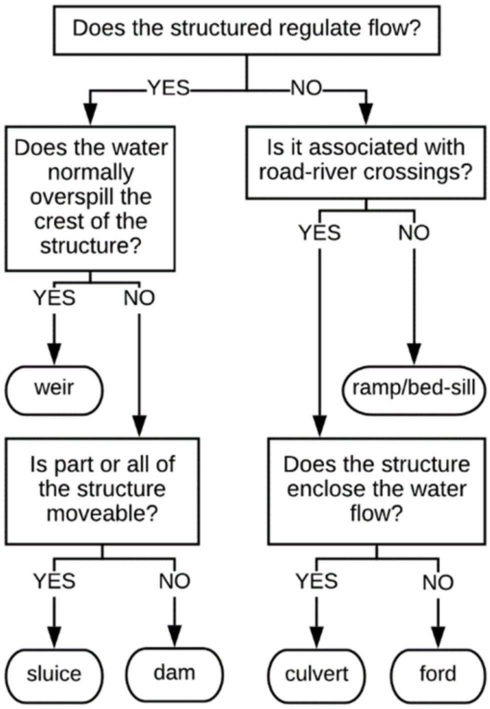
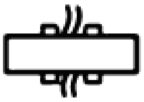
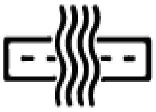

Does the structured regulate flow?
YES
NO
Does the water normally overspill the crest of the structure?
Is it associated with road-river crossings?
YES
NO
YES
NO
weir
ramp/bed-sill
Is part or all of the structure moveable?
Does the structure enclose the water flow?
YES
NO
YES
NO
sluice
dam
culvert
ford

Figure 2

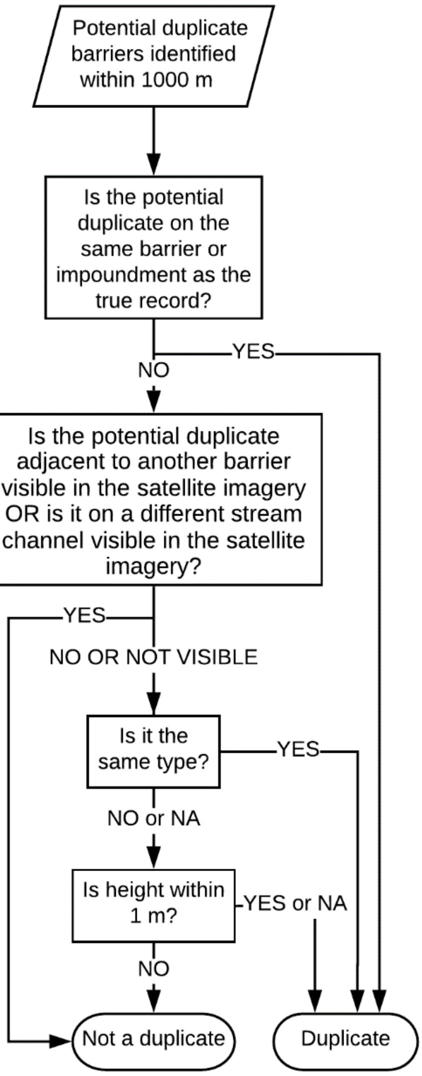
Dam  A dam is a barrier that regulates the flow of water and raises the water level, forming a reservoir. Dams come in many shapes and sizes but water does not normally overflow the crest. Dams are often used to generate hydropower or supply water for irrigation or drinking. They cause a significant alteration of river flow and disrupt the transport of sediments.	 Dam (Dora Baltea river, Italy). S. Bizzi (2017)
Weir  A weir is a barrier that raises the water level and regulates the water flow, but unlike a dam, water flows freely over its crest. Many weirs are old and many may be abandoned, revealing their former use abstracting water for watermills, sawmills, and foundries. They often have heights less than 5 m.	 Consolidation weir (Arno river, Italy. S. Bizzi (2017)

Sluice  Sluice A sluice is a barrier with one or more movable gates that are used to control water levels and flow rates. By opening or closing the sluice gate, water levels and flow rates can be altered. Sluices are used in river locks and canals, to allow boats to navigate over dams or overcome sudden changes in channel slope. They allow canals to be built over uneven landscapes.	 Tidal sluice gate (Netherlands). J. Van Deelen (2017)
Ford  Ford A ford is a low-head structure typically built in shallow streams for wading or crossing. Fords do not raise the water level or regulate the flow of water.	 Ford (Orco river, Italy). M. Micotti (2017)

Culvert  A culvert is a structure built to carry the stream flow at road crossigns. They are tipically built in small streams, under forest tracks or secondary roads. Unlike fords, culverts enclose the stream flow fully (pipe) or partially (half-pipe). They are often embedded in soil and may vary in shape from round and elliptical to box-shaped. Culverts do not raise the water level, but they can block the movement of organisms if they are perched, too shalow, or have too high water velocities.	 Culvert (Afan river, United Kingdom). J Jones (2019)
Ramp and bed-sill  A ramp or bed-sill is a structure designed to stabilize the channel bed. They are usually built in high energy streams to reduce channel erosion caused by channel straightening. They often have a height of less than 1-2m	 A) B) A) Bed sill (Marecchia river, Italy). B. Belletti (2017) B) Rock ramp (Switzerland). R. Bösiger (2018)

