

# Event Detection Module - Water Quality Supervision of Distribution Networks based on Machine Learning Algorithms

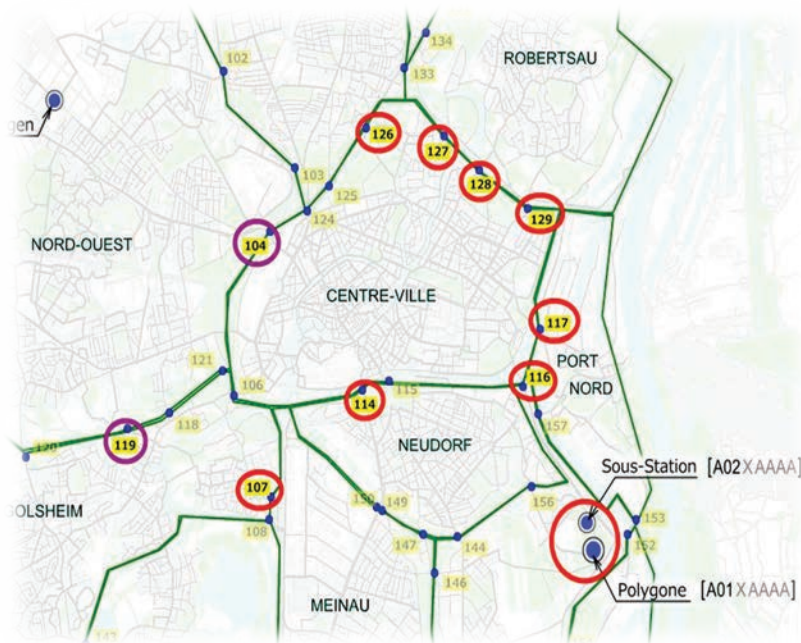
Christian Kühnert , Thomas Bernard, Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (IOSB), Karlsruhe, Germany

Idel Montalvo, 3S Consult, Karlsruhe, Germany

Reik Nitsche, DVGW-Technologiezentrum Wasser (TZW), Dresden, Germany

# Motivation

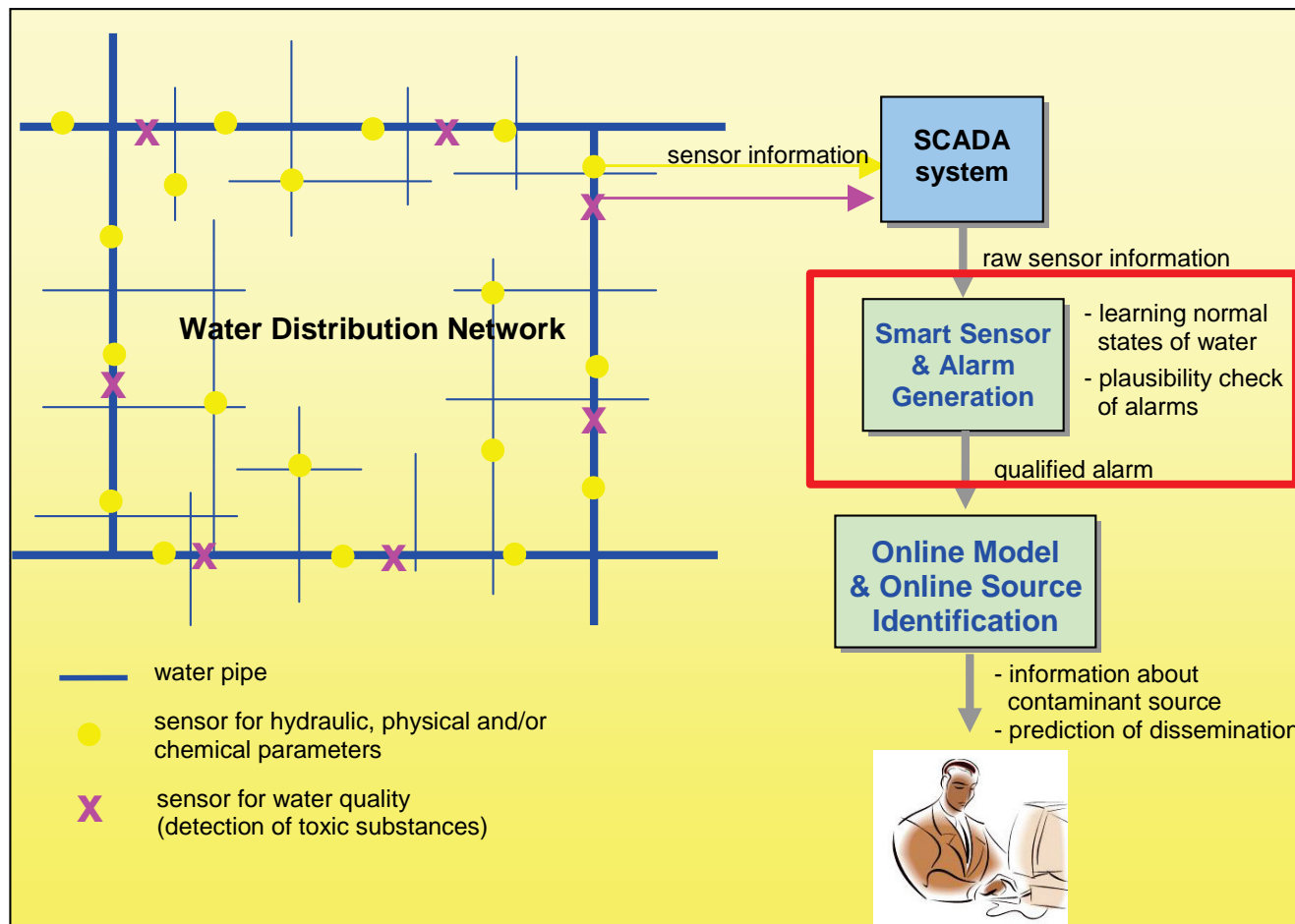
- Large amount of quality sensors placed in water distribution network (WDN)
- Sensors are used to monitor if a contamination has occurred in the network
- **Problem: Too many sensors for manual supervision by operators**  
Thresholds of quality parameters depend on e.g. water sources, operational modes
- **Solution: Data-driven event detection concept**



