Deriving Sampling Frequency Charts : example for Dissolved Organic Matter in Brittany, France

François Birgand[1] - Erin Bennett[1] - Gérard Gruau[2]

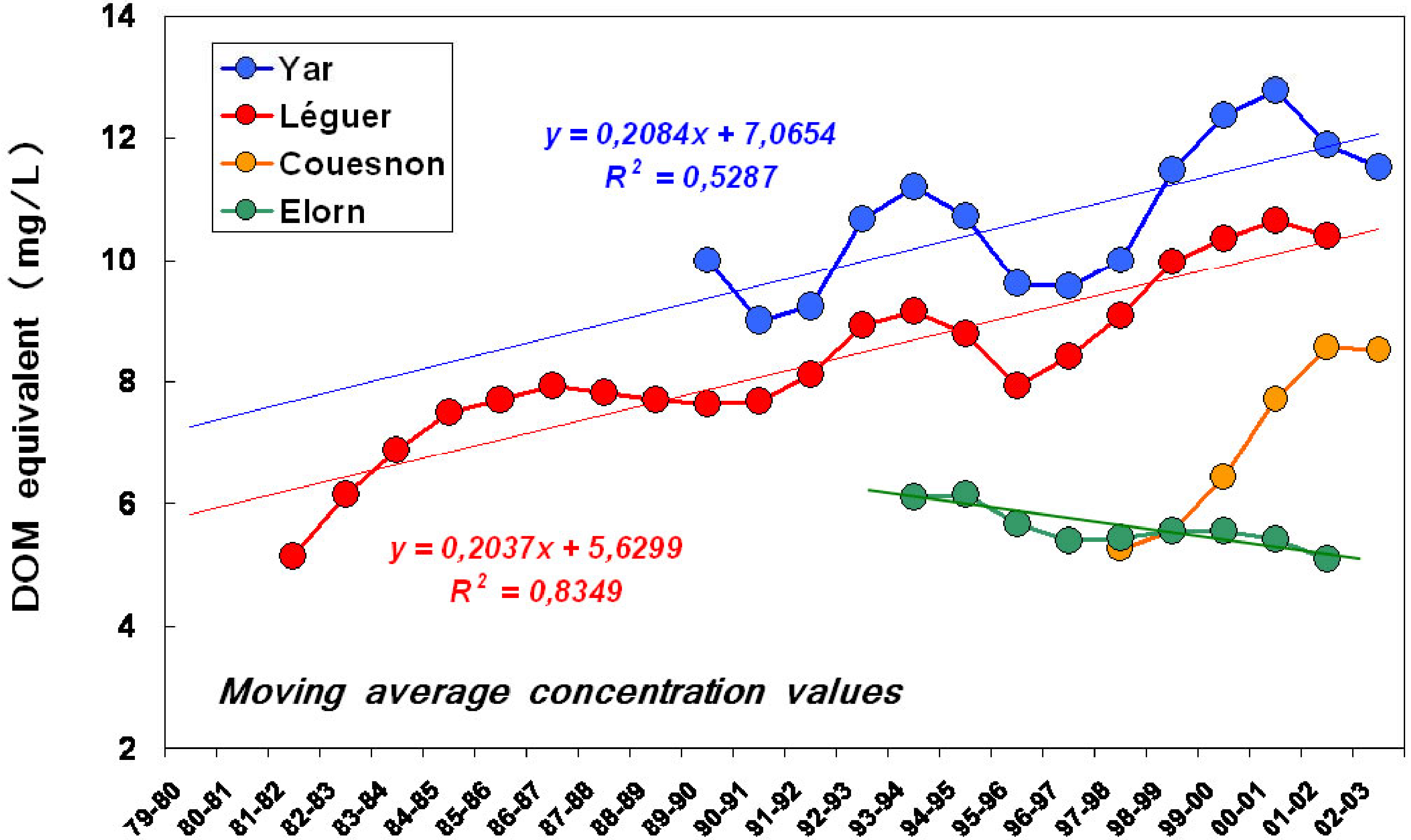
NC WRRI conference – Raleigh, NC March 31,2010

NC STATE UNIVERSITY [1]

UNIVERSITÉ DE RENNES 1 [2]

How often should one sample ?
For what uncertainty in the results?

The DOM problem

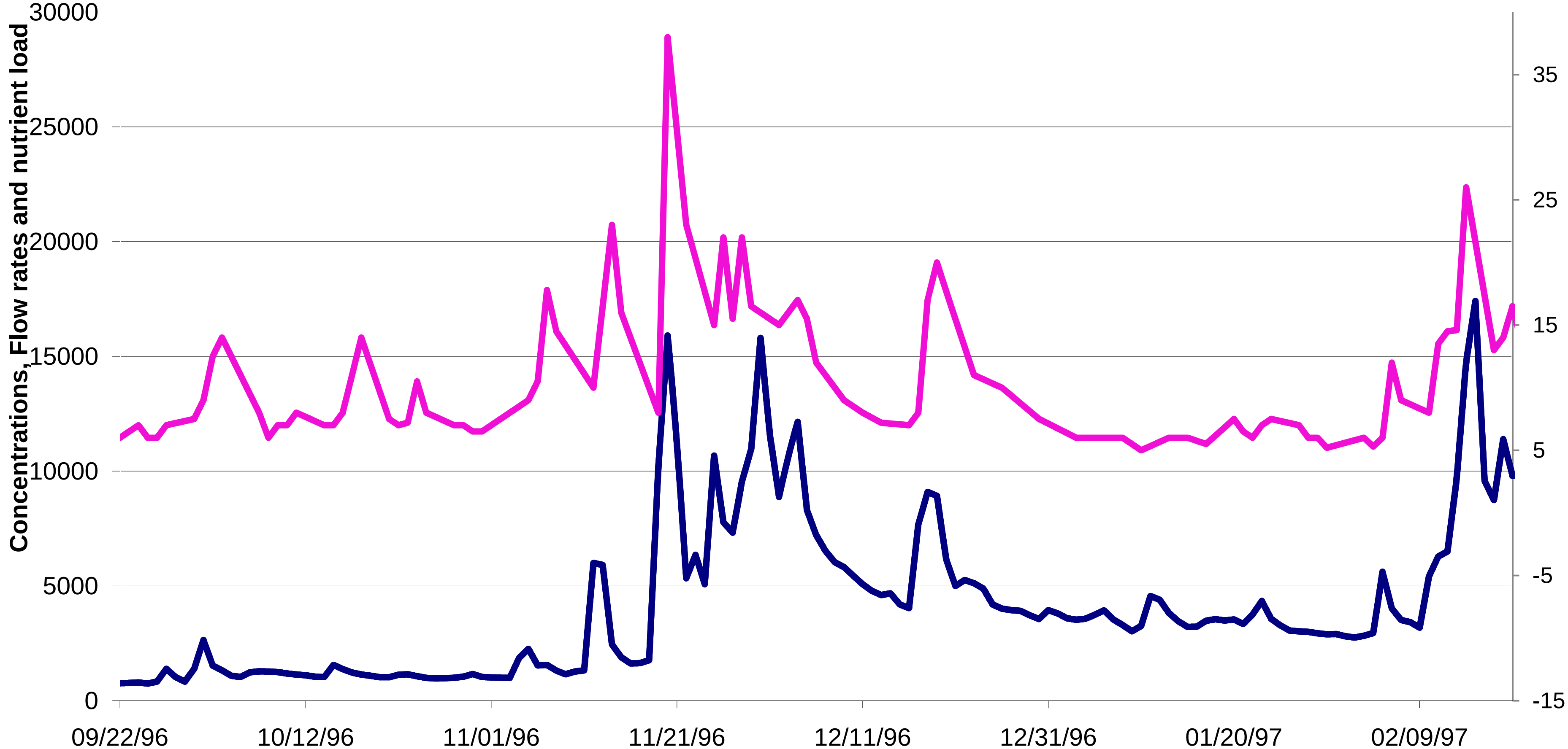


Sources of uncertainties while measuring nutrient loads and water quality indicators

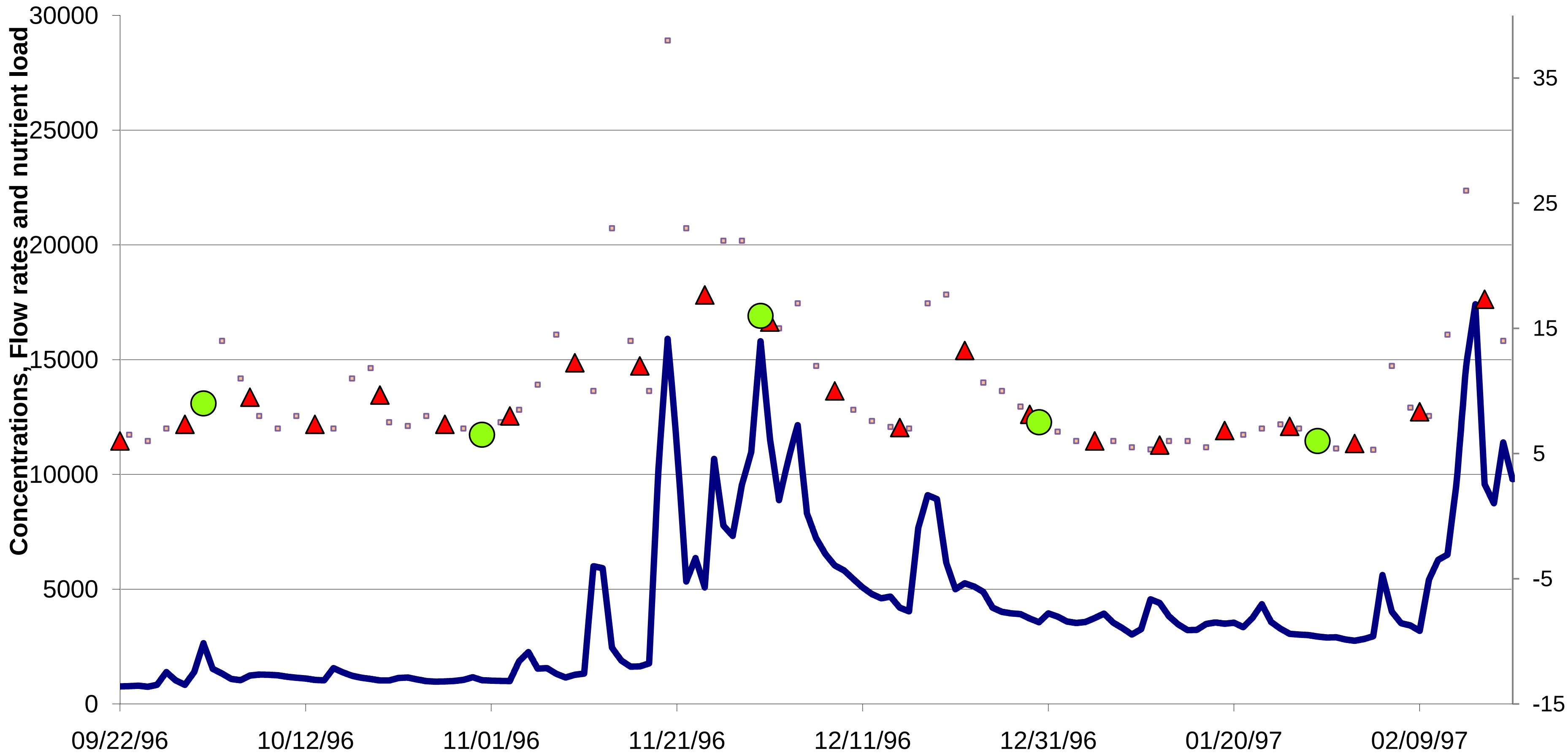
- Uncertainties on flow rates and cumulated flow
- Uncertainties due to the sampling location in the water column
- Uncertainties due to sample degradation between sampling and analysis
- Uncertainties of laboratory analyses

