Stream Restoration Monitoring: Discussion (2017)

Keith Eshleman (UMCES)

Outline

- Question(s) to be answered through monitoring
- Monitoring design issues
- The "nitty-gritty" (why, what, when, where, and how)

The "Nitty-Gritty" (why, what, when, where, and how)

- *Why* monitor?
  - To provide instantaneous measures of pollutant concentration (C) and discharge (Q) necessary for estimating **pollutant loads** ($L = \Sigma CQ$) and load reductions ($\hat{L}_2 - L_2$)
- *What* to measure?
  - Stage & discharge; rating curves; TSS; N; P; ancillary data (turbidity)?
- *When* to measure?
  - Baseflow & stormflow conditions
- *Where* to measure?
  - Upstream and downstream ends of reach(es)?
- *How* (and how frequently) to measure?

How (and how frequently) to measure?

- Summary
  - We discussed the following options:
    1) Hourly sampling ($$$$) or by deploying *in situ* probes for some constituents, e.g., turbidity, to reduce costs;
    2) Low frequency, e.g., weekly or bi-weekly fixed time-based sampling ($); and
    3) Hybrid sampling regimes, e.g., low-frequency for baseflow and high-frequency for targeted stormflow events ($$).
  - We also discussed errors in estimating loads, but stopped short of recommending specific designs that would be applied universally.

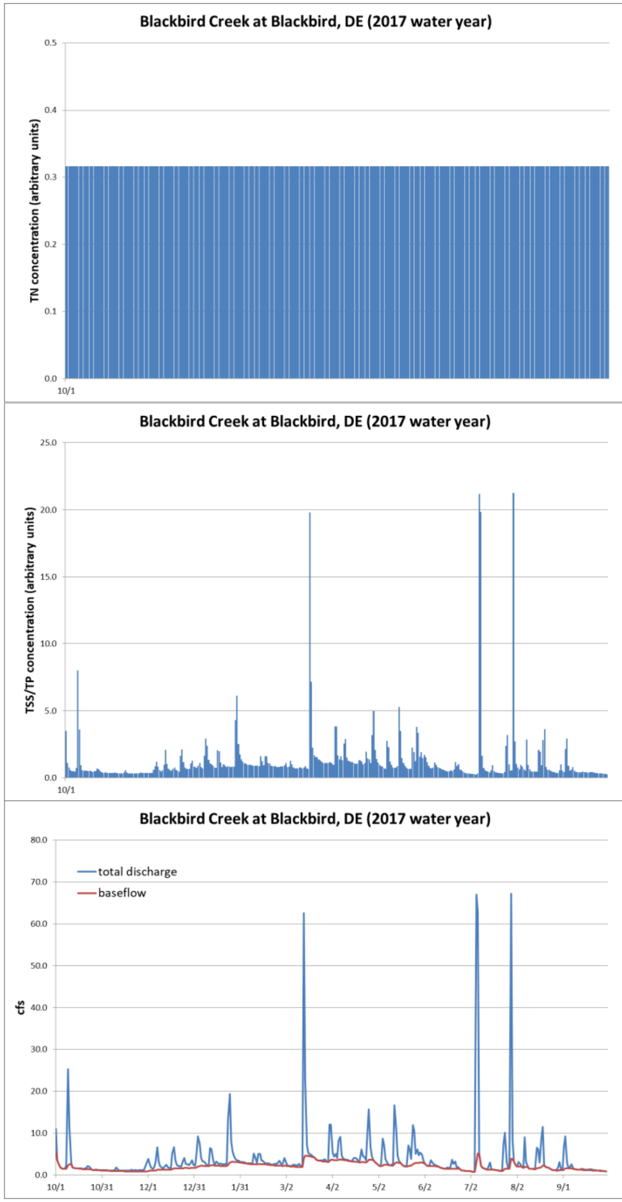
Follow-up questions:

- What are the primary sources of error in load estimation?
  - Biases
  - Random errors
- Does the magnitude of the error vary as a function of the constituent being measured?
- Given a focus on TSS, TP, and TN, how frequently should we sample in the Sassafrass River watershed?

