Load Estimation from Monitoring Data

TETRA TECH

complex world | CLEAR SOLUTIONS™

Introduction

Why measure pollutant load?

- Load drives impacts in receiving waters
- Best metric of source significance

  Which is the more significant source:

  - A WWTP discharging 10 mgd at 0.08 mg/L TP?
  - An agricultural watershed draining an average of 150 ft^3/sec at 0.8 mg/L TP?
- Integrates changes in response to both flow and concentration
- TMDL

TETRA TECH

Introduction

Monitoring for accurate load estimation is demanding and is not a trivial task that can be done as an afterthought:

- Continuous flow measurement
- High and carefully-designed sampling frequency to capture variability and important times of year
- The sampling regime needed for load estimation must be established in the initial monitoring design, based on quantitative statements of the precision required for the load estimate to meet project goals
- Load estimation is a statistical process and different monitoring and analysis procedures can yield very different load estimates
- Quarterly grab sampling will not give acceptable results

The Basics

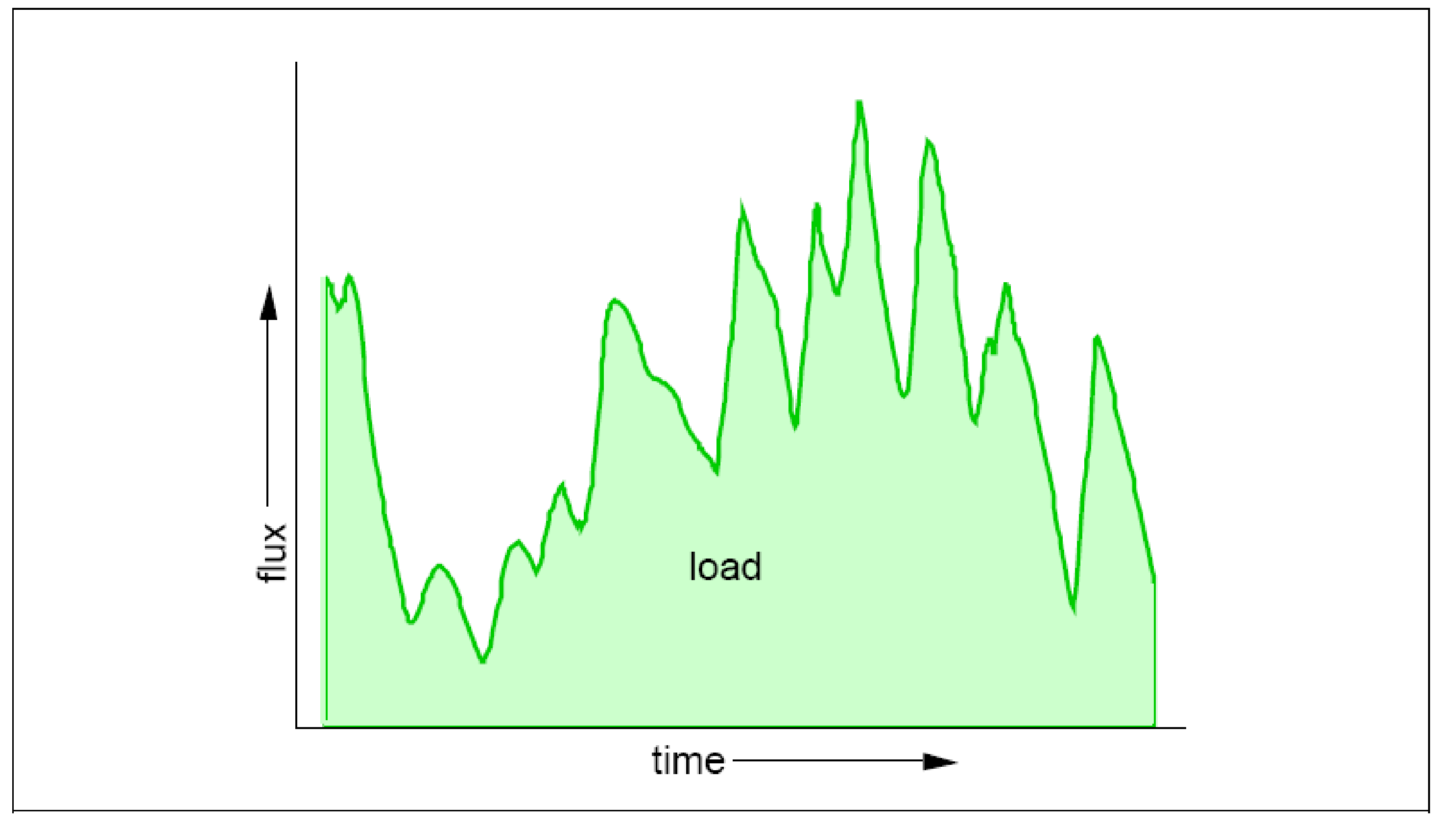
Flow = the instantaneous rate at which water is passing the reference point

Discharge = the volume of water that passes a cross-section of the river in a specific amount of time

Flux = the instantaneous rate at which the load is passing a point of reference on a river, e.g., a sampling station

Load = the mass or weight of pollutant that passes a cross-section of the river in a specific amount of time

The Basics



The Basics

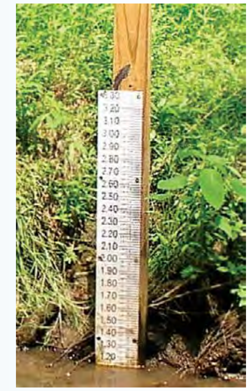
However, cannot measure flux directly, so calculate load as product of concentration and flow:

$$\text{Load} = k\int_t c(t)q(t)dt$$

Because we must almost always measure concentration in a series of discrete samples, estimation of load becomes sum of a set of products of flow and concentration:

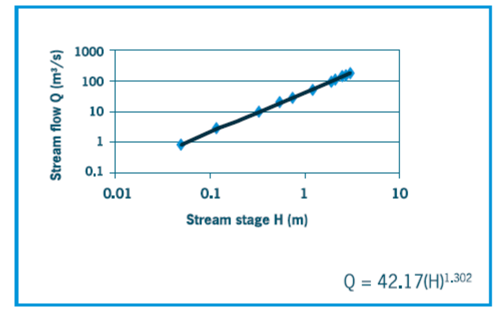
$$\text{Load} = k\sum_{i=1}^{n} c_i q_i Dt$$

Sources of Information



Flow data

Measure it yourself



- *Surface Water Flow Measurement for Water Quality Monitoring Projects* http://www.bae.ncsu.edu/programs/extension/wqg/319monitoring/TechNotes/technote3_surface_flow.pdf
- USGS *Techniques of Water Resource Investigation* *http://pubs.usgs.gov/twri*
- USDI Bureau of Land Reclamation. 2001. *Water Measurement Manual.* *http://www.usbr.gov/pmts/hydraulics_lab/pubs/wmm/*

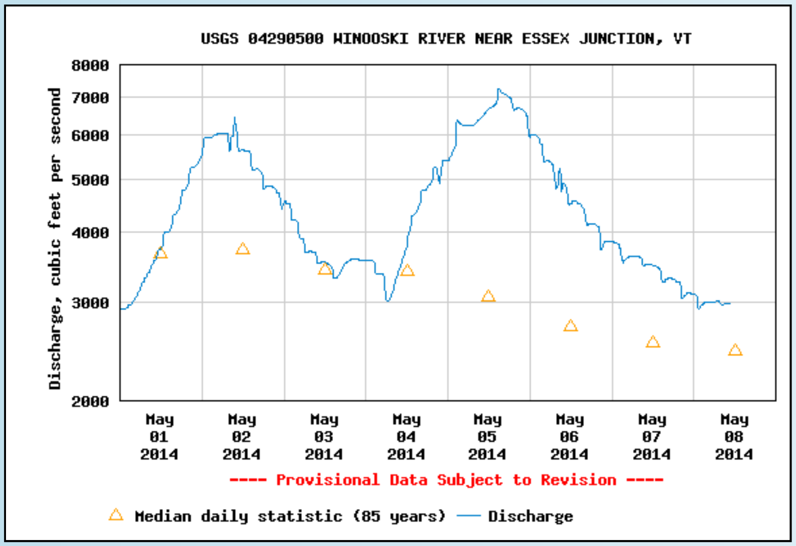
Sources of Information

Flow data

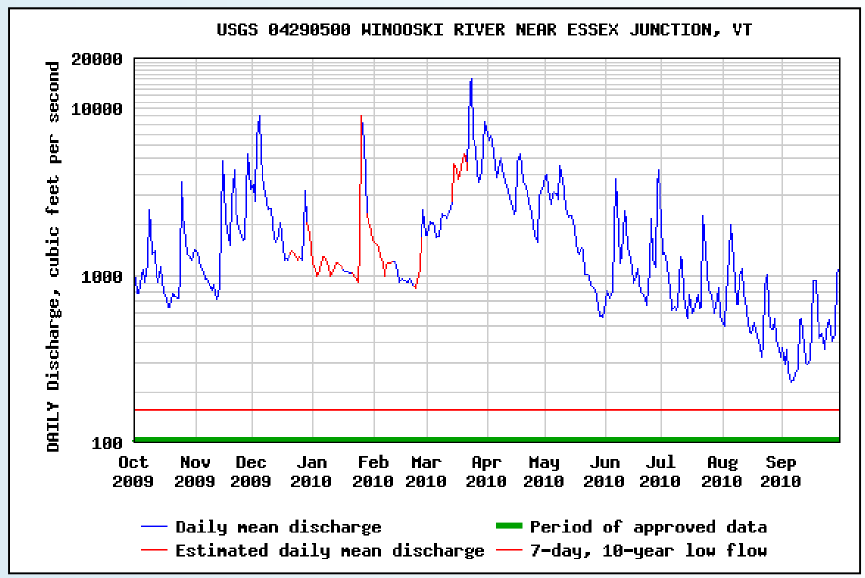
USGS National Streamflow Information Program (NSIP)

http://water.usgs.gov/nsip/

Real time data



Historical data



TETRA TECH

Sources of Information

- ► Concentration data
  - ▪ Project monitoring
  - ▪ Ongoing agency monitoring
  - ▪ Compliance/permit

Basic computations

At the most basic: load = concentration x flow

e.g.,

daily TP load = mean daily [TP] x mean daily Q

$= \text{mg/L} \times \text{ft}^3\text{/s} \times 28.32\ \text{L/ft}^3 \times 8640\ \text{s/day} \times 1\ \text{kg}/10^6\ \text{mg}$

$= 0.345\ \text{mg/L} \times 198\ \text{ft}^3\text{/sec}$

$= 16.7\ \text{kg/day}$

Can do the same procedure for daily, weekly, monthly data

Basic computations

- ► **Total load** is the load over the main period of interest, e.g., one year

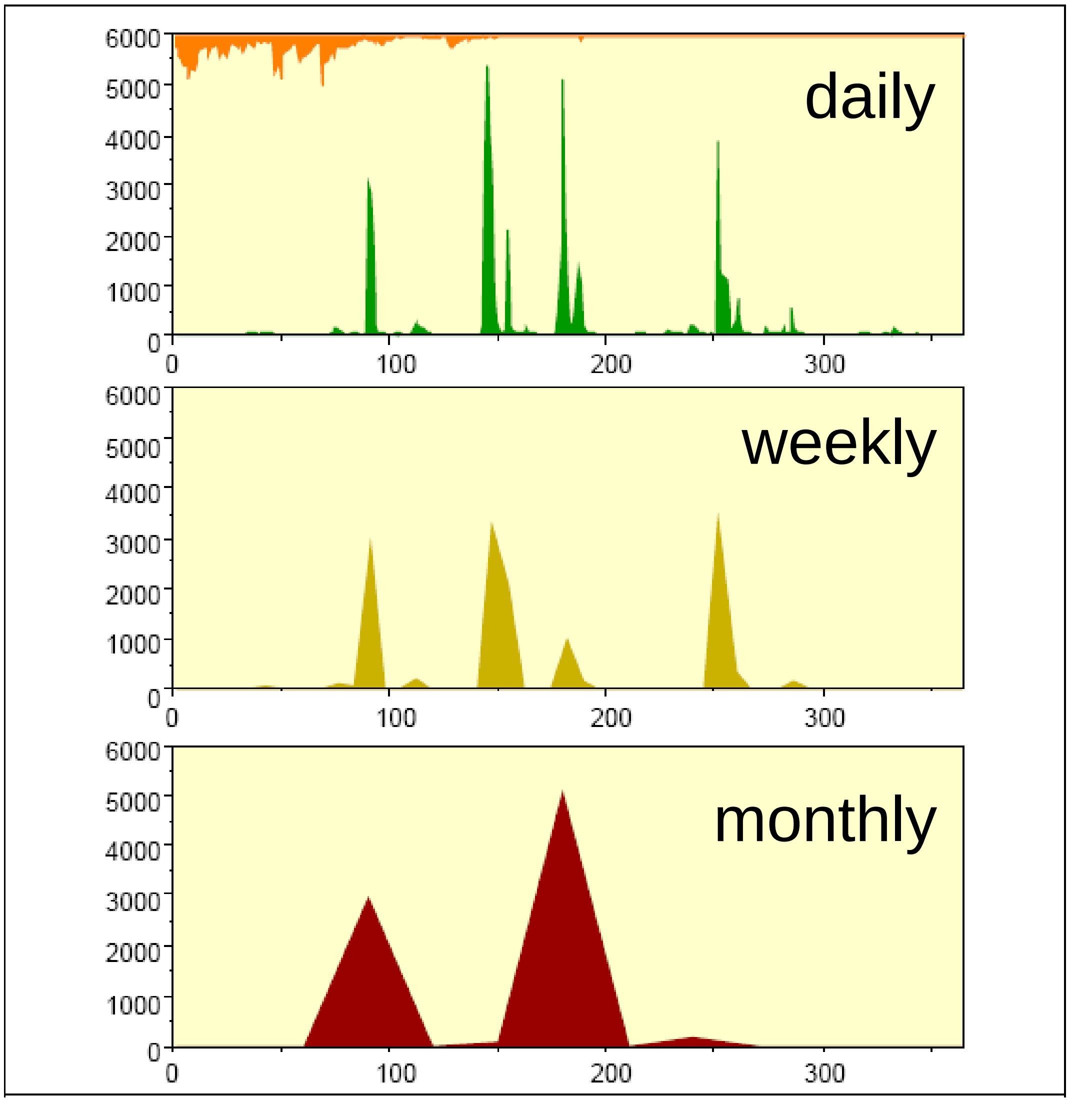
  is represented as the sum of

- ► **Unit loads**, i.e., individual calculations of load as product of concentration and flow over a smaller, more homogeneous time span.

The central problem is to accurately characterize all the unit loads;

adding them up to the total load is simple.