- <u>Uncertainties associated with infrequent sampling</u>

In a perfect world, continuous data…



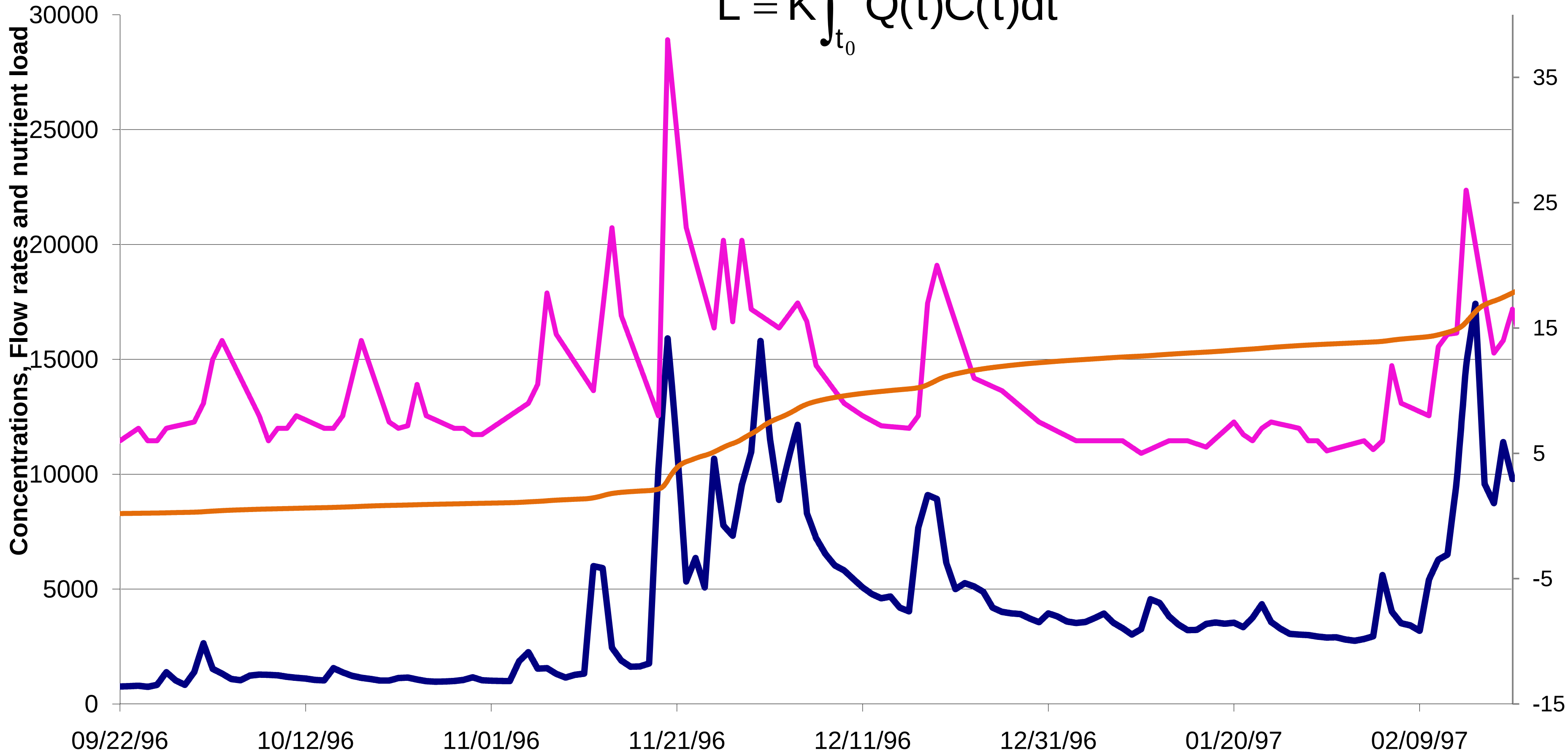
Discrete sampling…



Calculating nutrient annual loads

Definition : $$L = K\int_{t_0}^{T} L(t)dt$$

$$L = K\int_{t_0}^{T} Q(t)C(t)dt$$



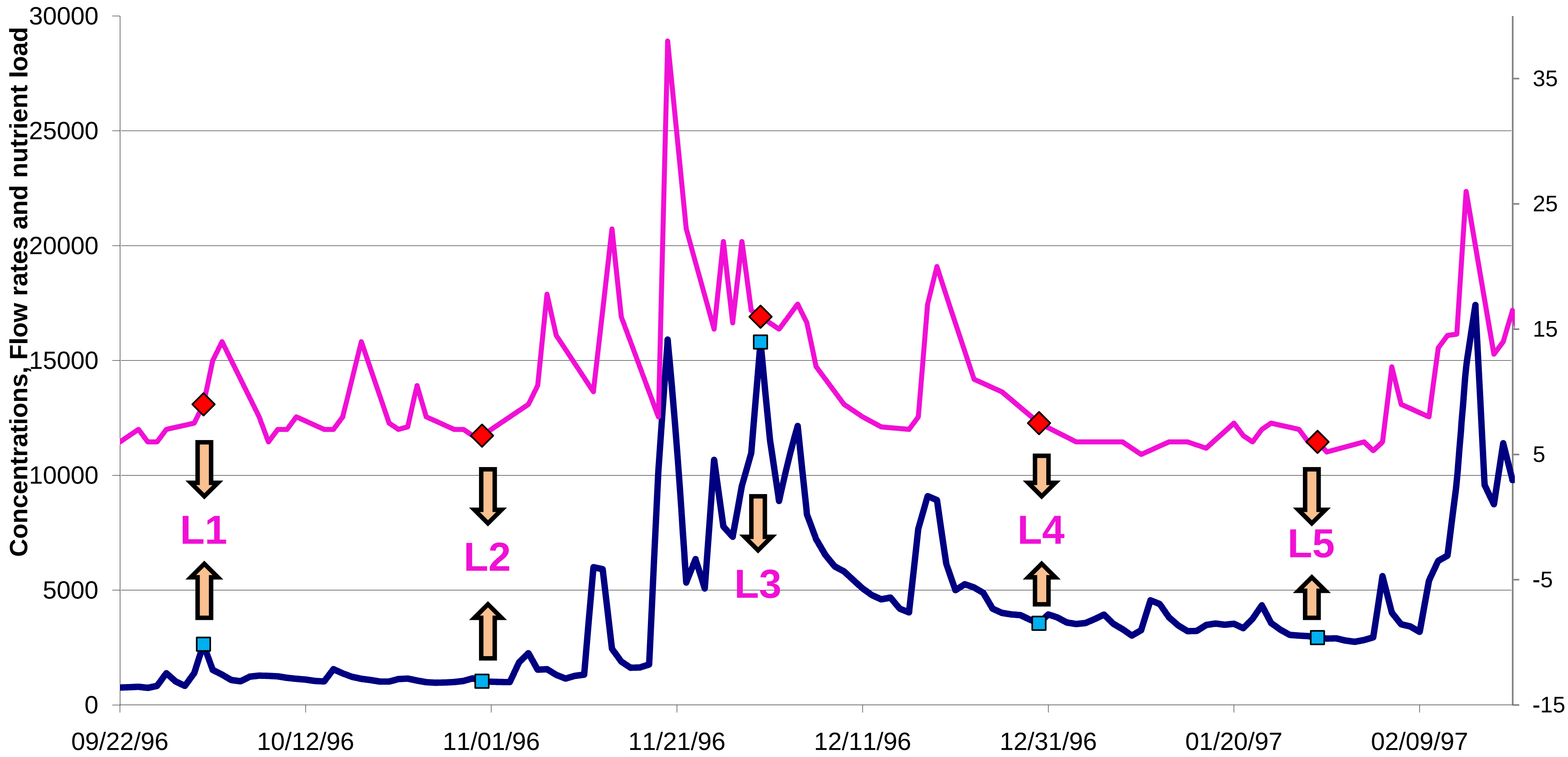
Objectives

- Investigate and document the level of uncertainties induced by infrequent sampling
- Provide tools to guide watershed managers: Sampling frequency charts