Part of Strasbourg's  
Water Distribution  
Network

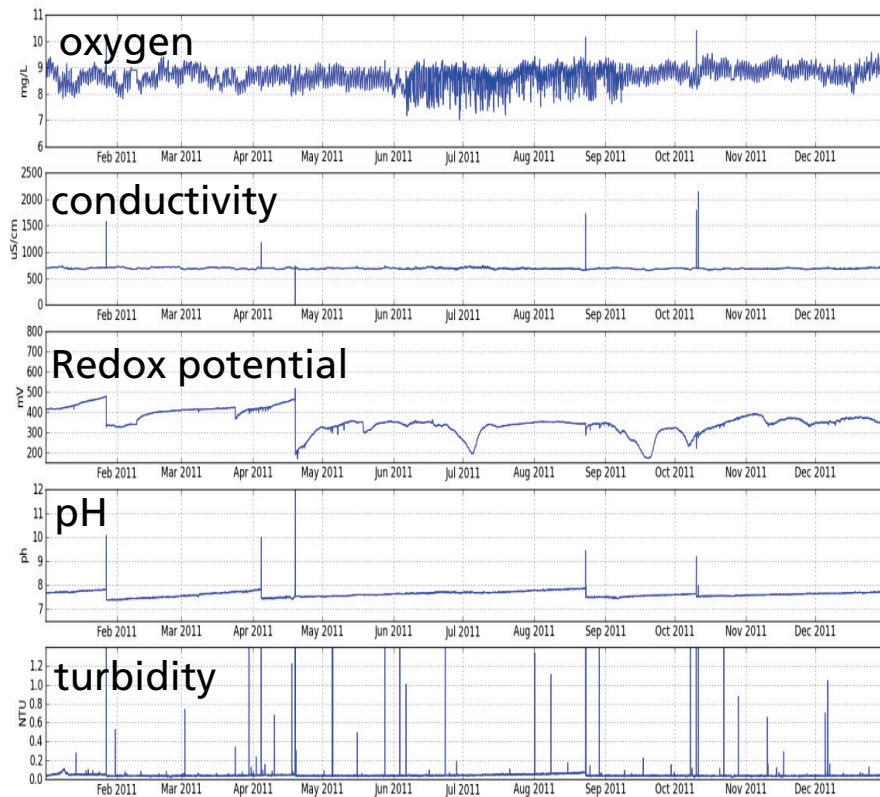
# Motivation

- Aim: Online detection of contamination events with low false alarm rate

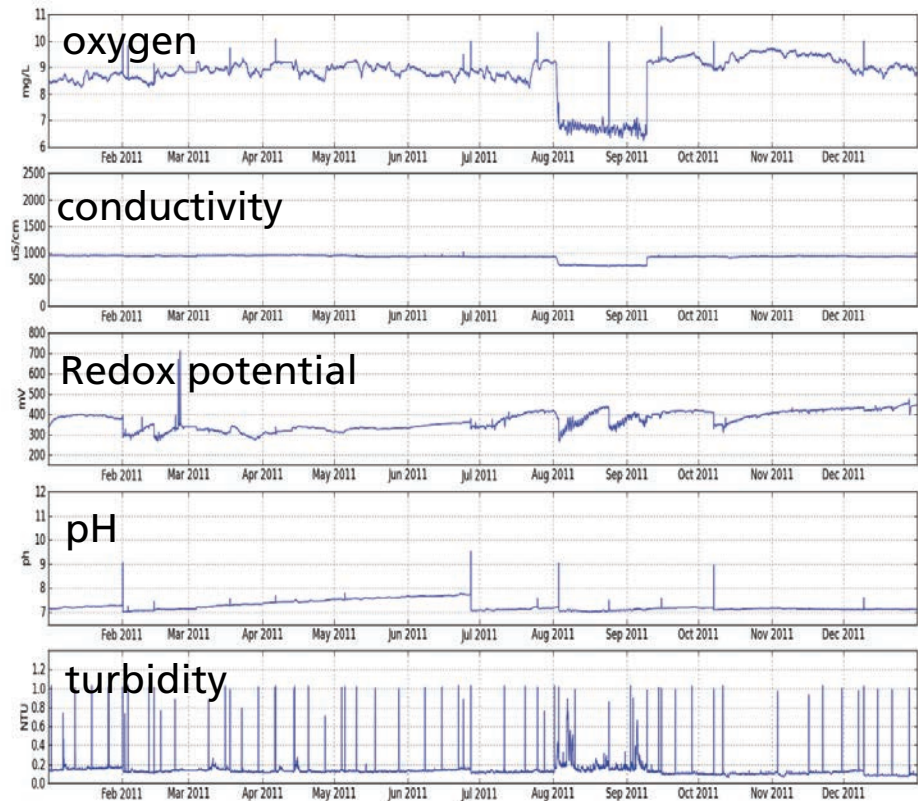


# Real Water Quality Data are complex....

## Berlin Station 1 (2011)



## Berlin Station 2 (2011)



---

# Expectations to Event Detection Module

- **Highly sensitive, but very low number of false alarms**
- Operator knowledge about water properties, sensor properties, operational modes can be included easily (**Operator Feedback**)

---

# Expectations to Event Detection Module

- Highly sensitive, but very low number of false alarms
- Operator knowledge about water properties, sensor properties, operational modes can be included easily (**Operator Feedback**)
- Consideration of several water quality parameters (**multivariate approach**)
- Only few parameters for fine tuning necessary (**self-learning approach**)

---

# Expectations to Event Detection Module

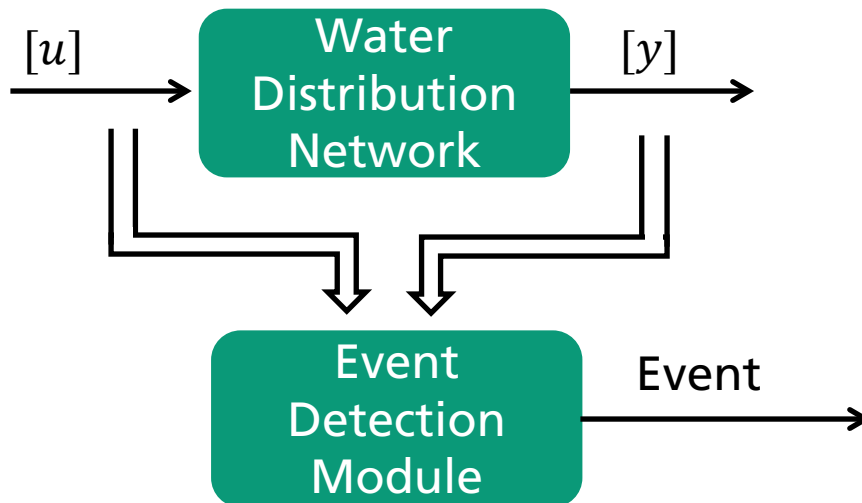
- Highly sensitive, but very low number of false alarms
- Operator knowledge about water properties, sensor properties, operational modes can be included easily (**Operator Feedback**)
- Consideration of several water quality parameters (**multivariate approach**)
- Only few parameters for fine tuning necessary (**self-learning approach**)
- Module can be easily applied to a large number of different measurement stations
- Detected events are interpretable (**no “black box” approaches**)

# Event Detection Module

## General approach of data-driven novelty detection

**Actuator Data**  
(pumps, valves ...)

**Sensor Data**



### Offline

- Training of multivariate process model with sensor data of "normal states"
- Store model as "normal" process condition

### Online

- An "Event" is a new, unknown state of the process
- Each time a new sample is acquired, the process model calculates if an event occurs



# Event Detection Module

## Step 1: Data Preprocessing (offline and online)

- Outlier filtering
- Smoothing of the measurements (e.g. by moving average filter)
- Normalization of the measurements

## Step 2: Training of “normal” Model (Offline)

- Calculating Principal Components (PCA)
- Generation of an alarm threshold

## Step 3: Calculation of an Alarm Index (Online)

- If alarm index is above threshold  
→ EVENT (has to be confirmed by operators!)

## Step 4: Re-Train Model after Operator Feedback (Offline)

- In case of false (positive/negative) alarm:  
Prepare new training data, go to Step 2

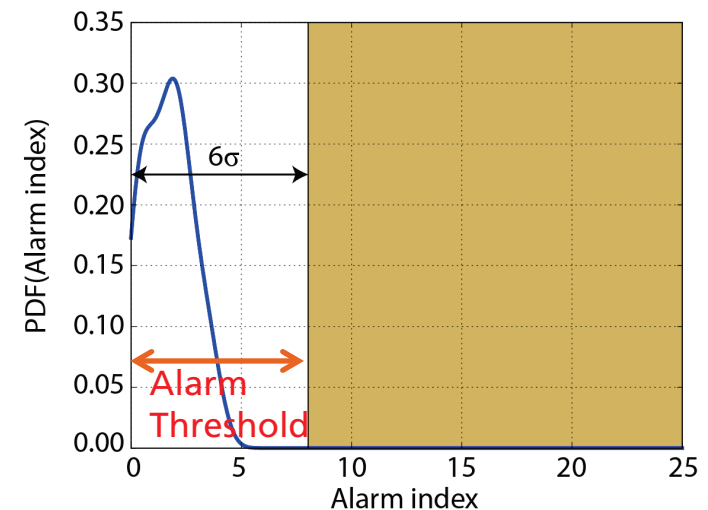
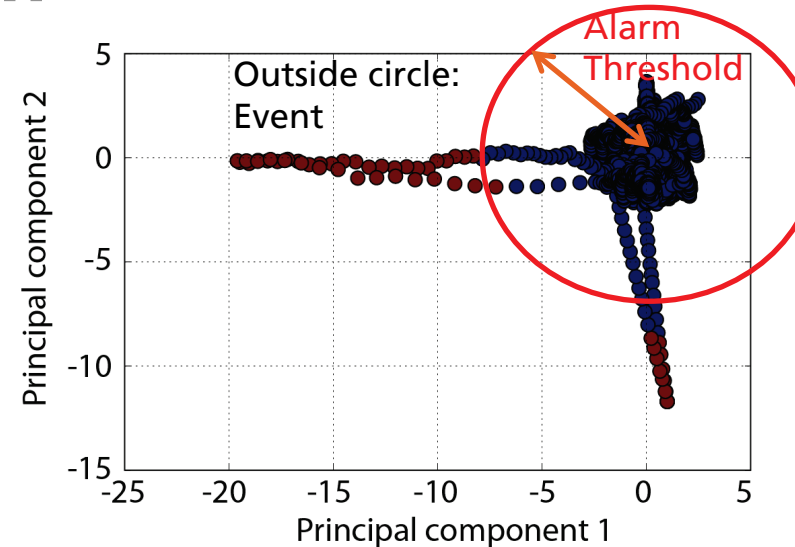
# Event Detection Algorithm