Basic computations

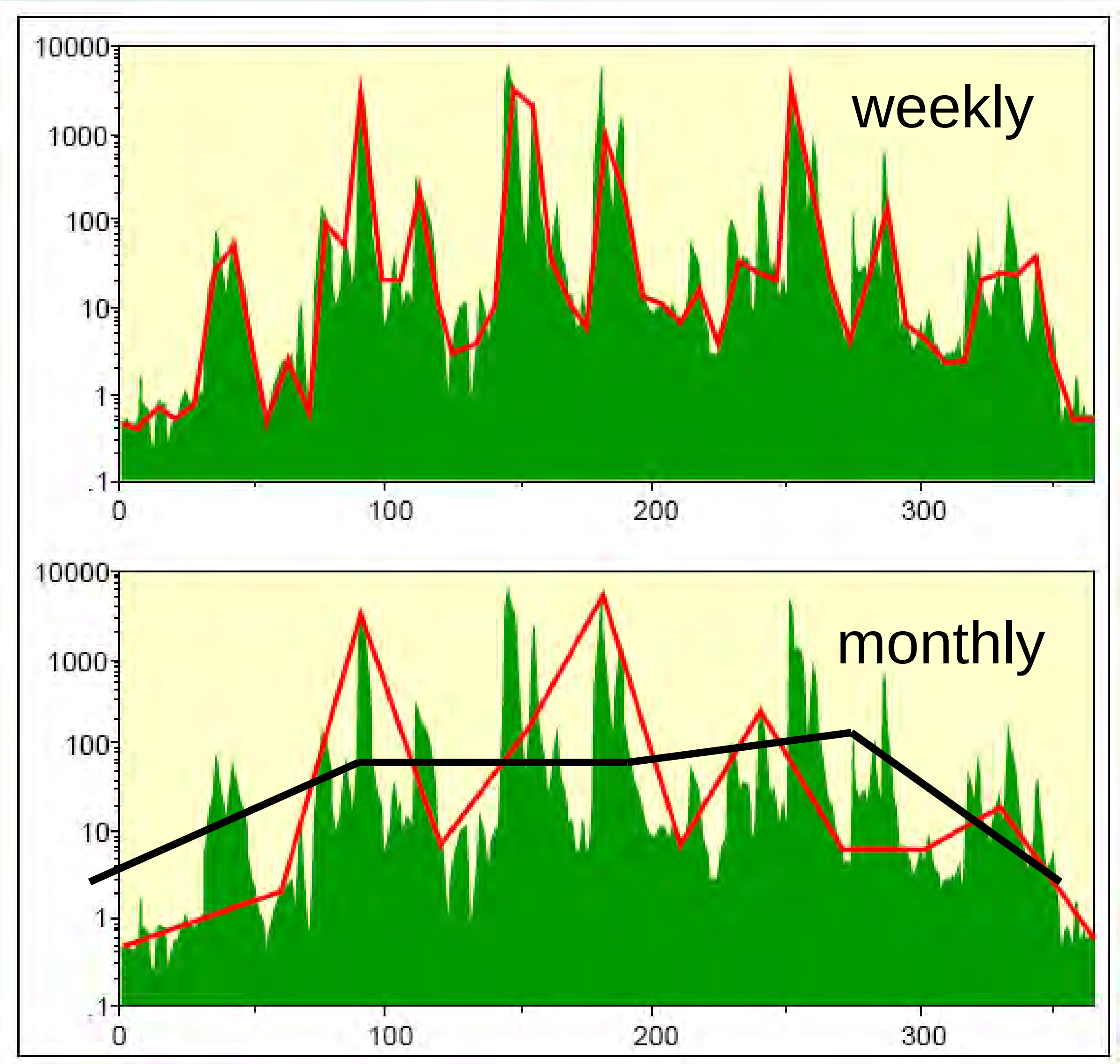


The central problem becomes how best to set up the discrete samples to capture complete information about pollutant concentration and give the most accurate estimate of load.

how many samples
and
when to take them

TETRA TECH

Basic computations



Because in NPS, most flux occurs during periods of high discharge (~80 – 90% of annual load in ~10 – 20% of time), when to sample is especially important.

Quarterly X

Monthly X

Weekly ?

TETRA TECH

Practical load estimation

Ideally, most accurate approach to load estimation is to measure flow continuously, sample frequently, and capture all the variability.

Flow is relatively straightforward to measure, but must be measured continuously and with acceptable accuracy → occasional instantaneous measurements or measurements at the time of grab sampling are not sufficient

Practical load estimation

Concentration is expensive to measure and in most cases impossible to measure continuously.

Must choose a sampling interval to give an appropriate characterization of concentration component.

Practical load estimation

In general, the accuracy and precision of a load estimate increases as sampling frequency increases.

Sample frequency determines the number of unit load estimates that go into our total load estimate and more unit loads mean we are more likely to capture variability across the year and not miss an important event

Because of autocorrelation, at some point, greater sample frequency will not improve load estimate

Practical load estimation

Minimum Detectable Change analysis

The minimum change in a pollutant concentration or load over a given period of time required to be considered statistically significant.

A function of

- Variability in concentration, flow, etc.
- Sampling frequency and duration

Practical load estimation

$$MDC = t_{(n_{pre} + n_{post} - 2)} \sqrt{\frac{MSE}{n_{pre}} + \frac{MSE}{n_{post}}}$$

Where: $t_{(n_{pre} + n_{post} - 2)}$ = one-sided Student's t-value with $(n_{pre} + n_{post} - 2)$ degrees of freedom.

$n_{pre} + n_{post}$ = the combined number of samples in the pre- and post-BMP periods

$s_{(\bar{X}pre - \bar{X}post)}$ = estimated standard error of the difference between the mean values in the pre- and the post-BMP periods.

$MSE = s_p^2$ = Estimate of the pooled Mean Square Error (MSE) or, equivalently, variance (s_p^2) within each period. The MSE estimate is obtained from the output of a statistical analysis using a t-test or ANCOVA with appropriate time series

http://www.bae.ncsu.edu/programs/extension/wqg/319monitoring/TechNotes/technote7_MDC.pdf

Practical load estimation

Grab samples – represent concentration at a single point in time

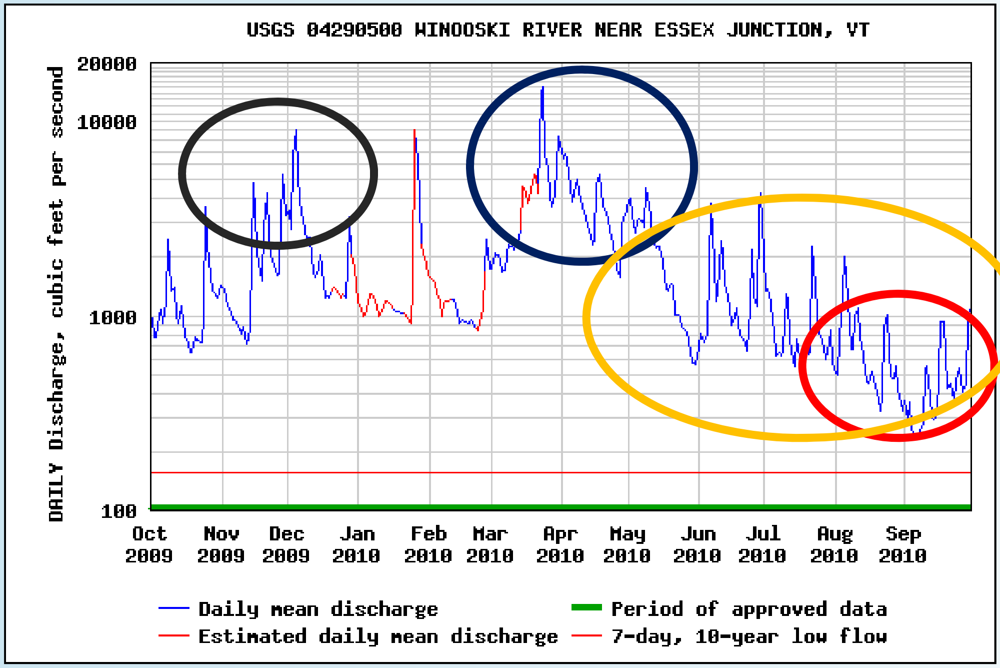
Fixed-interval (time-proportional) samples – poorly suited for load estimation because they ignore changes in flow that occur between samples and are usually biased toward low flows

Flow-proportional samples - ideally suited for load estimation, can provide a precise and accurate load estimate if the entire time interval is properly sampled.

TETRA TECH

Practical load estimation

- Timing of samples more complex than frequency
- Selecting when to collect samples for concentration determination = selecting when the unit loads that go into an annual load estimation are determined
- Consider sources of variability, e.g., season, flow, agricultural activities



Practical load estimation

1. Find a way to estimate "missing" concentrations to go with the flows observed at times when chemical samples were not taken.

2. Abandon most of the flow data and calculate the load using the concentration data and just those flows observed at the same time the samples were taken. X

3. Do something in between - find some way to use the more detailed knowledge of flow to adjust the load estimated from matched pairs of concentration and flow.

Practical load estimation

When decision to calculate loads is made after monitoring program is in place or data collected, little can be done to compensate for a data set that contains too few observations collected using an inappropriate sampling design

The sampling needed for load estimation must be established in the initial monitoring design, based on quantitative statements of the precision required for the load estimate to meet project goals.

Practical load estimation

- Is load estimation necessary or can project goals be met using concentration data – can you evaluate project effects using concentration data?
- Determine precision needed in load estimates – don't try to document a 25% load reduction from a BMP program with a monitoring program that may give load estimates ±50% of the true load.
- Decide what approach will be used to calculate the loads, based on known or expected attributes of the data.
- Use the precision goals to calculate the sampling frequency and timing requirements for the monitoring program.
- Compare ongoing load estimates with program goals and adjust the sampling program if necessary.

Practical load estimation

Someone may say that it's too expensive or complex to conduct a monitoring program sufficient to obtain good load estimates.

Is a biased, highly uncertain load estimate preferable to no load estimate at all?

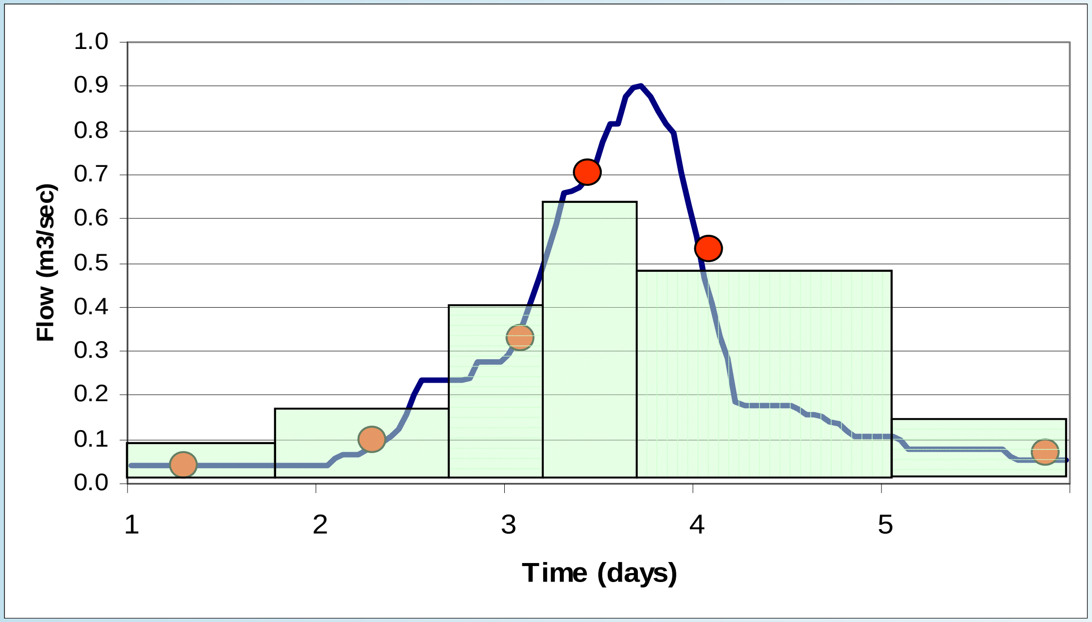
Approaches to Load Estimation

Numeric integration

$$\text{Load} = \sum_{i=1}^{n} c_i q_i t_i$$

c_i = concentration of i[th] sample
q_i = corresponding flow
t_i = time interval represented by i[th] sample

$$\frac{1}{2}(t_{i+1} - t_{i-1})$$



Approaches to Load Estimation

Numeric integration

Question becomes how fine to slice the pie – few slices will miss much variability, many slices will capture variability but at a higher cost/effort.

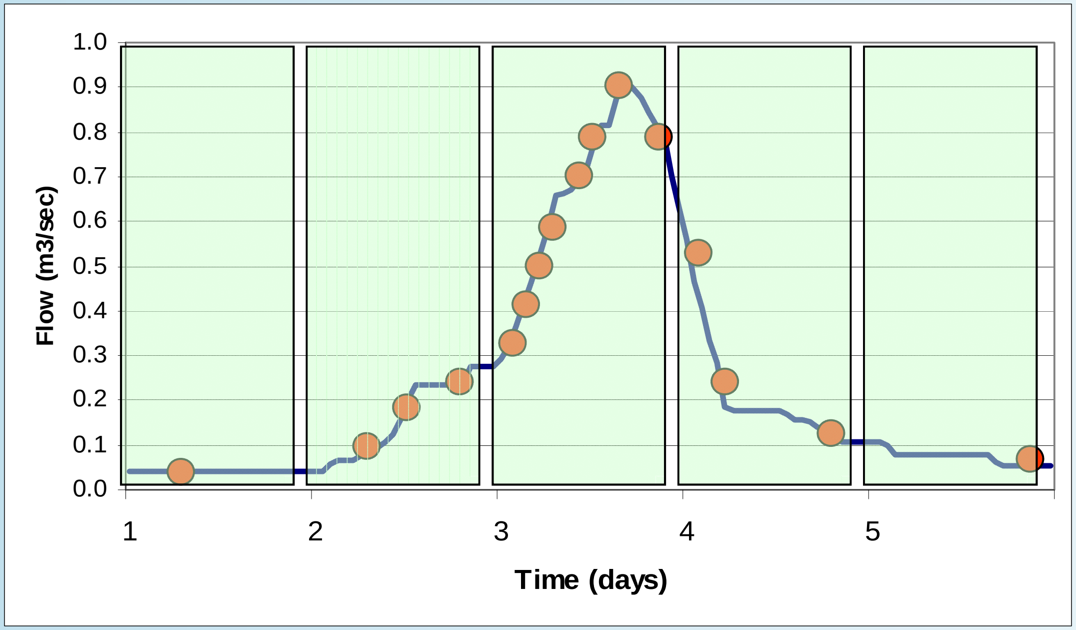
Numeric integration is only satisfactory if the sampling frequency is high - often on the order of 100 samples per year or more, and sufficiently frequent that all major runoff events are well sampled.