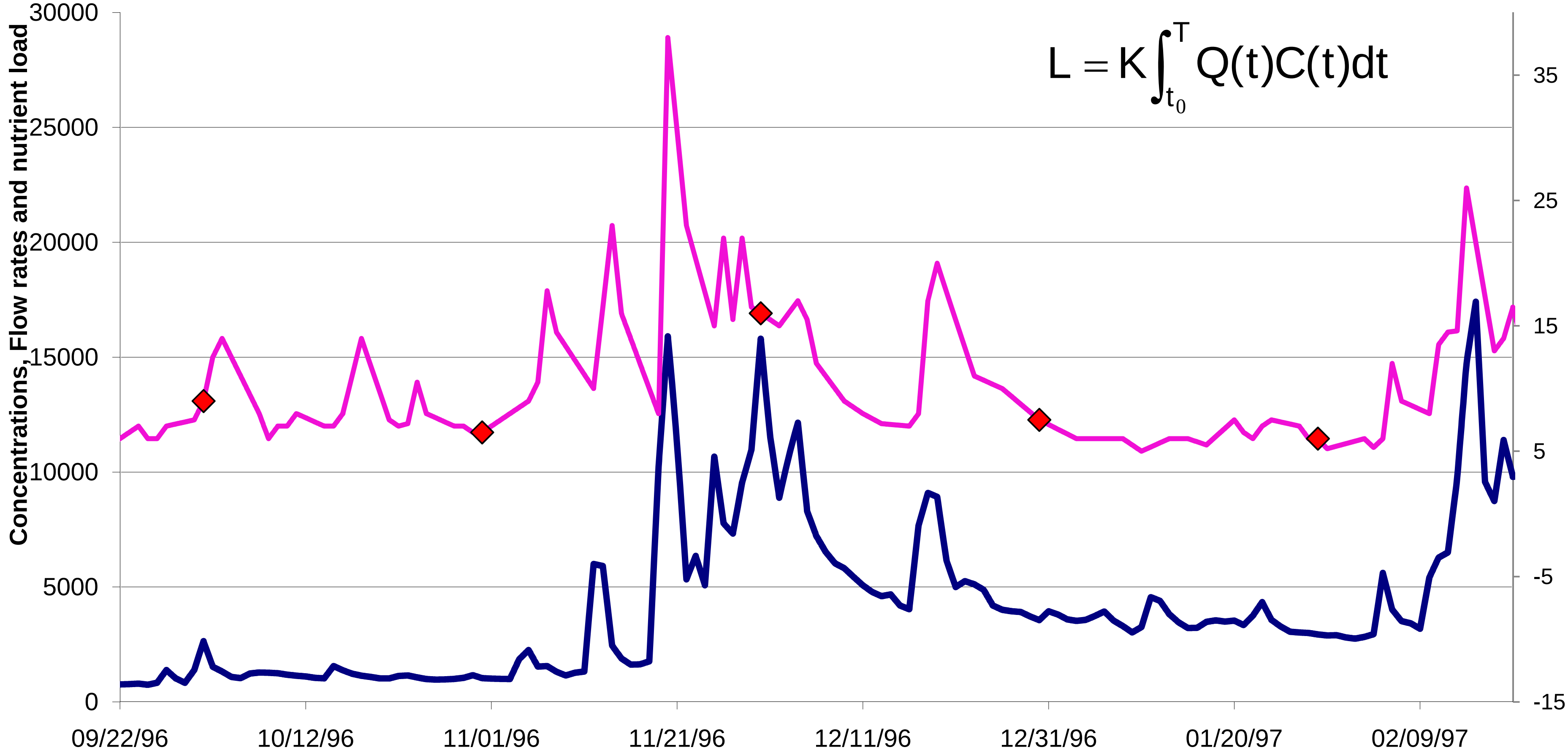
Method

- **4 reference watersheds (60 to 400 km^2)**
  - DOM
- **Numerical simulations of sampling strategies from reference watersheds**
- **16 flux calculation algorithms and sampling strategies tested**
  - **“Averaging” methods**
  - **Linear interpolation method**
  - **Regression methods**
  - **Composite sampling**

Fluxes: « Averaging » methods



Fluxes: « Averaging » methods

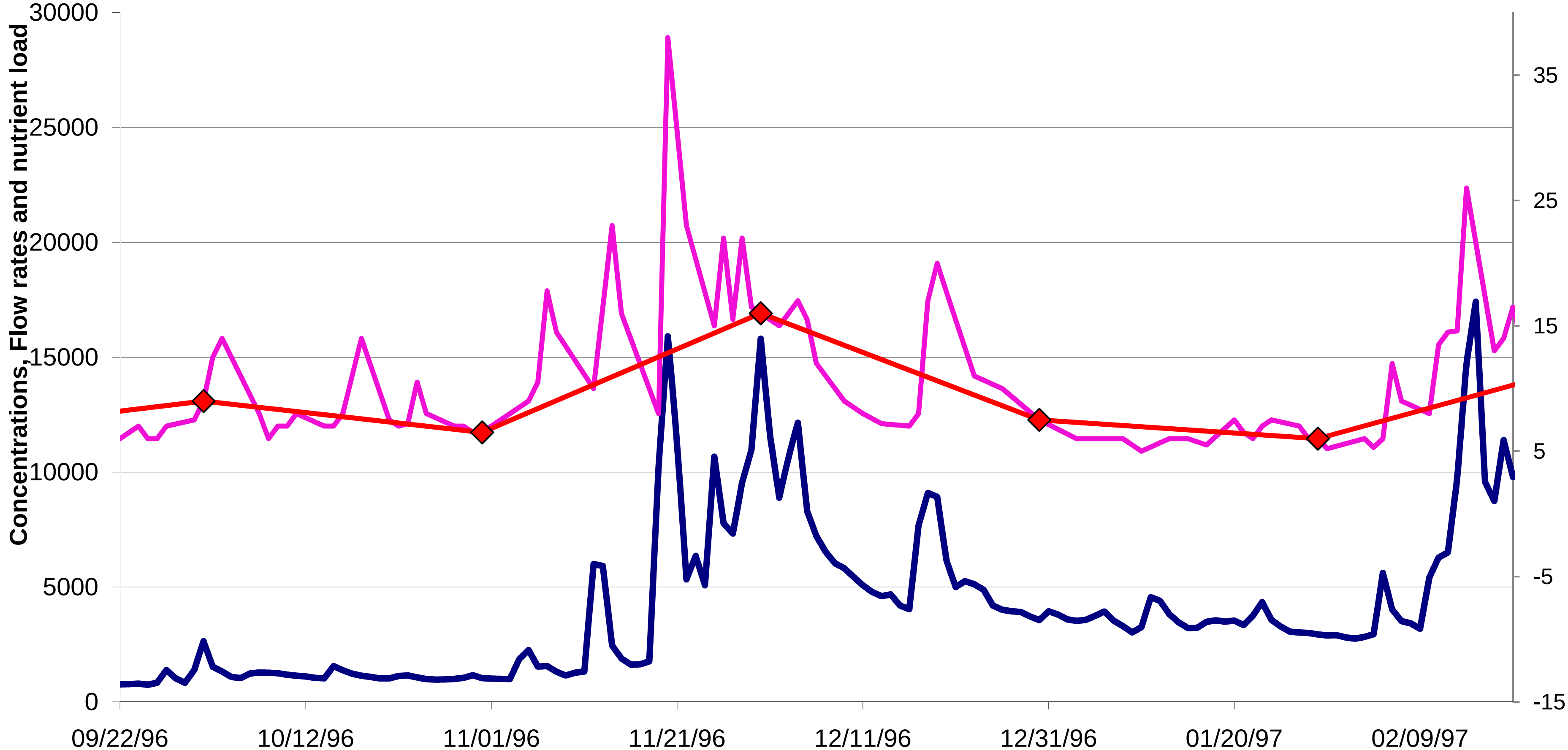


$$L = K\int_{t_0}^{T} Q(t)C(t)dt$$

« Method 5 »

$$\hat{L} = K\,\overline{Q}\,\frac{\sum_{i=1}^{n} C_i Q_i}{\sum_{i=1}^{n} Q_i}$$

Linear interpolation



$$\hat{L} = K \sum_{k=1}^{n} C_k^{int} Q_k$$

Linear interpolation

16 strategies and methods tested

1 $$\hat{L} = K\left(\sum_{i=1}^{n}\frac{C_i}{n}\right)\left(\sum_{i=1}^{n}\frac{Q_i}{n}\right)$$

[Preston et al., 1989]

2 $$\hat{L} = K\sum_{i=1}^{n}\frac{C_iQ_i}{n}$$

[Preston et al., 1989]

9 $$\hat{L} = K\sum_{i\in n} C_iQ_i\left(1+\frac{\sum_{j\in N\setminus n} Q_j^2}{\sum_{i\in n} Q_i^2}\right)$$

[Cooper, 2005]

3 $$\hat{L} = K\sum_{i=1}^{n} C_i\overline{Q_{i,i-1}}$$

[Meybeck et al., 1994]

4 $$\hat{L} = K\overline{Q}\sum_{i=1}^{n}\frac{C_i}{n}$$

[Shih et al., 1992]

5 $$\hat{L} = K\overline{Q}\frac{\sum_{i=1}^{n} C_iQ_i}{\sum_{i=1}^{n} Q_i}$$

[Littlewood, 1992]

8 $$\hat{L} = K\,\overline{Q}\,\frac{\bar{l}}{\bar{q}}\,\frac{1+\frac{1}{N}\frac{S_{lq}}{\bar{l}\bar{q}}}{1+\frac{1}{N}\frac{S_{q^2}}{\bar{q}^2}}$$ with $$\bar{l} = \sum_{i=1}^{n} C_iQ_i \qquad \bar{q} = \sum_{i=1}^{n} Q_i$$

$$S_{lq} = \frac{1}{n-1}\sum_{i=1}^{n}((Q_i-\bar{q})(C_iQ_i-\bar{l}))$$ and $$S_{q^2} = \frac{1}{n-1}\sum_{i=1}^{n}(Q_i-\bar{q})^2$$

[Cohn, 1995]

16 strategies and methods tested

$$\hat{L} = K\sum_{k=1}^{n} C_k^{int} Q_k$$

[Moatar and Meybeck, 2004]

$$\hat{L} = K\sum_{k=1}^{n} C_k^{ext} Q_k \qquad \ln(Ci) = a\ln(Qi) + b$$

[Cohn et al., 1989]

[Ferguson, 1986]

$$\hat{L} = K\sum_{k=1}^{n} C_k^{ext} Q_k \times \exp\left(\frac{s^2}{2}\right) \quad \text{with} \quad s^2 = \frac{\sum_{i=1}^{n}\left(\ln(C_i) - (a \times \ln(Q_i) + b)\right)^2}{n-2}$$

[Duan, 1983]

$$\hat{L} = K\sum_{k=1}^{n} C_k^{ext} Q_k \times \frac{\sum_{i=1}^{n}\exp\left(\ln(C_i) - (a \times \ln(Q_i) + b)\right)}{n}$$

$$\hat{L} = K\sum_{k=1}^{n} C_k^{ext} Q_k \times g_m\left(\frac{m+1}{2m} \times (1 - V) \times s^2\right)$$