Sources of Error

- Gaging errors: in development of *rating curves*, individual discharge measurements are likely only accurate to ±10% (USGS)
- Rating curves based on a large number of measurements will tend to average out random errors, however
- Gaging errors include various potential *biases*, however:
  - Inability to measure the mean vertical velocity in a channel
  - Presence of channel ice (cold climates only)
  - Shifts in rating curve due to channel scouring or deposition
  - Submergence of a critical depth meter
- Load bias will be directly proportional to gaging bias
- Significant load errors resulting from inability to adequately sample *concentrations*

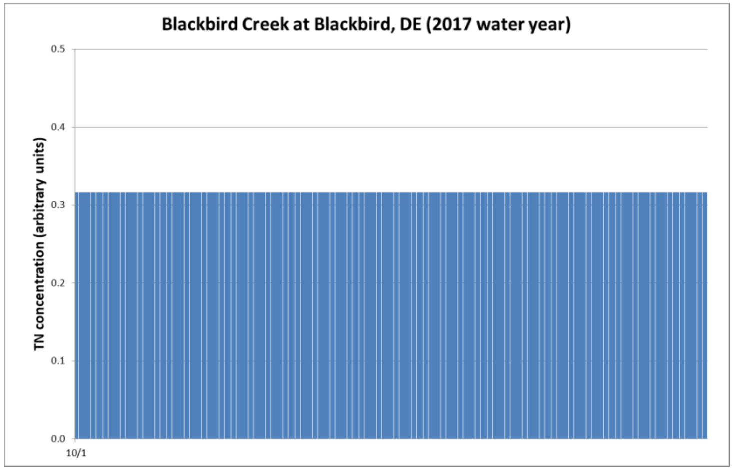
Concentration Errors

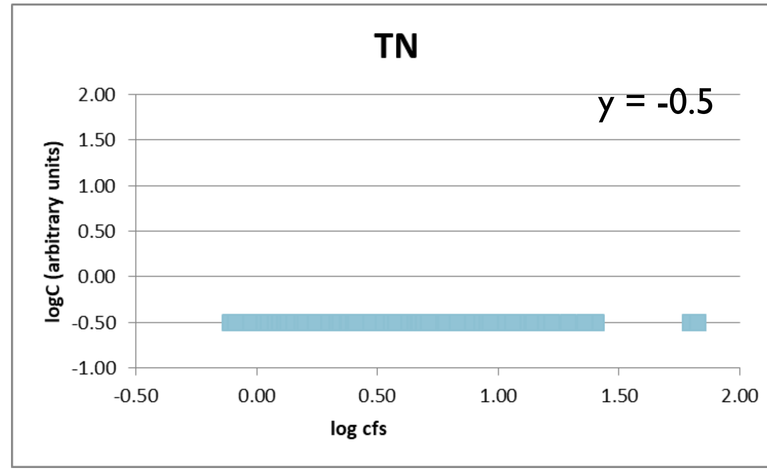


- Uniform distribution of concentrations across a wide variation in stream discharge = "chemostatic"
- Distribution of concentration is highly skewed and varies as a function of discharge = "chemodynamic"
  - Can exhibit either "concentration" (positive slope) or "dilution" (negative slope) behavior
- Normally based on graph of *logC* vs *logQ*