Selection of sample frequency and distribution over the year is critical – must focus on times when highest fluxes occur, i.e., periods of high discharge

Approaches to Load Estimation

Numeric integration

Flow-proportional sampling



Very efficient and cost-effective method of obtaining total load.

Requires reliable equipment and careful attention

No information available at resolution less than chosen period

Not compatible with other goals, such as monitoring for ambient concentrations that are highest at low flow

Daily load

Approaches to Load Estimation

Regression

Regression relationship developed between concentration and mean daily flow, based on the days on which samples are obtained.

Regression relationship used to estimate concentrations for each day on which a sample was not taken, based on mean daily flow.

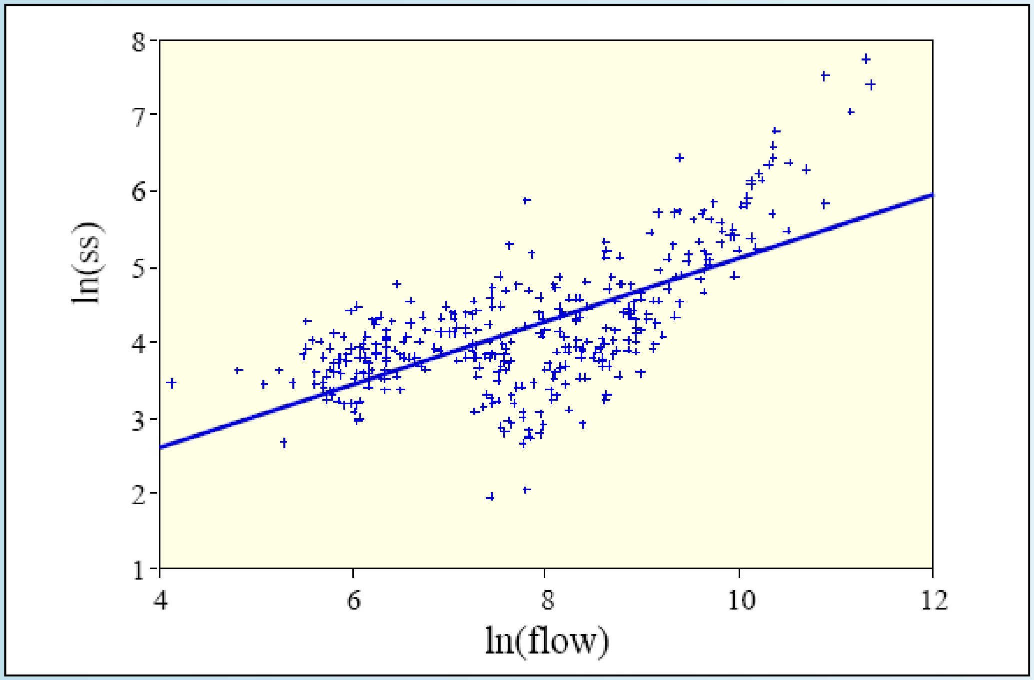
The total load is calculated as the sum of the daily loads, obtained by multiplying the measured or estimated concentration by the flow

Goal of chemical sampling becomes one to thoroughly characterize the relationship between flow and concentration. May be able to do this with ~20 samples a year, focusing on high-flow or critical season events

This approach is based on the past and may not necessarily be an accurate or appropriate approach to predicting the future or filling gaps

Approaches to Load Estimation

Regression



Cautions:

- Possible bias in back-transforming results if using log-log regression
- Must obtain statistically significant flow-conc. relationship, with residuals randomly distributed
- Must pay attention to potential changes or trends in flow-conc. relationships – especially where BMPs or other changing land management may influence
- Must manage sampling program to effectively capture range of flows/conditions – using data from fixed-interval time-based sampling is not appropriate

Approaches to Load Estimation

Ratio Estimators

On days on which samples are taken, the daily load is calculated as the product of concentration and flow, and the mean of these loads is also calculated.

The mean daily load is then adjusted by multiplying it by a flow ratio, which is derived by dividing the average flow for the year as a whole by the average flow for the days on which chemical samples were taken.

A bias correction factor is included in the calculation, to compensate for the effects of correlation between discharge and load.

The adjusted mean daily load is multiplied by 365 to obtain the annual load.

Approaches to Load Estimation

Ratio Estimators

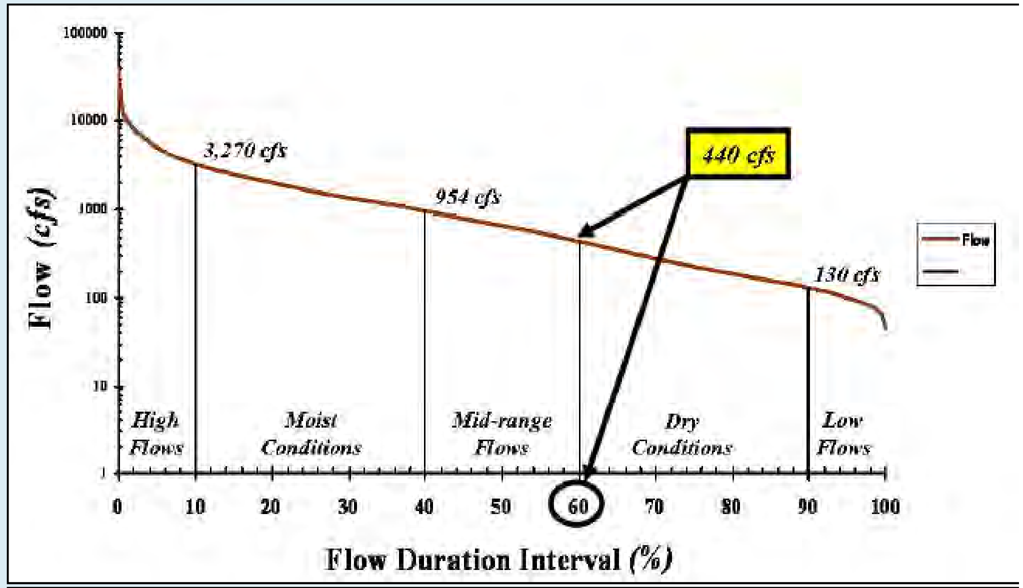
Stratification - division of the sampling effort or the sample set into two or more parts which are different from each other but relatively homogeneous within, e.g., growing season vs. winter vs. spring

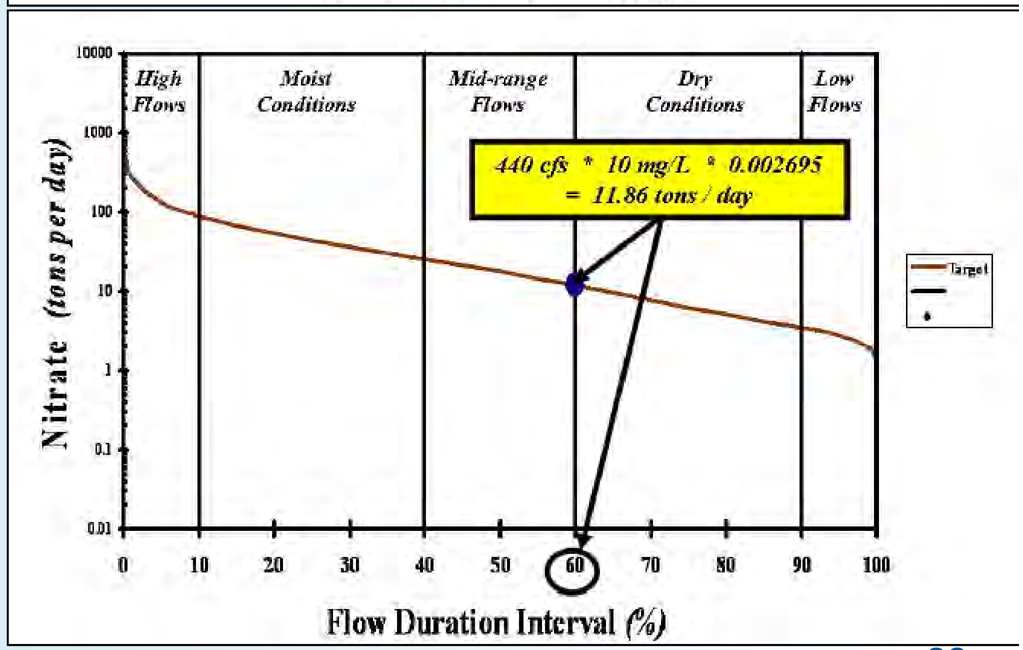
May improve precision and accuracy of load estimate by allocating more of the sampling effort to the aspects which are of greatest interest or which are most difficult to characterize because of great variability such as high flow seasons

Beale Ratio Estimator is one common technique; computer programs available to implement.

Load Duration Curves

- ► Flow duration curve
  - ▪ Cumulative frequency curve of historical flow
- ► Load duration curve
  - ▪ Multiply flows by target concentration to represent load target
  - ▪ Plot observed loads to compare to target



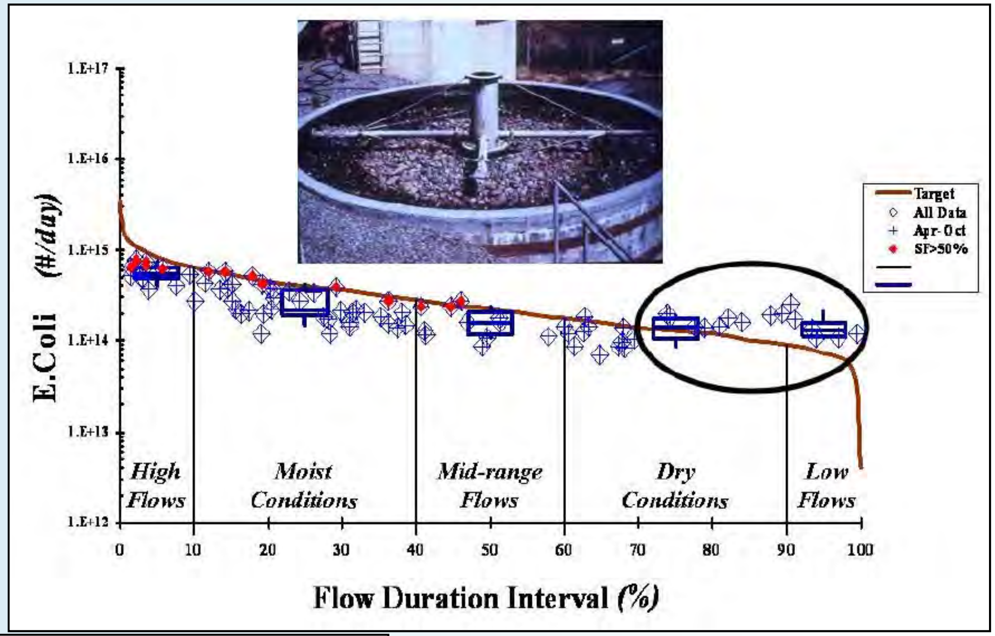


Source: USEPA (2007)

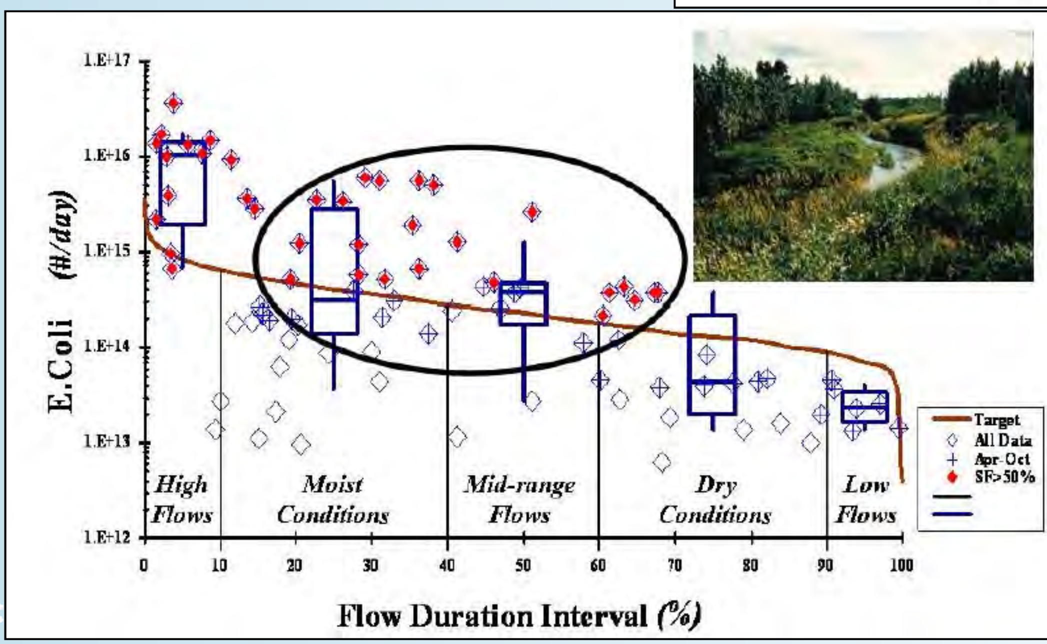
TETRA TECH

Load Duration Curves

- ► Point source
- ► Nonpoint source



Source: USEPA (2007)



TETRA TECH

Examples of load calculations

- Spreadsheet calculations
  - Simple loading calculations
    - Composite sampling
    - Numerical integration
  - Beale Ratio (see below)
- LOADEST—USGS
  - http://water.usgs.gov/software/loadest/
  - Regression
  - Program can run in Windows
- Purdue University LDC (Load Duration Curve)
  - https://engineering.purdue.edu/~ldc/pldc/
  - Includes LOADEST
- Beale Ratio
  - http://www.heidelberg.edu/academiclife/distinctive/ncwqr
  - National Center for Water Quality Research (NCWQR)
  - Currently working on a replacement using R