[Cohn et al., 1989]

$$\hat{L} = K\sum_{b \in B} L_b$$

[Schleppi et al., 2006]

$$L_b = \left(\frac{1}{180}\sum_{i=1}^{180} C_i\right) \times \sum_{i=1}^{180} Q_i$$

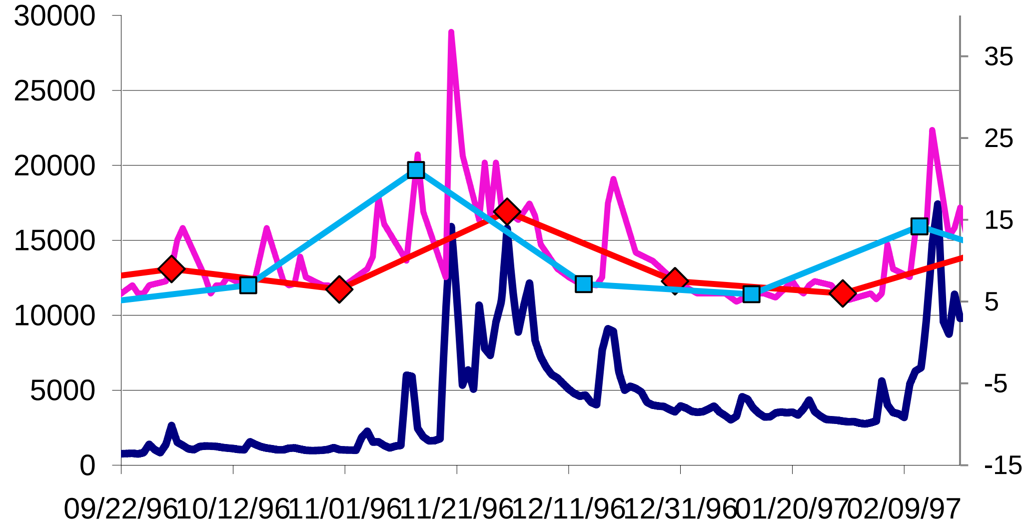
$$L_{int} = K' \times \sum_{k \in K} C_k^{int} Q_k$$

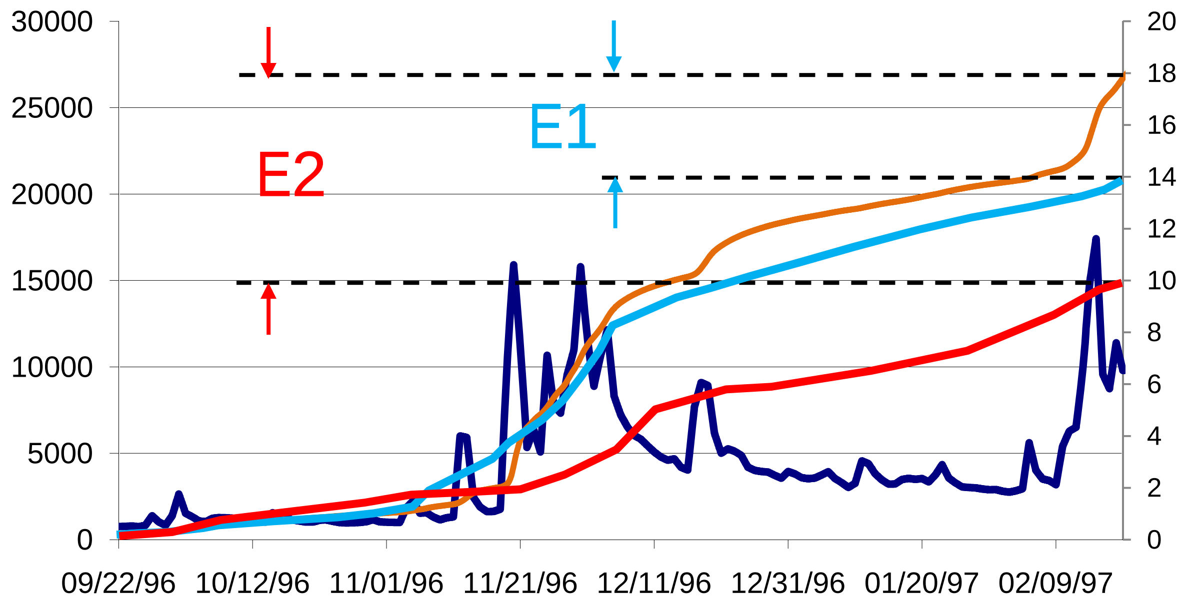
$$\hat{L} = \sum_{b \in B} L_b + \sum_{int=1}^{P} L_{int}$$

$$L_b = K \times \left(\frac{1}{n}\sum_{i=1}^{n} C_i\right) \times \sum_{i=1}^{n} Q_i$$

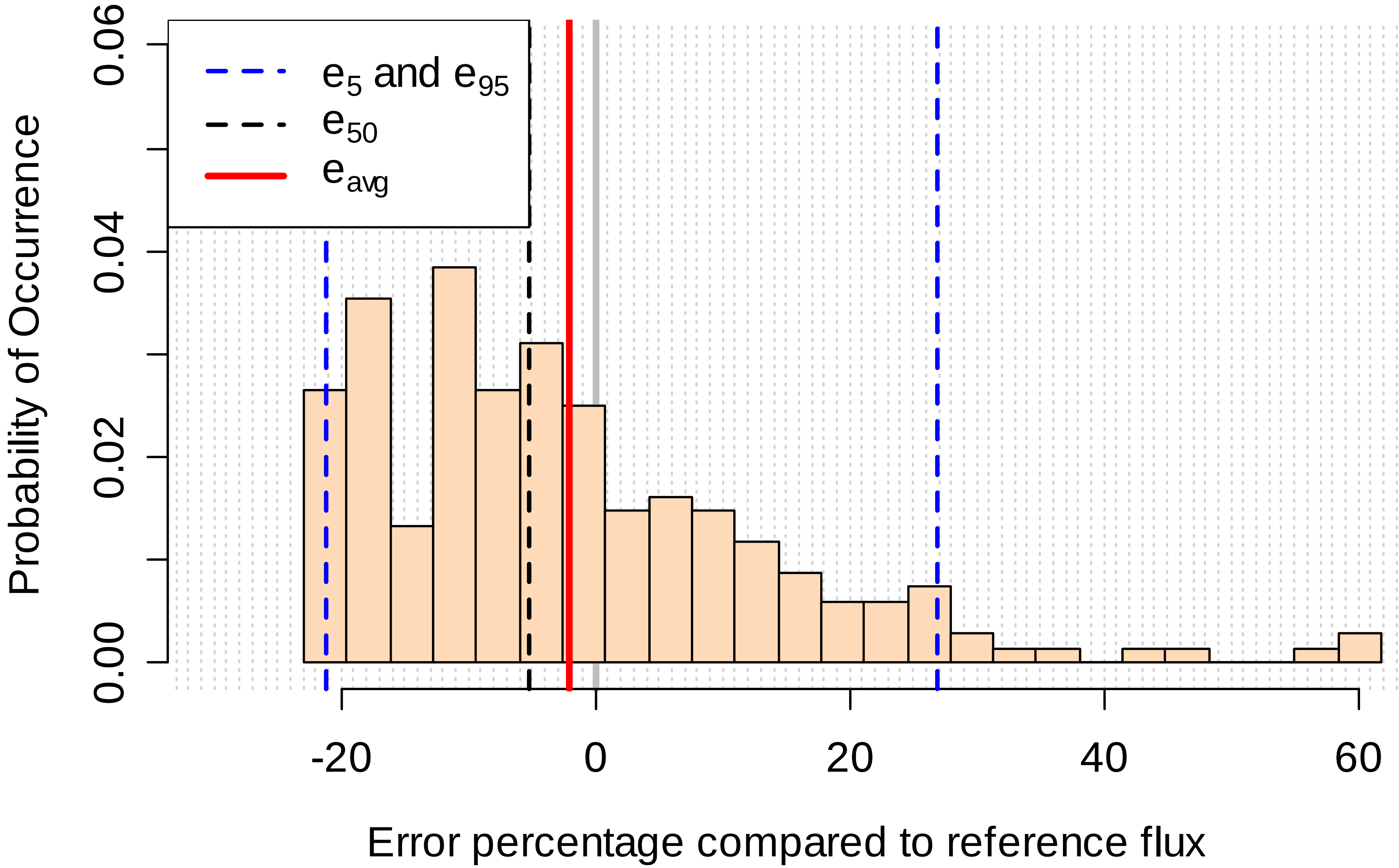
[Schleppi et al., 2006]