## Calculation of alarm index (offline / online)

- Based on Principal Components Analysis (PCA)
- First two principal components describe “state trajectory” of the measurement station
- Distance to center point describes value of alarm index

## Determination of alarm threshold (offline)

- Calculate alarm index using training data
- Calculate  $N$ - $\sigma$  environment of alarm index ( $\sigma$  : standard deviation,  $N$ : fit parameter, e.g.  $N = 6$ )
- Use value as threshold for event detection

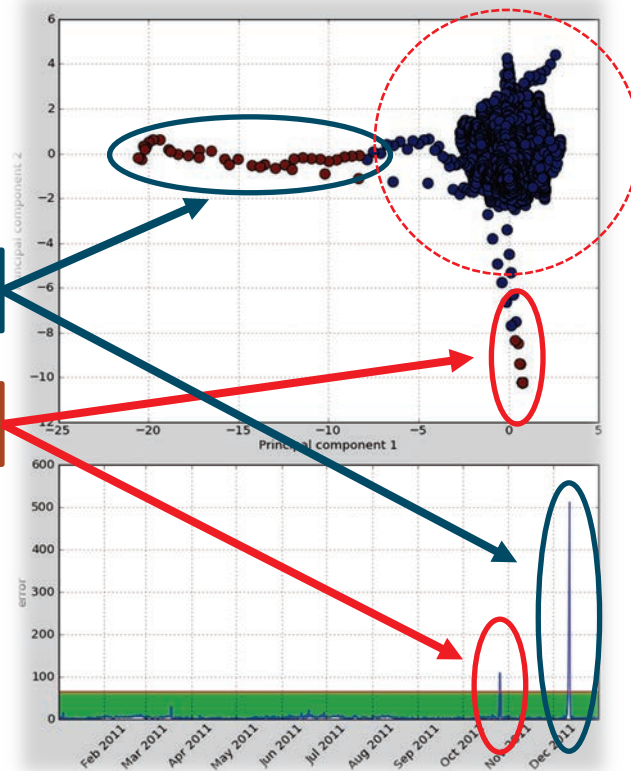


# Event Detection Algorithm

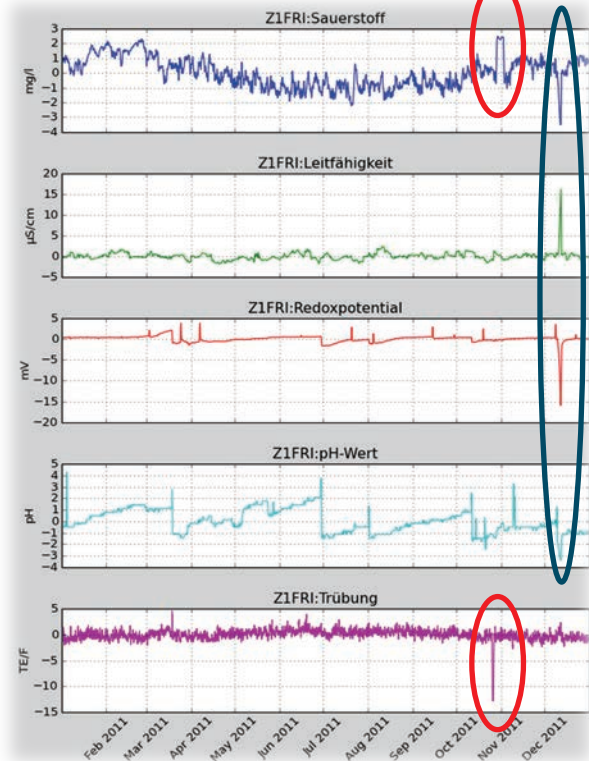
## Results from Berliner Wasserbetriebe (BWB)

Event 1

Event 2



Berlin-Friedrichshagen

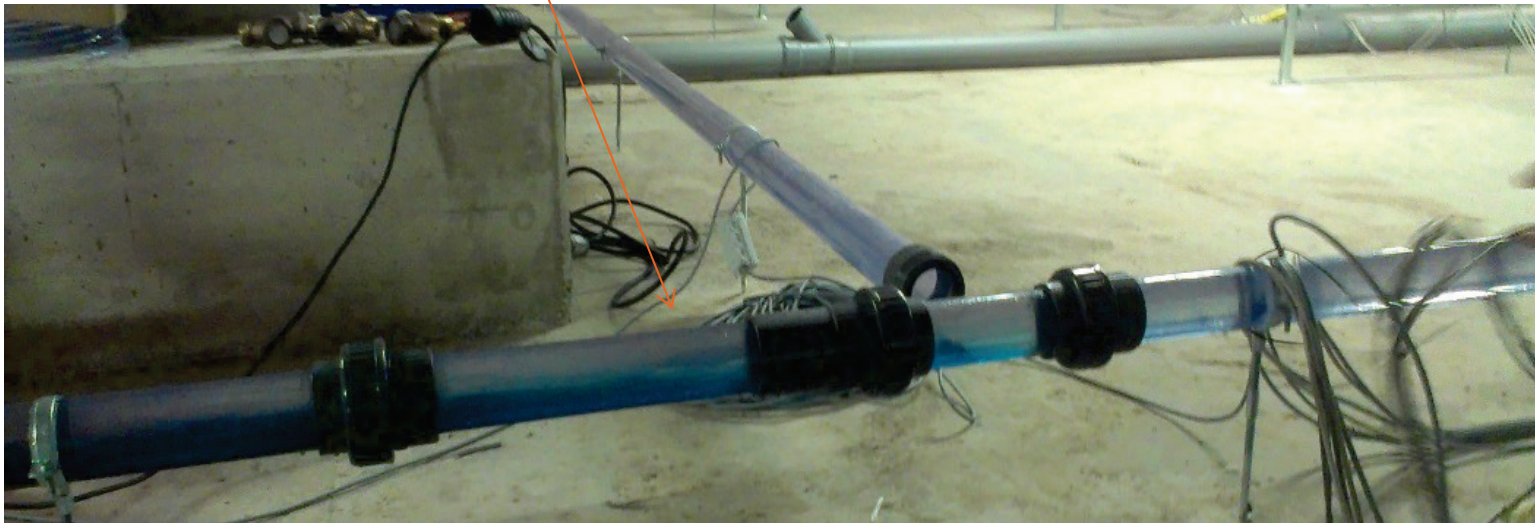


# Experimental Tests at Laboratory Plant

Test Network at TZW (Dresden)

**“Contaminant”: Indigo carmine with salt**

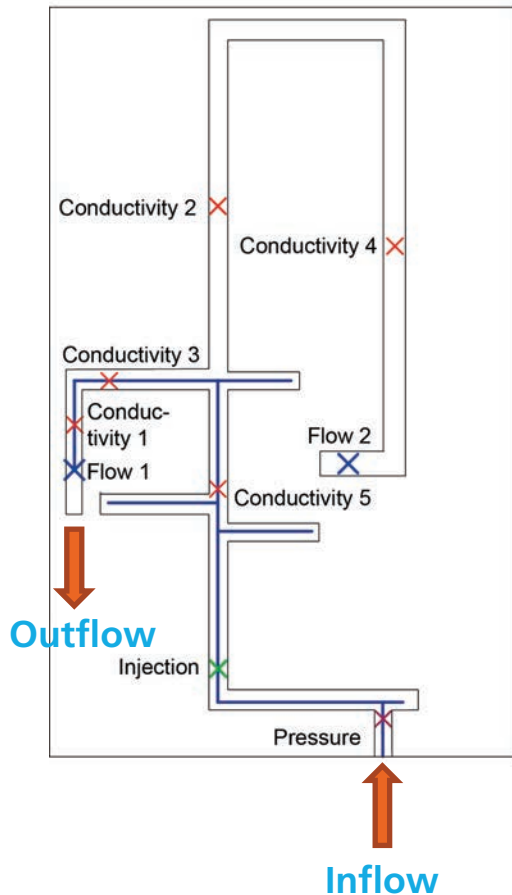
Injected “contamination”



