

Replicate Data: QC & Sample Replicates

DRAFT: March 28, 2013

Jo Opdyke Wilhelm, King County Department of Natural Resources and Parks

QC Reps

Quality Control replicates (QC reps) are sampled at up to 10% of sites for several projects with data stored in the Puget Sound Stream Benthos data management system (PSSB) (Table 1). In most cases, these samples are replicates collected on the same date as the primary sample to compare within site variability (error variance defined as the differences in B-IBI observed for samples taken on the same day at a site). I believe there are some projects that don't collect their QC reps on the same date, instead revisiting the site within a few weeks to collect a replicate sample. Unfortunately, the difference in methods for QC reps is not well documented or readily available. If this is a critical piece of information, we will have to reach out to some of our partners in the region.

Table 1. QC Rep data available in the PSSB as of March 27, 2013.

Agency	Project	2001	2002	2003	2005	2006	2007	2008	2009	2010	2011	2012	Total
Adopt-A-Stream	Jim Creek							1					1
Lake Forest Park	Benthic Invertebrates							1					1
Clallam Co	Streamkeepers						1						1
King Co - DNRP	Ambient		13	10	18	16	18	13	19	14	14		135
King Co - DNRP	Des Moines Ck Habitat									3			3
King Co - DNRP	WRIA08 WS Survey									5	5	5	15
King Co - Roads	CIP Support							1					1
King Co - Roads	ESA Water Quality						4	3	4	4			15
Kitsap Co	Stream Team	2	4	4									10
Kitsap Co	Navy's Envvest			3									3
Kitsap Co	Watershed Health										5		5
WA Ecology	Ambient Freshwater										1		1
WA Ecology	Deschutes Effectiveness										1		1
Grand Total		2	17	17	18	16	23	19	23	26	26	5	192

Four projects (King Co Ambient, WRIA 8 WS Surveys, King Co Roads ESA WQ, and Kitsap Co Stream Team) account for over 90% of the available QC data (Table 1). I thought Snohomish and Pierce Counties collected QC reps, however if they do those data are not stored in the PSSB or are not coded correctly to enable download. We could follow up with Snohomish and Pierce if more data are desired.

Most of the QC Rep data (>83%) are from projects that composite sample replicates for each site visit and from a total surface area of 3 ft² (Table 2).

Table 2. Number of site visits with QC rep data for different collection methods by total surface area collected.

Total Sq Ft	2001	2002	2003	2005	2006	2007	2008	2009	2010	2011	2012	Total
2 ¹							1					1
3	2	13	13	18	16	22	18	23	21	14		160
8									5	12	5	22
9		4	4			1						9
Total	2	17	17	18	16	23	19	23	26	26	5	192

[B-IBI data](#) were generated and subsequently downloaded from the PSSB with the following selected:

- QC Samples: Only events with QC samples
- Replicate handling: sum replicates' quantities, then calculate scores
- Taxa Attributes: Fore, Wisseman (2012)
- Taxa Resolution: Resolution used by lab
- Organisms per Visit: At most 500, subsampled when over
- Generation Time: Tuesday, March 26, 2013 4:11 PM

These data will be used for variability analysis such as signal to noise ratios (S:N) or minimal detectable differences (MDD). B-IBI scores are highly correlated ($r = 0.8828$, $p < 0.001$) for QC reps and the original replicate data (Figure 1).

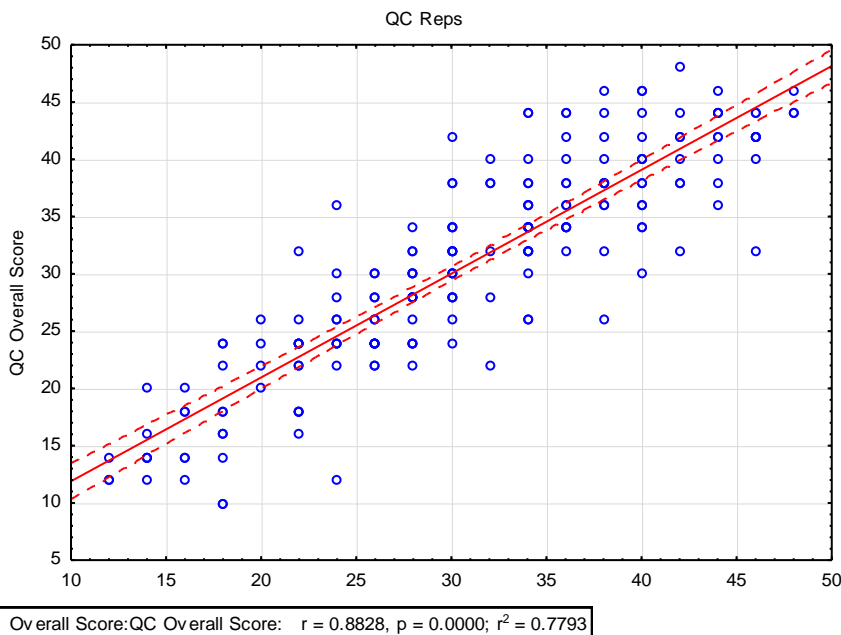


Figure 1. Overall B-IBI score versus QC rep B-IBI score.

Downloaded data can be found in [this spreadsheet](#).

¹ This two square foot sample may be mis-entered in the PSSB. It is for Lake Forest Park site [LyonLFPS35th for 2008](#). Three 1-sq ft samples were collected. It looks like rep 1 is labeled as a QC rep, and rep 2 and 3 are not. This likely should all just be part of a non-QC rep sample. I will follow up with the Lake Forest Park Streamkeepers and try to correct this information if it is indeed a mistake.