An infinite number of possible errors

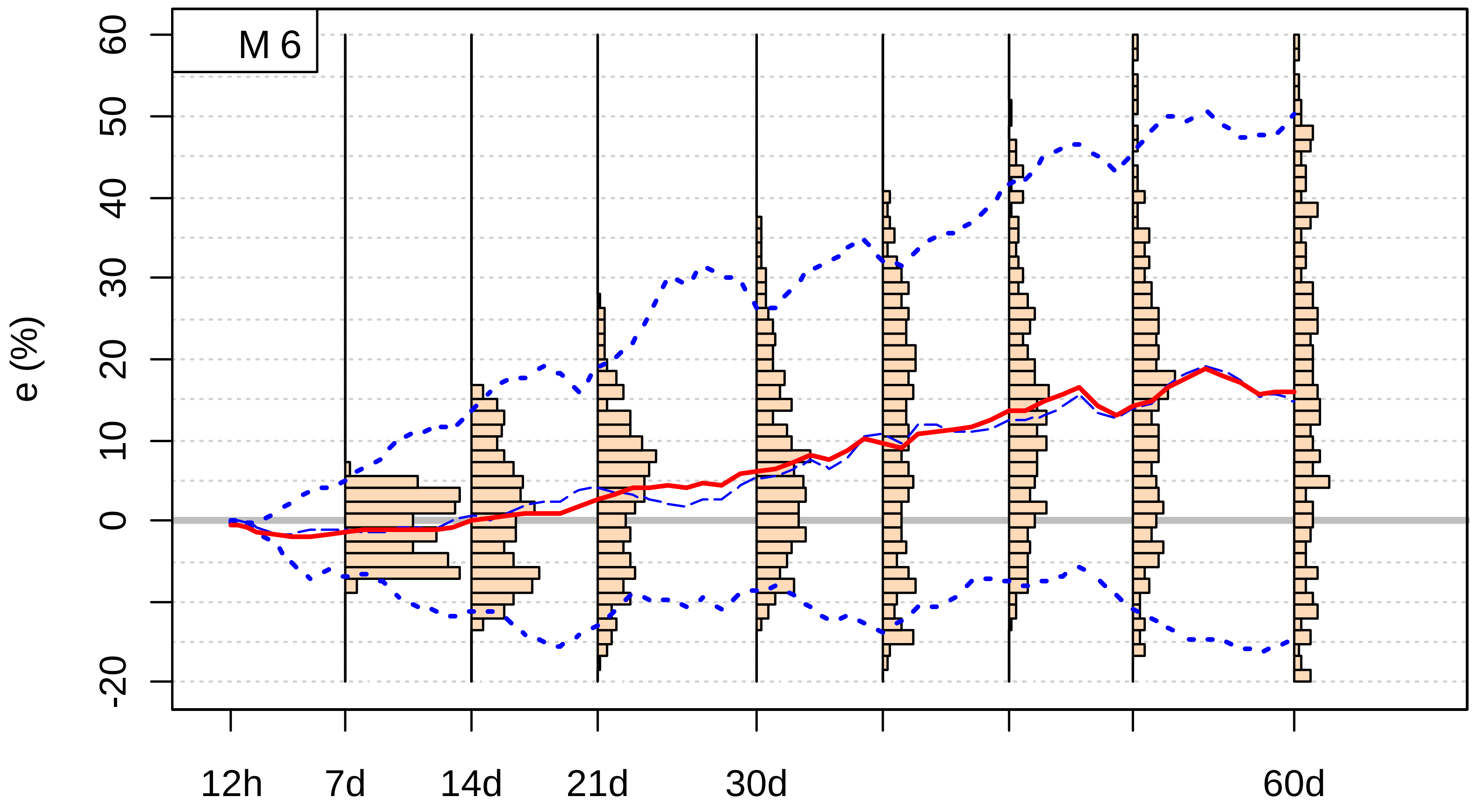




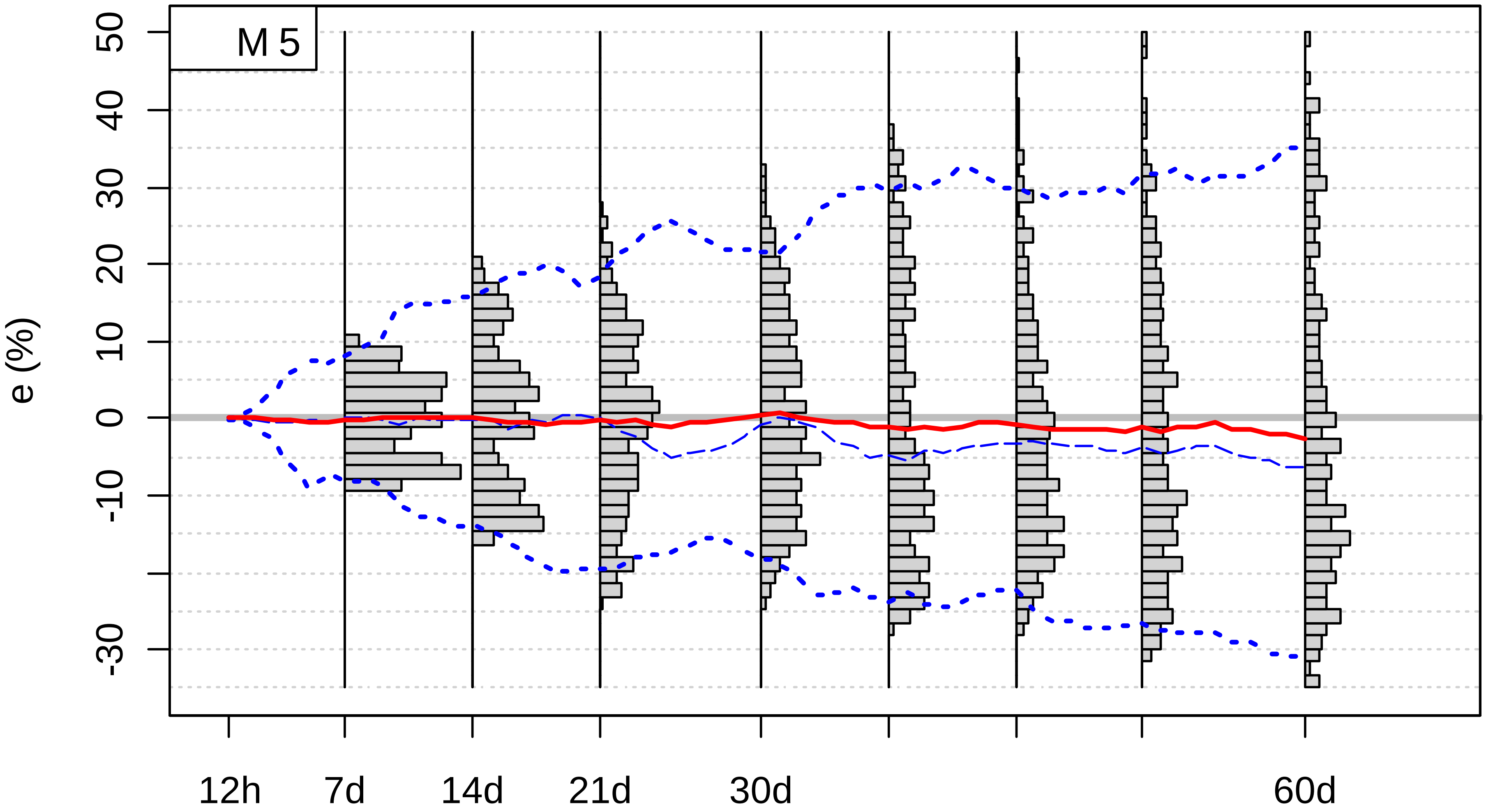
Error Distributions



Errors as a function of sampling frequency



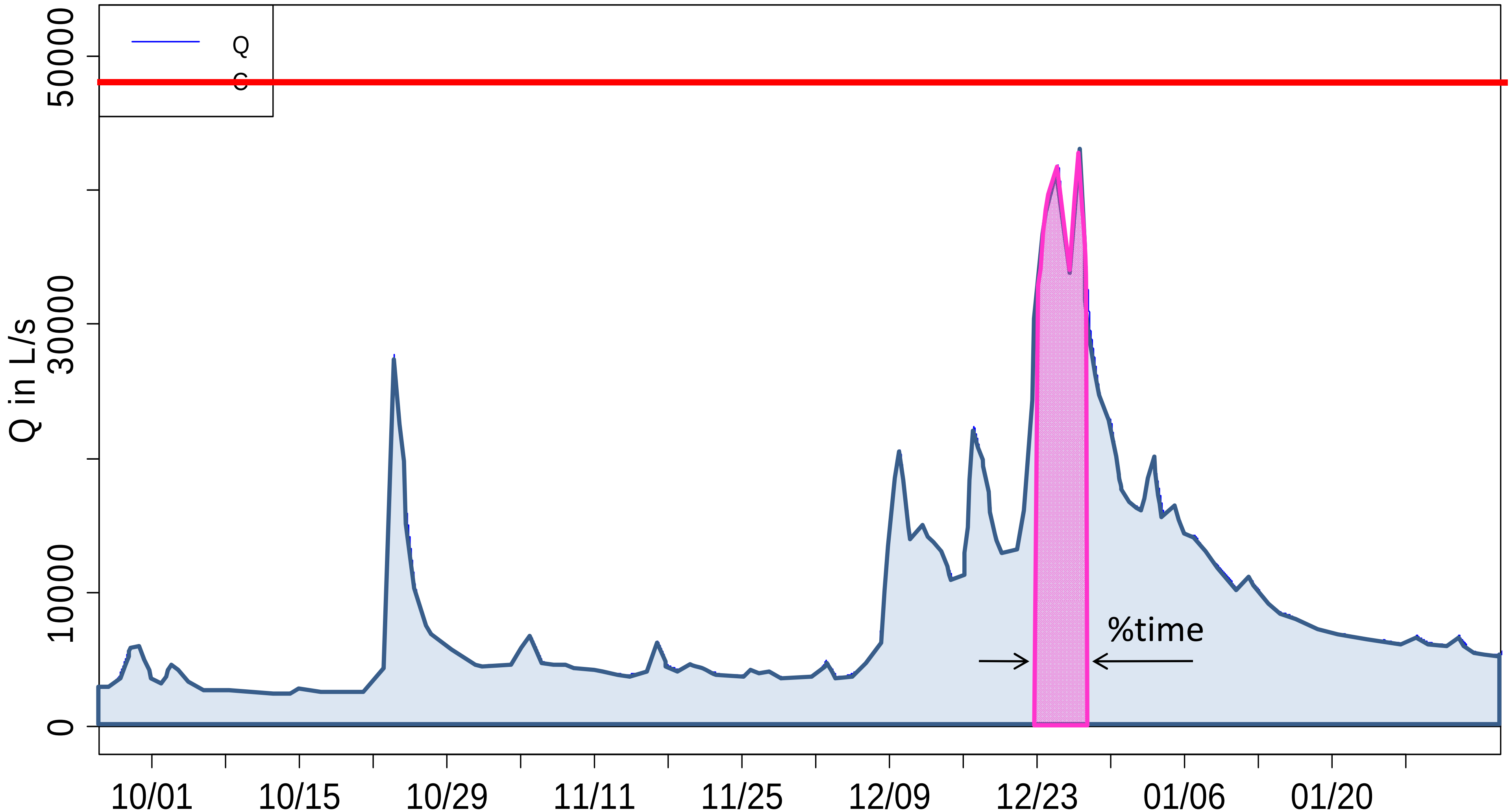
Flow weighted average: best method



