Designing Sustainable Landscapes:
Aquatic Barriers settings variable

A project of the Landscape Ecology Lab, Department of Environmental Conservation, University of Massachusetts, Amherst

Principals:

- William V. DeLuca, Adjunct Assistant Professor
- Bradley W. Compton, Research Associate
- Ethan B. Plunkett, Research Associate
- Kevin McGarigal, Emeritus Professor

With support from:

- US Fish and Wildlife Service, North Atlantic-Appalachian Region
- Northeast Climate Adaptation Science Center (USGS)
- University of Massachusetts, Amherst

UMass Amherst

Reference:

McGarigal K, Compton BW, Plunkett EB, DeLuca WV, and Grand J. 2020. Designing sustainable landscapes products, including technical documentation and data products. http://umassdsl.org/.

General description

Aquatic barriers is one of several ecological settings variables that collectively characterize the biophysical setting of each 30 m cell at a given point in time (McGarigal et al 2017). Aquatic barriers measures the relative degree to which road-stream crossings (i.e., bridges and culverts) and dams may physically impede upstream and downstream movement of aquatic organisms, particularly fish. It is derived from a custom algorithm (see below for details) applied to dams and derived road-stream crossings. Briefly, each dam has an aquatic barrier score based either on dam height or attributes indicating whether the dam has a partial/complete breach. Similarly, each road-stream crossing has an aquatic barrier score based either on an algorithm applied to field measurements of the crossing structure or predictions from a statistical model based on GIS data. Aquatic barriers is scaled 0-1, where dams and road-stream crossing are assigned values >0 (with 1=complete barrier) and all other cells (including terrestrial) are assigned 0 (**Fig. 1**).

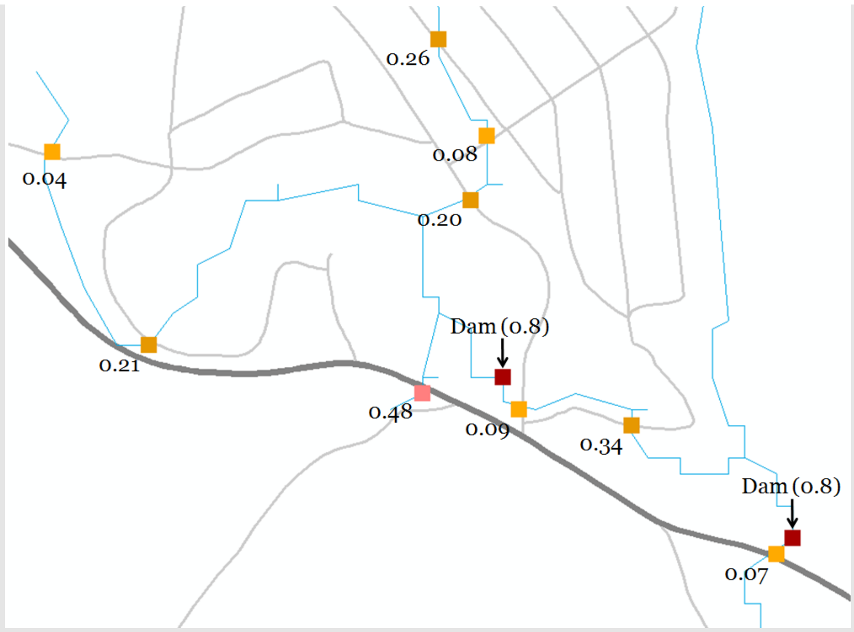

Figure 1. Aquatic barrier scores for dams and road-stream crossing with vector roads (gray scale) and streams (blue) in the background.

Use and interpretation of this layer

Aquatic barriers is used in the derivation of the aquatic connectedness metric in the context of the broader assessment of ecological integrity (see the technical document on integrity, McGarigal et al 2017). It is a measure of the degree to which a dam or road-stream crossing is predicted to be an impediment to movement of aquatic organisms, and its use should be guided by the following considerations:

- Aquatic barriers is formatted as a raster GIS data layer designed for use in the DSL Landscape Change, Assessment and Design (LCAD) model. It contains non-zero values only for cells classified as either dams or road-stream crossings; all other cells are assigned a value of 0. As such, it is a difficult layer to view since the eye is naturally drawn to the dominant matrix of zeros. For easier viewing and general purpose use, we also distribute two separate point shapefiles in vector format (dam removal impacts and culvert upgrade impacts) that contain the aquatic barrier scores along with many other statistics associated with the restoration potential of the structure, as described with those layers.

