Addressing the Uncertainties on Annual Nitrate Load Calculations in the Neuse River in North Carolina

Yimin Kao[1] - François Birgand[2] - Brian J. Reich[1] – William J. Showers[3]

NC WRRI conference – Raleigh, NC March 22,2011

Department of Statistics[1]

Department of Biological and Agricultural Engineering[2]

Department of Marine Earth and Atmospheric Sciences[3]

Introduction

- Annual nitrate load: $$L = \sum_{k=1}^{N} C_k Q_k$$

  Where C_k and Q_k are flow and concentration value in time k.
- For financial and technical reasons, nutrient loads in rivers are often calculated from **infrequent discrete water samples**.
  - Ex: Weekly, bi-weekly, or monthly samples
- The results are just **partial** information on the concentration variations.
- Therefore, this induces **uncertainty** on the loads calculations.

Sources of Uncertainties

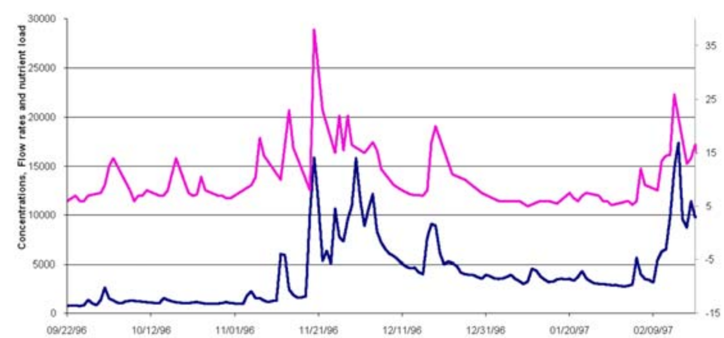
In practice, when we are measuring nutrients loads, there are sources of uncertainties:

1. Uncertainties on flow rates and cumulated flow.
2. Uncertainties due to the sampling location in the water column.
3. Uncertainties due to sample degradation between sampling and analysis.
4. Uncertainties of laboratory analyses.
5. Uncertainties associated with **infrequent sampling**

Uncertainties associated with infrequent sampling – (1)

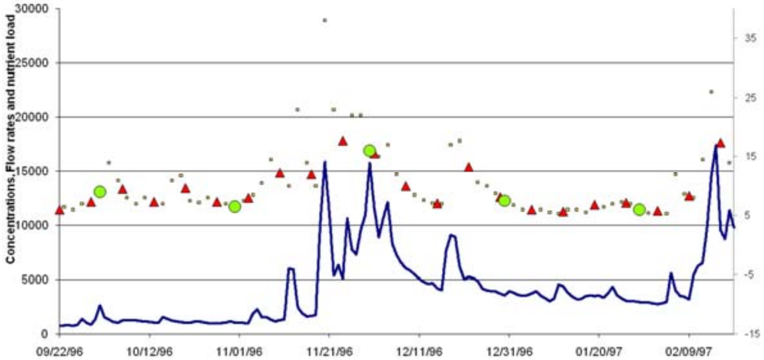
Question: Why we need sampling?

In a **perfect** world, we have continuous data:



Uncertainties associated with infrequent sampling – (2)

However, in **reality**, we possible be able to have continuous data for the flow, but **not** for the concentration:



General Objectives

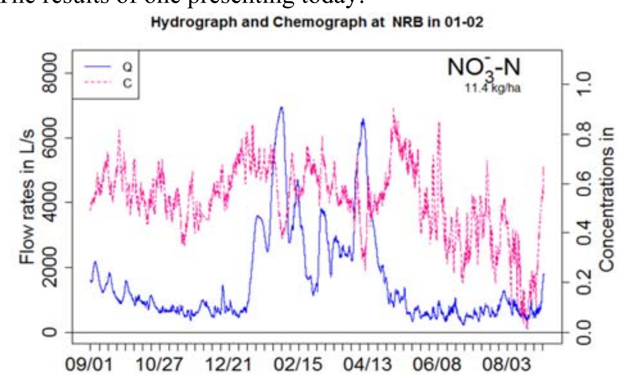
- Estimate the level of **uncertainties** for the existing nutrient loads on North Carolina rivers.
- Provide **guidance** for watershed managers on their sampling strategies and frequencies according to the potential level of uncertainties.
- Provide statistical methods and models to **lower** the level of uncertainties.