Other Other types of barriers that can impact on longitudinal connectivity include fish traps and lateral groynes or wing dykes built perpendicuiar to the river bank to divert the flow of water and reduce flooding or bank erosion, such as the one shown in the picture.	 Other (Dora Baltea river, Italy). B. Belletti (2017)

782 **Figure 3**



783

Figure 4

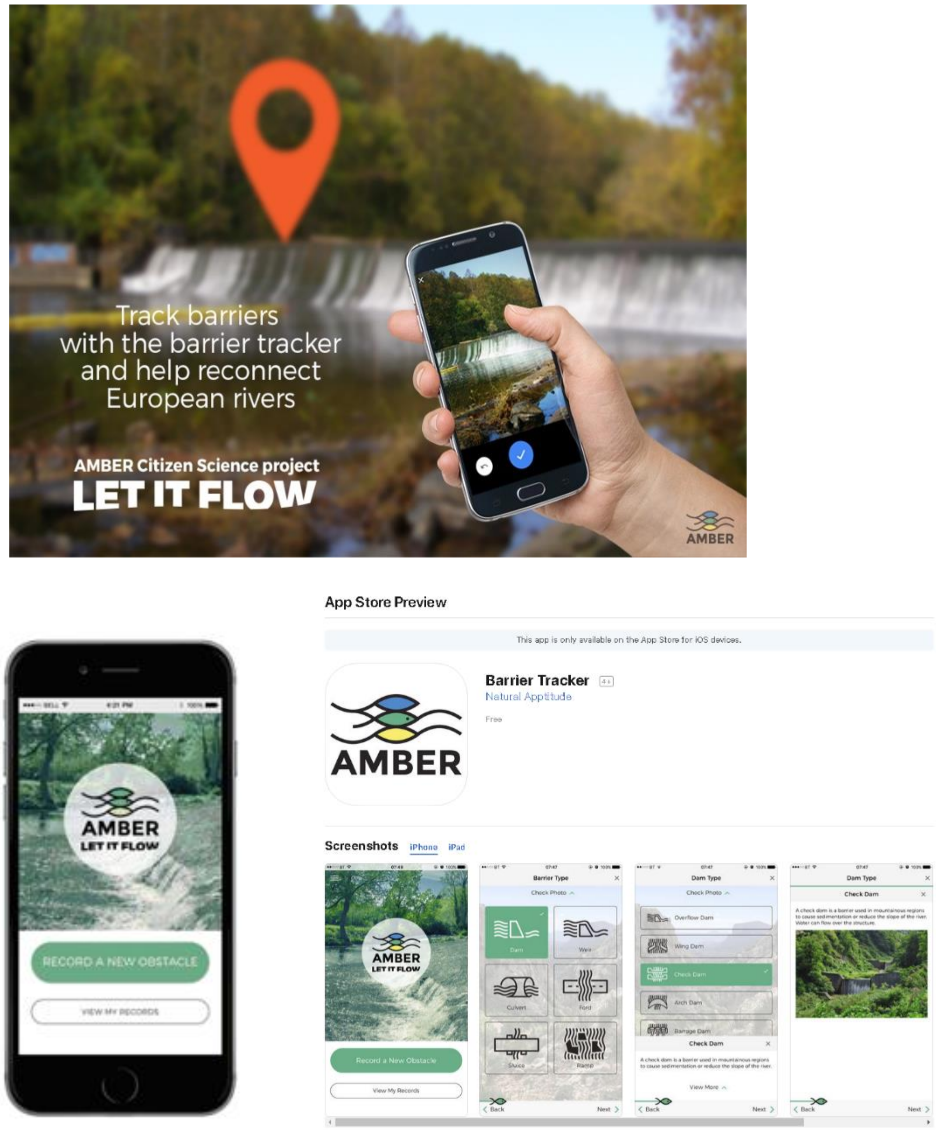
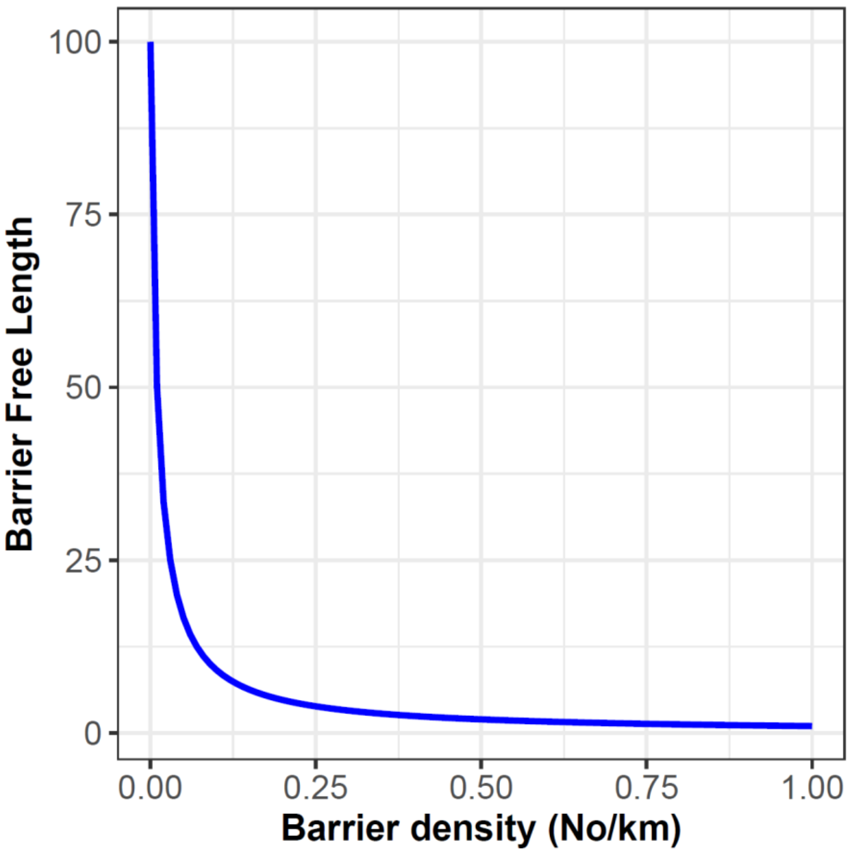
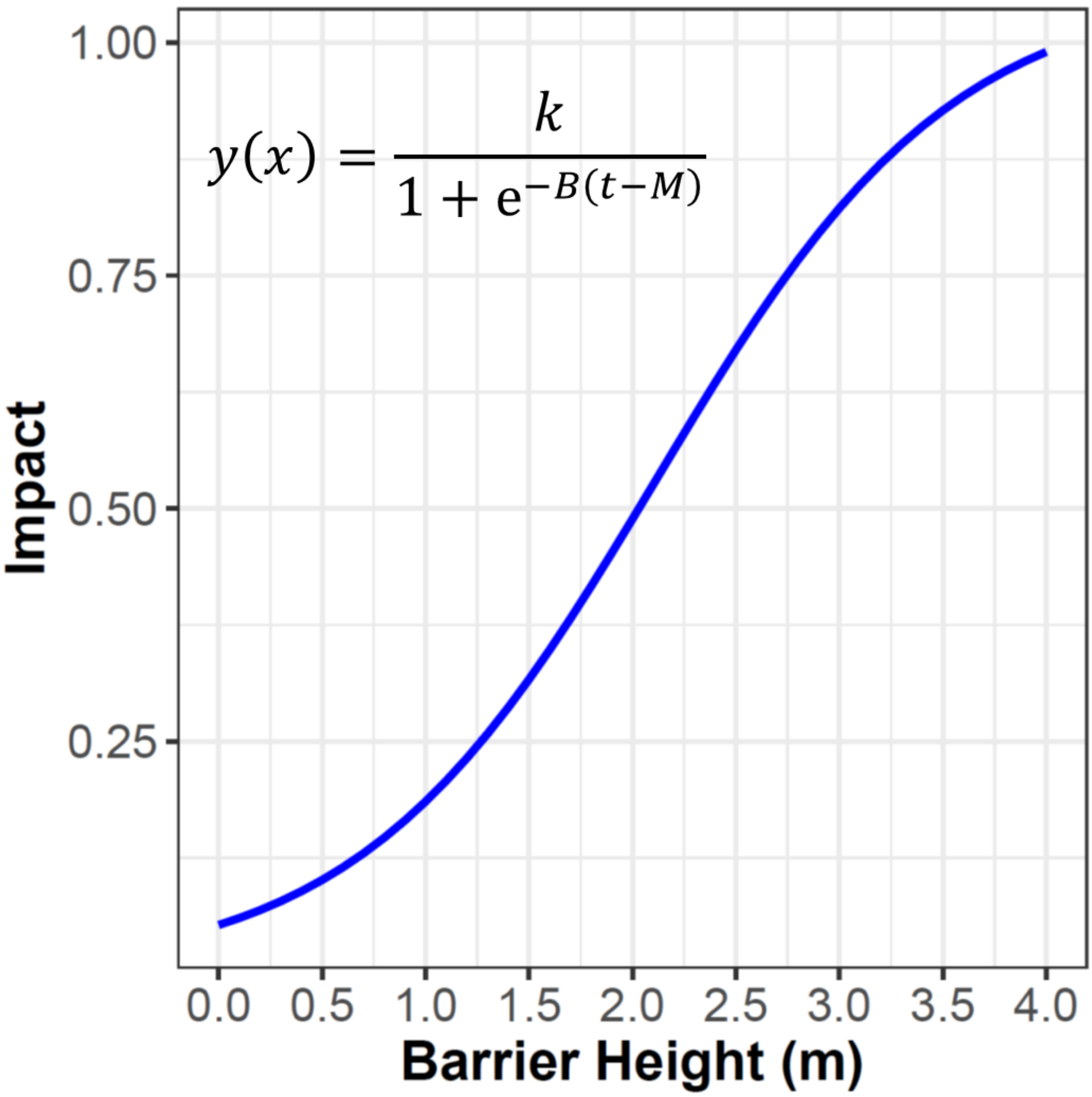