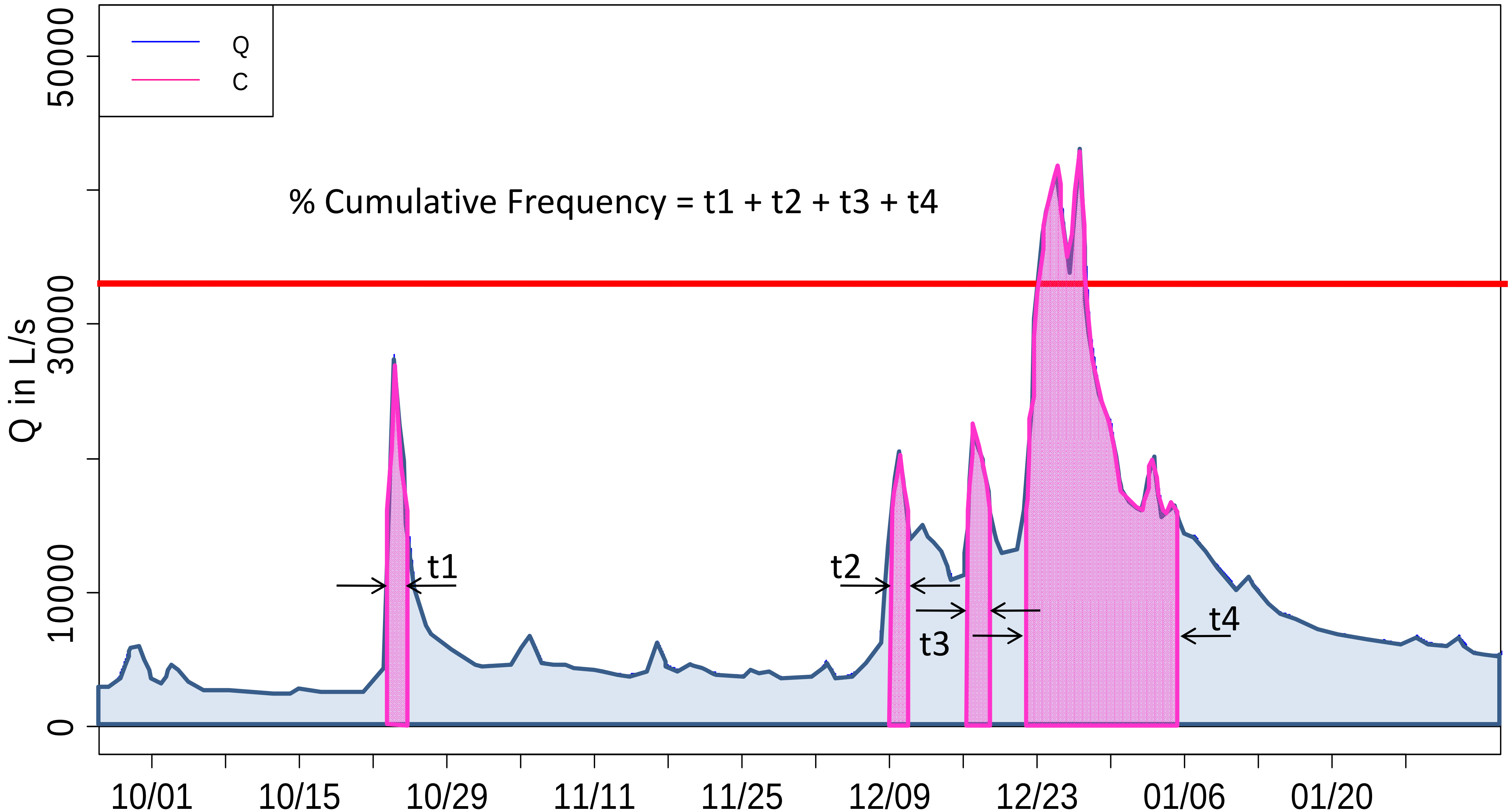
COULD ONE PREDICT THE MAGNITUDE OF ERRORS?

CASE STUDY: DISSOLVED ORGANIC MATTER IN BRITTANY, FRANCE

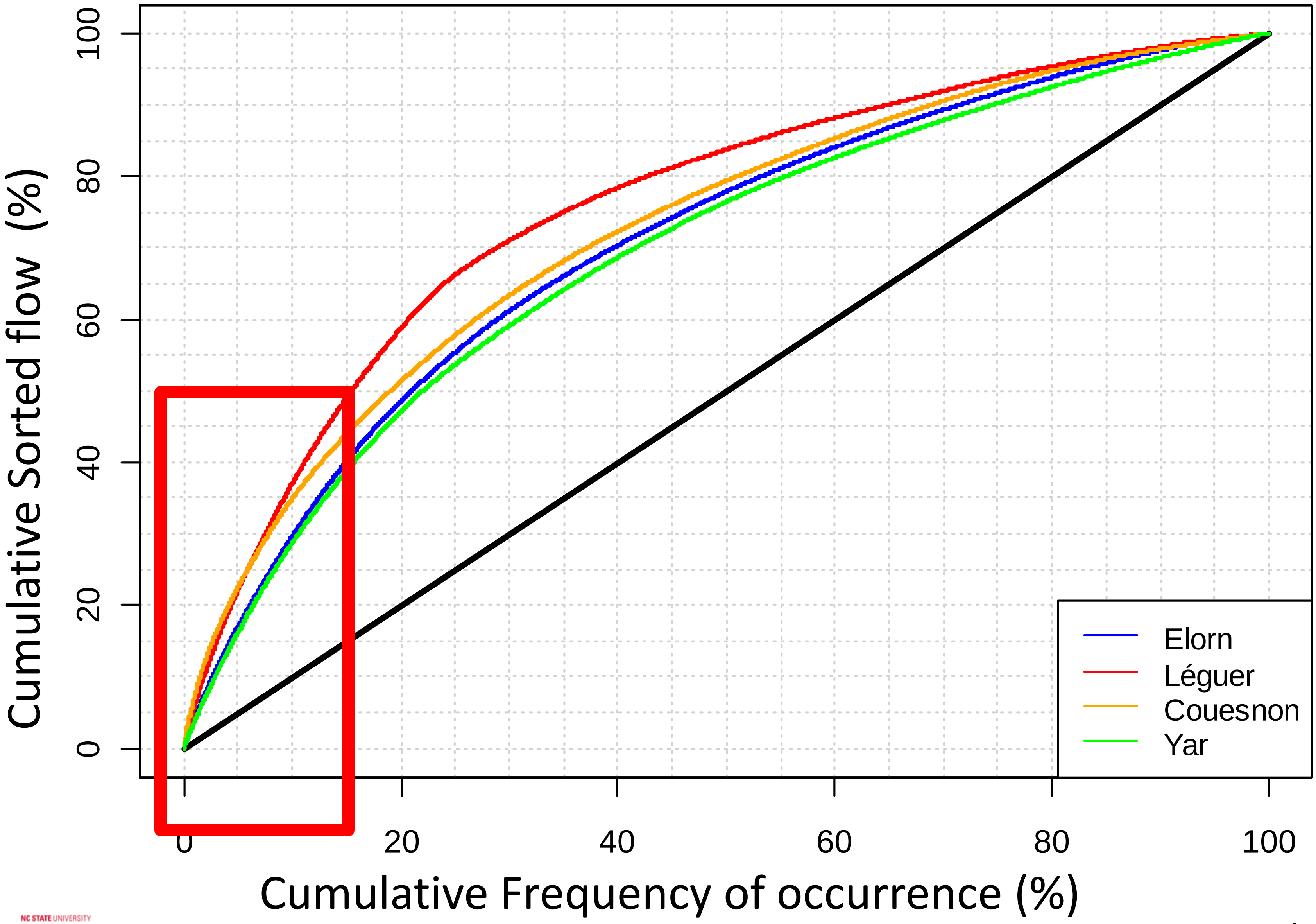
Deriving watershed reactivity indicators



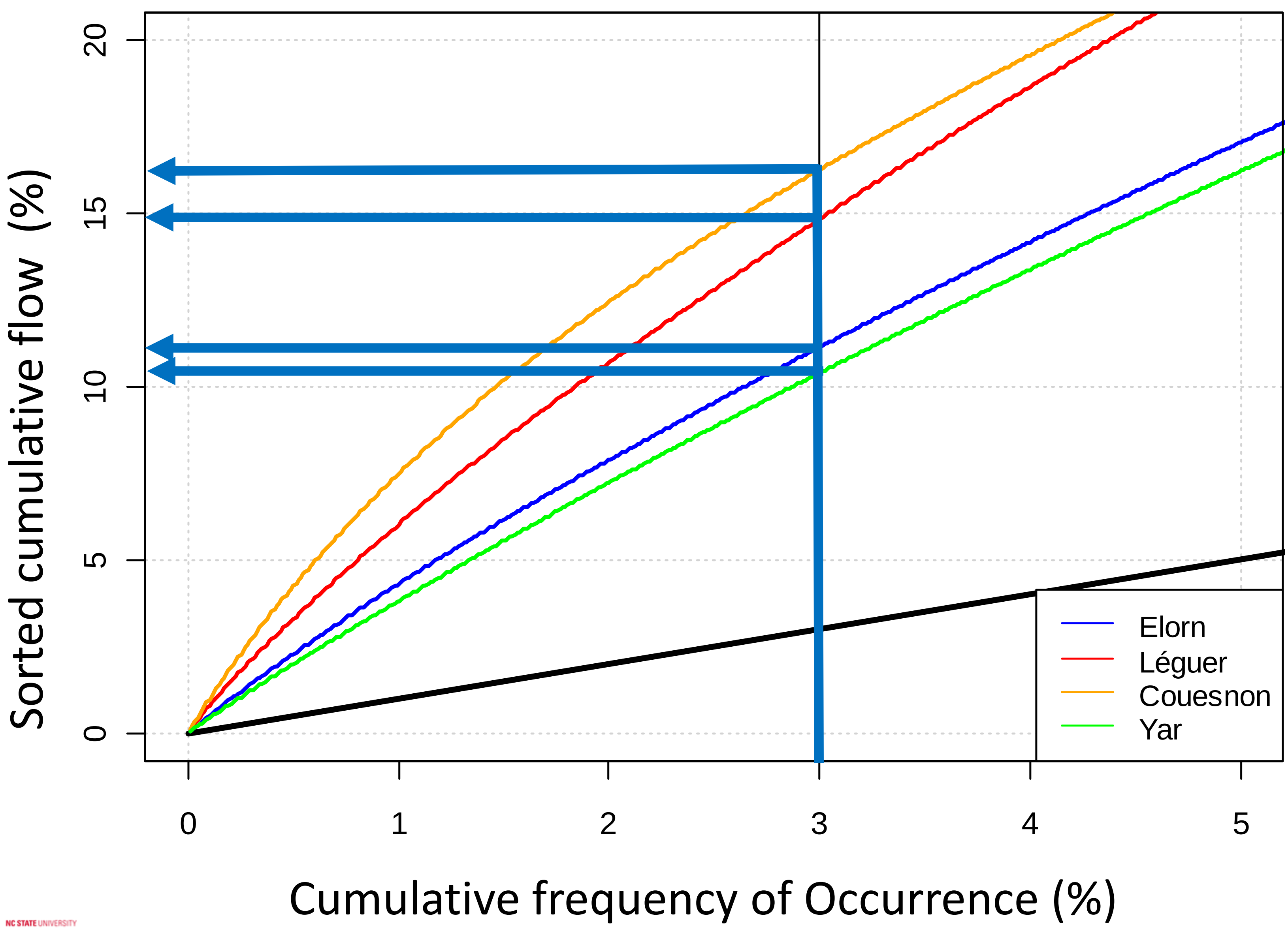
Deriving watershed reactivity indicators