Data Sets

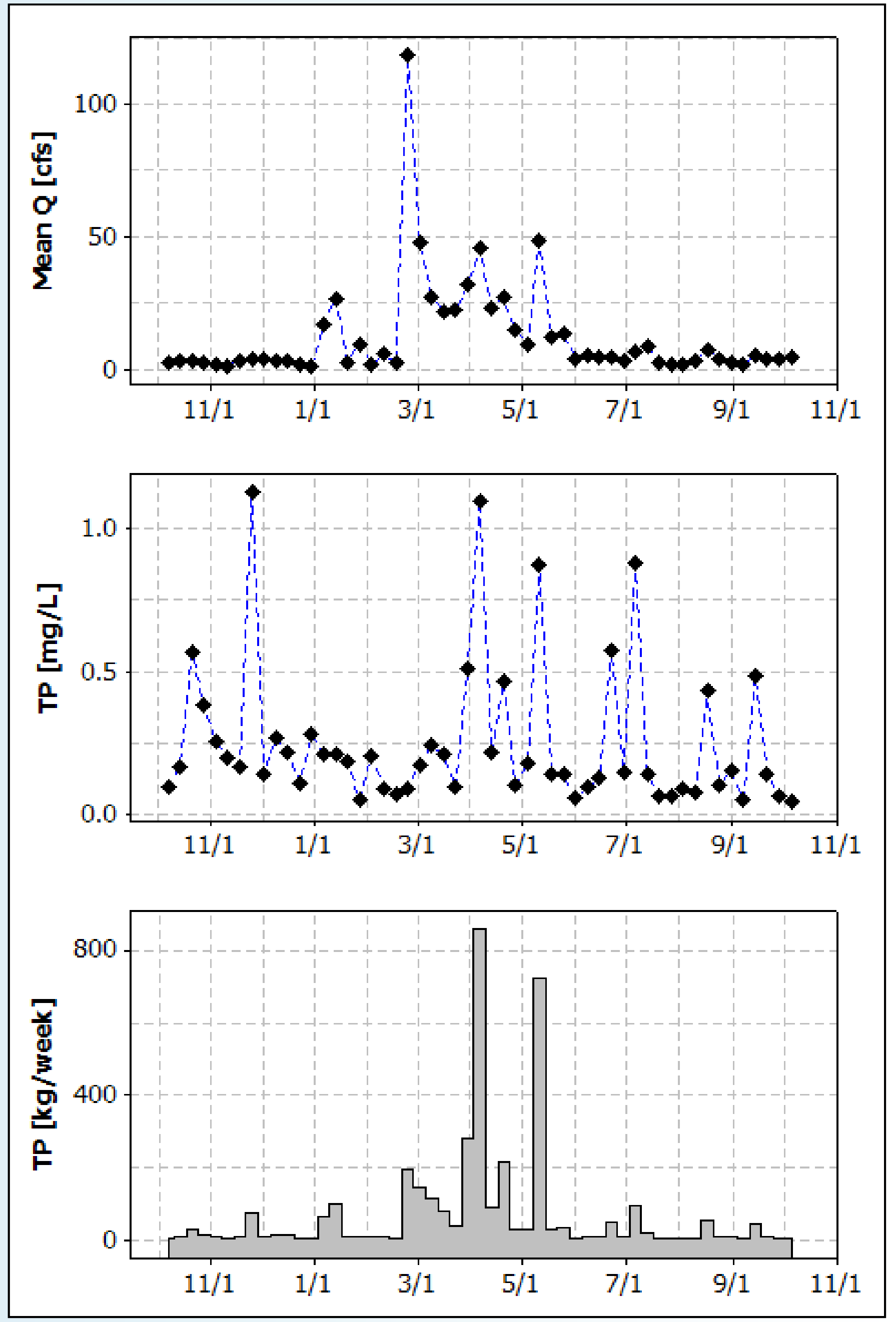
- Vermont NMP (~52 observations)
  - 1 year
  - Continuous stage/flow
  - Weekly flow-proportional composite samples for TP
- Sandusky River (~5700 observations)
  - 13 years (2000-12)
  - Matched flow data
  - Grab samples for SRP
- Little Calumet East Branch (~150 observations)
  - 13 years (2000-12)
  - Daily flow
  - Monthly TP grab sampling

Vermont NMP

1. <u>Stage measured continuously</u>
   - Flow calculated from site-specific rating equation
   - Total weekly discharge tracked by summing flow at 5-min intervals
2. <u>Concentration determined from weekly flow-proportional composite sample</u>
   - Equivalent to EMC
3. Weekly load = K **x** Q **x** TP EMC

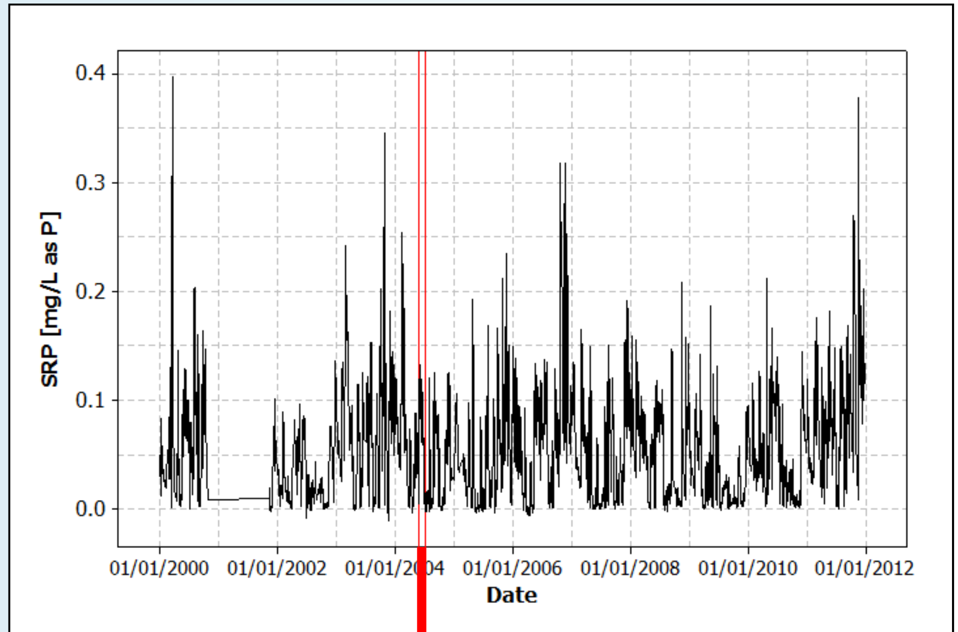
   Annual Load → 3,600 kg/yr

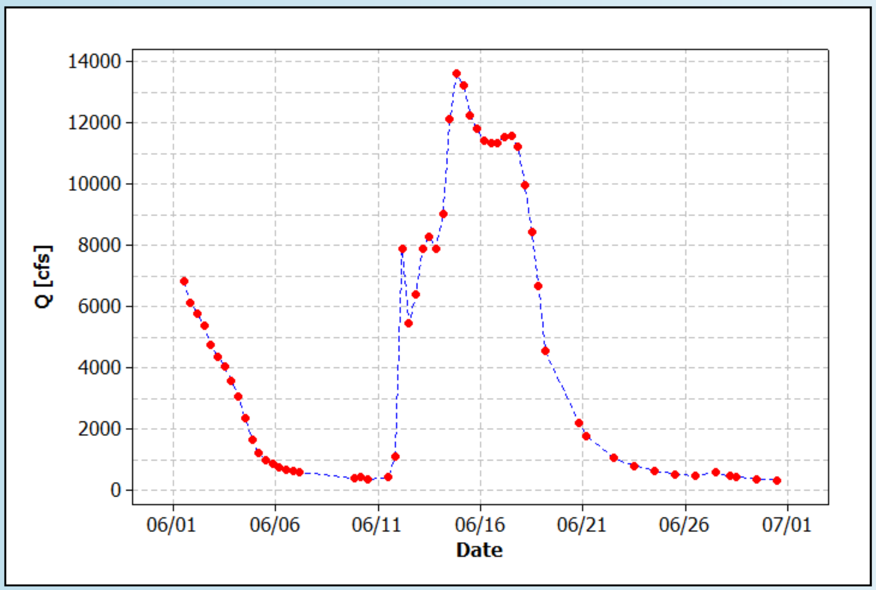
Weekdate	Mean Q [cfs]	TP [mg/L]	TP [kg/week]
7-Oct-99	2.2	0.099	3.7
14-Oct-99	2.9	0.167	8.3
21-Oct-99	2.9	0.568	28.2
...	...	...	...
21-Sep-00	3.9	0.145	9.7
28-Sep-00	4.1	0.067	4.7
5-Oct-00	4.4	0.050	3.8
Annual Load [kg/year] -->			3596.2

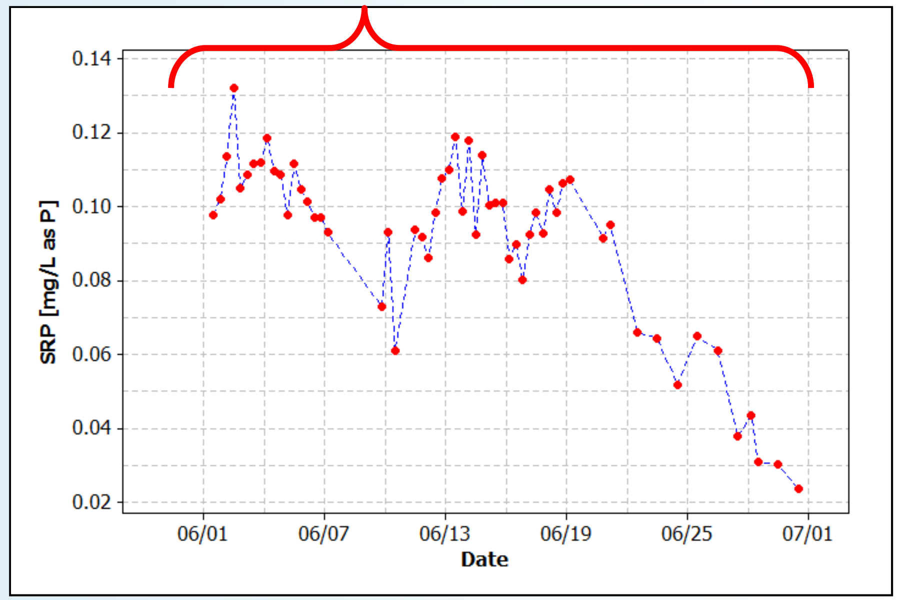


Sandusky River

- **Soluble reactive phosphorus**
  - 2000-12: ~5,700 observations
  - June 2004: 57 observations





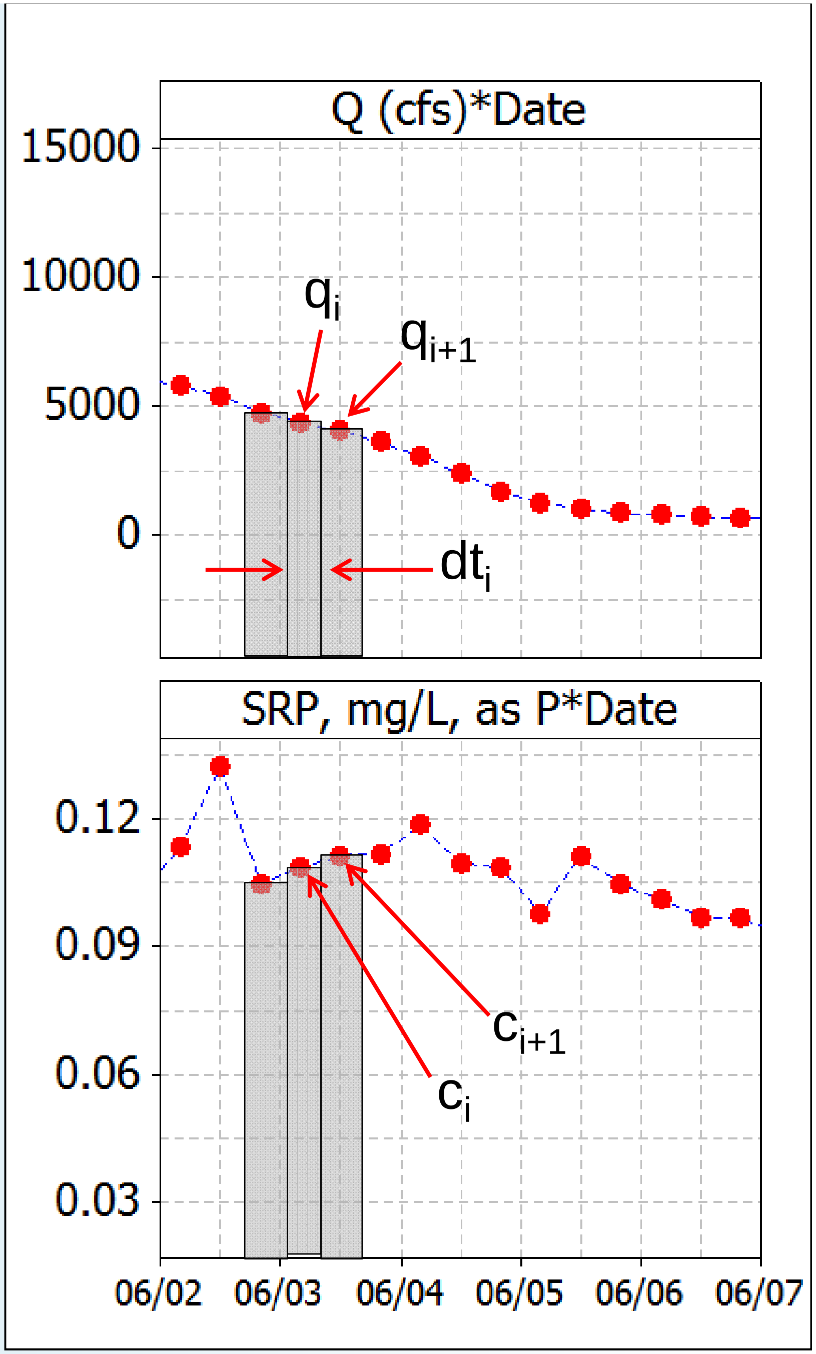


Sandusky River

- Numerical Integration
  - Discrete Time Intervals
- Load
  - $load_i = K\, q_i\, c_i\, dt_i$

Date	Begin Per.	End Per.	dt(i) Period [days]	q(i) Q [cfs]	c(i) SRP [mg/L as P]	load(i) SRP [kg/period]
06/01 12:00	06/01 00:00	06/01 16:00	0.67	6830.2	0.098	1,090.6
06/01 20:00	06/01 16:00	06/02 00:00	0.33	6143.0	0.102	511.0
06/02 04:00	06/02 00:00	06/02 08:00	0.33	5774.6	0.114	535.4
...	...	...	...	...	...	...
06/28 12:00	06/28 08:00	06/29 00:00	0.67	432.4	0.031	21.8
06/29 12:00	06/29 00:00	06/30 00:00	1.00	357.4	0.030	26.4
06/30 12:00	06/30 00:00	07/01 00:00	1.00	323.8	0.024	18.6
					June 2004 Load (kg) -->	25,470.6

June 2004 Load → 25,000 kg



Beale Ratio—Richards (1998)