Track barriers
with the barrier tracker
and help reconnect
European rivers
AMBER Citizen Science project
LET IT FLOW
AMBER
AMBER
LET IT FLOW
RECORD A NEW OBSTACLE
VIEW MY RECORDS
App Store Preview
This app is only available on the App Store for iOS devices.
AMBER
Barrier Tracker
Natural Apptitude
Free
Screenshots
iPhone
iPad
AMBER
LET IT FLOW
Record a New Obstacle
View My Records
Barrier Type
Dam Type
Dam Type
Check Dam
Check Dam
Back
Next

Figure 5



Barrier Free Length
100
75
50
25
0
0.00
0.25
0.50
0.75
1.00
Barrier density (No/km)

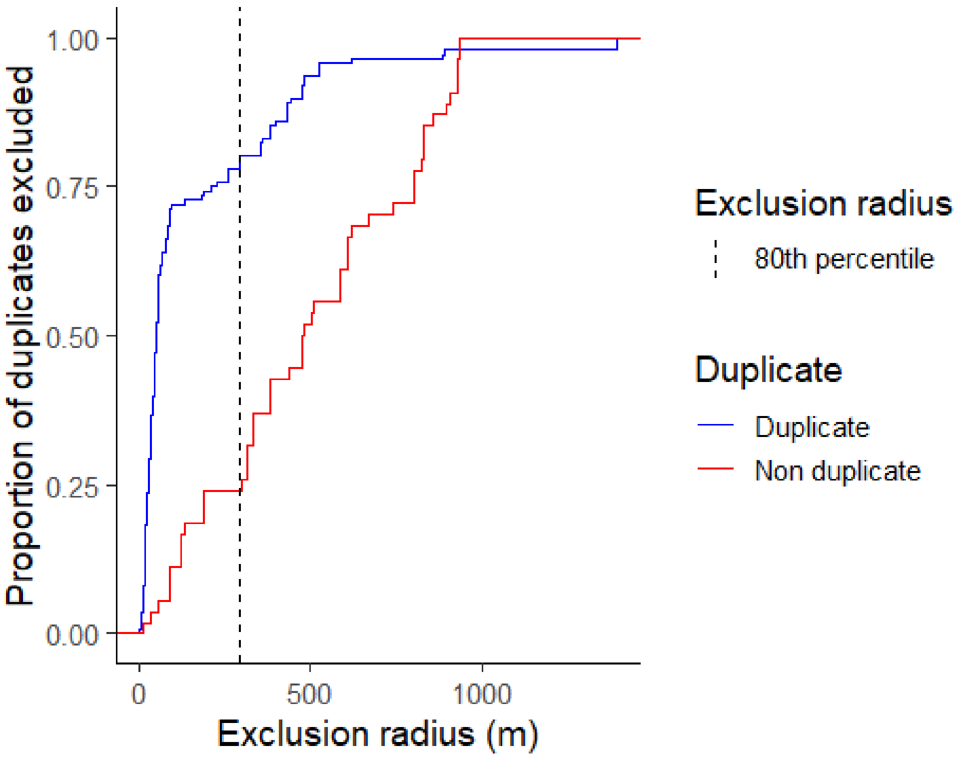
Figure 6



$y(x) = \frac{k}{1 + e^{-B(t-M)}}$
Impact
1.00
0.75
0.50
0.25
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Barrier Height (m)

Figure 7

(A)


Proportion of duplicates excluded
1.00
0.75
0.50
0.25
0.00
0
500
1000
Exclusion radius (m)
Exclusion radius
80th percentile
Duplicate
Duplicate
Non duplicate

(B)

