

The Scoring Process

Author: James Develle, King County
Updated: 2/21/2014

Introduction

This article describes the steps taken to generate scores from sampling data on the Puget Sound Stream Benthos website <https://pugetsoundstreambenthos.org>. This does not discuss specific metrics and B-IBI calculation which is covered elsewhere (see <https://pugetsoundstreambenthos.org/About-BIBI.aspx>). Rather, this article covers the steps before metrics are calculated (filtering visits, combining sample replicates, optionally subsampling, applying taxonomic filters), and the steps after the metrics are calculated (aggregating and presenting scores in various formats).

Outline

Retrieving Sampling Data.....	2
Applying Filters.....	2
Hydrologic Areas	2
Projects	2
Keywords.....	2
Sample Non-QC and QC Replicates.....	3
Taxa	3
Number of Organisms.....	3
Date Filters.....	3
Chironomid Resolution	3
Processing Each Visit.....	4
Subsampling.....	4
Replicate Handling	4
Average replicates' overall scores.....	4
Average replicates' quantities, calculate scores	4
Combine replicates, then calculate scores (recommended)	4
Processing Taxonomic Composition	4
Excluding Taxa.....	4
Taxonomic Resolution: Rolling Up To Coarser Taxonomic Ranks.....	5
Determining Taxonomic Attributes	5
Considering Uniqueness	5
Counting and Calculating Metrics.....	6
Averaging Replicates Metric Quantities.....	6
Calculating Metric Scores.....	6
Averaging Replicates Metric Scores.....	6
Calculating an Overall Score	6
Aggregating Visit Scores.....	7
Specifying a Hydrologic Aggregation Level	7
Clustering Visits by Hydrologic Level	7
Averaging During Aggregation and Clustering.....	7
Resulting Formats	7

Retrieving Sampling Data

The first stage of creating scores is to get all the released sampling data that matches the specified criteria.

Applying Filters

Hydrologic Areas

The hydrologic area (also called just “area”) can be any name specified in the hydrologic hierarchy of streams. Streams are organized by stream name, sub-basin, basin, and WRIA. By default, the area specified is “All Streams.”

Rivers are also included in this list. By design, river data are never returned in the same query as stream data, and they are not part of the hydrologic hierarchy of streams. There is also an “All Rivers” option.

Sites need not have a standardized stream or river: A site may have a stream entered as free form text (rivers are never entered this way) or it may temporarily not have a stream at all. Data from these sites are returned **only** when the “All Streams” default option is used.

Projects

Result may be filtered by a specified project, or all projects of a specified agency. Some projects and agencies are not designated for public access. If a user wants to query private data, he or she must log in to select the private project. Once logged in, the user can select and view any project in his or her agency, but still cannot see private data from other agencies.

The system is designed such that private data and public data are never combined. If an agency has both public and private projects, then all the data from the public projects may be returned together, but data from private projects must be queried one project at a time.

Keywords

A list of keywords may be entered; if a site matches any of the keywords listed, its data are returned. The * wildcard is allowed at the end of the word. Adding the word “not” removes the matches from the results.

The fields searched are the following:

- City
- County
- Site name
- Site code

Additionally, some labeled searches are allowed. These are typically not specified in a list because here the results must match all the criteria; wildcards and the “not” word are not applicable. These are examples of labeled searches:

- sample-code: HylbFW99_2005R1
- sample-id: 1132
- visit-id: 999
- site-id: 549
- sample-tag: MyTag

Sample Non-QC and QC Replicates

Normally QC replicates and deficient samples are omitted, however, scores of the QC replicates are available through downloads and the visit details page.

Taxa

The user may specify taxa (up to 5) that must exist at a visit by visit at the visit. The results will omit visits that don't have all the specified taxa, and for taxa specified with “not,” it will omit visits that have any of those.

A taxon will match even if it is a finer level than the specified filter; for example, if the user filters for Genus G, its species G a, G b, etc., will also match.

Number of Organisms

Samples may be flagged or filtered by the number of organisms counted. When applying these thresholds at visit level, all the non-QC, non-deficient sample replicates of the visit are combined and the filter is applied to the total number of organisms. When applying at sample level, only those samples whose counts are not within the thresholds are omitted.

Date Filters

Visits can be filtered by year. Along with an optional range of years, one of the following two options can be applied.

- When the user specifies “All [Visits per Site]”, all visits occurring in each year in the range are included in the results, including when multiple visits occurred for the same site in the same year.
- When the user specifies “Latest per Site”, only the latest visit at each site for each year is included in the results: For each site and year, there is only one visit. This filter is required when the requested format requires one visit per site per year (as when plotting a map, tabulating by year, or charting). This option is automatically selected for such formats.