Sample Replicates

Some collection methods specify keeping individual sample replicates separate with multiple samples being sent to a taxonomic laboratory for each site visit. The most common collection protocols with sample replicates are summarized in Table 3 and include sampling from 3-, 8-, or 9- total ft². Some data indicate 2-, 6-, or 16- total ft² sampled, but these are likely data entry errors or represent cases where one replicate was lost or invalidated for some reason.

Table 3. Collection methods and surface area sampled for data in the PSSB with sample replicates.

Total Ft ²	# Site Visits	Method Summary	Agencies or Projects
3	588	3 1-ft ² samples	Federal Way (2002-08), Kirkland, Lake Forest Park, Kitsap Co (1998-2003), Pierce Co, King Co: UPD, Biosolids, CAO
8	51	4 2-ft ² samples	Bellingham
9	350	3 3-ft ² samples	Bainbridge, Bellevue, Federal Way (2009-onward), Seattle (2006), Clallam (2002-2007), Thurston County, Kitsap Co (2006)

B-IBI score calculation is influenced by replicate handling, which has three options on the PSSB:

- Sum replicates' quantities, then calculate scores (SumRQ)
- Average replicates' quantities, then calculate scores (AvgRQ)
- Average replicates' overall scores, then calculate scores (AvgR)

Replicate handling is also influenced by the type of subsampling selected on the PSSB:

- Sample: each individual sample is first subsampled to the desired count (500S)
- Visit: taxa for the entire visit are subsampled to the desired count; therefore when sample replicates are present each sample will likely have variable representation (500V)
- None: data used as reported in the PSSB with no sub sampling (NoSS)

To compare the influence of these different calculation options, nine different combinations were downloaded (Table 4).

Table 4. Data were downloaded from the PSSB nine different ways. Codes used in the spreadsheet are presented here for each of the 9 download combinations.

Treatment	Sample	Visit	No Subsample
Sum reps	SumRQ_500S	SumRQ_500V	SumRQ_NoSS
Avg reps qty	AvgRQ_500S	AvgRQ_500V	AvgRQ_500NoSS
Avg reps	AvgR_500S	AvgR_500V	AvgR_NoSS

B-IBI scores from various replicate and subsampling combinations are highly correlated, but they do appear to influence B-IBI scores (Figure 2). Summing the replicates' quantities yields higher B-IBI scores than averaging the replicates' scores or quantities. Subsampling by visit when averaging replicates' scores or quantities (AvgR_500V and AvgRQ_500V) results in lower B-IBI scores.

Scatterplot of multiple variables against SumR_500V
Reps_Combined 9v*1007c

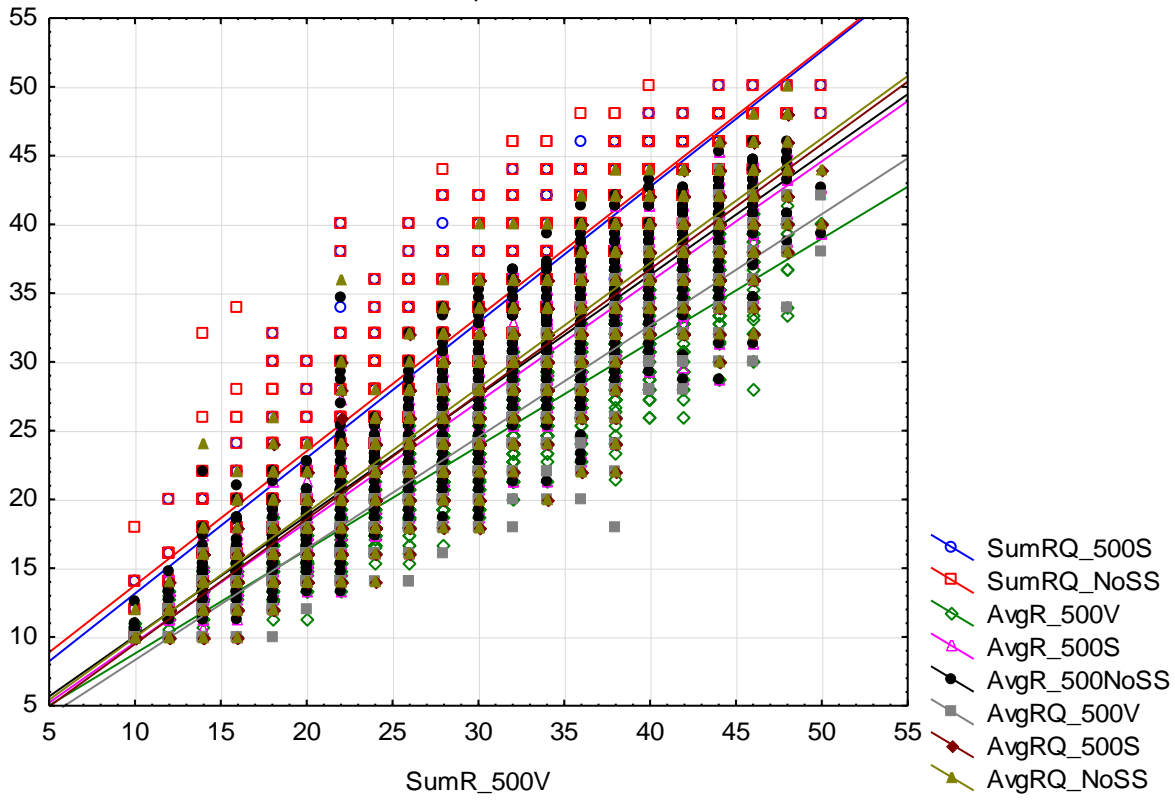


Figure 1. Overall B-IBI score based on replicate and subsampling handling.

Further analyses will be done to determine whether B-IBI scoring criteria need to be adjusted to account for replicate handling. Downloaded data can be found in [this spreadsheet](#).