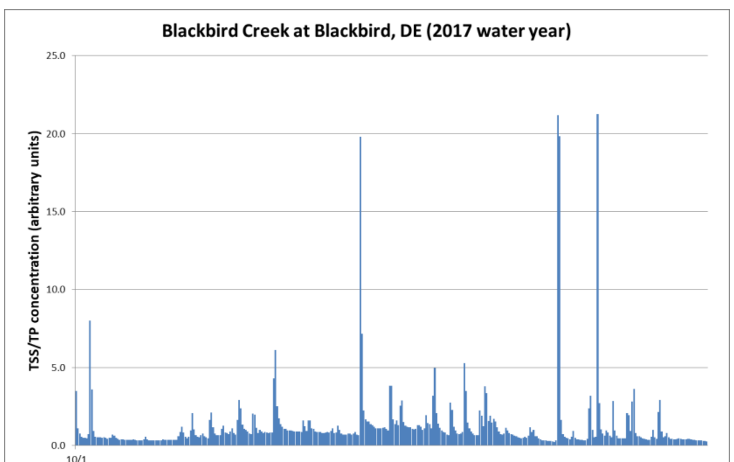
Concentration Errors

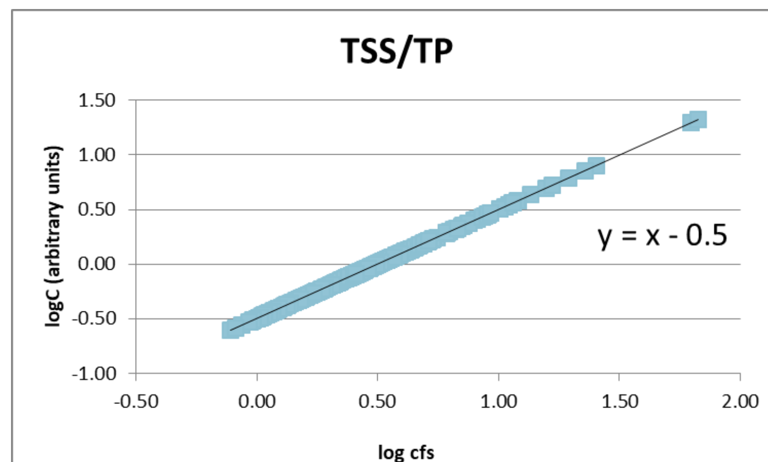
Chemostatic case



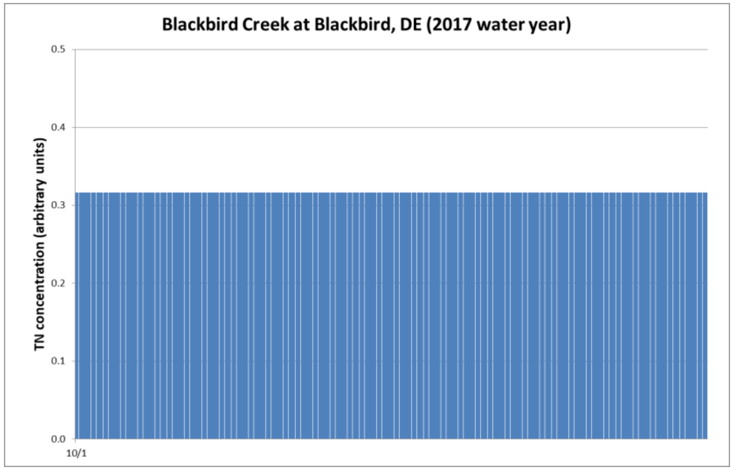


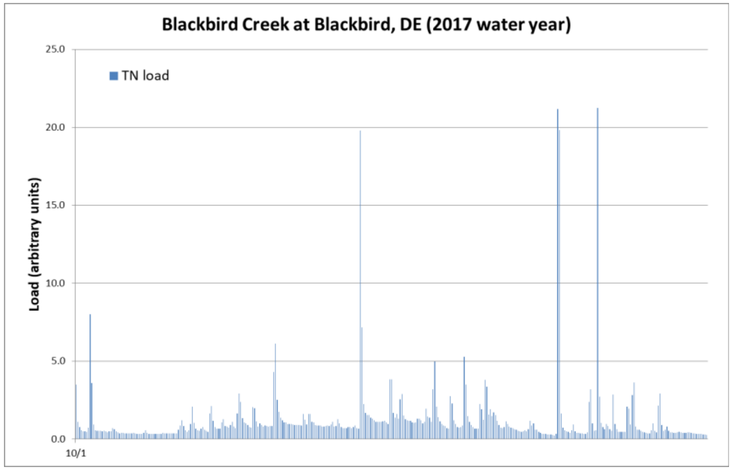
Chemodynamic case



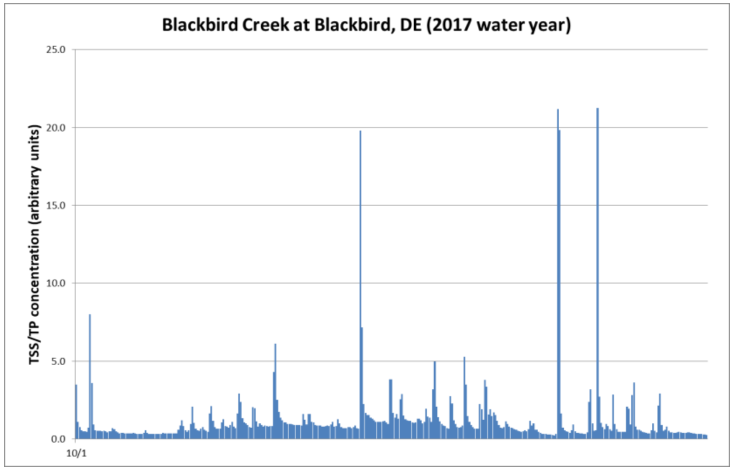


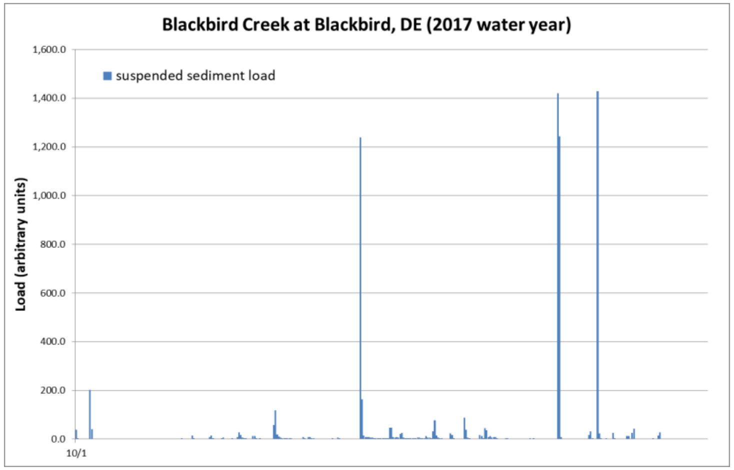
Chemostatic case





Chemodynamic case





Method

- "Sampled" the *measured* discharge and *synthetic* TSS/TP and TN concentration time series on monthly, biweekly, and weekly basis
- Estimated annual loads from the sample data from:

$$L = \frac{\sum_{i=1}^{n}(C_i Q_i)(Q_A)}{\sum_{i=1}^{n} Q_i}$$

where C_i and Q_i are instantaneous concentration and discharge values measured at time *i*, and Q_A is the mean discharge for the estimation period