- Ratio of load to flow for the entire year should equal ratio on dates where concentration was measured
- Days with flow and concentration
  - Mean observed load, $\bar{l}_o$
  - Mean observed flow, $\bar{q}_o$
- Days with no concentrations
  - Mean flow, $\bar{q}_a$
- Bias correction factor
  - Account for correlation between load and flow

$$\frac{\bar{l}_a}{\bar{q}_a} = \frac{\bar{l}_o}{\bar{q}_o}$$

$$\bar{l}_a = \bar{l}_o \frac{\bar{q}_a}{\bar{q}_o}$$

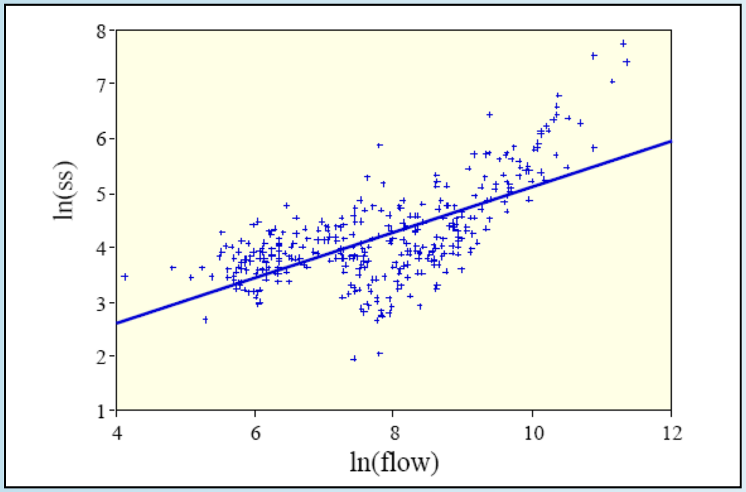
$$L = 365\bar{l}_a$$

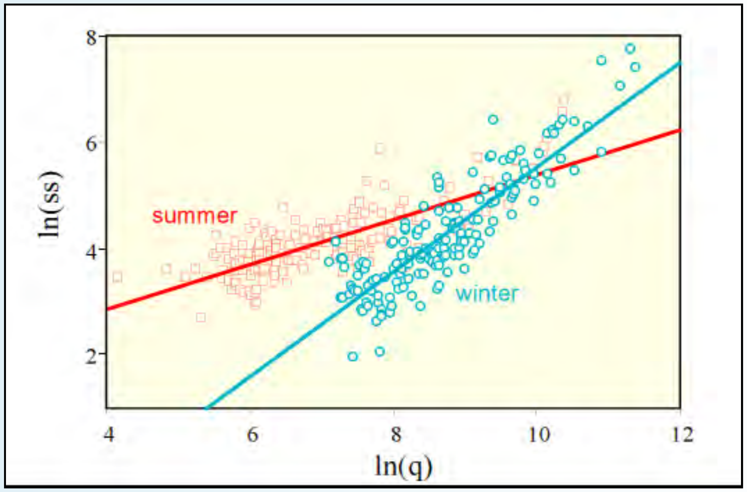
$$\bar{l}_a = \bar{l}_o \frac{q_a}{q_o}\left[\frac{1+\left(\frac{1}{n}-\frac{1}{N}\right)\frac{s_{lq}}{\bar{l}_o\bar{q}_o}}{1+\left(\frac{1}{n}-\frac{1}{N}\right)\frac{s_{qq}}{\bar{q}_o^2}}\right]$$

TETRA TECH

Beale Ratio—Richards (1998)

► Stratification – Season, Flow, Season and Flow

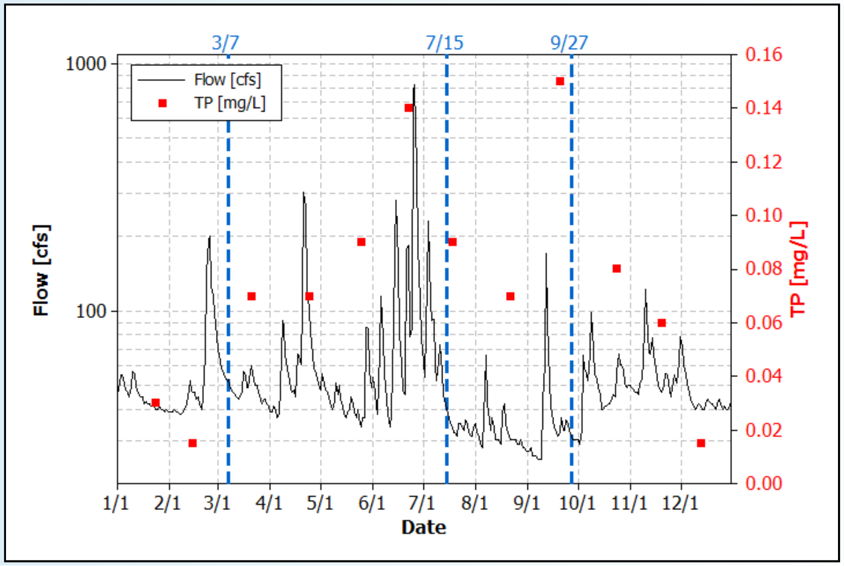




Source: Richards (1998)

► Example Calculations

- Little Calumet East Branch
- Year: 2000, 12 samples
- For demonstration
  - 4 seasons



Beale Ratio

- Sample Data
  - ❶ Identify strata
    1. Jan 1-Mar 7
    2. Mar 8-Jun 15
    3. Jun 16-Sep 27
    4. Sep 28-Dec 31
  - ❷ Compute flux
- Sample Statistics
  - Number of Days
  - Mean flow and flux
  - ❸ Flux and flow covariance
    - =covariance.s(flux,flow)
  - Flow variance
    - =var.s(flow)

Sample Data

❶ Stratum	Date	TP [mg/L]	Flow [cfs]	Flow [cms]	❷ Flux [kg/day]
1	20000124	0.03	40	1.133	2.935
1	20000215	0.015	47	1.331	1.725
2	20000321	0.07	60	1.699	10.274
2	20000424	0.07	102	2.888	17.466
2	20000525	0.09	34	0.963	7.485
2	20000622	0.14	184	5.210	63.015
3	20000718	0.09	34	0.963	7.485
3	20000822	0.07	30	0.849	5.137
3	20000920	0.15	32	0.906	11.742
4	20001024	0.08	57	1.614	11.155
4	20001120	0.06	47	1.331	6.898
4	20001213	0.015	41	1.161	1.504

Sample Statistics

	Sample Calculations (days with observed concentration)				
	n_o	Q_o	L_o	❸ $S(L_oQ_o)$	$S(Q_oQ_o)$
Stratum	Number of Days	Mean Flow [cms]	Mean Flux [kg/day]	Covariance of flux and flow	Variance of Flow
1	2	1.23	2.33	-0.12	0.020
2	4	2.69	24.56	46.38	3.451
3	3	0.91	8.12	0.07	0.003
4	3	1.37	6.52	1.08	0.052
	12				

TETRA TECH

Beale Ratio

- 1 Sample Statistics
- Stratum Loading
  - Number of days
  - Mean flow
  - Loading

2 Biased

$$\bar{l}_a = \bar{l}_o \frac{\bar{q}_a}{\bar{q}_o}$$

3 Unbiased

4

$$\bar{l}_a = \bar{l}_o \frac{\bar{q}_a}{\bar{q}_o}\left[\frac{1+\left(\frac{1}{n}-\frac{1}{N}\right)\frac{s_{lq}}{\bar{l}_o\bar{q}_o}}{1+\left(\frac{1}{n}-\frac{1}{N}\right)\frac{s_{qq}}{\bar{q}_o^2}}\right]$$

- Strata Loading
- Annual Load → 4,400 kg/yr

Sample Statistics

1 Sample Calculations (days with observed concentration)

Stratum	n_o Number of Days	Q_o Mean Flow [cms]	L_o Mean Flux [kg/day]	$S(L_oQ_o)$ Covariance of flux and flow	$S(Q_oQ_o)$ Variance of Flow
1	2	1.23	2.33	-0.12	0.020
2	4	2.69	24.56	46.38	3.451
3	3	0.91	8.12	0.07	0.003
4	3	1.37	6.52	1.08	0.052
	12				

Stratum Loading

2 Stratum Loading Calculations

4 3

Stratum	N_a Number of Days	Q_a Mean Flow [cms]	Q_a/Q_o Flow Ratio	$L_o (Q_a/Q_o)$ LOADING: Biased Estimate [kg/day]	BCT BIAS Correction Term	Bias Correction [kg/day]	$L_o (Q_a/Q_o)$ BCT LOADING: Un-biased Estimate [kg/day]	STRATA LOADING: Un-biased Estimate [kg/strata]
1	67	1.572	1.28	2.97	0.974	-0.078	2.895	193.9
2	130	2.299	0.85	20.99	1.049	1.026	22.014	2,861.9
3	74	1.018	1.12	9.13	1.002	0.015	9.145	676.7
4	95	1.439	1.05	6.86	1.030	0.205	7.062	670.9
	366						Annual Load [kg/year] -->	4,403.4
							Annual Load [kg/day] -->	12.0

LOADEST (USGS)

- http://water.usgs.gov/software/loadest/
- Runkel, R.L., Crawford, C.G., and Cohn, T.A., 2004, **Load Estimator (LOADEST): A FORTRAN Program for Estimating Constituent Loads in Streams and Rivers**: U.S. Geological Survey Techniques and Methods Book 4, Chapter A5, 69 p.

$$ln(l) = a_0 + a_1 \ln Q + a_2 \ln Q^2 + a_3 \sin(2\pi dtime) + a_4 \cos(2\pi dtime) + a_5 dtime + a_6 dtime^2$$

- Runkel, R.L., 2013, **Revisions to LOADEST, April 2013** (MOD48)
  - Bias diagnostics and residual analysis

LOADEST (USGS)

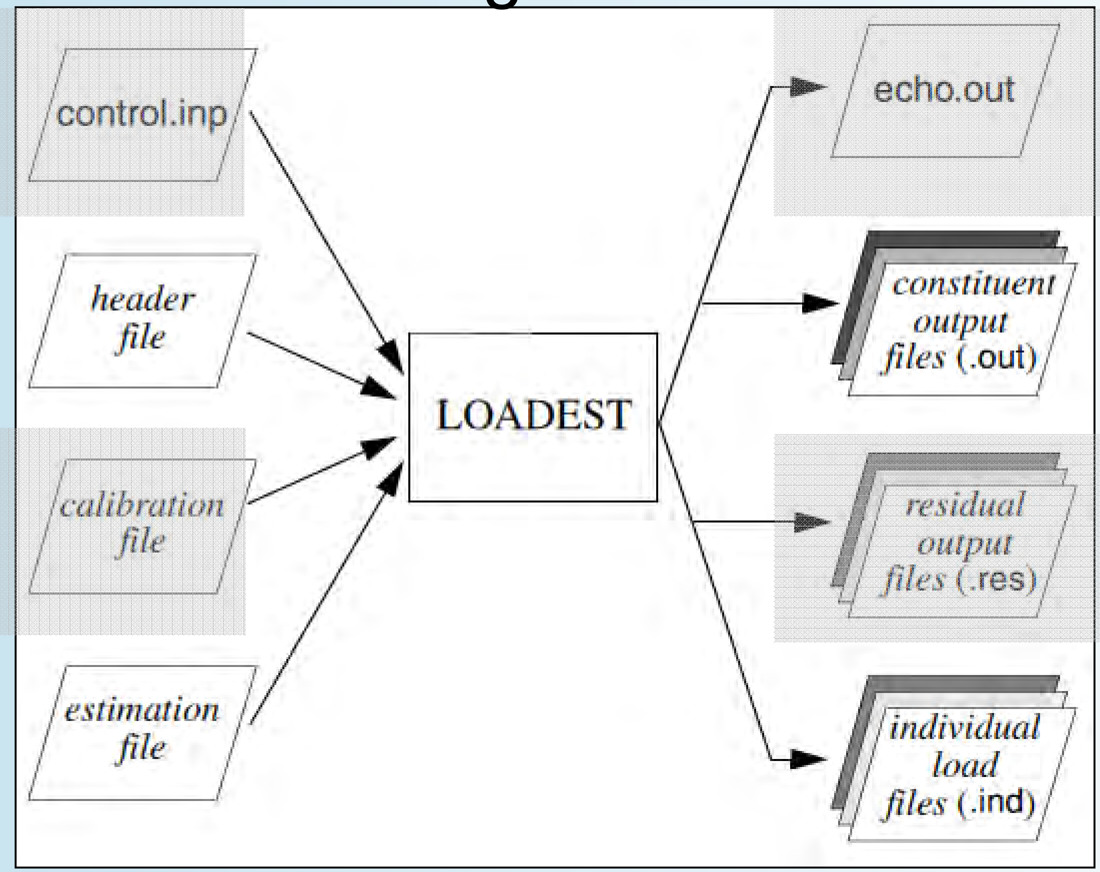
3 Primary Computation Methods

- ► MLE: Maximum likelihood estimation
- ► AMLE: Adjusted maximum likelihood estimation
  - Preferred option for data with censored observations
- ► LAD: Least absolute deviation
  - Preferred option when residual are not normally distributed
  - Cannot be used for data with censored observations

LOADEST (USGS)—Input/Output Files

Input → Program → Output

Input		Output	
Input file names	control.inp	echo.out	Model specs. and matched data
Model specifications	header file	constituent output files (.out)	Model results, diagnostics, and loads
Matched data	calibration file	residual output files (.res)	Residuals
Discharge data	estimation file	individual load files (.ind)	Daily Load



Source: Runkel et al (2004)

LOADEST (USGS)—Input Files

Windows Explorer:

1

Active library
Once
Name
- loadest.exe
- control.inp
- calibration.inp
- q.inp
- header.inp

2

```
# control.inp
header.inp
calibration.inp
q.inp
```

3

```
# calibration.inp
#
# Fields:
# – Date [YYYYMMDD]
# – Time [HMM]
# – Discharge [cfs]
# – TP [mg/L]
#
20000124 1650 40 0.03
20000215 1100 47 <0.03
20000321 1130 60 0.07
20000424 1500 102 0.07
20000525 0845 34 0.09
...
20120223 0900 96 <0.03
20120320 0830 73 0.04
20120419 0840 52 0.04
```

4

```
# q.inp
1
20000101 1200 46
20000102 1200 48
20000103 1200 55
20000104 1200 55
20000105 1200 51
...
20001229 1200 40
20001230 1200 41
20001231 1200 43
```

