

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 production : biomass ratios generated from taxa with distinguishable cohorts to calculate secondary production for taxa without distinguishable cohorts. The primary assumption of this method is that the borrowed production : biomass ratios reflect the actual production : biomass ratios of the taxa without distinguishable cohorts.

Note that this estimate of secondary production does not carry a unit of time (e.g., daily, monthly, or yearly) because it is calculated using a production : biomass ratio (which does not carry a unit of time) rather than a growth rate (which carries a unit of time; daily, monthly, or yearly). You will need to divide the number from this calculation by the unit of time over which the production : biomass ratio was calculated. For example, if the production : biomass ratio was estimated over 3 months, then you could divide by 3 to estimate mg dry mass m⁻² mo⁻¹ or by 90 to estimate mg dry mass m⁻² d⁻¹.

- 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

PB_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 PB_v1.csv.
- GENUS = name assigned to taxa
- TOTBIO = the total biomass (mg dry mass m⁻²) contained in each replicate sample.
- 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.

PB_v1.csv description:

- ID_IGR = a unique number used to link this data table to PB_no_cohort_v1.csv.
- ID_TAXA = name assigned to the average production : biomass ratio for the conglomerate group
- AVG_PB = average production : biomass ratio calculated from all taxa with distinguishable cohorts

Calculating secondary production

2oProd_PB_no_cohort_v1.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 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_PB resample respective data sets with replacement.

Step 2

- N_resample defines the number of resampling events. We conducted 1000.
- Loop 1 begins by running the Resample_function_biomass and Resample_function_PB functions.

Step 3

- Habitat-specific secondary production. Calculations in this step are HABITAT-SPECIFIC.
- Step 3.1 - Calculates average biomass per taxa per habitat.
- Step 3.2 – Calculates total secondary production (P) by multiplying by the randomly sampled PB ratio (Resample_PB; AVG_PB) and the average biomass per taxa. Note that this estimate of secondary production does not carry a unit of time (e.g., daily,

monthly, or yearly) because it is calculated using a production : biomass ratio (which does not carry a unit of time) rather than a growth rate (which carries a unit of time; daily, monthly, or yearly). You will need to divide the number from this calculation by the unit of time over which the production : biomass ratio was calculated. For example, if the production : biomass ratio was estimated over 3 months, then you could divide by 3 to estimate mg dry mass $\text{m}^{-2} \text{mo}^{-1}$ or by 90 to estimate mg dry mass $\text{m}^{-2} \text{d}^{-1}$.

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.