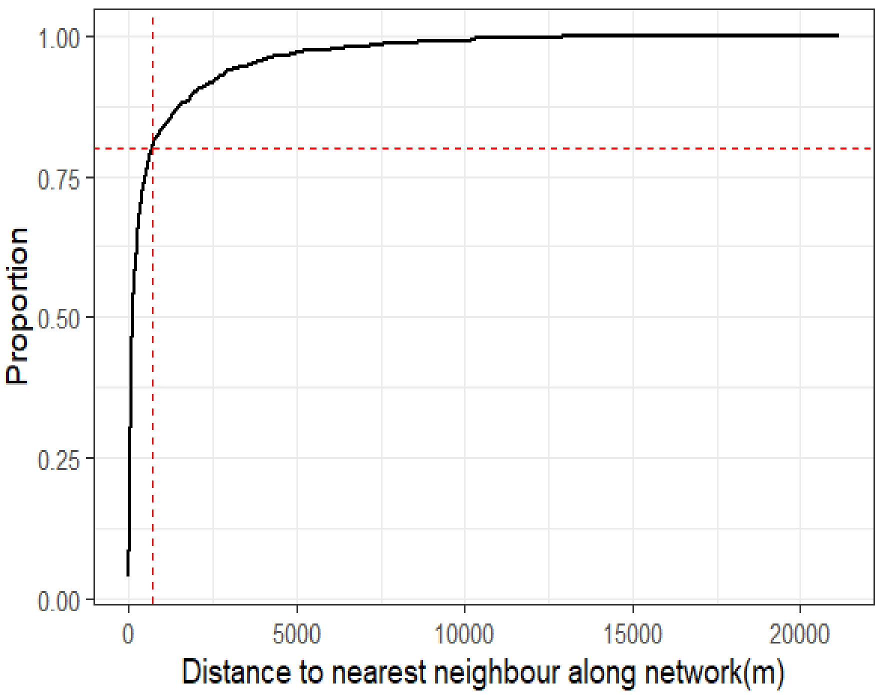

Proportion
1.00
0.75
0.50
0.25
0.00
0
5000
10000
15000
20000
Distance to nearest neighbour along network(m)

Figure 8

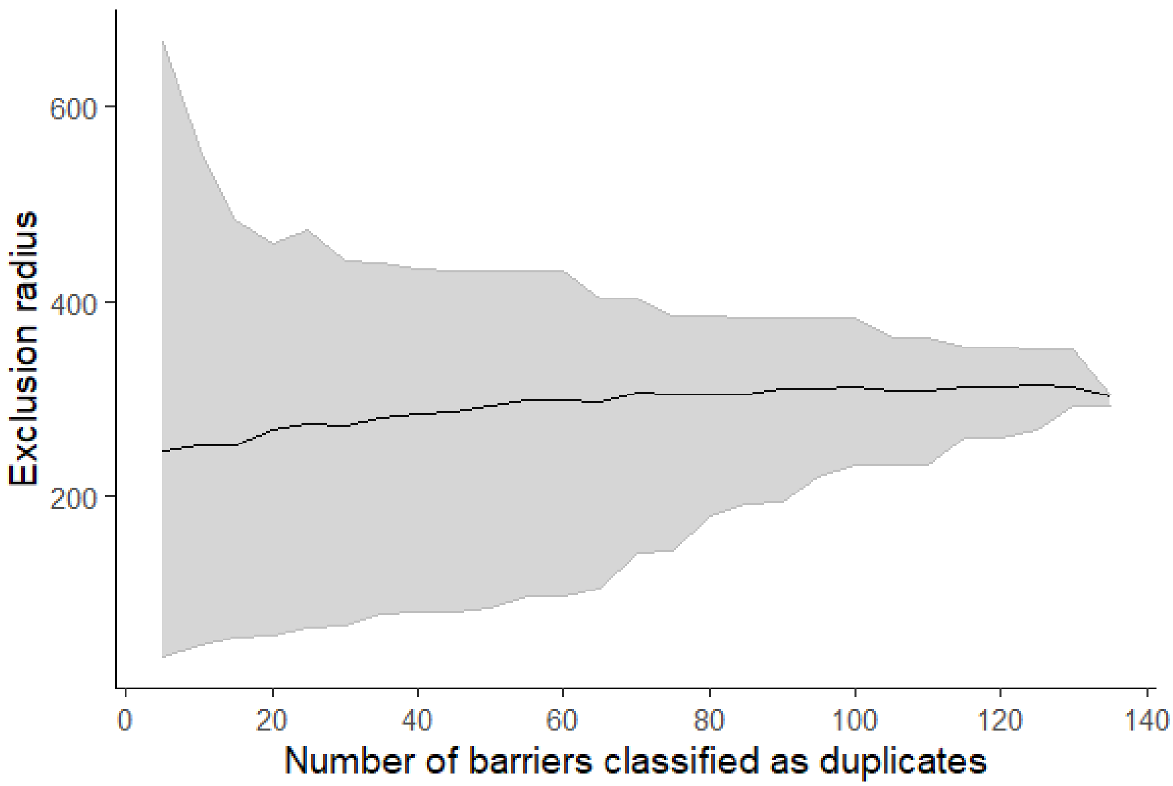

600
400
200
Exclusion radius
0
20
40
60
80
100
120
140
Number of barriers classified as duplicates