Injection pump

# Setup of two Flow Configurations

## Flow Configuration 1

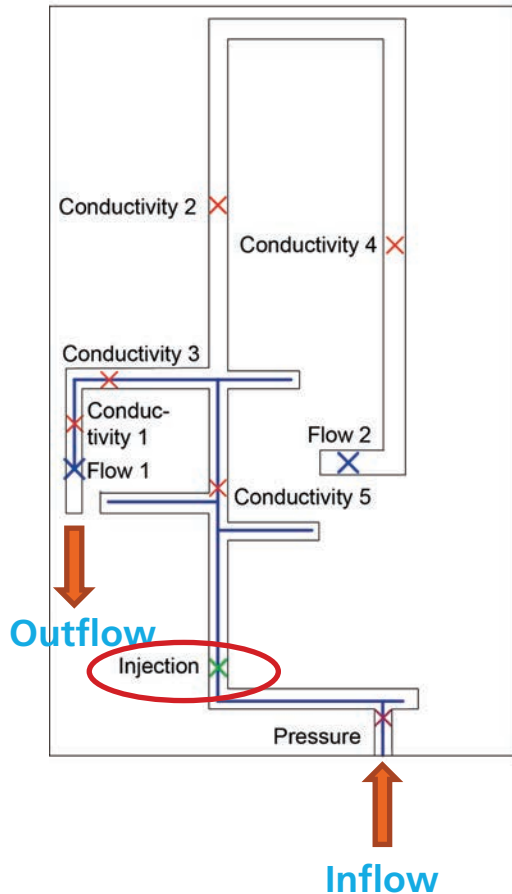


## Sensor Equipment

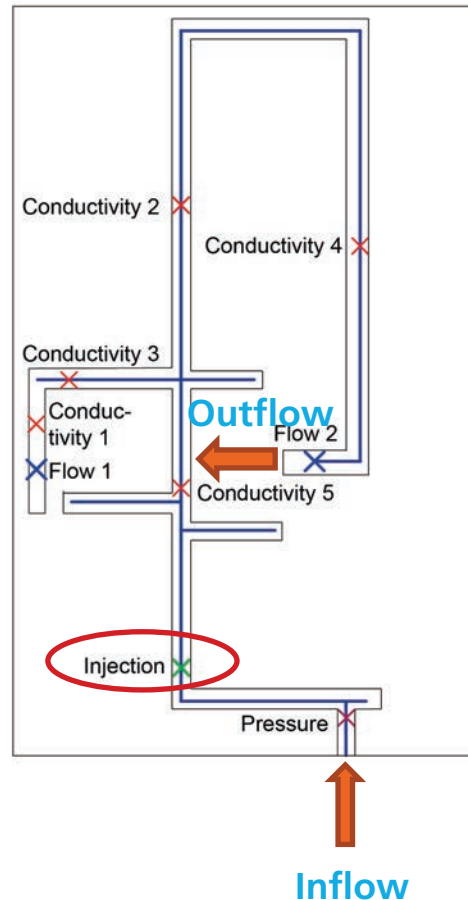
- 5 Conductivity sensors
- 1 Pressure sensor
- 2 Flow sensors