Can include additional columns for additional constituents

```
# calibration.inp
#
# Fields:
# – Date [YYYYMMDD]
# – Time [HMM]
# – Discharge [cfs]
# – TP [mg/L]
# – TSS [mg/L]
#
20000124 1650 40 0.03 4
20000215 1100 47 <0.03 5
20000321 1130 60 0.07 15
20000424 1500 102 0.07 19
20000525 0845 34 0.09 25
...
20120223 0900 96 <0.03 5
20120320 0830 73 0.04 17
20120419 0840 52 0.04 13
```

LOADEST (USGS)—Input/Output Files

Windows Explorer:

Active library
Once
Name
- loadest.exe ← 2
- control.inp
- calibration.inp
- q.inp
- header.inp → 1

```
# header.inp
Little Calumet River TP & TSS Load Estimation
1         |  PRTOPT (0/1: do not/do print ind loads
3         |  SEOPT (3: compute exact standard error for AM
3         |  LDOPT (3: compute load for entire period, use
2         |  NSEAS (col.1-5)
0101  0731
0801  1231
0         |  MODNO (Model type -> 0: automatic)
1         |  Rec type 12: # constituents
###################################
#
#  Unit Conc: 1-mg/L, 2-µg/L
#  Flag Load: 1-kg/day, 2-g/day, 3-lbs/day, 4-tons/day
#
###################################          Unit Flags
#CNAME                                       Conc Load
TP                                           1    3
```

Active library
Once
Name
- loadest.exe
- control.inp
- calibration.inp
- q.inp
- header.inp
- echo.out → 4
- TP.ind
- TP.out
- TP.res

3

```
                         LOADEST
             A Program to Estimate Constituent Loads
         U.S. Geological Survey, Version: MOD48 (March 2013)
         ---------------------------------------------------

HEADER FILE      : header.inp
CALIBRATION FILE: calibration.inp
ESTIMATION FILE : q.inp
```

TETRA TECH

LOADEST (USGS)—Input/Output Files

Windows Explorer:

Active library
Once
Name
- loadest.exe
- control.inp
- calibration.inp
- q.inp
- header.inp

Active library
Once
Name
- loadest.exe
- control.inp
- calibration.inp
- q.inp
- header.inp
- echo.out
- TP.ind
- TP.out
- TP.res

1

```
Individual Load Estimates

                                    Loads Estimated by:

Date       Time    Flow         AMLE          MLE
-----------------------------------------------------------------
20000101  1200    4.600E+01   1.0594E+01   1.0597E+01
20000102  1200    4.800E+01   1.1168E+01   1.1171E+01
20000103  1200    5.500E+01   1.3403E+01   1.3407E+01
20000104  1200    5.500E+01   1.3336E+01   1.3340E+01
20000105  1200    5.100E+01   1.1955E+01   1.1959E+01
20000106  1200    4.800E+01   1.0949E+01   1.0952E+01
20000107  1200    4.700E+01   1.0590E+01   1.0593E+01
20000108  1200    4.500E+01   9.9371E+00   9.9405E+00
20000109  1200    4.900E+01   1.1114E+01   1.1117E+01
...
```

Would include LAD estimates if no censored data

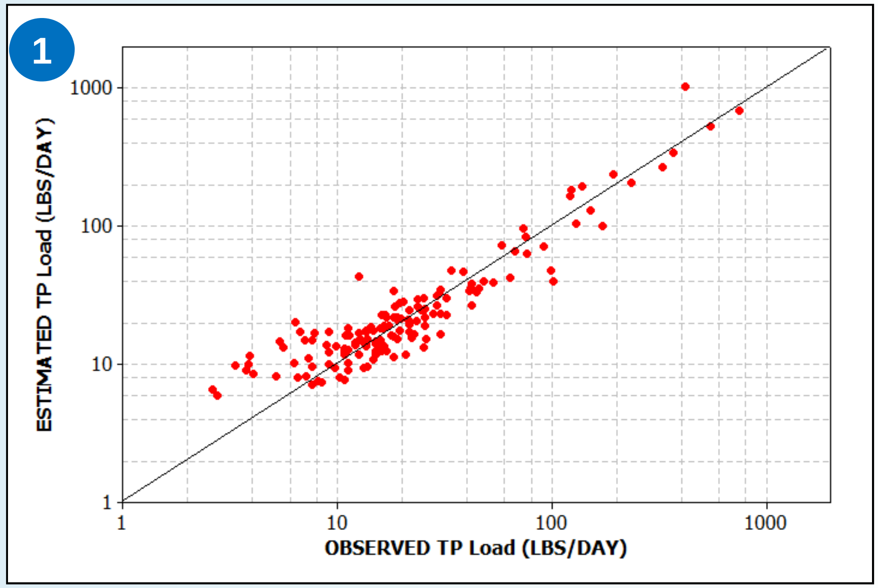
2

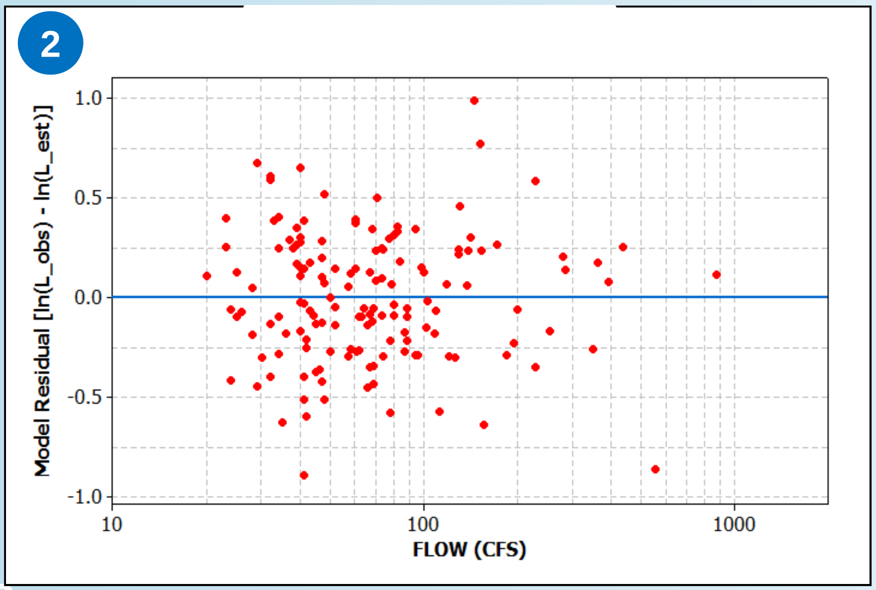
```
#
#  Residual output file
#
#  notes
#  -----
#  DTIME      decimal time minus "center" of decimal time
#  LN(CFLOW)  natural log of (uncentered) streamflow
#  F          flag indicating observation is censored (C) or uncensored (U)
#  CCONC      observed concentration for F=U; 1/2 of the observed concentration for F=C
#  CCONCAML   estimated concentration
#  YHATC      estimated natural log of concentration
#  CLOAD      observed load for F=U; 1/2 of the observed load for F=C (units dependent on ULFLAG)
#  CLOADAML   estimated load (units dependent on ULFLAG)
#  YHAT       estimated natural log of load (where load is in kg/d)
#  RESID      difference between observed and estimated values of log load (or log concentration)
#  Z          z-score for residual
#
#
#DATE     TIME  DTIME         LN(CFLOW)    F  CCONC       CCONCAML     YHATC        CLOAD        CLOADAM
#
20000124 1650 -6.10880E+00  3.68888E+00 U 3.00000E-02 3.73698E-02 -3.33955E+00 6.47246E+00 8.06248
20000215 1100 -6.04935E+00  3.85015E+00 C 1.50000E-02 3.95646E-02 -3.28277E+00 3.80257E+00 1.00298
20000321 1130 -5.95367E+00  4.09434E+00 U 7.00000E-02 5.08415E-02 -3.03227E+00 2.26536E+01 1.64535
20000424 1500 -5.86037E+00  4.62497E+00 U 7.00000E-02 8.58062E-02 -2.50833E+00 3.85112E+01 4.72071
20000525  845 -5.77639E+00  3.52636E+00 U 9.00000E-02 7.39155E-02 -2.65784E+00 1.65048E+01 1.35551
...
```

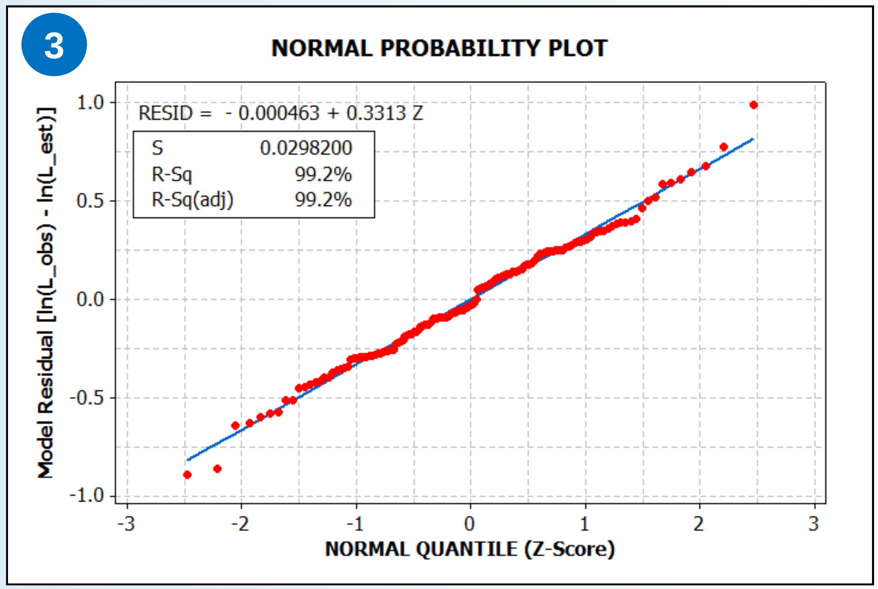
TETRA TECH

LOADEST (USGS)—Residuals

1. Estimated TP load vs. observed TP load
2. Residuals vs. Flow
3. Residual vs. Z-score



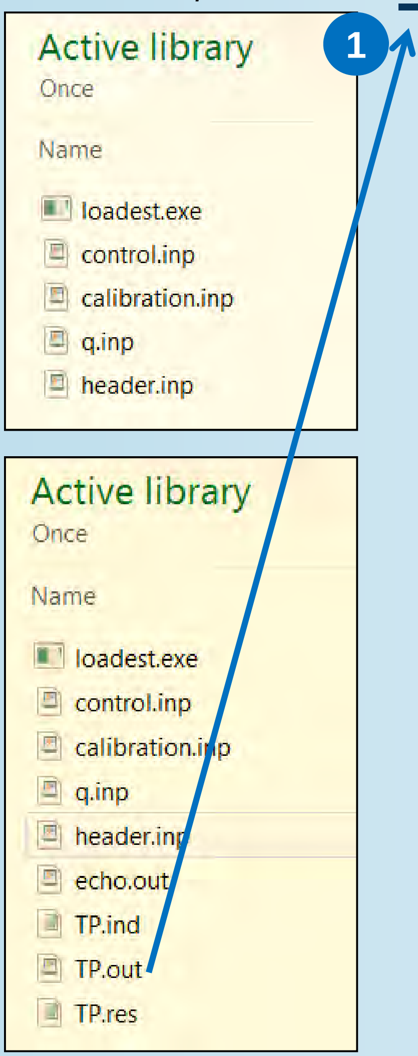




TETRA TECH

LOADEST (USGS)—Input/Output Files

Windows Explorer:



TP.out

- Selected model and coefficients
- Regression diagnostics, R^2, probability plot correlation coefficient (PPCC)
- 2 Bias diagnostics and 3 Load Estimates

Summary Stats: Est. and Obs. Loads in [LBS/DAY]

	Min.	25th Pct	Med.	75th Pct	90th Pct	95th Pct	99th Pct	Max.
Est.	5.94E+00	1.29E+01	1.75E+01	3.05E+01	8.70E+01	2.03E+02	8.73E+02	1.04E+03
Obs.	2.59E+00	1.12E+01	1.68E+01	3.01E+01	9.90E+01	1.85E+02	6.57E+02	7.53E+02
Est/Obs	2.29	1.16	1.04	1.01	0.88	1.10	1.33	1.38

Est/Obs > 1 indicates overestimation; Est/Obs < 1 indicates underestimation

2

Bias Diagnostics

Bp [%]	7.756
PLR	1.078
E	0.680

AMLE Load Estimates

3

	N	Mean Load	95% Conf.Intervals Lower	95% Conf.Intervals Upper	Std Error Prediction	Standard Error
Est. Period	366	40.71	31.76	51.40	5.02	3.69
Jan. 2000	31	9.85	8.14	11.81	0.94	0.70
Feb. 2000	29	19.13	15.10	23.90	2.25	1.30
...						