Figure 9

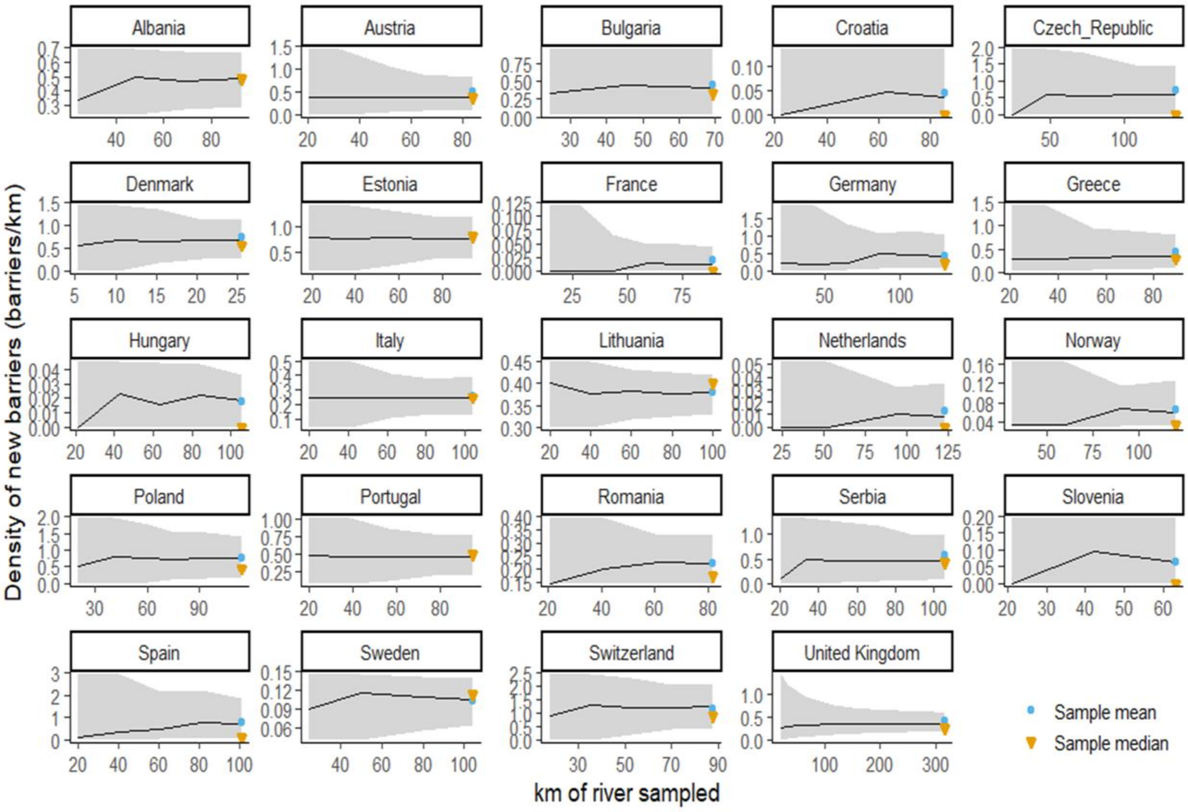

Albania
Austria
Bulgaria
Croatia
Czech_Republic
Denmark
Estonia
France
Germany
Greece
Hungary
Italy
Lithuania
Netherlands
Norway
Poland
Portugal
Romania
Serbia
Slovenia
Spain
Sweden
Switzerland
United Kingdom
Sample mean
Sample median
Density of new barriers (barriers/km)
km of river sampled

