20-March-2017

Below are step-wise descriptions of example data sets and Program R code for calculating secondary production for taxa without distinguishable cohort structure. This approach borrows instantaneous growth rates from taxa with distinguishable cohorts to calculate secondary production for taxa without distinguishable cohorts. The primary assumption of this method is that instantaneous growth rates reflect the actual instantaneous growth rates of the taxa without distinguishable cohorts.

- Email questions, comments, and ideas on how to improve either the analysis itself or the methodological description to Michael Venarsky (mvenarsky@gmail.com). Meaningful contributions will be credited to the person.

Data files

IGR_no_cohort_v1.csv description:

- DATE = sampling date
- HABITAT = type of habitat samples. Pool is depositional habitat dominated by fine substrates. Riff is erosional habitat dominated by large cobbles and boulders. Wood is sticks and logs submerged in the stream.
- ID_IGR = a unique number used to link this data table to IGR_v1.
- GENUS = name assigned to conglomerate group.
- 1 to 23 = Size classes (mm). Note that the largest individual is 13mm. The excess size classes (14 to 23) will be removed by the code during secondary production analyses. These columns contain biomass (mg dry mass) per unit area (m^{-2}).
- Each line is a replicate. Note that replicates are NOT uniquely identified (e.g., Pool-1, Pool-2, etc.). Doing this will cause the code to return a "resampled data table" that is identical to the original data set.

IGR_v1.csv description:

- ID_IGR = a unique number used to link this data table to IGR_no_cohort_v1.csv.
- ID_GENUS = name assigned to the average instantaneous growth rate for the conglomerate group
- IGR = average instantaneous growth rate calculated from taxa with distinguishable cohorts.
- INTERCEPT = average intercept for instantaneous growth rate calculated from taxa with distinguishable cohorts.

Calculating secondary production

2oProd_IGR_no_cohort_20Mar2017.R:

Pre-analysis setup:

- Check code 20-March-2017
- Constructed using R version 3.3.1 (2016-06-21) -- "Bug in Your Hair". Copyright (C) 2016 The R Foundation for Statistical Computing Platform: x86_64-w64-mingw32/x64 (64-bit).
- Michael Venarsky (mvenarsky@gmail.com) wrote code using dplyr, tidyr, lubridate, and zoo packages on 28-March-2015. Several stackoverflow.com users provided assistance to Michael Venarsky. See questions by stackoverflow user Vesuccio.
- For those not familiar with dplyr or tidyr, consider loading the swirl package which is an interactive tutorial for dplyr and tidyr implemented in R studio.
- Links to dplyr and tidyr cheat sheets:
  - http://www.rstudio.com/wp-content/uploads/2015/02/data-wrangling-cheatsheet.pdf
  - https://www.rstudio.com/resources/cheatsheets/
- Include three folders your working directory: data, results, check. CSV files will be written to each of these folders. Note that CSV files are written to the "check" folder only when you delete the "#" at the beginning of the line of code.
- Clear all data from previous runs, load required packages, set working directory and import all data files.

Step 1

- Resample_function_biomass and Resample_function_growth resample respective data sets with replacement.

Step 2

- N_resample defines the number of resampling events. We conducted 1000.
- Beginning of Loop 1, which is a very long loop. Loop 1 begins by running the Resample_function_biomass and Resample_function_growth functions.
- Size_class and Loop 2 remove size classes with 0 individuals from Resample_biomass.

Step 3

- Habitat-specific secondary production. Calculations in this step are HABITAT-SPECIFIC.
- Step 3.1 - - Calculates average biomass per size class per taxa per habitat per date.

- Step 3.2 – Calculates total biomass per taxa per habitat per date and then calculates days elapsed between sampling dates.
- Step 3.3 – Calculates interval secondary production (PINT) by averaging biomass between sampling dates then multiplying by the randomly sampled instantaneous daily growth rate (Resample_growth; IGR) and the number of days between sampling dates (DAYSB). Units are mg dry mass m^{-2} d^{-1}
- Step 3.4 – Calculates total secondary production (P) and mean biomass (B) per HABITAT

Step 4

- N_resample defines the number of times Loop 1 is run. We conducted 1000. If the N_resample is 1, the output from the analyses are exported. If the N_resample is >1, then the results from each resampling are sequentially assembled into a table and the exported as a CSV file.

Step 5

- Calculate 95% quantiles.