Production Workflow

Derek H. Ogle

Data Preparations

General (from Original Files)

1. Load files in 'data/original/', prep as described in `Data_Prepper.R`, and output prepped files to 'data/prepped/'. Once properly prepped with the final original data, `Data_Prepper'R` should not need to be run again. A simple log of each run of `Data_Prepper.R` is saved in 'data/prepped/dataPrepper_logs/'. These logs should be carefully examined each time new prepped files are made to compare processes to see if anything obviously went awry.

Computing All Weight-Length Regressions

1. Loaded prepped `len_wt.csv` file.
2. Computed all regresions and extracted coefficients (loga and b), sample size (n), coefficient of determination (r^2), and range of lengths by
   a. Each WBIC_YEAR.
   b. Each WBIC.
   c. Each lake class.
   d. All fish regardless of any classification.
3. Combined all regressions into a single data.frame.
4. Created a `use` variable that is set to `yes` if the regression is "valid" to use and `NO` if it is not valid to use. A regression was considered valid to use if it met the following criteria (see code in `calcLWRegs.R` for specific criteria).
   a. Sample size was above some minimum threshold.
   b. r^2 was above some minimum threshold.
   c. b was between 2 and 4 (Froese (2006) showed empirically that most MEAN (by species) values of b were between 2.5 and 3.5. Individual values would likely be a little wider; thus, the value set here.)
   d. **QUESTION** – Do we want to include a criterion based on the range of lengths in the regression. There are a handful of regressions with a fairly small range of values (~8 regression, of those remaining after above, with a range less than 200 mm).
5. Output this data.frame to 'data/prepped/' as `LWRegs.csv` so that this work would not have to be run each time and so that the weight-length regressions used could be more easily observed.

Computing All Age-Length Keys

1. Age-Length Key
    1. Use WBIC_YEAR ALK
    2. If not, then use WBIC ALK
    3. If not, then use regional ALK (east and west in ceded territory)
    4. Do not worry about sex at this point
    5. Come up with some sort of cutoff for what will be considered a useful ALK (something like less than 30 ... but come up with a rationale for the 30 ... try doing we expect this many per 0.5 inch category times this many possible categories on average).
    6. Decide whether to use modeled (multinomial) or completely empirical ALK. [LIKELY THE WAY TO GO]
    7. After this reduce to just age-3 fish.
2. Weight-Length Regression
    1. Use WBIC_YEAR W-L
    2. If not, then use WBIC W-L
    3. If not, then use regional W-L (east and west in ceded territory)
    4. Do not worry about sex at this point
    5. Cutoff may be something like reasonable sample size across reasonable range of lengths with a certain r^2 value.
    6. Estimate a weight for each length in the fmdb fish.

Follow-Up

- Analysis
- Kobe Plots