Methods

1. Use **reference** continuous flow and concentration data to calculate the **reference loads**.
2. We numerically **simulate** discrete sampling.
3. Calculate the loads from simulated samples using **algorithms**.
4. Calculate the **error** as relative difference between the reference loads and the simulated loads.
5. Derive the **expected level of uncertainty** for North Carolina conditions.

Reference Data

- **Continuous** flow and nitrate concentration data from the RiverNet.
- Five stations on the Neuse River.
  - The results of one presenting today:



Algorithms for calculating fluxes

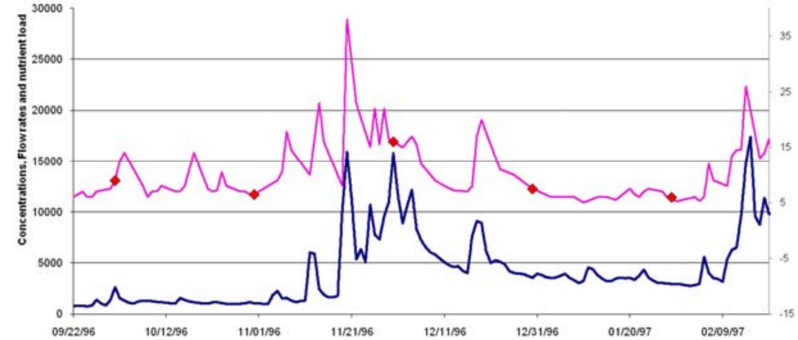
- There are plenty of algorithms for calculating fluxes.
- We present the top two algorithms that have the best estimation:

1. « Averaging method » (M5)
2. « Linear interpolation method»

Averaging Method

Formula:

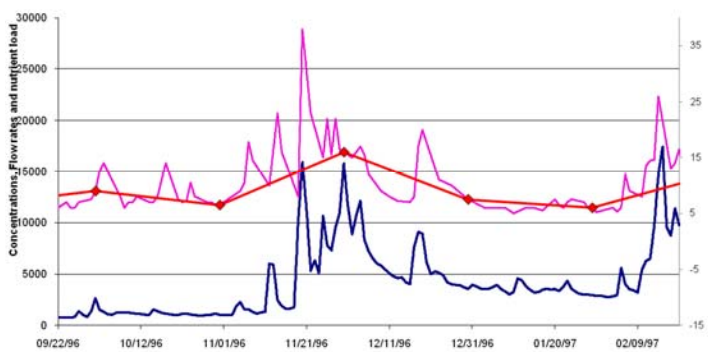
$$\hat{L} = K\overline{Q}\frac{\sum_{i=1}^{n} C_i Q_i}{\sum_{i=1}^{n} Q_i}$$



Linear Interpolation method

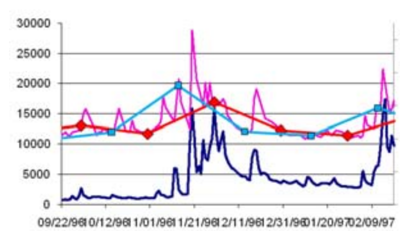
Formula:

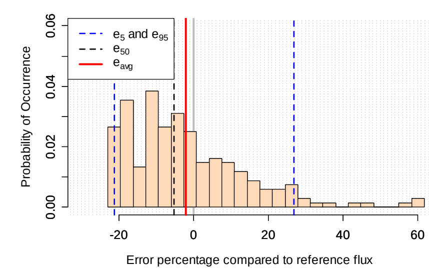
$$\hat{L} = K\sum_{k=1}^{N} C_k^{\text{int}} Q_k$$



Error Distribution

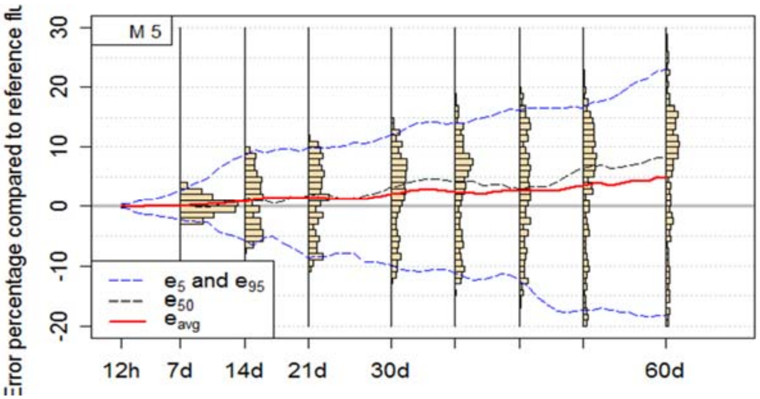
- In our study, we have the reference loads, so we can **calculate the error** that made by the estimation.
- There are **infinite number** of possible errors for "each" sampling frequency.
- Therefore, we can look at the **distribution** of the errors.