- Computed mean load, bias (% of actual load), and relative mean standard deviation (RMSD, %) for both constituents for all three sampling frequencies

Results

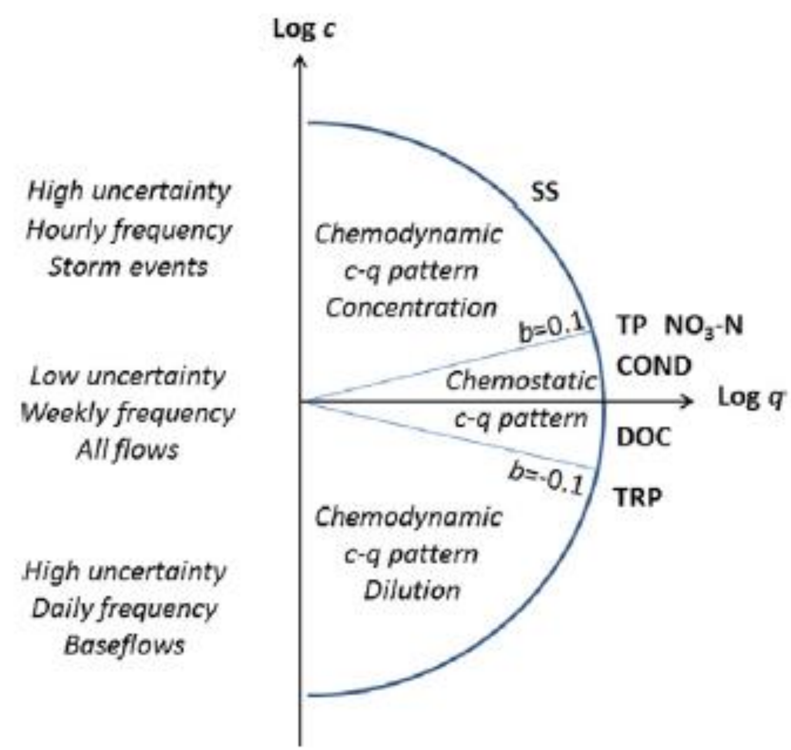
Estimator	TSS/TP			TN		
	M	BW	W	M	BW	W
True load	7598	7598	7598	455	455	455
Est. load	4898	6087	5726	455	455	455
Bias (%)	-36	-20	-25	0	0	0
RMSD (%)	87	46	42	0	0	0

- As expected, errors associated with loads of constituents that are *perfectly chemostatic* are strictly gaging errors
- For *chemodynamically-positive* constituents:
    - Routine sampling produces *negative bias* in load estimates and *random errors* that are likely too large for assessment purposes
    - RMSD's were reduced somewhat by higher frequency sampling, however
- What about "real" constituents (*logC* vs. *logQ* slopes between 0 and 1)? See 2018 paper by MZ Bieroza *et al.* (*STOTEN 630*:738-749)