Figure 10

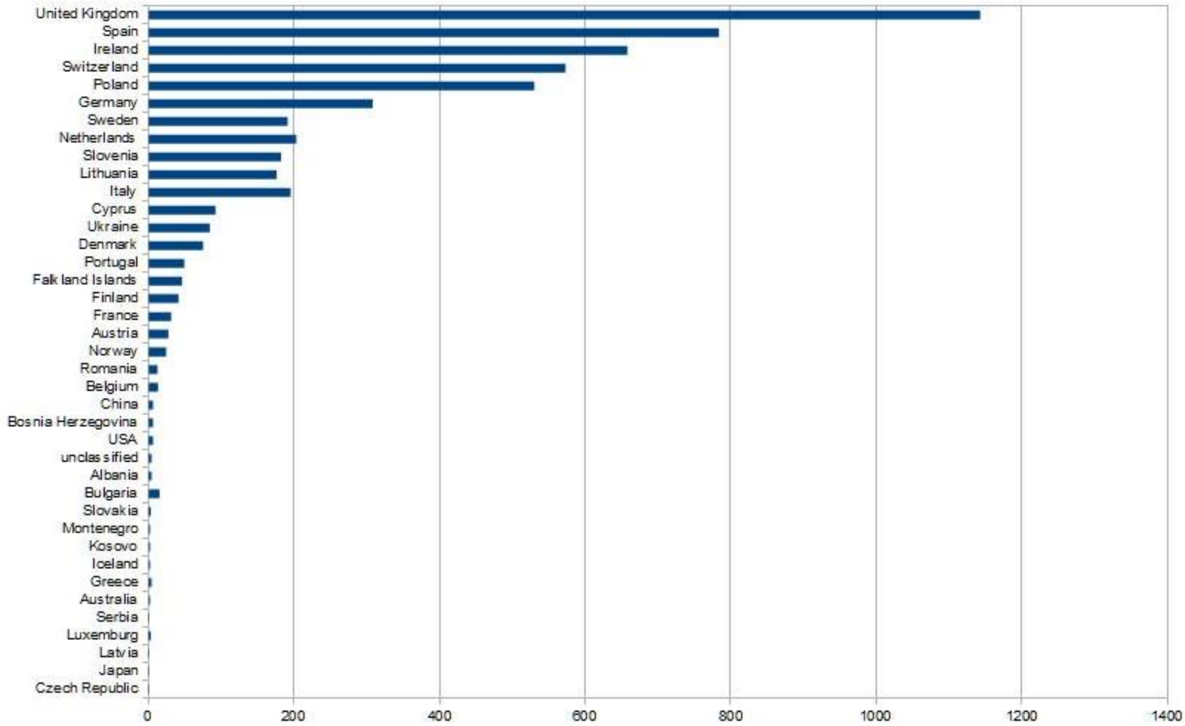
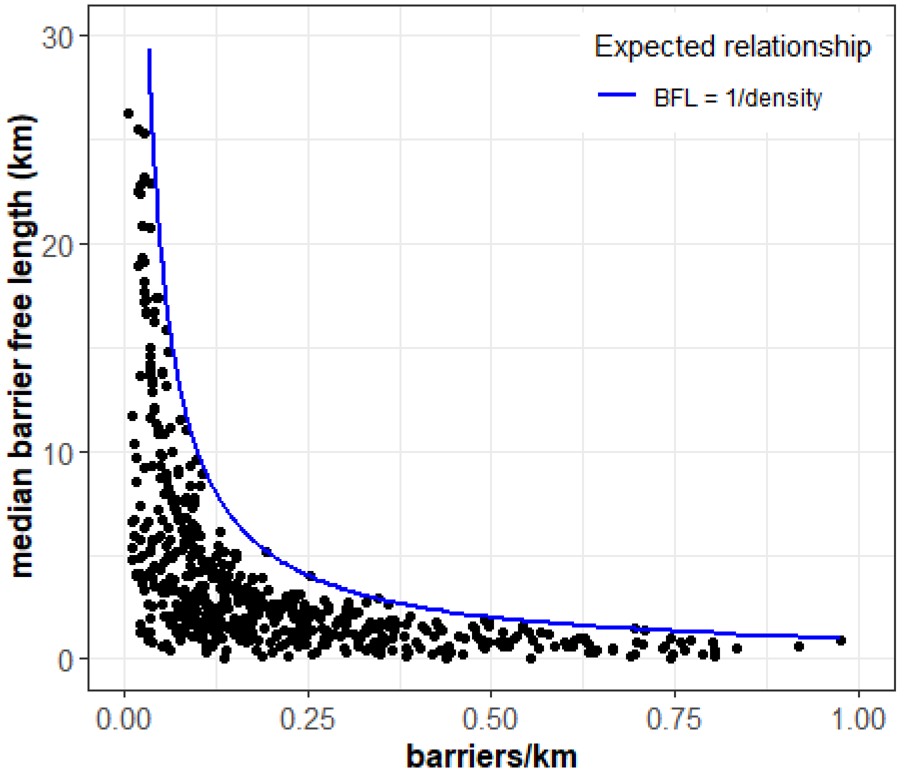

United Kingdom
Spain
Ireland
Switzerland
Poland
Germany
Sweden
Netherlands
Slovenia
Lithuania
Italy
Cyprus
Ukraine
Denmark
Portugal
Falkland Islands
Finland
France
Austria
Norway
Romania
Belgium
China
Bosnia Herzegovina
USA
unclassified
Albania
Bulgaria
Slovakia
Montenegro
Kosovo
Iceland
Greece
Australia
Serbia
Luxemburg
Latvia
Japan
Czech Republic
0
200
400
600
800
1000
1200
1400

Figure 11

(A)



Expected relationship
BFL = 1/density
median barrier free length (km)
30
20
10
0
0.00
0.25
0.50
0.75
1.00
barriers/km

(B)