Relationship between Sampling strategy and Error distribution

- Main Idea: "More often we sampled, less errors we got."
  - Data from Neuse River Basin watershed.

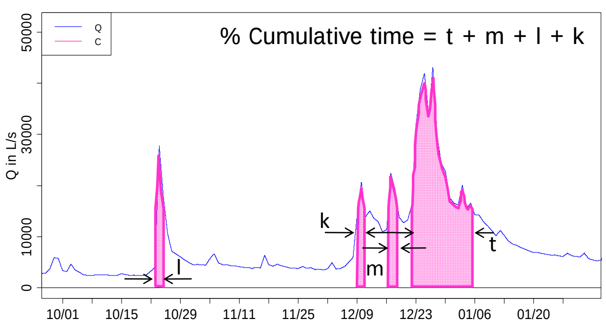


Extrapolating results

- In this stage, we have uncertainty obtained from reference watersheds, but we are interested in uncertainty from **other watersheds**.
- Can we relate level of uncertainty with some types of **watershed characteristics**? Yes!! It has been done already in previous researches!!
- Then use this relationship to predict the level of uncertainty for **non-reference watersheds**.

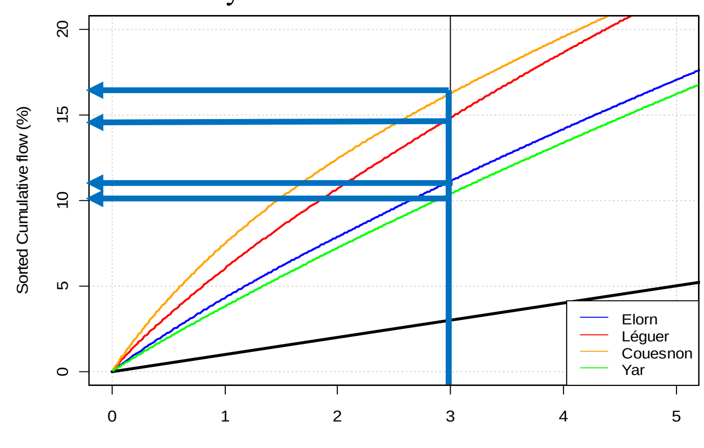
Reactivity Indicator-(1)

- A characteristic to describe the behavior of the flow in a period of time.
- $V_{w\%}$ → Percentage of the "sorted cummulative flow" under w% of "cummulative time".
- Sorted cummulative flow:

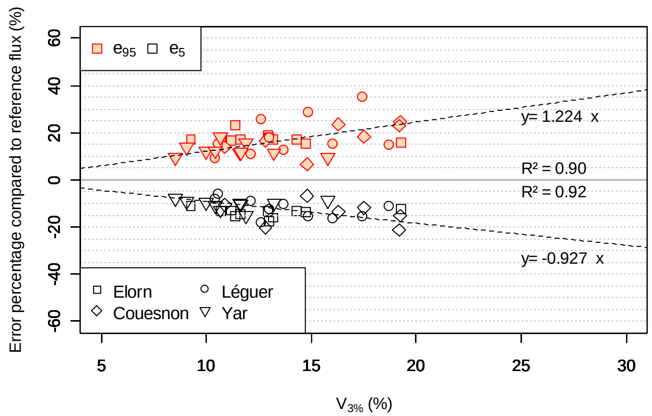


Reactivity Indicator-(2)

- What does Reactivity Indicator tells us?