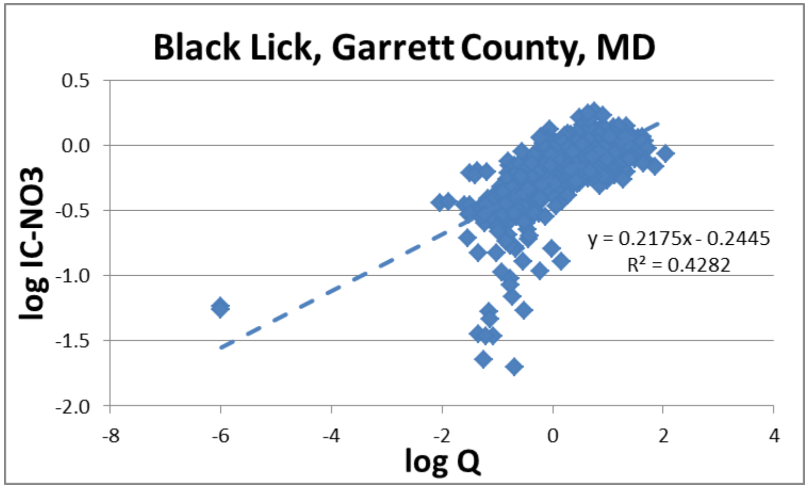
MZ Bieroza *et al.* (2018)

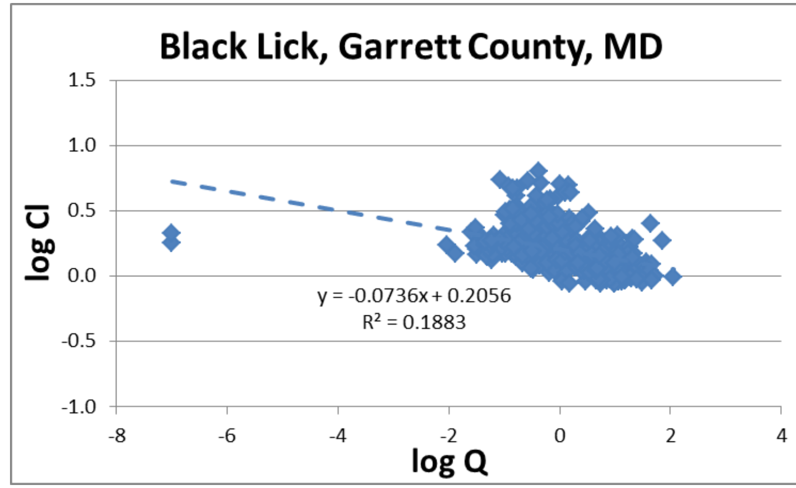
- Similar (but far more elegant and complete) statistical analysis using *real* high- and low-frequency water quality data from gaged agricultural watersheds in Europe
- Relatively large suite of water quality constituents
- Findings:
  - Uncertainty in load estimates increases with reduced sampling frequency as a function of the *logC-logQ* slope
  - As predicted, some constituents are highly chemostatic (e.g,. conductivity) while others are highly chemodynamic (e.g., TSS)

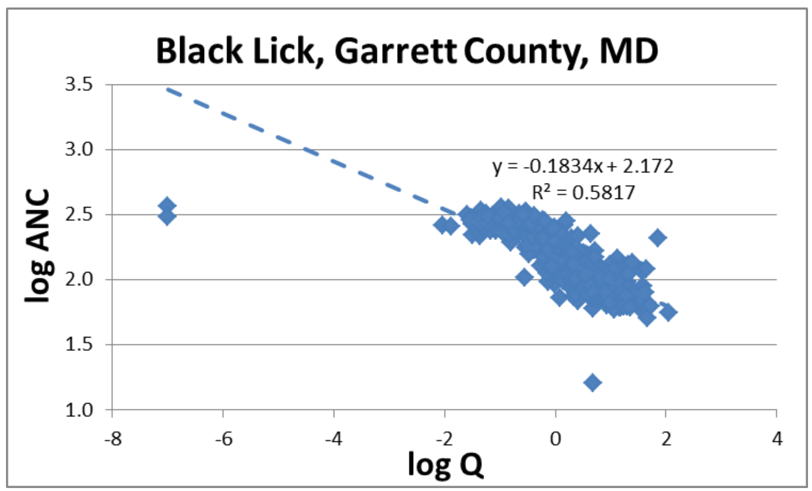
GRAPHICAL ABSTRACT



Unpublished Data (Eshleman *et al.*)







Conclusions

- For chemodynamically-positive constituents like many (e.g., TSS, TP) that we are interested in reducing through stream restoration and watershed management, sampling should principally target stormflow conditions and be performed at ~hourly sampling frequency
- For ~chemostatic constituents (e.g., nitrate, conductivity, etc.), sampling can be performed at a much lower frequency across all flows
- Therefore, a *hybrid sampling design* was proposed that involved a combination of low-frequency (weekly or bi-weekly) sampling and targeted stormflow sampling relying on programmable automatic samplers
- Such as design would be reliable, while reducing sampling costs, lab analysis costs, random errors, and biases to a reasonable extent