# Procedure

## Background

Fish entrained through hydroelectric facilities are exposed to turbine passage mortality stressors. Mortality through hydroelectric turbines has been well studied, with mathematical models able to predict the probability fish will get struck by a turbine blade (Von Raben 1957, Franke et al. 1997). The rate at which fish are entrained (fish per million [M] cubic feet [ft3] of water) through hydroelectric facilities is also a well-studied phenomenon, with results from field trials contributing to an entrainment database compiled by the Electric Power Research Institute (EPRI 1997). The 1997 EPRI database contains observations of 70 species at 43 facilities east of the Mississippi River. The EPRI dataset is particularly useful for quantitative analysis based on the assumption that when entrainment counts are standardized by discharge across facilities and holistically observed, the database will provide a reasonable estimate of entrainment rates for a watershed of a given size that are suitable for decision making purposes. Also, by describing entrainment rates with statistical distributions and simulating with Monte Carlo methods, it is possible to estimate average daily entrainment and mortality with measures of certainty, as well as estimating the likelihood an event of a given size will occur.

Simulated fish migrate through a hydroelectric project where passage routes are described with a mathematical network. We assume all simulated fish will move downstream as they approach the project. If fish survive their current node, they can move to the next one. If there is more than one node available at their current location, then Monte-Carlo role of the dice and a priori determined transition probabilities control their movement. The simulation ends for a fish when it arrives at the last node in the network or dies.

For fish passing via entrainment, individuals are exposed to turbine blade strike, which is modeled with the Franke et al. (1997) equations. For fish that pass via passage structures or spill, mortality is assessed with a roll of the dice using survival metrics determined a priori, sourced from similar studies, or from expert opinion. The Franke et al. (1997) equations calculate the probability a fish of a given length will get struck by a turbine runner blade. With these equations, if we know how long a given fish is, the amount of discharge of through the turbine, the type of turbine, how many blades, and how fast it is rotating, then we can calculate with certainty the probability of being struck. Therefore, the only morphometric parameter needed to assess blade strike is length. All other input parameters are sourced from technical drawings of the facility.

This procedure will guide the end user through an assessment of entrainment effects at a hydroelectric facility. Stryke has two modes of operation for two different types of desktop entrainment studies: (1) it can be used to assess survival of anadromous species as they migrate past a facility (or series of facilities), and (2) estimate entrainment rates and survival of native species to assess population level impacts attributed to entrainment through a hydroelectric station.

## Create a Project

The Stryke source code is hosted on GitHub. To create a project, you must first create a directory on your desktop (we will upload later). Next, clone the repository (<https://github.com/knebiolo/stryke>) into your new directory, with GitHub Desktop. Then, create a new project in Spyder that points to the project directory you just created. After cloning, open up the ‘Stryke’ folder that you will now find in the project directory, and then open up the Interface Spreadsheet to enter model parameters before running any Python code.

## Flow Scenarios

This section contains instruction on setting up the ‘Flow Scenarios’ tab in the interface spreadsheet, with directions for native and anadromous entrainment assessments. Entrainment events often occur on a seasonal cycle and are a function of the amount of water discharged through a facility. Therefore, it is important for our model to recreate potential hydrologic conditions and the operating scenarios for such conditions. For facilities with multiple units, it is assumed that a single unit would be operated up until its most efficient flow. At that point, water will then begin to flow through other units up until their most efficient flow or until the hydrologic capacity of the facility is met. Any more discharge is then spilled over the dam. If we assume fish proportionally follow the flow, we can estimate the rates at which fish will pass via each passage route. Thus, if we know the river discharge, we can simulate passage through the facility.

### Flow Release Scenarios for Anadromous Configuration

For anadromous flow release scenarios, a single discharge is used to model an entrainment event. The required columns to be completed on the Flow Scenarios tab are: Scenario, Flow, Season, *Min\_Op\_Flow*, *Env\_Flow*, Hours, and Months columns. If an assessment of route of passage is not required, set *Min\_Op\_Flow* and *Env\_Flow* to zero. Table XX lists how to parameterize Stryke for assessments of anadromous species.

|  |  |
| --- | --- |
| Column | Data Input |
| Scenario | Text, user defined scenario name |
| Flow | Integer, user defined river discharge (cfs) |
| Season | Text, user defined season name, must match season column on Population tab. |
| Min\_Op\_Flow | Integer, user defined minimum operating discharge |
| Env\_Flow | Integer, user defined environmental flow release |
| Hours | Integer, user defined hours operation, Run of River facilities run for 24 hours |
| Months | List of integers separated with comma that correspond with season column. |

If the end user does not know flow scenarios *a priori*, Stryke can estimate discharges at specific exceedance probabilities using a novel regression based approach. Stryke gathers data from the 100 nearest USGS stream gages to the hydroelectric facility of interest. With a GIS, identify the 100 nearest [USGS stream gages](https://www.sciencebase.gov/catalog/item/577445bee4b07657d1a991b6) and export a table with all joined attributes. In the export table, there must be a column, ‘Drainage\_A’, that contains the drainage area of the project in square miles. If the end user uses The National Inventory of Dams (2005) as a source data file, the shapefile attributes will contain this column. Next, open the script hydrology.py found in the stryke folder in the project directory and edit the following variables: dataWS, NID\_to\_gage, exceedance, season\_dict, and the flow.seasonal\_exceedance function call.