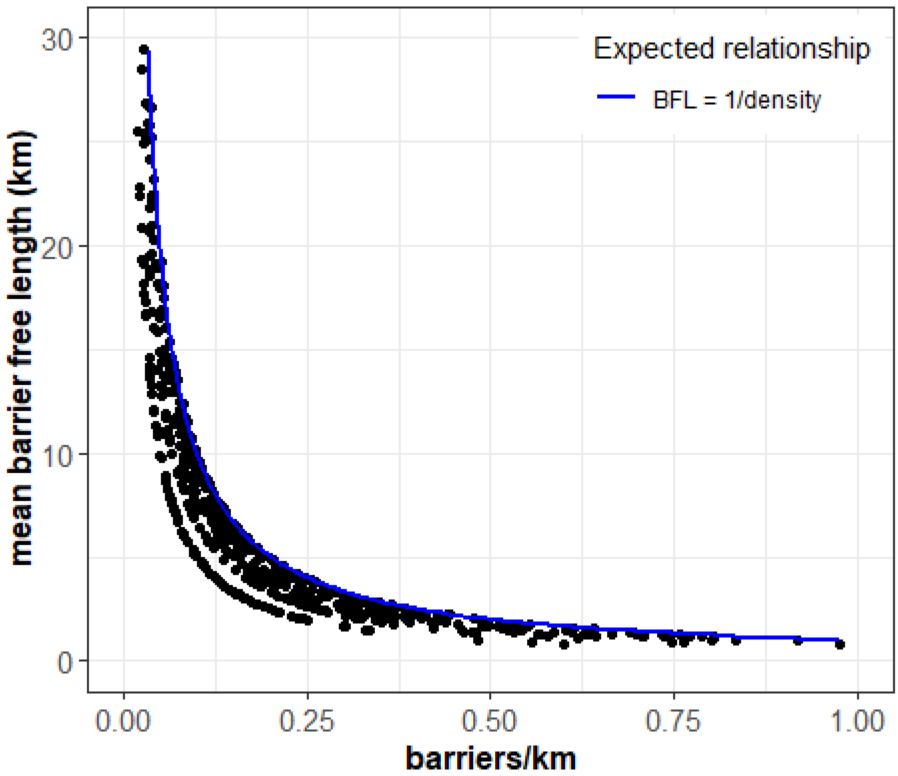
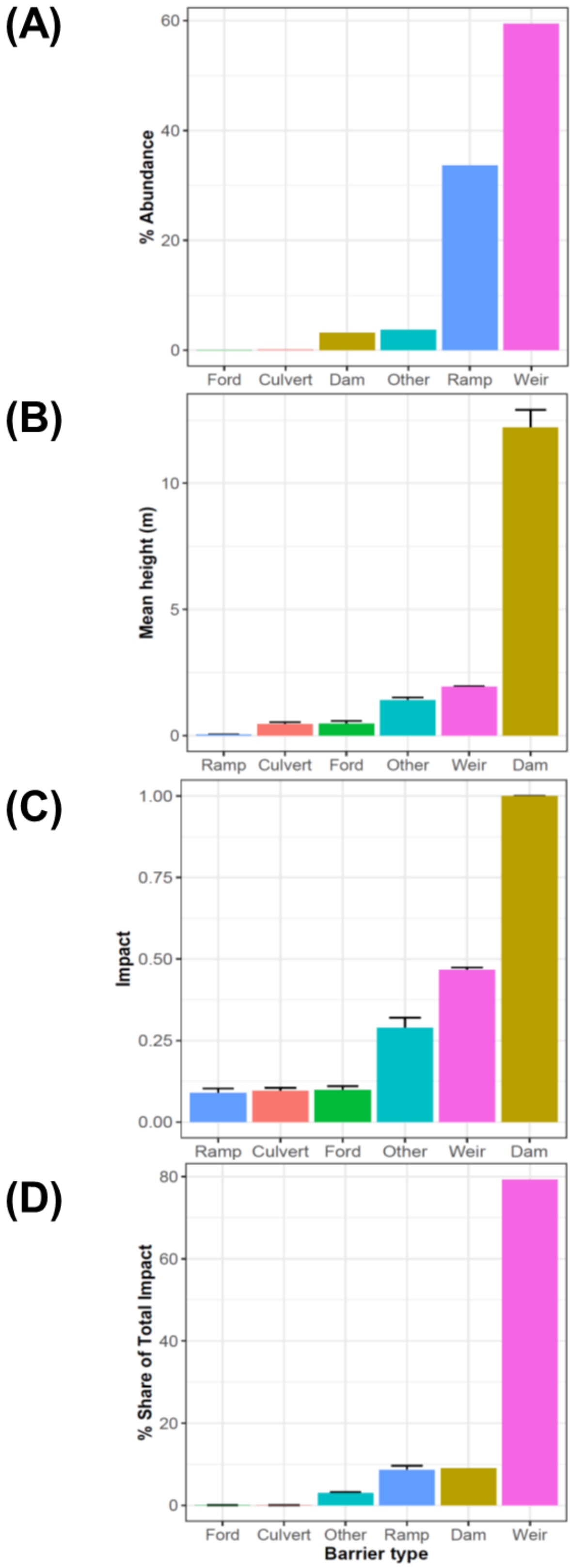

Expected relationship
BFL = 1/density
mean barrier free length (km)
30
20
10
0
0.00
0.25
0.50
0.75
1.00
barriers/km

Figure 12



(A)
% Abundance
0
20
40
60
Ford
Culvert
Dam
Other
Ramp
Weir
(B)
Mean height (m)
0
5
10
Ramp
Culvert
Ford
Other
Weir
Dam
(C)
Impact
0.00
0.25
0.50
0.75
1.00
Ramp
Culvert
Ford
Other
Weir
Dam
(D)
% Share of Total Impact
0
20
40
60
80
Ford
Culvert
Other
Ramp
Dam
Weir
Barrier type