LOADEST (USGS)

Bias Diagnostics

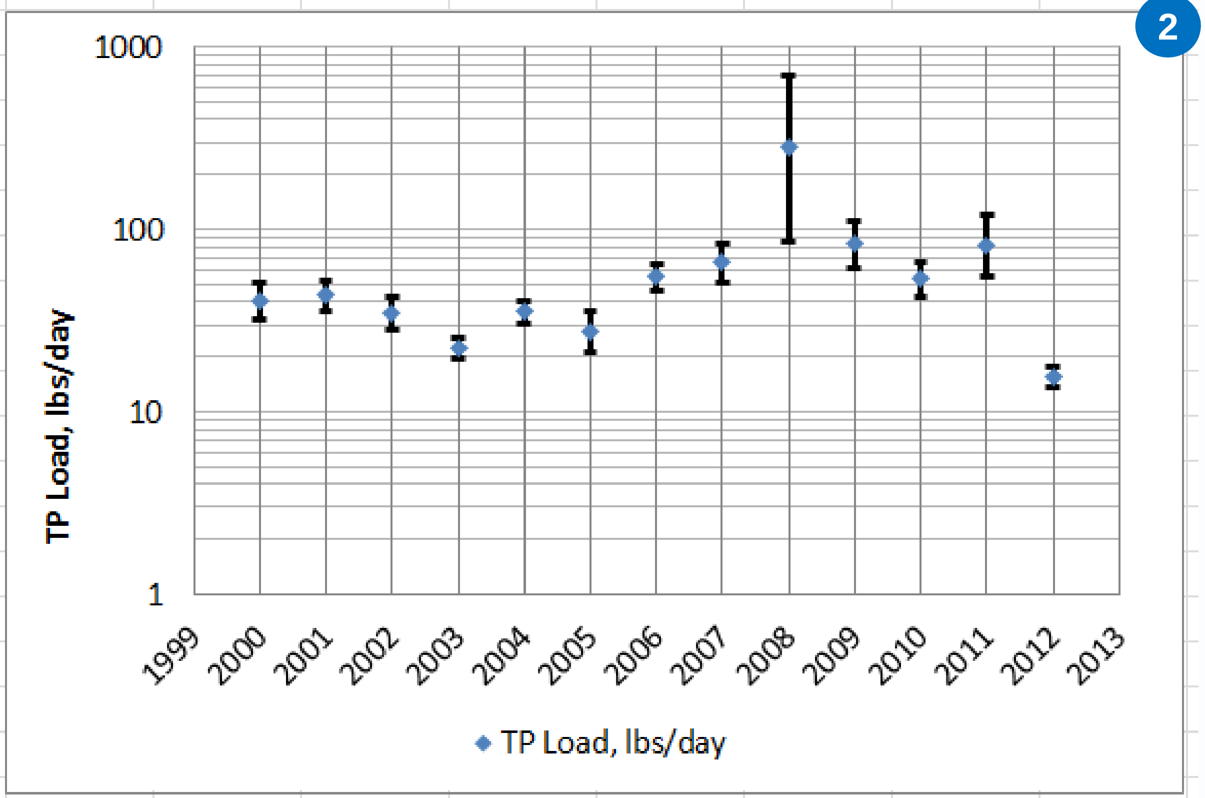
- ► Partial Load Ratio (PLR)
  - ∑ Estimated Loads / ∑ Observed Loads
  - Load Bias in Percent (B_p) = (100xPLR) – 100
  - USGS Recommendation: Ok for $B_p \leq \pm 25\%$
- ► Nash-Sutcliffe Efficiency Index, E
  - Ranges from -infinity to 1.0
  - E = 1—a perfect fit to observed data.
  - E = 0—model estimates are as accurate as the mean of observed data.
  - E < 0—the observed mean is a better estimate than the model estimates.
  - USGS Recommendation: Don't use when E<0

LOADEST (USGS)—Results: Annual TP Loads

1 Year	TP Load, lbs/day	L95% CI	U95% CI	StErr Pred.	Std Error	Mu-L95	U95-Mu
2000	40.71	31.76	51.4	5.02	3.69	8.95	10.69
2001	43.32	35.43	52.45	4.35	3.57	7.89	9.13
2002	35.01	28.38	42.72	3.66	2.76	6.63	7.71
2003	22.36	19.76	25.2	1.39	0.94	2.6	2.84
2004	35.38	30.51	40.79	2.62	1.96	4.87	5.41
2005	27.72	21.24	35.56	3.66	2.98	6.48	7.84
2006	54.93	46.03	65.05	4.86	4.1	8.9	10.12
2007	65.63	51.14	82.95	8.13	7	14.49	17.32
2008	281.49	86.17	693.01				
2009	83.37	60.77	111.65				
2010	53.58	42.81	66.23				
2011	82.27	55.05	118.35				
2012	15.36	13.47	17.44				

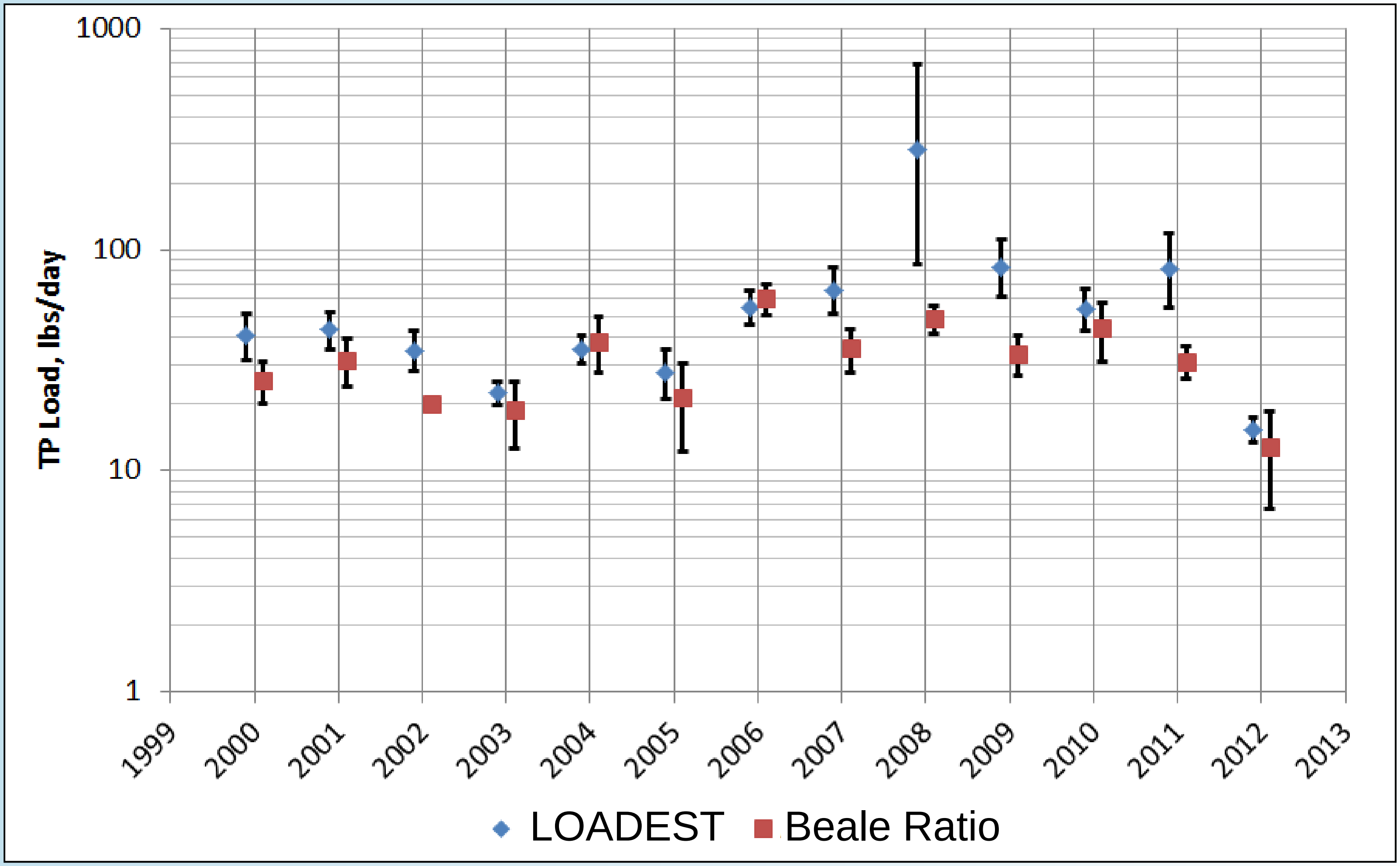
1 Compile 2000-2012 annual TP loads

2 Graphic (lbs/day)

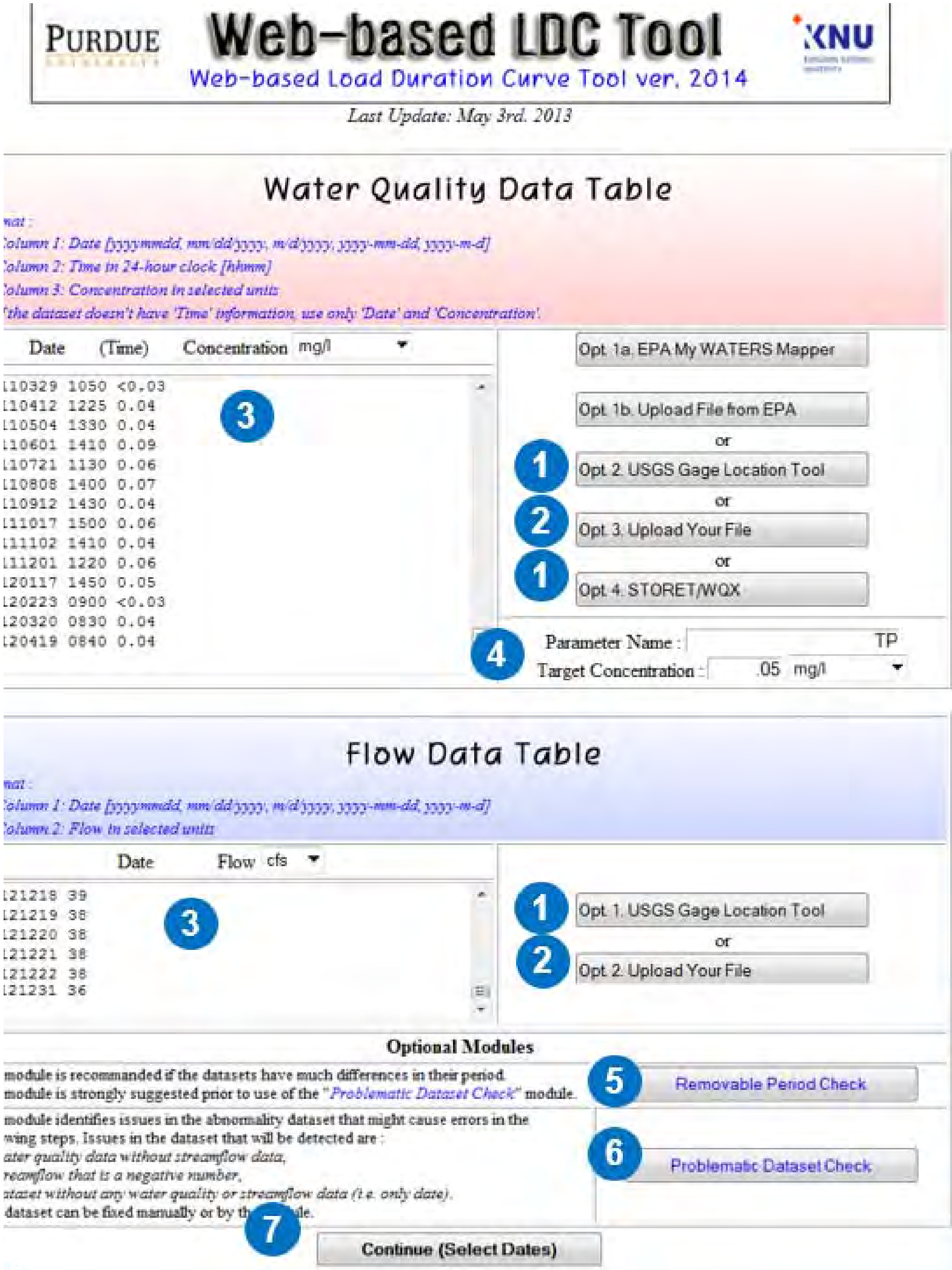


TETRA TECH

Comparison of LOADEST and Beale Ratio
(w/ 95% confidence interval)



TETRA TECH



Purdue Web Site

https://engineering.purdue.edu/~ldc

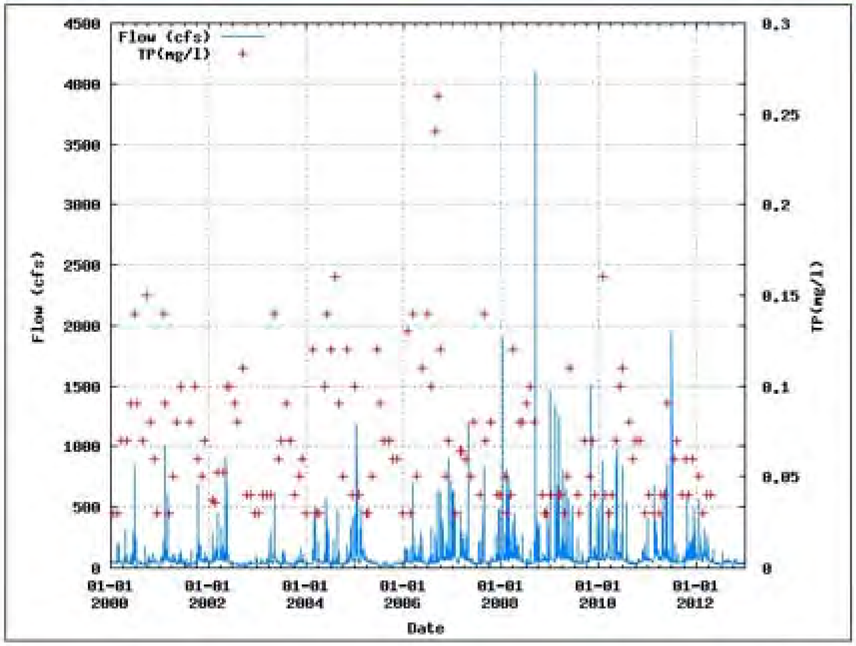
- Data sources
  - (1) Direct import from STORET or USGS
  - (2) Upload your file
    - Cut & paste (Ctrl-C & P)
- Browse data
  - (3) Edit down to desired data
    - Cut & paste into a text file for safe keeping
- (4) Parameter Name & Target Concentration
- (5) Removable period check
- (6) Problematic dataset check
- (7) Click continue

Purdue Web Site

https://engineering.purdue.edu/~ldc

- Data sources
  - Direct import from STORET or USGS

Purdue Web Site



	Beginning Date			Ending Date		
FlowData	2000 - 01 - 01			2012 - 12 - 31		
Water Quality Data	2000 - 01 - 24			2012 - 04 - 19		
1 Load Estimation Period	2000 YEAR	01 MONTH	01 DAY	2012 YEAR	12 MONTH	31 DAY

Optional Inputs

☑ Plot 50 % Surface Flow (2)

Select and enter BFI value for separation of Surface Runoff : (3)