# Setup of two Flow Configurations

## Flow Configuration 1



## Flow Configuration 2



## Sensor Equipment

- 5 Conductivity sensors
- 1 Pressure sensor
- 2 Flow sensors

## Four Data sets (time series):

*Normal water configuration 1*

*Contaminated water config. 1*

*Normal water configuration 2*

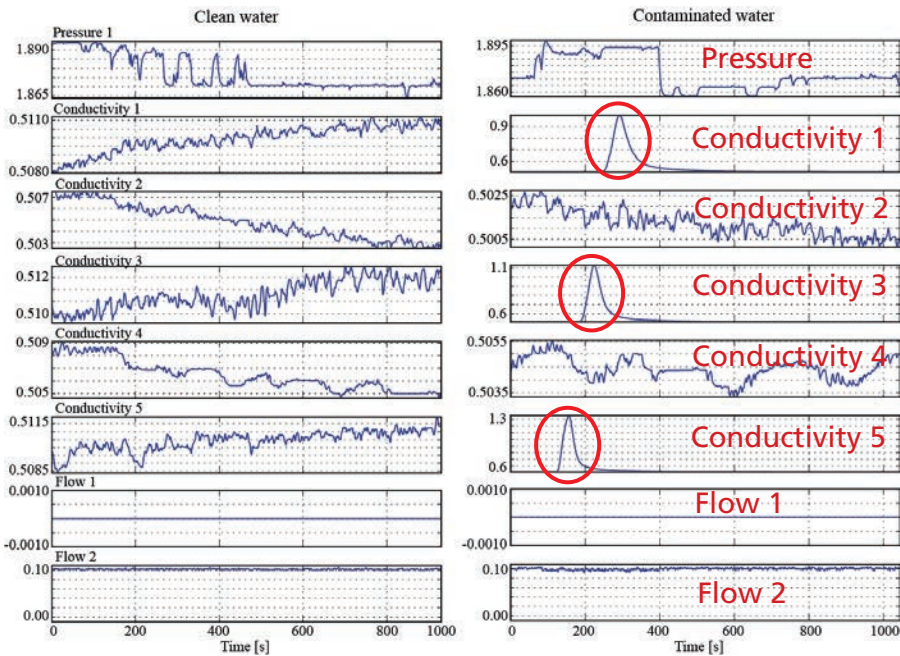
*Contaminated water config 2*



## Flow Configuration 1

*Normal  
Water 1*

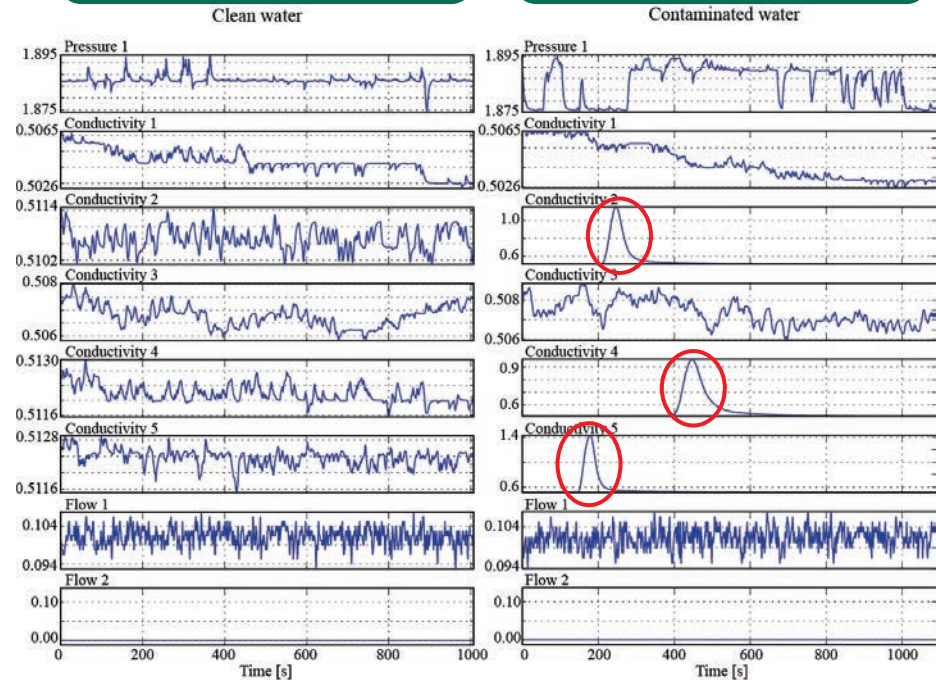
*Contaminated  
Water 1*



## Flow Configuration 2

*Normal  
Water 2*

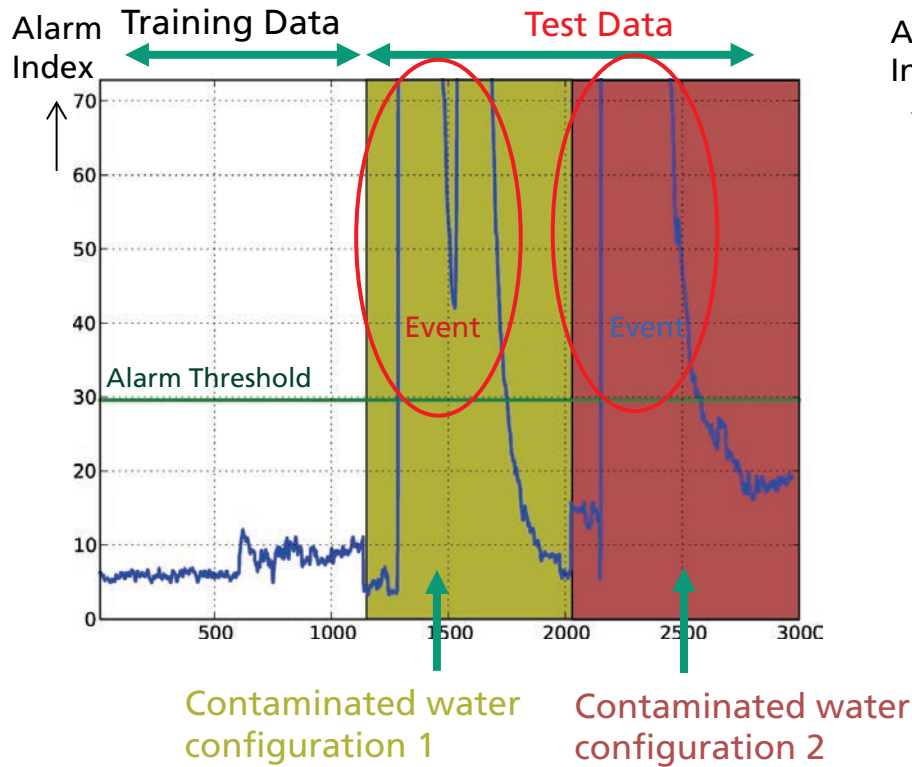
*Contaminated  
Water 2*



## Training/Test Scenario 1

- Training Data: *Normal 1, Normal 2*
- Test Data: *Contaminated 1, Contam. 2*

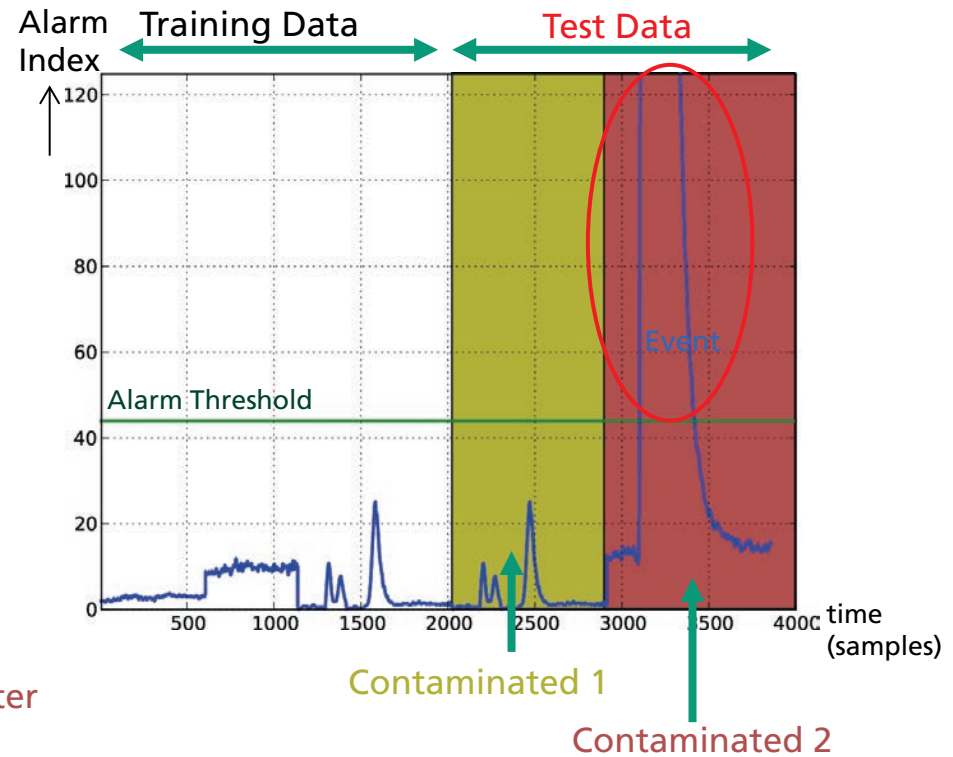
→ Event found at contamination 1 + 2



## Training/Test Scenario 2

- Training Data: *Normal 1, Normal 2, Contaminated 1*
- Test Data: *Contaminated 1, Contam. 2*

→ Event found only at contamination 2

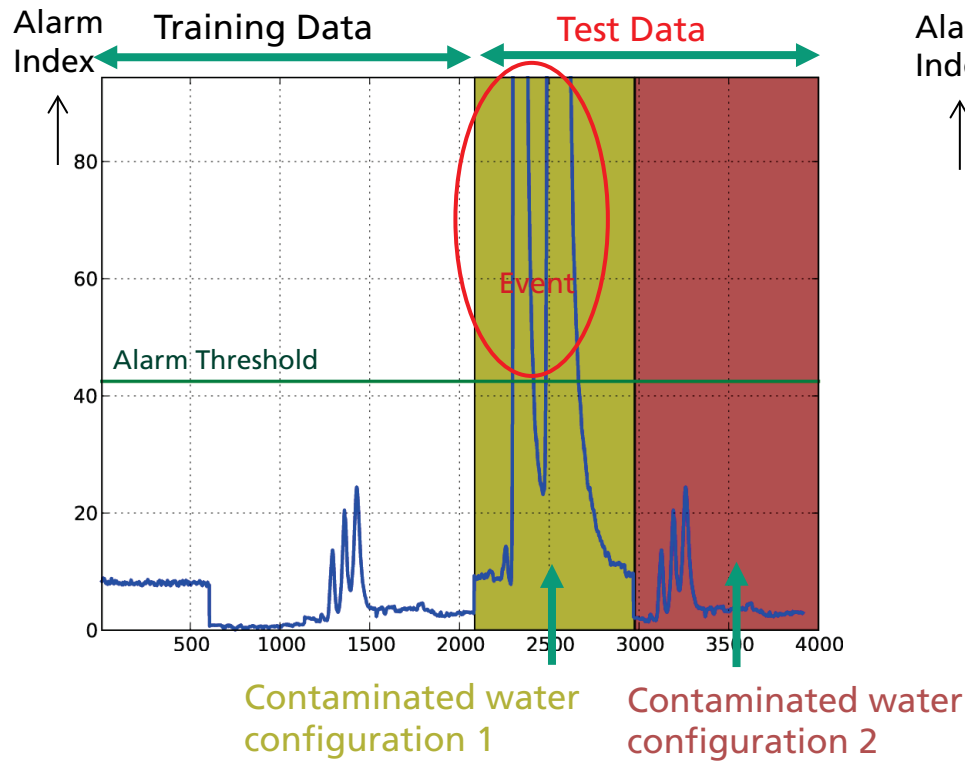




## Training/Test Scenario 3

- Training Data: *Normal 1, Normal 2, Contaminated 2*
- Test Data: *Contaminated 1, Contam. 2*