- It is important to recognize the relative nature of the aquatic barrier scores. A score of 0 means that the structure (dam, bridge, or culvert) is predicted to have no effect on aquatic passability, and a score of 1 means that the structure is predicted to be a complete barrier to most aquatic organisms, particularly fish. However, intermediate values represent an index of the relative degree of obstruction to the movement of aquatic organisms, such that a 0.4 score is predicted to confer roughly twice the degree of impediment to movement than a 0.2 score. Because the score is a relative index, the absolute value does not have a simple interpretation. Moreover, because the score is an index to passability for all aquatic organisms, but emphasizing fish passage, it does not have an specific interpretation for any single species. Increasing barrier scores should indicate fewer species that can pass, and, in general, fewer individuals of a particular species that can pass. However, because aquatic organisms vary widely in their vagility and their abilities to pass different types of barriers and data as to the exact nature of each barrier are unavailable, interpretation must be general.

- It is important to acknowledge that the aquatic barrier scores are derived from a model, and thus subject to the limitations of any model due to incomplete and imperfect data, and a limited understanding of the phenomenon being represented. In particular, the GIS data on dams and road-stream crossings are imperfect; they contain errors of both omission (e.g., missing dams) and commission (e.g., derived road-stream crossings that don't exist in the real world). Consequently, there will be many places where the model gets it wrong, not necessarily because the model itself is wrong, but rather because the input data are wrong. In addition, the scores themselves are derived from a model based on expert opinion of the factors affecting passability for aquatic organisms, and while the model incorporates many of the factors known or believed to affect passability, it is almost certainly an incomplete and imperfect representation of the real-world factors affecting passability. This model has not been extensively tested against empirical studies of passability in the field. Moreover, the majority of road-stream crossings (93.8%) have not been surveyed in the field, and their predicted aquatic barrier scores are based on an even simpler and less perfect model derived from GIS data (as so many of the determinants of culvert and bridge passability are idiosyncratic, and unrepresented by GIS data). Thus, aquatic barriers should be used and interpreted with caution and an appreciation for the limits of the available data and models.

- While aquatic barriers has a wide variety of potential uses, perhaps its most significant application is to aid in the assessment of aquatic connectivity, for example via incorporation into the DSL aquatic connectedness metric and the assessment of aquatic ecological integrity and critical linkages (i.e., prioritization of dams for removal and road-stream crossings for culvert upgrades). Outside of these DSL applications, aquatic barriers could be used on its own to help direct conservation actions to restore aquatic connectivity.

Derivation of this layer

Aquatic barrier scores were assigned separately for dams, surveyed road-stream crossings and unsurveyed road-stream crossings, as follows:

1. Aquatic barrier scores for dams

We obtained dam locations from the Northeast Aquatic Connectivity Assessment Project (NACAP) version 2 (Martin and Levine, 2017). The NACAP dams were located on NHDPlus (1:100,000) flow lines, while DSL uses a cleaned-up version of the NHD (1:24,000) streams which contain considerably more detail. Thus, we needed to snap the dams to our stream network. Additionally, the NACAP dams included dams that were on the NHDPlus network (use = 1 in the attributes) as well as dams that were not on the network (use = 2).

For use 1 dams (on the NHD flowlines) we initially snapped the dams to our stream network and then compiled a set of metrics to describe the snap. These metrics included the snap distance, the distance to the nearest stream junction, the ratio of the snap distance to junction distance, the difference in flow accumulation between the NHDPlus flow accumulation at the original location and the DSL flow accumulation at the snapped location. Next, we selected a random subset of the dams and based on orthophotos and our stream linework classified the snap as correct or erroneous. Then we used the sampled data to build a statistical model of snapping accuracy. Finally, for all snaps that were modeled as less than 95% accurate we inspected the original and snapped location against an orthophoto with the DSL stream network and on a case by case basis determined where the dam should on our network (and in some cases dropped the dam completely).

For Use 2 dams we retained the dam only if the snap distance was less than 40 meters, the ratio of the snap distance to the nearest junction was less than 0.2, and had a flow accumulation of less than 2800 30 m cells (252 ha). These thresholds were chosen after looking at the distributions of each metric and were designed to be fairly strict; thus we only retained use 2 dams that snapped to small streams close to their location when we could be fairly confident that the snap was correct.

Aquatic barrier scores for dams were based on dam height (**Fig. 2**) given by the following algorithm:

Passability = 0.2 × LOGISTIC (height, inflection=1.5, scale=-0.2), where:

$$LOGISTIC = \frac{1}{1 + exp\left(\frac{-(height - inflection)}{scale}\right)}$$

and height is in meters.