Perennial streams with porous aquifers ▾ - BFI value : 0.80

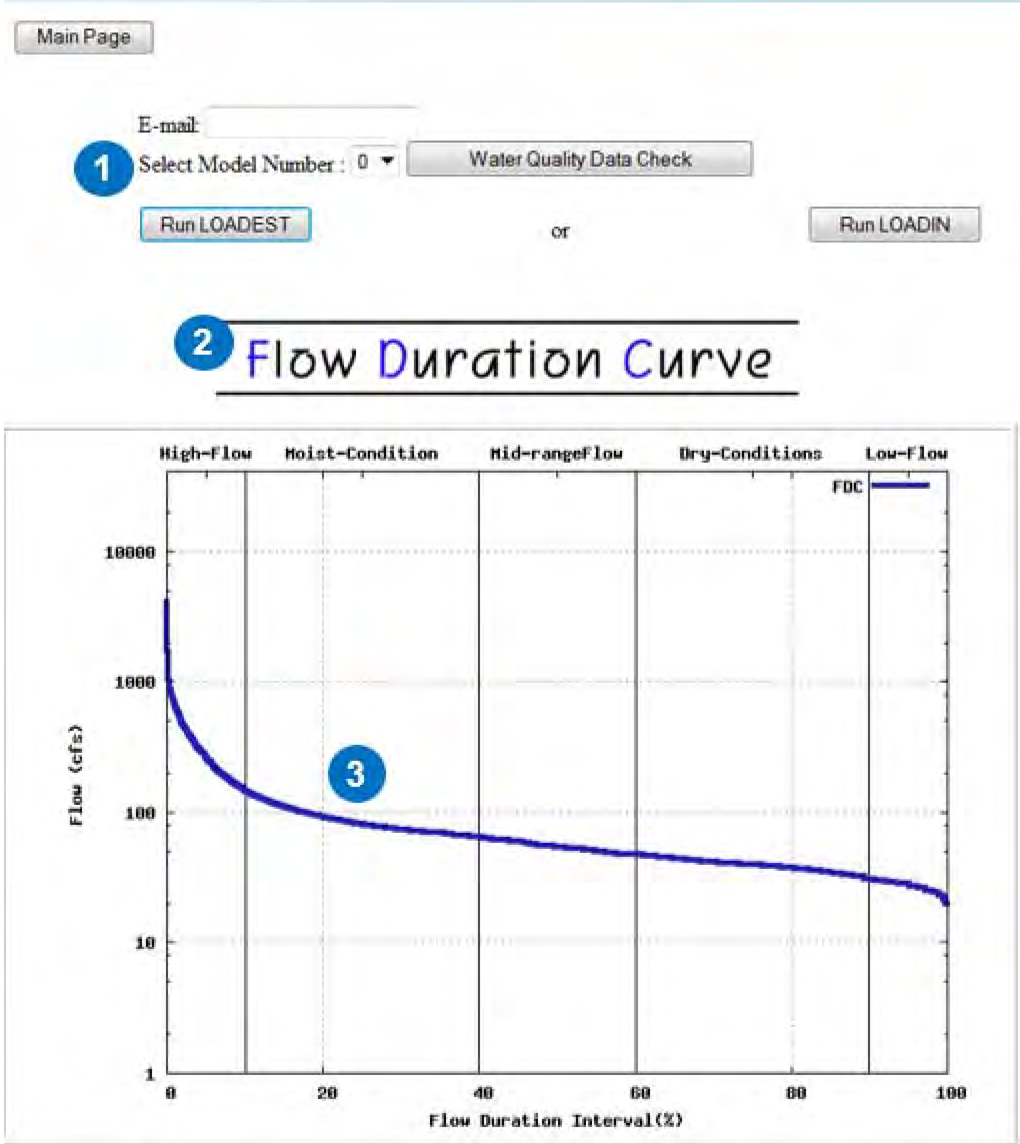
☑ Plot Growing Season

Select Period : From Apr ▾ To Oct ▾ (4)

(5) Plot LDC

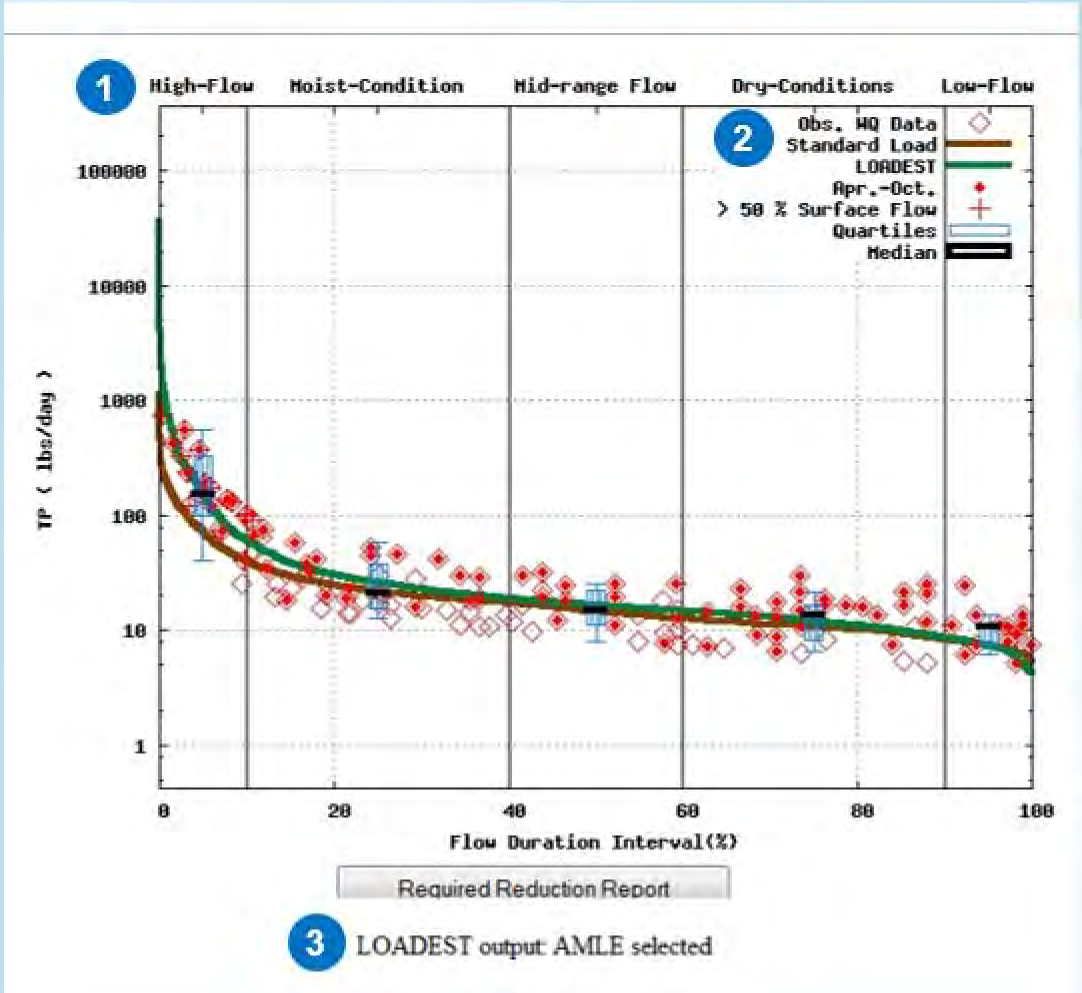
- Load Estimation Period
  1. Set to a full year
- Identify "storm flow"
  2. % Surface Flow
  3. Baseflow Separation Index (BFI)
     - 0.80—Perennial streams with porous aquifers
     - 0.50—Ephemeral streams with porous aquifers
     - 0.80—Perennial streams with hard rock aquifers
  4. Growing season
  5. Click Plot LDC

Purdue Web Site

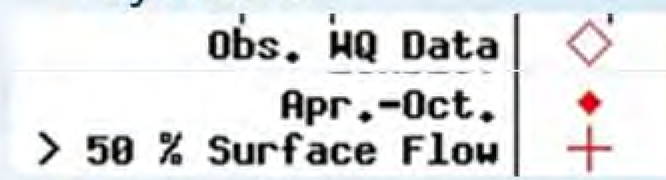


1. LOADEST
   - Enter email
   - Select model number (0, for auto select.)
   - Click Run LOADEST
   - ~30 minute wait (or come back to results sent to email)
2. Flow Duration Curve
   3. From 2000-2012, flow is greater than 100 cfs 20% of the time

Purdue Web Site

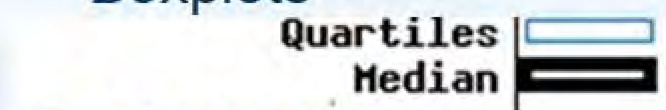


1. Load Duration Curve
   - Based on constant concentration set earlier
2. Legend
   - Symbols



   - Lines

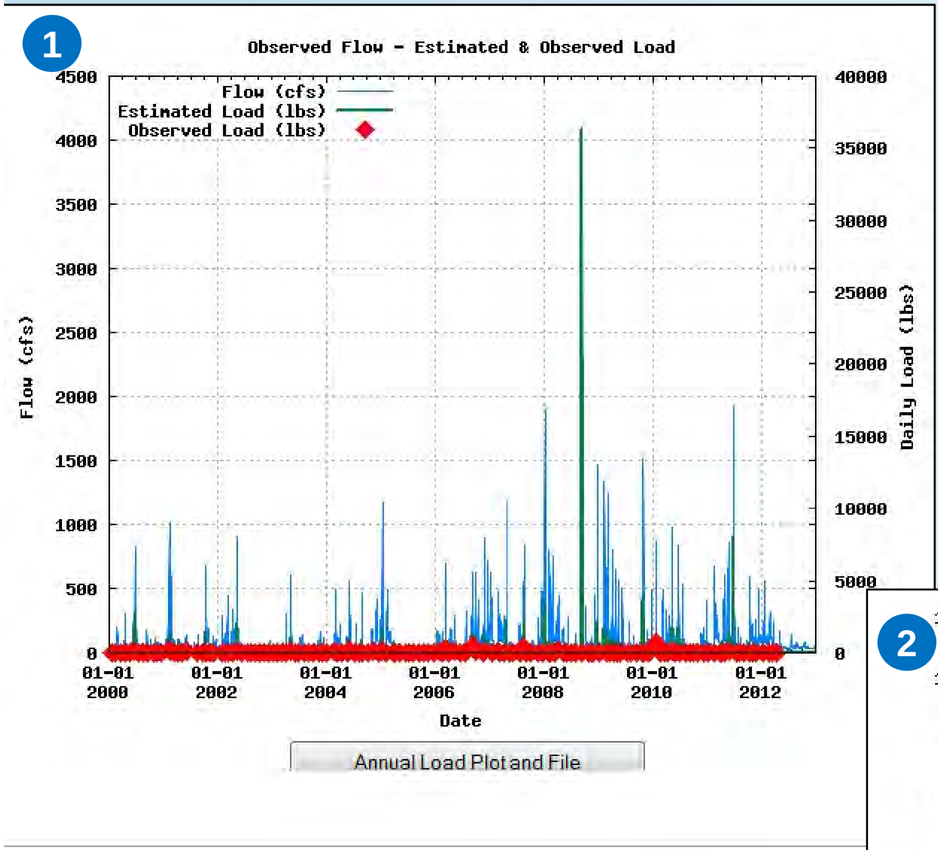
   - Boxplots

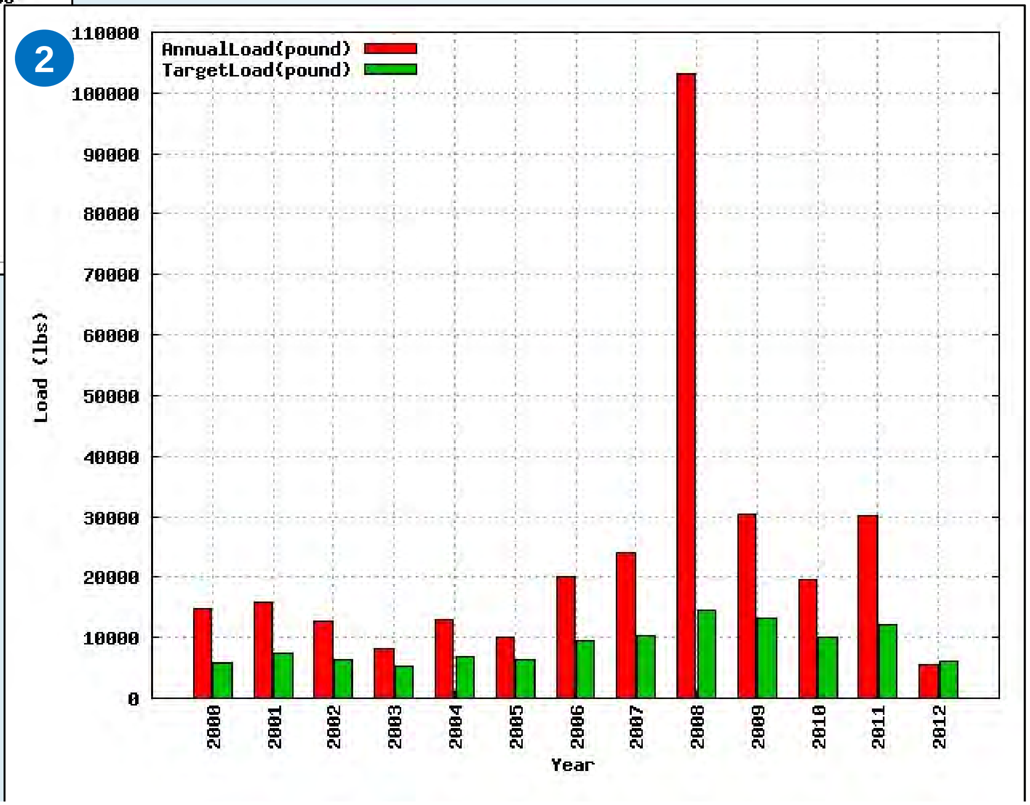


3. LOADEST option

Purdue Web Site

1. Time Series Plot
2. Annual Load Plot





Purdue Web Site

File	Description
1 Combined Data	The file includes flow (cfs) and concentration (given unit) on time step. *(Format: Date, Time, Flow, Concentration)*
1 Flow Percentile	The file contains sorted data based on flow percentile. *(Format: Percentile, Flow)*
1 Observed Load Percentile	The file includes target and observed loads. *(Format: Percentile, Flow, Target Load, Observed Load)*
3 LOADEST IN & OUT	The ZIP file contains input and output of LOADEST model.
2 Annual Load	The file is annual target and estimated loads file. *(Format: Year, Estimated Annual Load, Target Annual Load)*
2 Simulated Load Percentile	The file contains percentile-based flow, target load, estimated load, and observed load. *(Format: Percentile, Flow, Target Load, Estimated Load, Observed Load)*

File Download

- Flow/Load Duration Curve
  - 1 Exports to CSV or directly in to Excel
- LOADEST
  - 2 Exports to CSV or directly in to Excel
  - 3 Zipped File
    - Input Files
    - Output Files

calib.inp
control.inp
echo.out
est.inp
gnuAnnualLoad.inp
header.inp
LIPcalib.inp
LIPest.inp
TP.ind
TP.out
TP.res

Some Citations

- National Nonpoint Source Monitoring Program TechNote
  - http://www.bae.ncsu.edu/programs/extension/wqg/319monitoring/TechNotes/technote8_load_estimation.pdf
- LOADEST—USGS
  - http://water.usgs.gov/software/loadest/
- Purdue University LDC (Load Duration Curve)
  - https://engineering.purdue.edu/~ldc/pldc/
  - https://engineering.purdue.edu/~ldc/pldc/help.html
- P. Richards' nomograph—1998
  - http://www.heidelberg.edu/sites/default/files/jfuller/images/Load_Est1.pdf

Also out there

- ▶ USGS R packages
  - https://github.com/USGS-R
  - rloadest—R version of LOADEST
  - dataRetrieval—Download data from NWIS and STORET
  - EGRET—Exploration and Graphics for RivEr Trends
  - WRTDS—Weighted Regressions on Time, Discharge, and Season