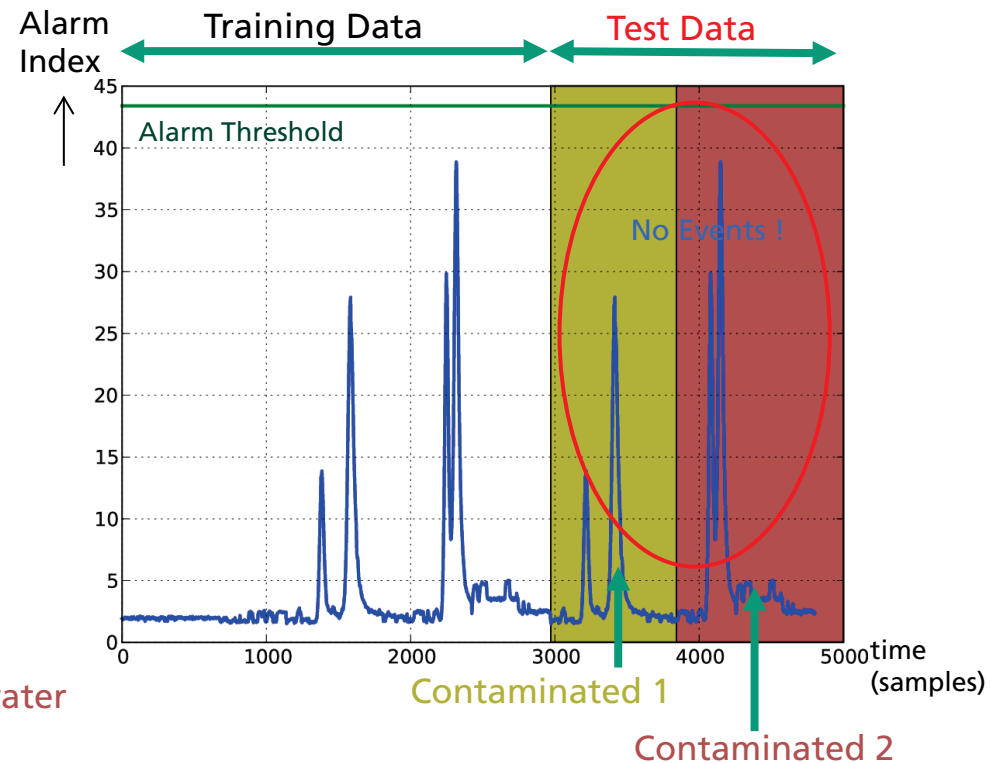
→ Event found only at contamination 1



## Training/Test Scenario 4

- Training Data: *Normal 1, Normal 2, Contaminated 1, Contaminated 2*
- Test Data: *Contaminated 1, Contam. 2*

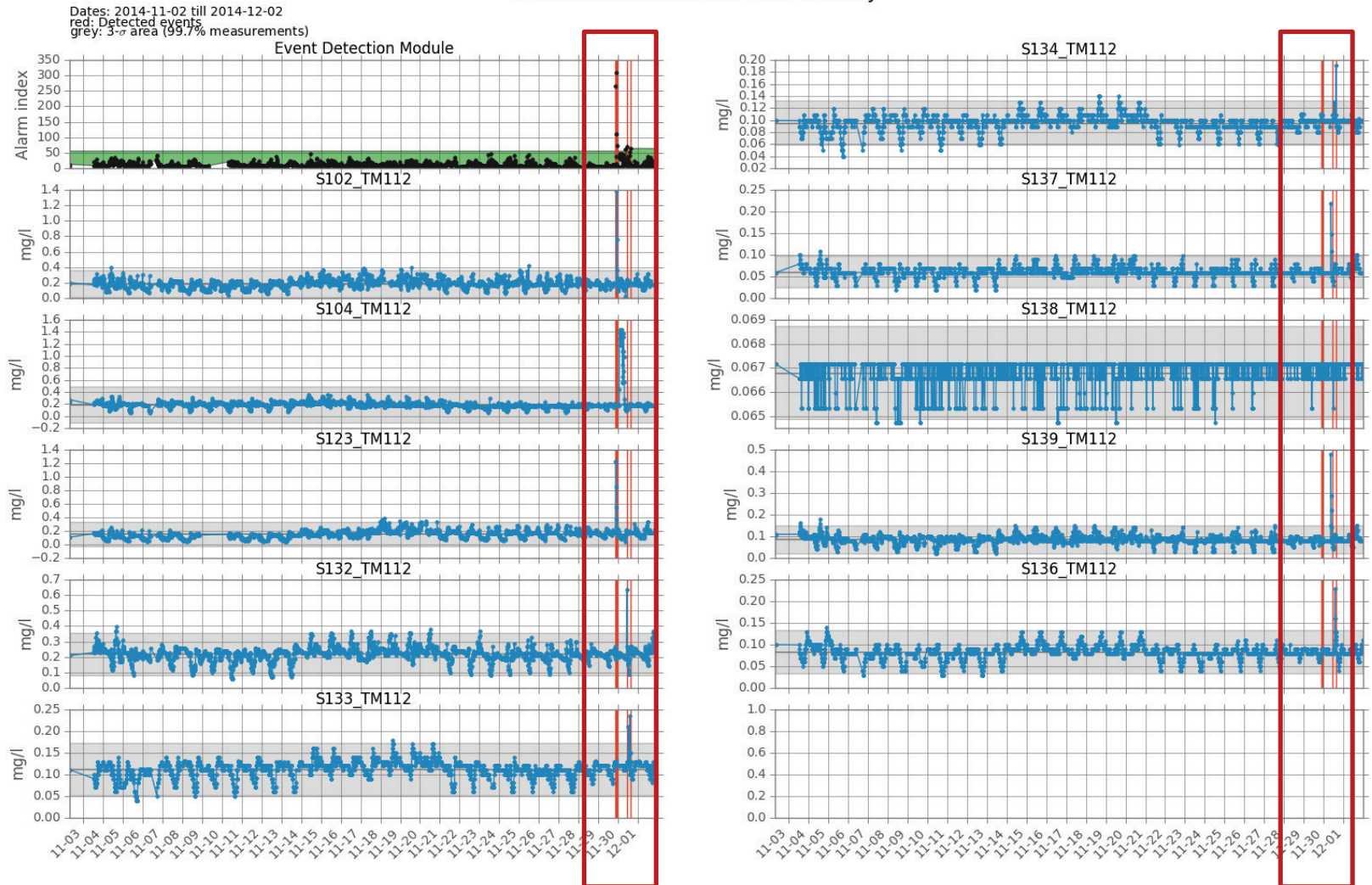
→ No Event found!



# Event Detection in Real World Data (1)

- Data from Strasbourg (CUS), 11/2014
- Propagation of chlorine peak through the network