Chironomid Resolution

You can filter the results by the way the lab counted chironomids during its analysis of a visit's samples. This filter is applied at the visit level, not the sample level.

Processing Each Visit

Subsampling

Samples can have a wide range in their number of organisms counted. This range can be the result of differing counting protocols (e.g., using a Caton subsampler) or because of environmental conditions reducing the total number of organisms. To better compare samples with differing counting protocols, samples with many organisms can be randomly subsampled to smaller subsamples, with fewer organisms.

Subsampling can be done sample-by-sample or visit-by-visit (after all the samples in a visit are combined). Combining all samples in a visit is the best option when the replicate handling is combining all replicates at each visit before calculating (discussed below).

The randomization of organisms is stored in the database so that a specific subsampling query will return the same results each and every time.

Replicate Handling

When there are multiple samples replicated within a visit, those replicates are aggregated to produce a single visit score in one of three ways. (Note QC replicates are never included.)

Average replicates' overall scores

Each replicate is processed independently, and then a final, single visit score is created. Each final sub-score is based on the average of the respective sub-scores of the replicates. The final overall score is based on the average of the overall scores of replicates.

Average replicates' quantities, calculate scores

Each replicate is processed independently, and then a final, single visit score is created. The average of the replicate quantities is used to calculate each sub-score. The final overall score is based on the sum of the final sub-scores.

Combine replicates, then calculate scores (recommended)

First all the replicates are combined, and then quantities are counted and sub-scores and the overall score are calculated. This is the default.

Processing Taxonomic Composition

Each taxon record is further processed in the following stages before it is counted.

Excluding Taxa

The taxon is used to look up a matching record in a table of excluded taxa. If a match is found in the table, the searching stops. The record has a field that specified whether or not the taxon is to be excluded from scoring. Usually the exclusion record specifies that the taxon is to be excluded from scoring, but in some instances it is explicitly specifies to include the taxon in scoring.

- If an exclusion record is found and it specifies the taxon is to be excluded, the original taxon ignored and processing skips to the next taxon in the sample.

- If no entry is found in this process, or an exclusion record is found and specifies the taxon is to be included, the original taxon is passed to the next stage or processing.

Taxonomic Resolution: Rolling Up To Coarser Taxonomic Ranks

The Taxonomic Resolution that the user requests will specify a single taxonomic resolution table to use in this next step. This is also referred to as “Standard Taxonomic Effort” or “STE.”

The taxon is used to look up an entry in specified table of taxonomic resolutions, using each coarser rank until an entry is found. This process is exactly like the recursive look-up used for Excluding Taxa.

- If an entry is found, it specifies the taxon rank and TSN to use for metric calculation of the scoring process for that row. Usually the rank is the same rank as where the match occurred, but occasionally it is a coarser rank.
- If an entry is not found, the original taxon used for metric calculation of the score processing.

Determining Taxonomic Attributes

The taxonomic attributes (characteristics of the organism) are loaded from a source table specified by the user. The taxon resulting from the previous step is used to look up a set of attributes associated with this taxon, using each coarser rank until an entry is found. This process is exactly like the recursive look-up used for Excluding Taxa.

Considering Uniqueness

When a lab analyst identifies a taxon in a sample and records it in the system, it may flag the record as unique (which is the default) or not unique within the sample. When interpreting this field in the context of scoring, it is important to consider several issues:

- Labs may ignore this field and thus it will always be set to unique.
- Even though a taxon is marked as unique, it may not be: there may be another entry of that taxon within the same sample. For example, one entry could be marked as adult and another as pupa. By default, if they are the same taxon, only one will be counted as unique, unless both are marked as not unique. If both are marked as not unique, then neither will be counted as unique.
- In the process of scoring, the given taxon may be rolled up to a coarser taxonomic rank, and then be grouped with other records.
- Replicates may be combined within a visit, and what was unique for a sample may not be unique for all the sample replicates from a site visit.
- When subsampling is specified, organisms are randomly selected to meet the subsample threshold with a bias towards the entries marked unique, i.e., all unique organisms will be selected before non-unique organisms. This segregation allows uniqueness to be counted correctly.

A more accurate (and awkward) definition is that, this field determines whether or not the taxon “may be unique” when calculating diversity metrics:

- If “Is Unique” is true, it may be unique. The taxon is considered unique and increases diversity only if no other such taxon has been yet found within the sample or visit, with respect to the replicate handling, after any statistical subsampling is applied.
- If “Is Unique” is false, it is never considered unique and never increases diversity.