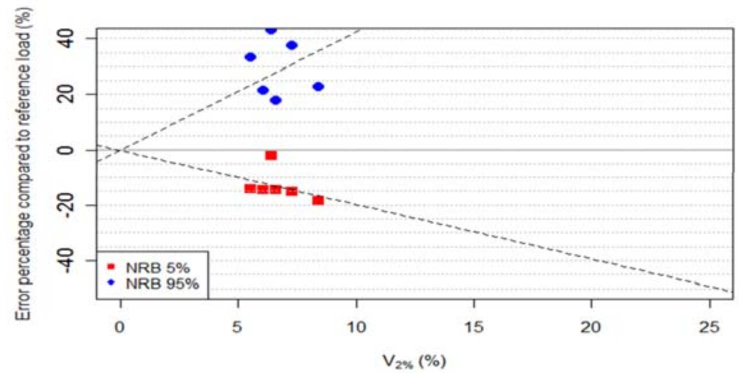
Relationship between Error distribution and Reactivity Indicator



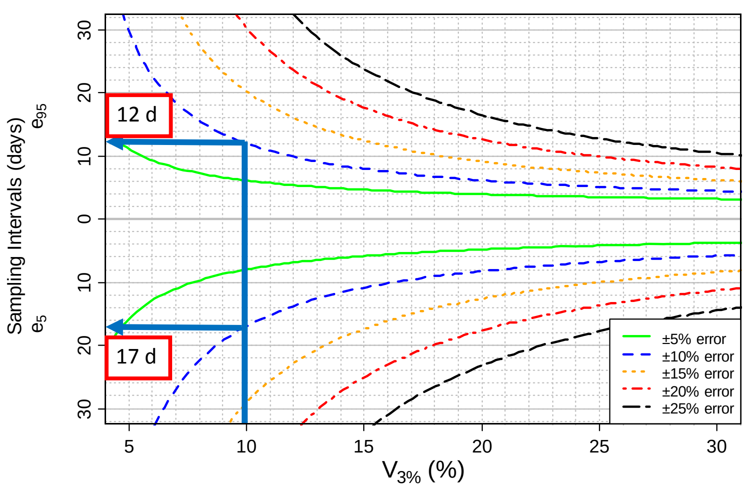
Note: This is for one specific sampling frequency.

Case study: Neuse River Basin in North Carolina

- This is one watershed with complete six hydrological years data from Neuse River Basin:



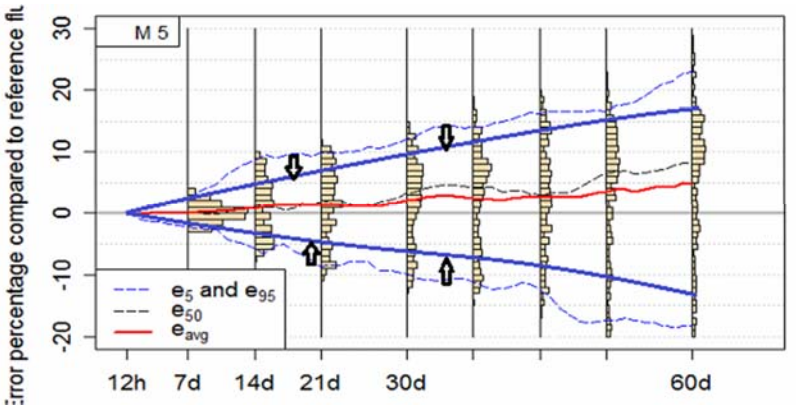
Sampling Frequency chart



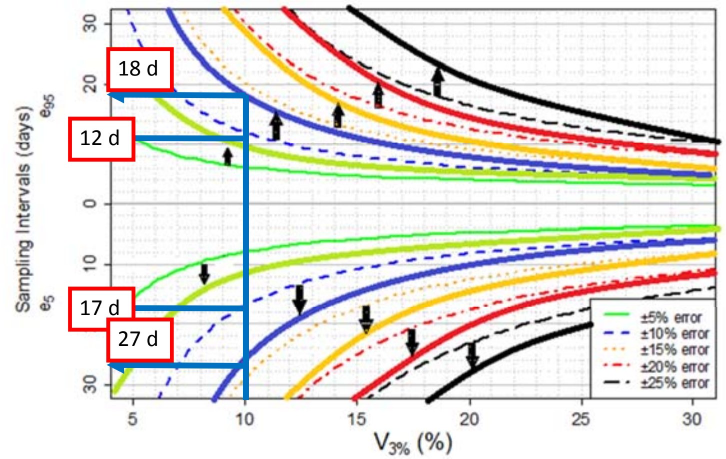
Future Works

1. More watersheds data to complete the sampling frequency chart for North Carolina.
2. Improve the algorithms of calculating fluxes.
   - Introduce new covariates, such as **"Turbidity", "SpCond", and "Water Temperature"**, in the algorithm.
   - Introduce **time correlations** in the models.
   - Consider **Spatial correlation** between different watersheds.

After improving algorithms-(1)



After improving algorithms-(2)



Thank you! Questions?