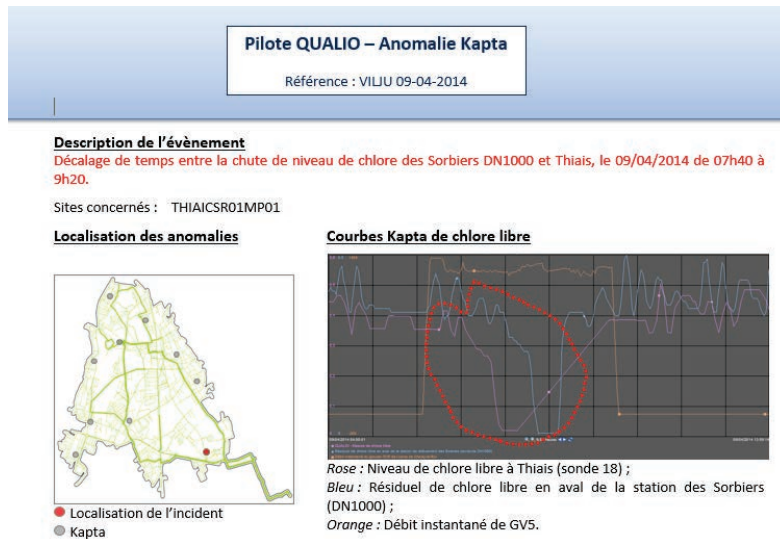
Event Detection Sector Vert Monthly



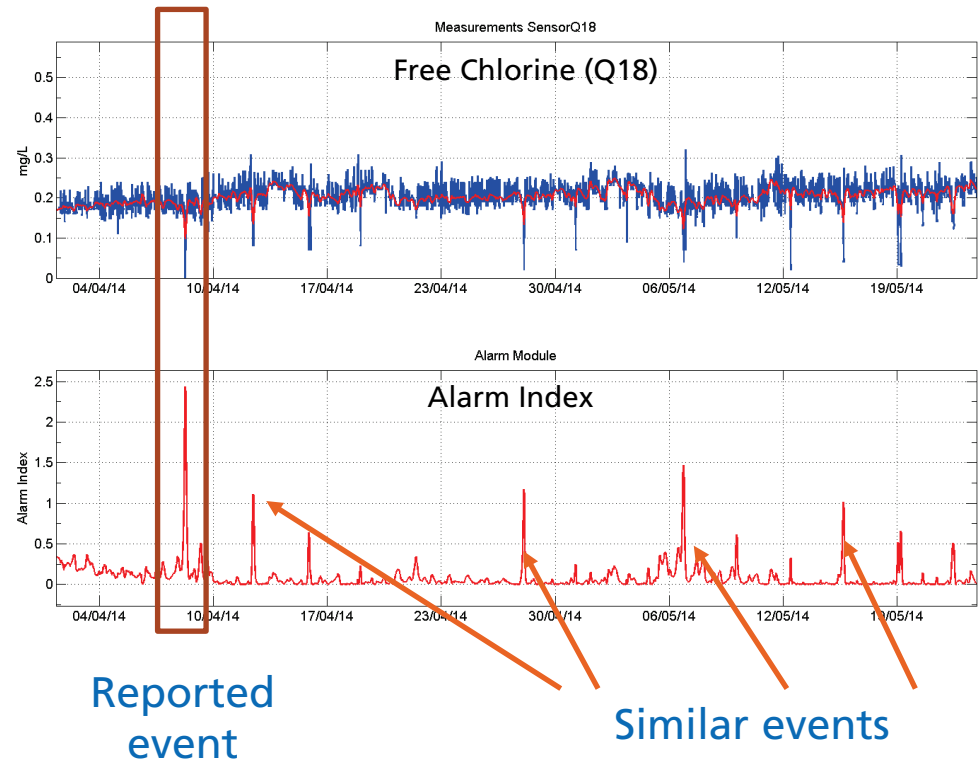
# Event Detection in Real World Data (2)

- Data from Paris (SEDIF), 9.4.2014
- Strong decay of free chlorine (due to change of water source)

## Reported event

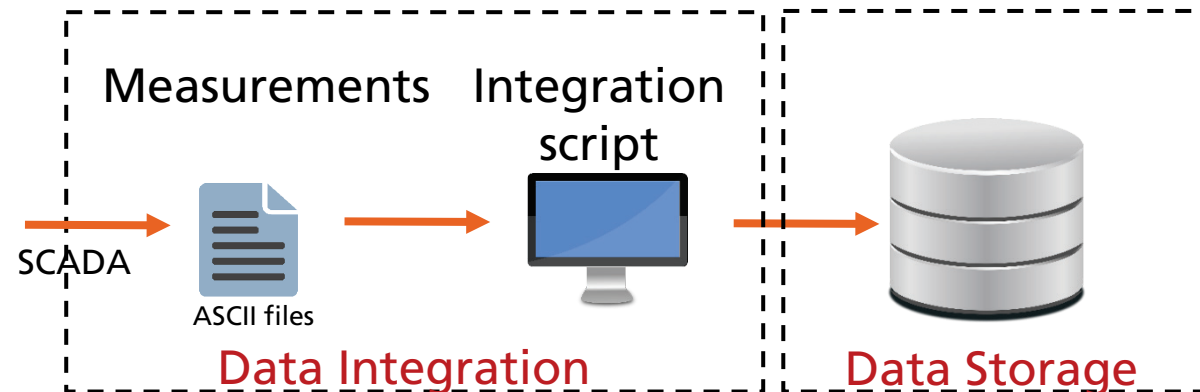


## Measurements / Resulting alarm index



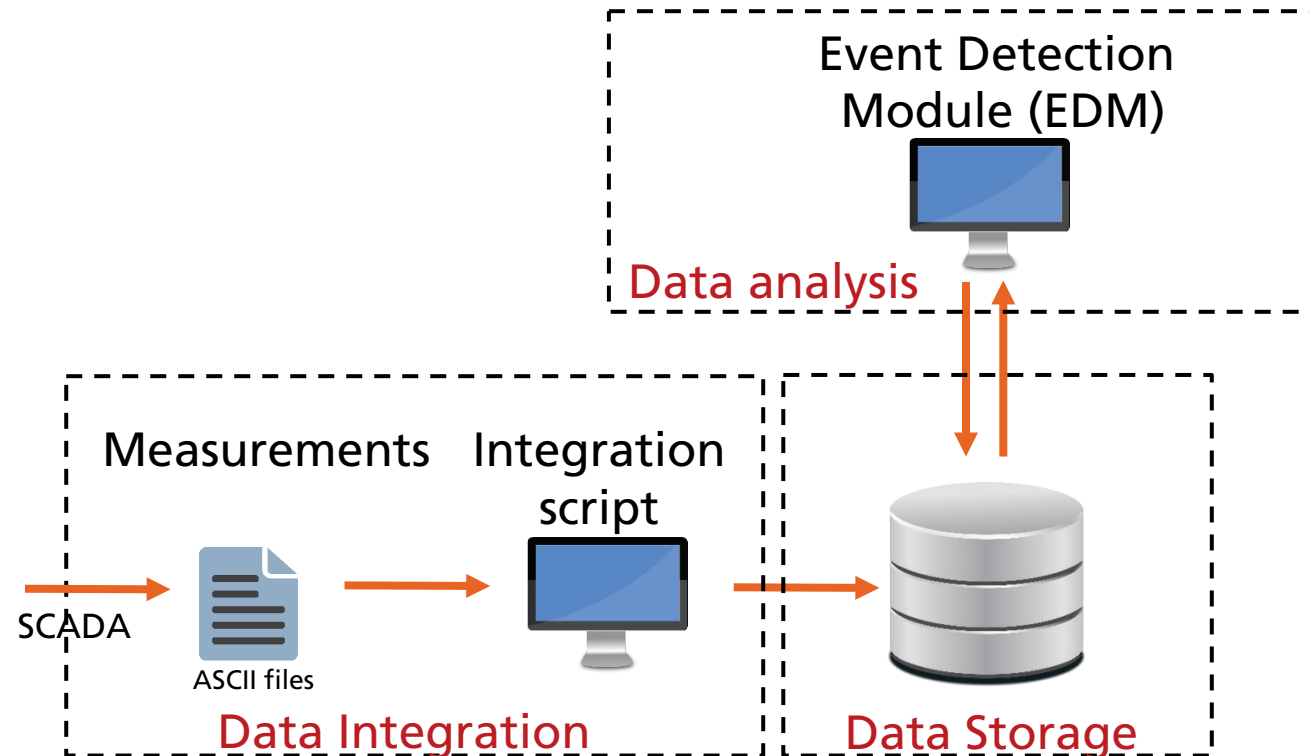
# Implementation as real-time Software (1)

- Automated data integration and plot generation (data+events)
- Results of Event Detection Module (events) are stored in database



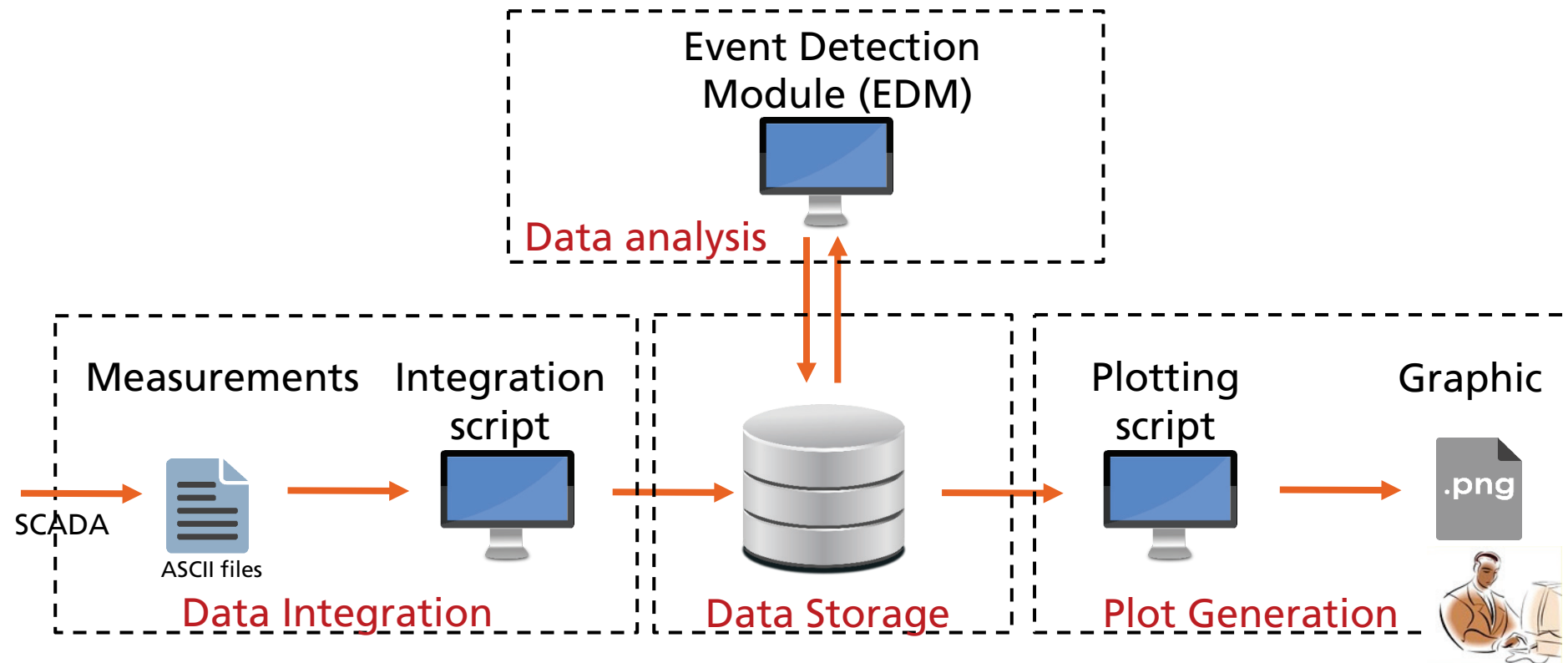
# Implementation as real-time Software (1)