To make subsampling work with respect to the “Is Unique” flag, the random selection of organisms is adjusted slightly such that all organisms from the unique taxa are selected before the organisms from non-unique taxa.

Counting and Calculating Metrics

The taxon resulting from the taxonomic rollup step, its taxon attributes, and its “Unique” flag are used for calculating quantities and scores for the metrics. The metrics are best described in these two web pages:

- <https://pugetsoundstreambenthos.org/About-BIBI.aspx>.
- <https://pugetsoundstreambenthos.org/BIBI-Scoring-Types.aspx>

In 2013, a project was dedicated to analyzing and updating the B-IBI. You can read about the B-IBI Recalibration project and its results here:

- <https://pugetsoundstreambenthos.org/Projects/BIBI-Recalibration.aspx>

Averaging Replicates Metric Quantities

Replicate quantities may be averaged at after all the replicates of a visit are processed. See the Replicate Handling section for details.

Calculating Metric Scores

All the metrics that are specified for the selected Score Type are mapped to score values using the metrics’ quantities. The quantities tested against ranges specified for the given score type and metric, and the associated score is used.

Averaging Replicates Metric Scores

See the Replicate Handling section.

Calculating an Overall Score

The overall score is the sum of each metric score, with respect to the specified Replicate Handling.

Aggregating Visit Scores

Visits may be aggregated by year and/or hydrologic level (WRIA, Basin, Sub-Basin, Stream, or River.) The user may specify one of these levels. Also, scores are “cluster” aggregated to limit the number of scores that are plotted on the map <https://pugetsoundstreambenthos.org/Biotic-Integrity-Map.aspx>.

Specifying a Hydrologic Aggregation Level

When one of the hydrologic levels is specified for aggregation, then only river or stream data will be included, with respect to the user choice. Visits that have no standardized stream or river name specified will be excluded; these are the sites where the stream field is blank or entered as free-form text.

Clustering Visits by Hydrologic Level

Clustering is used to reduce the total number of scores shown on the map, thereby making the map less crowded, more evenly distributed, and minimizing page load times. To these ends, the areas with more scores are aggregated to a coarser hydrologic level as needed to reduce the number of marker to be plotted.

Clustering may be turned off with an option on the bottom of the map page. This will show all the scores, but your browser may have script timeout issues.

Unlike user-specified aggregate levels, visits may be cluster aggregated to varying levels, or not aggregated at all, within the same result set. It may also include sites where the stream is blank or entered as free-form text; these scores will not be aggregated, but will still be displayed.

If the user specifies an aggregation level, clustering may still be needed to reduce the number of score plotted on the map. In this case, cluster aggregation happens after the user specified aggregation has completed; scores will be aggregated at the user specified level or at a coarser level.

Averaging During Aggregation and Clustering

After quantities, sub-scores, and overall scores are created for all the visits in a given hydrologic area, a final set of values is created for the area. Each final sub-score is based on the average of the respective sub-scores of the visit in that hydrologic area. The final overall score is based on the average of the overall scores of replicates. Likewise with quantities, final metric quantities are based on averages of the visits.

Resulting Formats

Scores are present on color-coded maps, tables, and for trends, charts and interactive tables. There are various ways to download the scores to use them in other applications and databases.

The interactive by-year pivot table (see “Tabulate Trend” on the website) can be exported (downloaded) in PDF and Excel formats. These are tables with locations in row headers, and years in column headers. A raw data pivot table is also available in the download screen (see “Download,” then “Download by Year.”)

A year-by-year chart (see “Chart Trend” on the website) provides a more graphic view of the same time series data. These are standard PNG-format images that can be copied and downloaded from within any browser.

From the download screen, you can download a table of scores with more data attributes than found elsewhere on the website. There are also options for including scores generated from QC replicates.

Download Format

While the downloads are intended to be imported (or pasted) into MS Excel, MS Access and other database software, the details below may make working with the downloaded data easier.

For all download options, the first row contains the column headers, and the rest of the rows are the data. The columns are comma delimited.

If the scores downloaded are visit scores (not aggregates), that score table can be joined to a sample table on the surrogate key VisitID (as opposed to the natural key composed of Agency, Site Code, and Visit Date). Sites can be identified with the surrogate key SiteID (as opposed to the natural key composed of Agency, and Site Code). These surrogate key provide a convenient way for databases to relate tables.