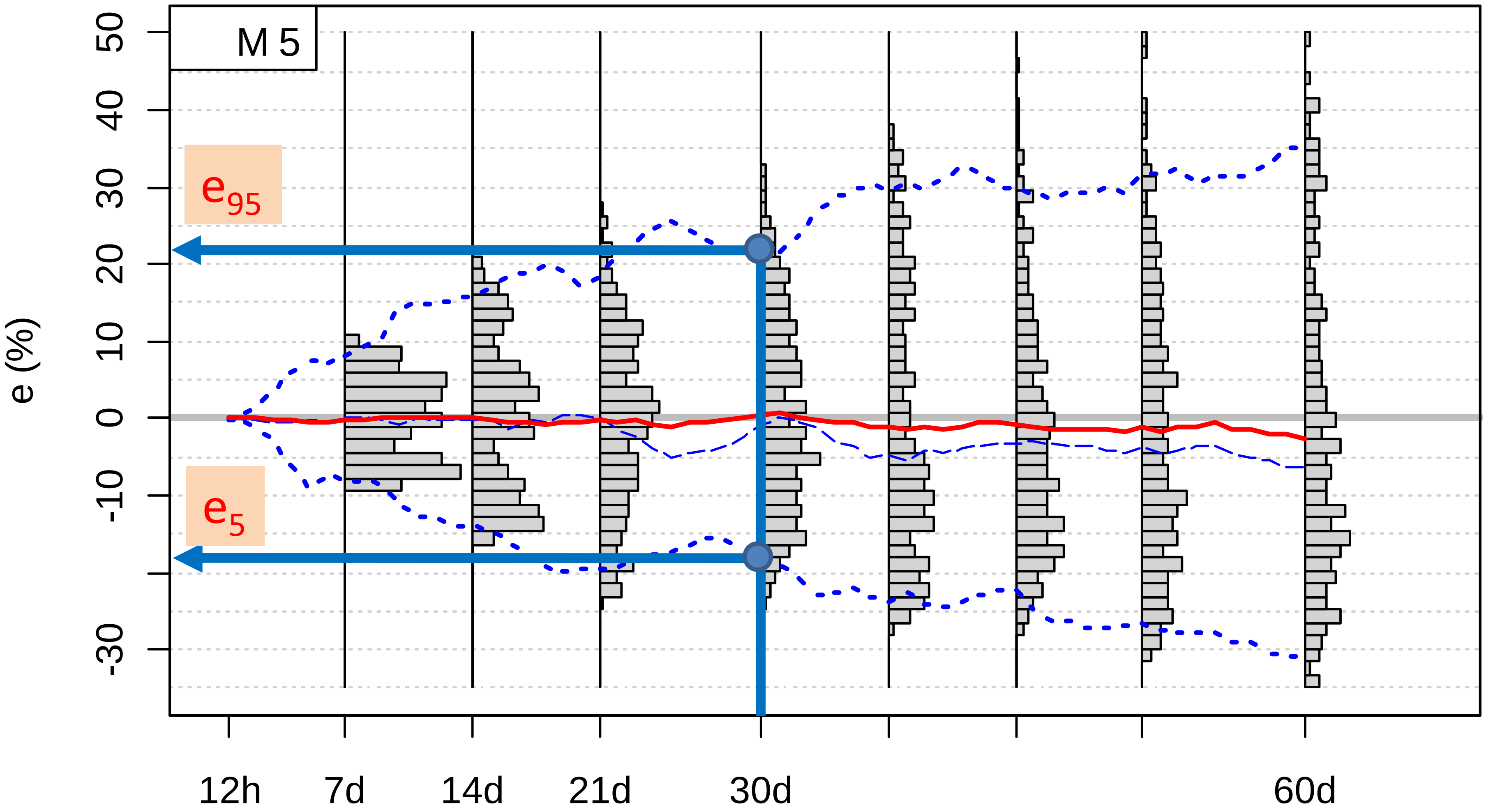
Sorted cumulative flow frequency



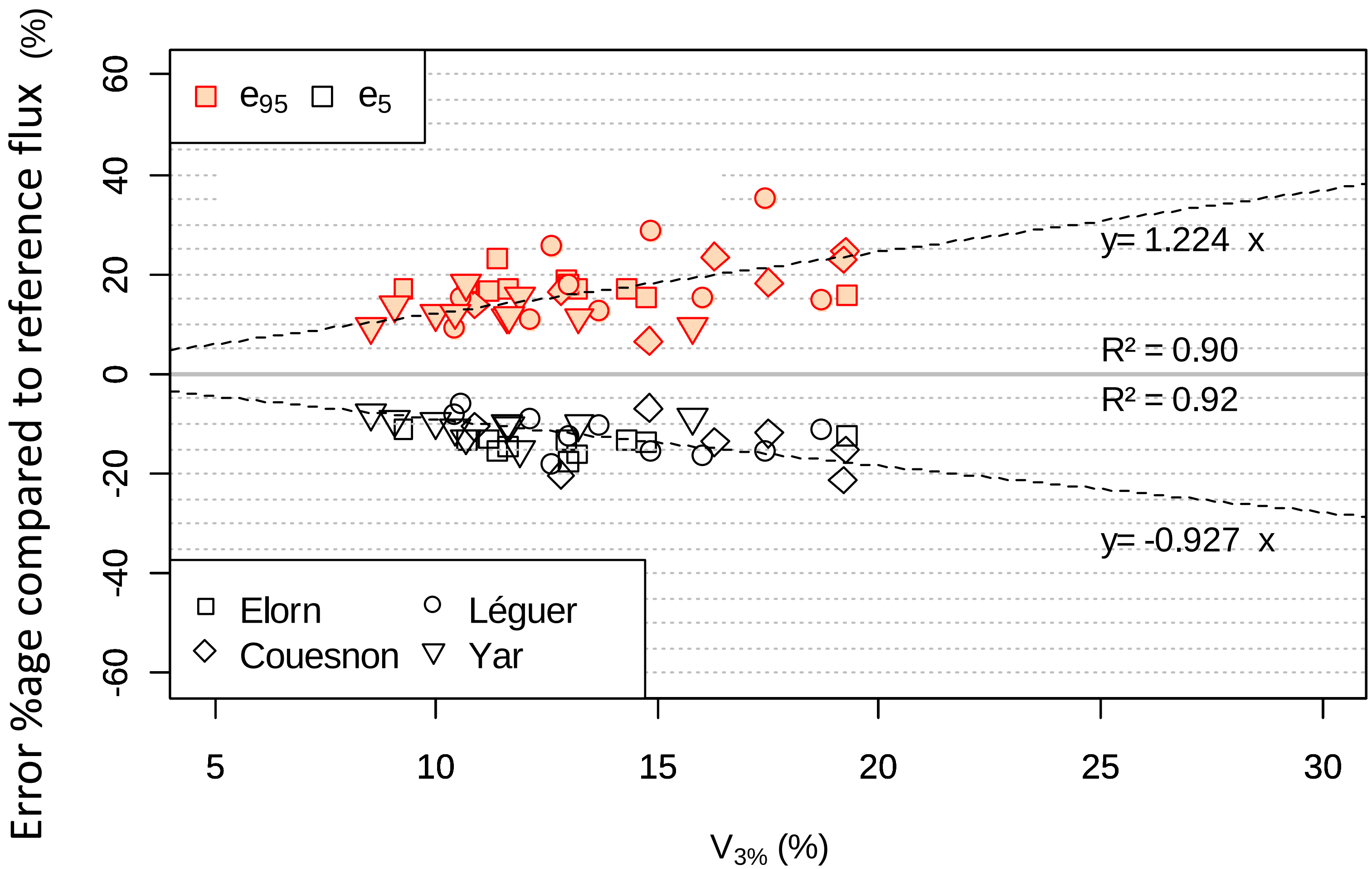
Indicator $V_{3\%}$



Errors on the annual flux estimator

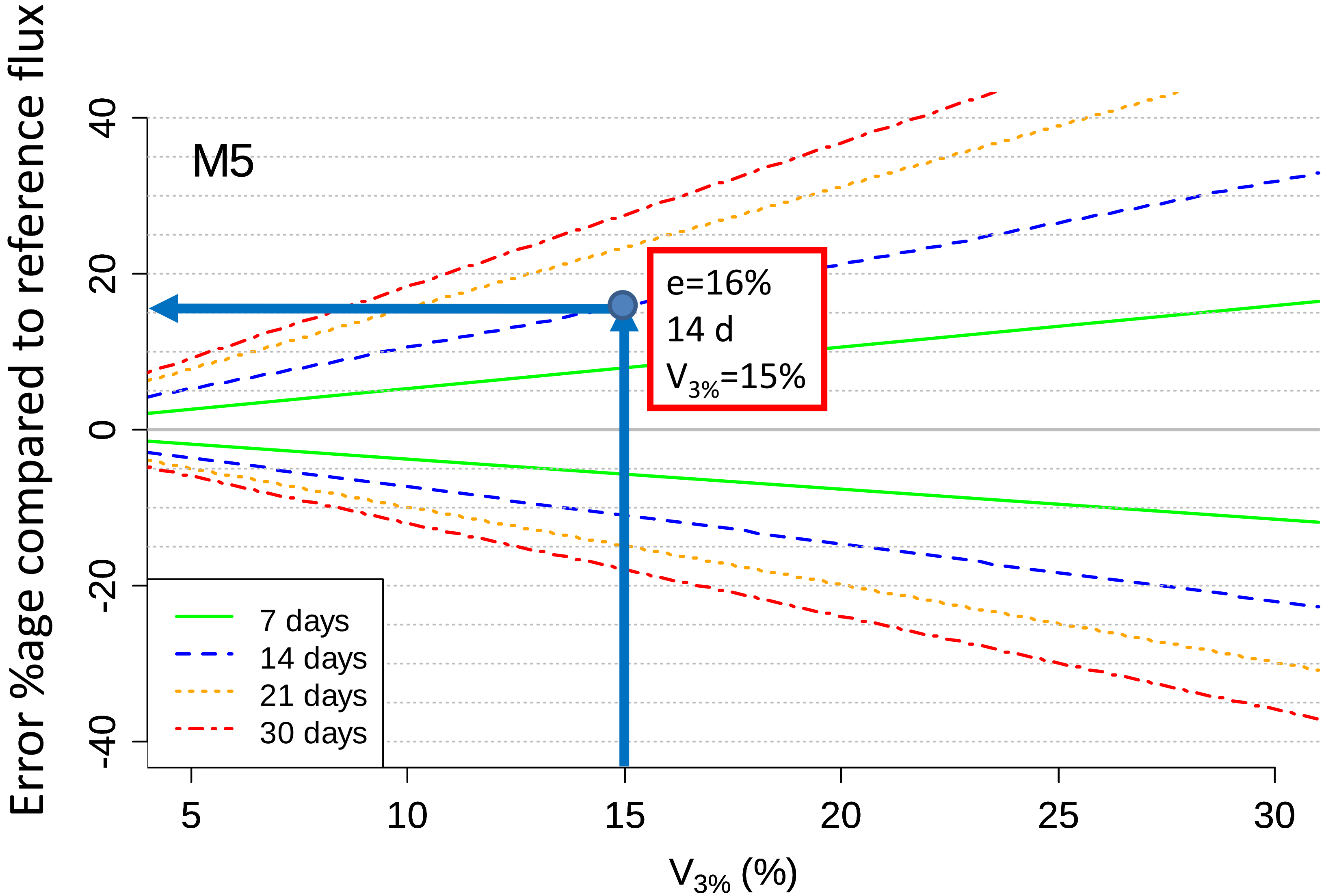


Link between $V_{3\%}$ and estimator errors

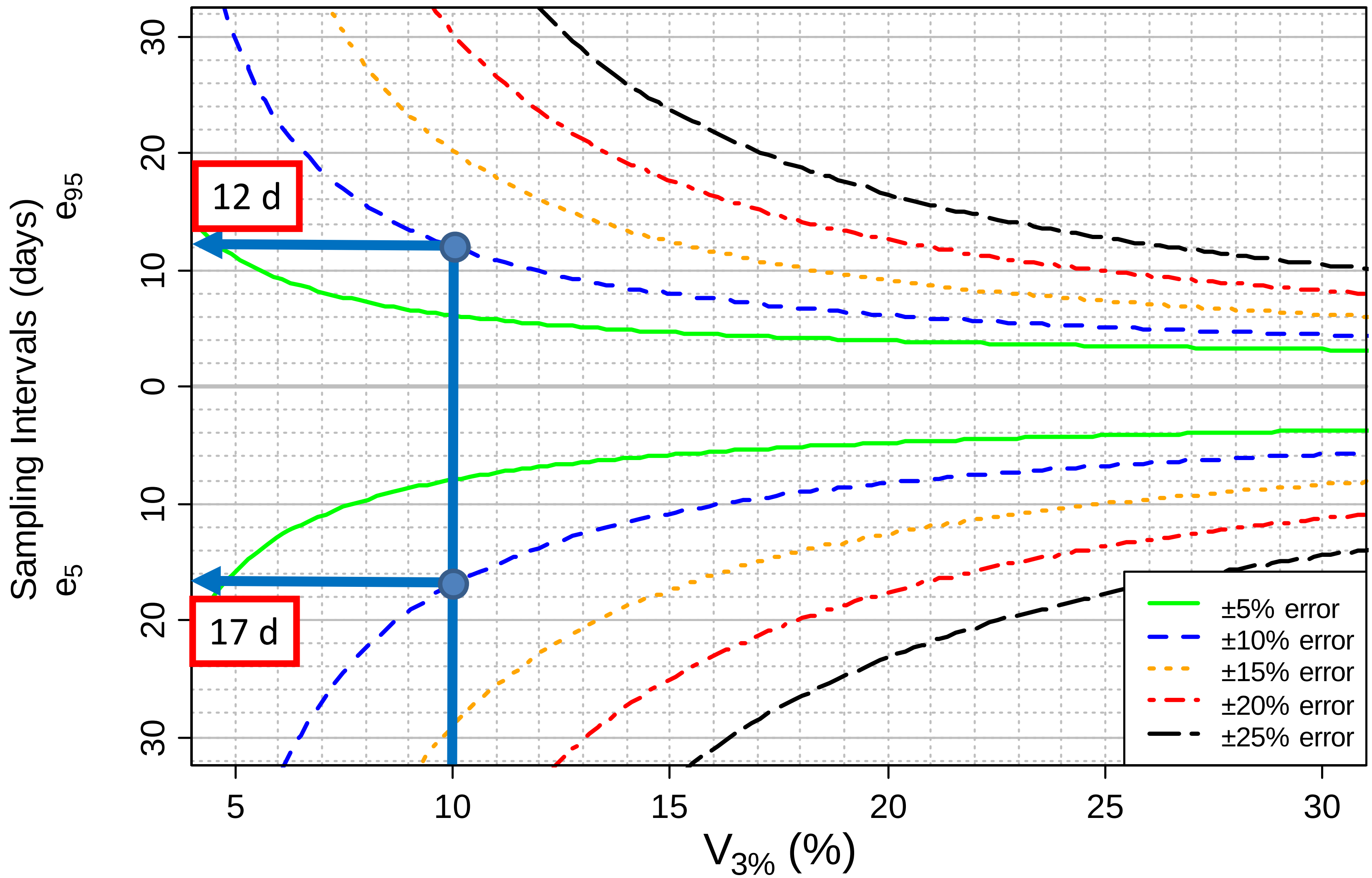


Percentage of the annual volume flowing in 3% of the time corresponding to the highest flow

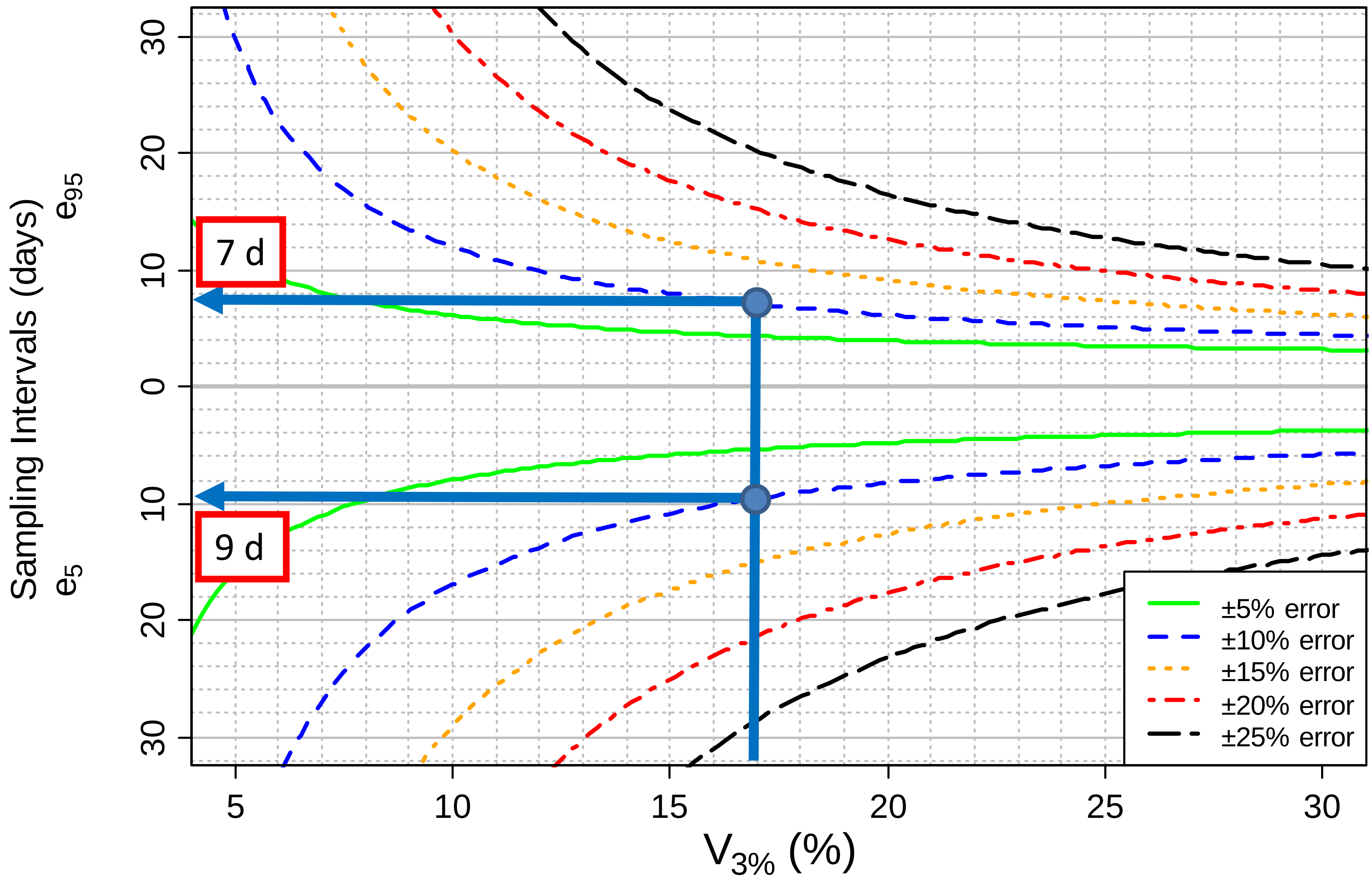
Unique link between errors and $V_{3\%}$



Guideline curves for DOM in Brittany, France



Guideline curves for DOM in Brittany, France



Conclusions

- Infrequent sampling does imply errors on water quality indicators such as annual fluxes and concentration indicators
- Flow weigthed concentration method best for this region
- Uncertainties depend largely on watershed hydrological reactivity

Conclusion (Con'd)

- Annual flux for bi-monthly sampling: -13% to +17%, slightly biased ($V_{3\%}$ = 15%)
- Errors on monthly sampling are rather low compared to other pollutants such as TSS (e.g. 200%) but higher than for nitrate
- Apply the same and other statistical methods to North Carolina data
- Join the study, share your data!!

Thank you for your kind attention!

francois_birgand@ncsu.edu