The hydrology.py script imports stryke, which has convenience functions that calculate exceedance probabilities from the last 10 years of data at each gage with the Python modules; *hydrofunctions*, *Pandas*, and *statsmodels*. A linear relationship is evident when we graph exceedance flow as a function of drainage area. Stryke uses ordinary least squares regression to estimate seasonal exceedance flows. Figure 2 shows one such example of the 50-percent spring exceedance flow at Allegheny Lock and Dam (red dot) on the Allegheny River in Pennsylvania. The estimate of 22,887 cfs is then entered into the Flow Scenario tab, Flow column.

Chart, scatter chart

Description automatically generated

### Flow Release Scenario Native Species Configuration

An entrainment risk assessment for native species requires a yearly entrainment estimate, which is a function of the daily discharge through a facility. The end user will identify representative flow years (e.g. low flow = 2005, high flow = 1996) from a USGS gage of interest to simulate daily discharge, which influences daily entrainment rates. The required columns to be completed on the *Flow Scenarios* tab is: Scenario, Flow, Gage, FlowYear, Prorate, Season, *Min\_Op\_Flow*, *Env\_Flow*, Hours, and Months columns. If an assessment of route of passage is not required, set *Min\_Op\_Flow* and *Env\_Flow* to zero. Table XX lists the required columns and required data to configure Stryke for an assessment of anadromous species.

|  |  |
| --- | --- |
| Column | Data Input |
| Scenario | Text, user defined scenario name |
| Flow | Text, required: ‘hydrograph’ |
| Gage | Text, USGS stream gage ID |
| FlowYear | Integer, user defined representative calendar year |
| Prorate | Float, user defined prorate factor used to scale USGS watershed to project watershed |
| Season | Text, user defined season name, must match season column on Population tab. |
| Min\_Op\_Flow | Integer, user defined minimum operating discharge |
| Env\_Flow | Integer, user defined environmental flow release |
| Hours | Integer, user defined hours operation, Run of River facilities run for 24 hours |
| Months | List of integers separated with comma that correspond with season column. |

## Population

The population tab includes information on the starting population, which can be modeled as a fixed number of obligate migrants, or as an entrainment event sampled from a statistical distribution that describes the adjusted entrainment rate in units of . The number of iterations is the number of times Stryke will simulate the species-discharge scenario. If your study is not using the EPRI (1997) entrainment database, fish length is modeled with a normal distribution with parameters for mean length and standard deviation (units in inches) and the end user must enter a Starting Population size.

If the end user is using the EPRI entrainment database to also inform on the frequency and magnitude of entrainment events, then run the script epri\_database.py. This script assists with describing the presence rate of fish in entrainment samples and the magnitude of entrainment events. The only parameters required for the end user to pass onto Stryke are database query commands.

The EPRI database can be queried many different ways. Note, all function arguments are optional named arguments, meaning the end user doesn't have to query the database at all. In this instance the returned distribution parameters will be representative of the entire dataset.

A list of all function arguments is found in the table below.

|  |  |
| --- | --- |
| Species | list like object of Scientific Species names or single string object |
| Genus | list like object of multiple Genus names or string object |
| Family | list like object of Family names or single string object |
| Feeding | List like object of abbreviated feeding guilds or single string object. Possible guilds include:  CA = carnivore  FF = filter feeder  HE = herbivore  IC = insectivorous cyprinid  IN = invertivore  OM = omnivore  PR = parasite |
| Habitat | List like object of abbreviated habitat types or single string object. Possible habitat types include:  BEN = benthic  BFS = benthic fluvial specialist  FS = fluvial specialist  Lit = littoral (near cover/shorelines)  Pel = pelagic  Pool = pool (minnows)  RP = run/pool (minnows)  RRP = riffle/run/pool (minnows) |
| HUC02 (Hydrologic Unit Code) | List like object of hydrologic units. Possible hydrologic units include:  1 = New England  2 = Hudson River and Mid Atlantic Drainage  3 = Southeast United States  4 = Great Lakes and St. Lawrence  5 = Ohio River  7 = Mississippi River  19 = Pacific Northwest |
| Month | List like object of month numbers of user defined seasons. |

The script then returns a figure to assist with entrainment event distribution selection, prints the presence rate, entrainment event distribution parameters, and length distribution parameters. The end user then enters

## Nodes and Edges

The Nodes tab identifies physical locations within the migratory network, while the Edges tab describes the possible routes between them. Together, Stryke uses these two pages to create a directed acyclic graph. That means direction is one-way with edges always describing the from and to nodes in that order. Please note, Stryke requires specific keywords for locations (Table XX). On the node tab, you will also note columns Surv\_Fun and Survival. For those rows indicating turbine locations, Surv\_Fun should list the type of turbine while Survival remains empty. For all other locations, Surv\_Fun should remain “a priori” while Survival is a float describing the survival probability at that location within the migratory network. The spelling of the Edges must match the Nodes tab. All weights should default to 1.0. Weights are adjusted in the simulation according to the daily discharge.

|  |  |
| --- | --- |
| Location Keyword | Notes |
| river\_node\_0, river\_node\_1, …, river\_node\_n | Stretches of river between reaches affected by hydropower projects. |
| forebay | Forebay |
| U1, U2, …, Un | Turbines |
| spill | Spillway |
| tailrace | Tailrace |

## Unit Params

The Unit Params tab contains information on each individual unit within the powerhouse(s).

## Python

Python stuff

## Model Interpretation