Barrier score = 1 - Passability.

2. Aquatic barrier scores for road-stream crossings

We derived road-stream crossings in the landscape based on the intersection of the cleaned and trimmed vector National Hydrology Dataset (NHD) streams and Open Street Map (OSM) roads and railroads. Each of these point crossings was then moved to the nearest crossing pixel in the raster representation of the streams and roads for representation in the aquatic barriers layer. However, we retained both the original (vector) and moved (cell) locations for subsequent use (see below). We assigned an aquatic barrier score to each crossing in the raster representation, but the derivation of the score depended on whether the crossing was surveyed in the field or not, as follows.

2.1 *Surveyed road-stream crossings*

We used the aquatic passability scores assigned by the North Atlantic Aquatic Connectivity Collaborative (NAACC). NAAC maintains a database of surveyed crossings, survey metrics, and passability scores derived from them. We used 1 – passability as assigned by NAACC as the barrier score for the surveyed crossings. The scores are based on the November 10, 2015 scoring algorithm (NAAC, 2015) extracted from a January 2, 2020 export from the database for a set of 34,707 crossings after considerable filtering of the crossings (see **Appendix**) to ensure correspondence with our derived road-stream crossings.

2.2 Unsurveyed road-stream crossings

To assign aquatic barrier scores for those crossings that had not been assessed in the field (i.e., unsurveyed crossings), we used GIS data and crossing scores from the filtered set of 34,707 crossings (see **Appendix**) to create a statistical model to predict aquatic barrier scores, as follows.

1. We assembled a suite of predictors to be used in the model either by sampling grids at the cell location of the crossing or by analysis of a window centered on the crossing (**Table 1**). For the scale-dependent variables, we calculated their values in square windows with sides of 90, 150, 210, 270, 330, 390, 450, 510, 570, and 630 meters.
2. We then performed additive stepwise variable selection with a Random Forest model to find the set of variables that resulted in a Random Forest with the highest R-squared between the field survey-based aquatic passability score and the out-of-bag prediction from the model. Note, Random Forest is a non-parametric method that is effective at optimizing reliable predictions.
3. We fit similar models from the same suite of variables to estimate whether the crossing was a bridge or not.
4. Note, for the Connecticut River watershed Landscape Conservation Design pilot (CTR LCD) we used the predicted bridge status of the crossing to assign the mean terrestrial passability score of crossings with the same status from the surveyed crossings to the unsurveyed crossings. Thus, all unsurveyed crossings predicted to be bridges were assigned the mean passability of the surveyed bridges, and all unsurveyed crossings predicted not to be bridges (including, e.g., culverts, fords, open-bottom arches) were assigned the mean passability of the surveyed crossing there were not bridges.

However, for the Northeast regional product that we are distributing, the terrestrial barrier scores reflect the predicted passability scores from the Random Forest model. Note, in the culvert upgrade impacts shapefile we include both the modeled score and the mean score, for those that prefer to use the latter, along with many other statistics associated with the restoration potential of the structure, as described for that layer.

5. Lastly, the aquatic barrier score for unsurveyed road-stream crossings was given as the complement of the aquatic passability score (i.e., 1 - passability).

GIS metadata

This data product is distributed as a geoTIFF raster (30 m cells). The cell value is the aquatic barrier score, which ranges from 0 (all cells not mapped as either a dam or road-stream crossing) to 1 (maximum barrier score; i.e., likely to be a complete barrier to most aquatic organisms, particularly fish). This data product may be obtained at McGarigal et al (2020).

Literature Cited

Martin, E. H. and J. Levine. 2017. Northeast Aquatic Connectivity Assessment Project - Version 2.0: Assessing the ecological impact of barriers on Northeastern rivers. The Nature Conservancy, Brunswick, Maine. http://maps.freshwaternetwork.org/northeast/

McGarigal K, Compton BW, Plunkett EB, DeLuca WV, and Grand J. 2020. Designing sustainable landscapes products, including technical documentation and data products. https://scholarworks.umass.edu/designing_sustainable_landscapes/

North Atlantic Aquatic Connectivity Collaborative (NAACC). 2015. Scoring Road-Stream Crossings as Part of the North Atlantic Aquatic Connectivity Collaborative (NAACC). https://streamcontinuity.org/sites/streamcontinuity.org/files/projects/images/Aquatic_Passability_Scoring.pdf

Table 1. Variables used to predict whether a road-stream crossing was a bridge or culvert and the aquatic passability score for the structure.

