



Integrating fault detection in wastewater quality monitoring



Romain Philippe^{1,2}, Cyril Garneau^{1,2} and Peter A. Vanrolleghem^{1,2}

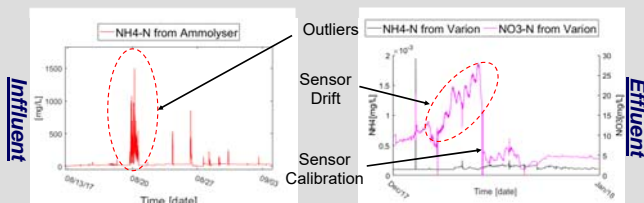
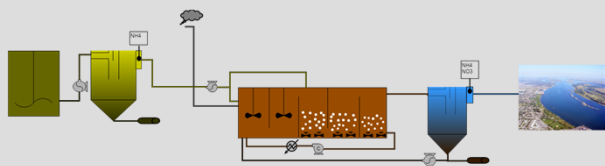
¹ modelEAU, Département de génie civil et de génie des eaux, Université Laval, Pavillon Adrien-Pouliot,

² CentrEau, Université Laval, Pavillon Adrien-Pouliot,

1065 Avenue de la Médecine, Québec (QC) G1V 0A6, Canada

Problem Statement

Nowadays, several on-line sensors are used in the wastewater field (WRRFs, sewers, rivers) to monitor wastewater quality.



Monitoring must be **reliable** and provide **good quality data**

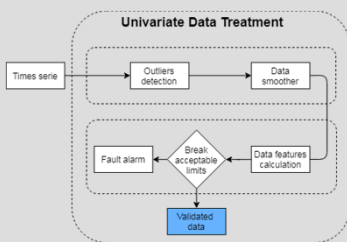
Objectives

This project proposes a **full workflow** allowing to collect raw data and to transform them into actionable information (i.e. control or process monitoring).

The **full modular framework** was applied to real time series from **sewers and WRRFs**.

Method

The fault detection and data filtering framework uses four steps to automatically validate time series from online sensors:



Time series formatting: Data with various formats are converted from their sensor dependent format to a framework-compatible format through modular functions.

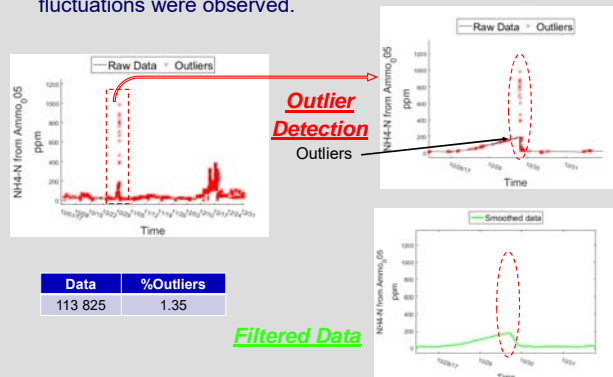
Outlier detection: Outliers are data points far from the main signal that corrupt the time series. A method based on an exponentially weighted moving average (EWMA) was used here, although other outlier detection methods can be integrated into the framework.

Data filtering: A moving average filter allows smoothing the data.

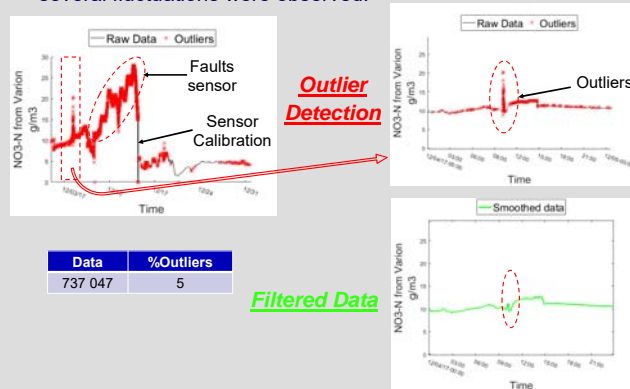
Fault feature calculation: Up to four scores (Slope, Standard deviation, upper and lower limits based on realistic ranges, Autocorrelation) can be computed for fault and event detection.

Results

- The framework was applied to 2 time series. A time series (3 months) taken from an ammolyser NH_4 sensor in the influent of WRRF containing 114 000 data points. In these data, several fluctuations were observed.



- A time series (1 month) taken from a Varion NO_3 sensor in the effluent of a WRRF containing 740 000 data points. Again, several fluctuations were observed.



Conclusion

The proposed framework offers a **modular structure** capable of providing **automated data filtration and validation**. Initial results suggest that consistent results can be achieved for different time series with a variety of sensors.

TAKE HOME MESSAGE

- Automatic validation of data is performed by a **Full Modular Framework**
- Improvement of the **reliability and quality of data** for different sensors from different locations