Variable	**Description**
d8accum	Number of cells that flow into the crossing.
gradient	Stream gradient at crossing.
elevation.range.[scale]	The range of elevation observed in a window of [scale] dimension (in meters).
incisement.[scale]	Within a window of [scale] meters centered on the crossing cell, the difference between the mean elevation of the water cells and the mean elevation of all other cells.
elevation.sd	The standard deviation in elevation within a window of [scale] dimension centered on the crossing.

Appendix

The following is a detailed description of the process for filtering the crossing records in the source database obtained from NAACC in order to include only those records and unique surveys that we could reliably associate with one of our derived road-stream crossings.

We began with source data from NAACC (https://www.streamcontinuity.org/cdb2) contained in two databases, one with older data migrated from the original UMass Stream Continuity Project and one with newer data in a revised format settled on by NAACC. The crossings in these two databases were scored based on the algorithm dated November 10, 2015 (https://streamcontinuity.org/pdf_files/Aquatic_Passability_Scoring.pdf). Filtered records were exported from both databases, then cleaned, converted into a standard format, merged, and filtered further, as follows:

Database 1 export (from Stream Continuity Project):

1. We started with 13,206 records for 11,626 unique surveys at 10,721 unique crossings.
2. We dropped 1,580 duplicated records (probably due to the crossing having multiple structures each with its own line in the data export), leaving 11,626 records for 11,626 unique surveys at 10,721 unique crossings.
3. We dropped 18 records that were on a list of "bad" records provided by Scott Jackson, leaving 11,608 records for 11,608 unique surveys at 10,712 unique crossings.
4. We dropped 322 records where the GPS location was greater than 200 meters from the crossing location (GPS is a field measure), leaving 11,286 records for 11,286 unique surveys at 10,437 unique crossings.

5. We dropped 739 records with missing location data, leaving 10,547 records for 10,547 unique surveys at 9,706 unique crossings.
6. We dropped 490 records where the aquatic scores were NA, leaving 10,057 records of 10,057 unique surveys at 9,239 unique crossings.
7. We dropped 818 records with duplicate crossing codes (repeat surveys of the same crossing), keeping the most recent survey, leaving 9,239 records of 9,239 unique surveys at 9,239 unique crossings.

Database 2 export (from NAACC):

1. We started with 52,627 records for 46,553 unique surveys at 45,452 unique crossings.
2. We dropped 6,074 duplicated records (due to the crossing having multiple structures each with its own line in the data export), leaving 46,553 records for 46,553 unique surveys at 45,452 unique crossings.
3. We dropped 3,117 records were the crossing was listed as inaccessible in the database, leaving 43,436 records and surveys for 42,519 unique crossings.
4. We dropped 624 records where the GPS location was greater than 200 meters from the crossing location (GPS is a field measure), leaving 42,812 records and unique surveys at 41,942 unique crossings.
5. We dropped 2 records that were missing location information, leaving 42,810 records and surveys at 41,941 unique crossings.
6. We dropped 2,105 records where the aquatic crossing scores were, leaving 40,705 unique records and surveys at 39,991 unique crossings.
7. We dropped 714 records with duplicate crossing codes (repeat surveys of the same crossing), keeping the most recent survey, leaving 39991 unique records, surveys, and crossings.

Merge of two survey databases:

1. We dropped 879 records from dataset 1 that had crossing codes identical to those in dataset 2, and dropped 3 additional crossings with inconsistent coring type and no crossing fields, and merged the two datasets, resulting in 48,348 records for 48,348 unique surveys at 48,348 unique locations, of these 2,987 of the surveys indicate that there was no crossing at the location.

Combination with GIS derived crossing locations.

1. We dropped 11,428 surveyed locations because they were greater than 30 meters from the nearest GIS derived location or because another surveyd location was closer to the closest GIS location. The resulting full dataset includes 71,7519 locations of which 36,920 (5.1%) were surveyed. This full dataset includes trimmed streams where flow accumulation was too low to be retained in our stream network, crossings on tracks which were dropped from our landcover, and surveyed non-crossings. Of the 36,920 surveyed locations, 2,213 were surveyed non-crossings, leaving 34,707 surveyed

crossings matched to GIS derived road stream crossings that were used to fit models of passability at non-surveyed crossings.

2. The selected dataset includes only crossing that correspond to our final landcover and that were included in the raster barrier layer. This excludes crossing on tracks (which we dropped because many of the tracks were erroneous), and trimmed portions of streams (dropped because of inconsistent mapping of low flow streams), and surveyed non-crossings. It contains 556,246 crossings of which 34,707 (6.2%) were surveyed.