- Automated data integration and plot generation (data+events)
- Results of Event Detection Module (events) are stored in database



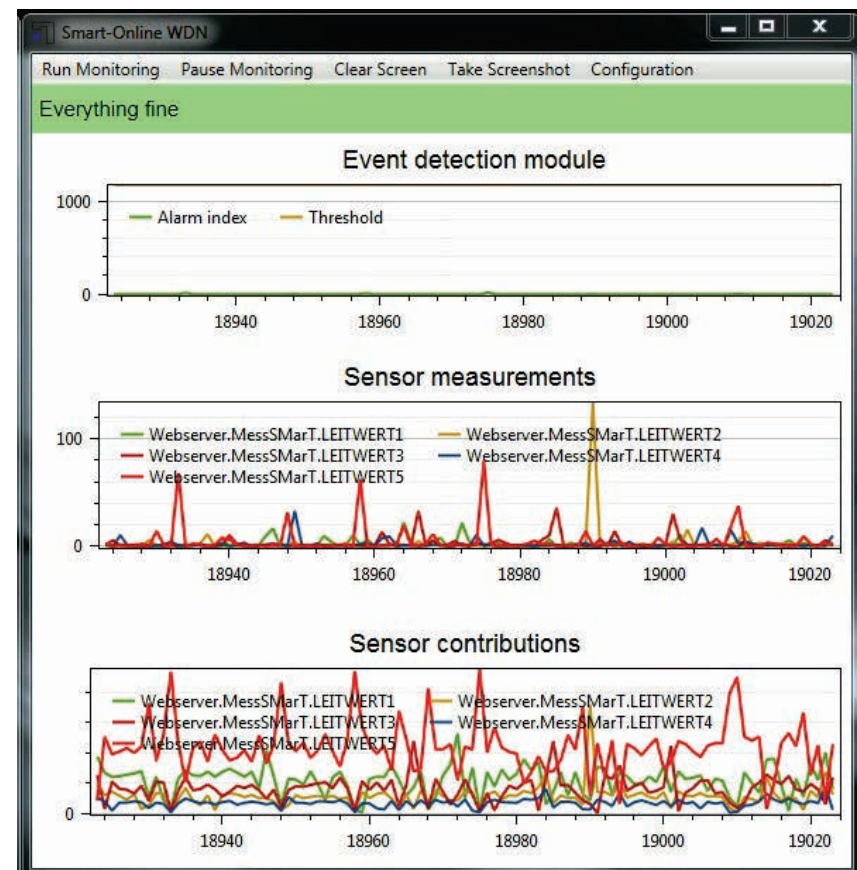
# Implementation as real-time Software (1)

- Automated data integration and plot generation (data+events)
- Results of Event Detection Module (events) are stored in database



# Implementation as real-time Software (2)

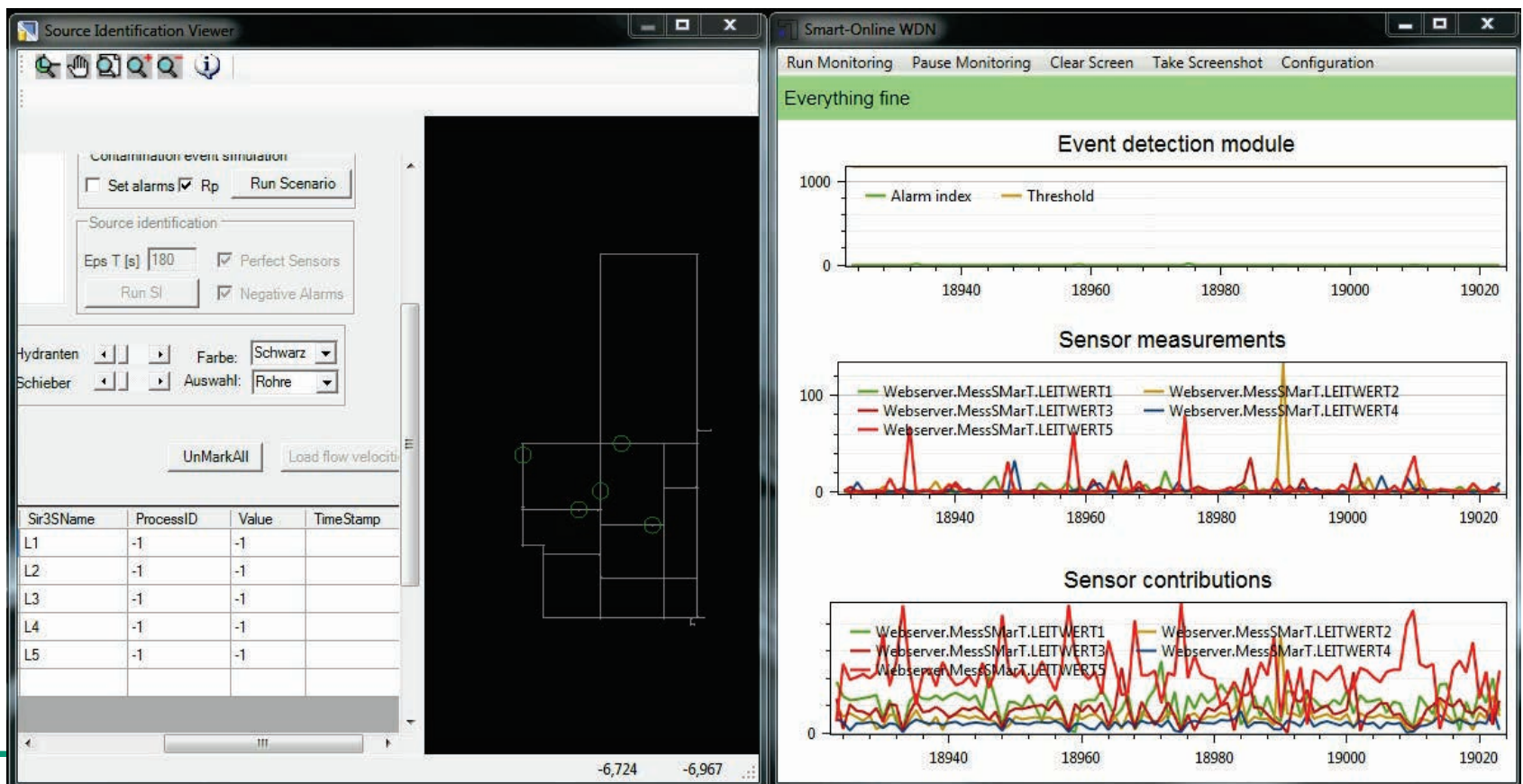
- Software communicates over OPC server
- Activates other modules can be started (contamination source identification, look ahead simulation, planning of mitigation measures ...)





# Implementation as real-time Software (2)

- Software communicates over OPC server
- Activates other modules can be started (contamination source identification, look ahead simulation, planning of mitigation measures ...)





# Summary

- Data-driven real-time event detection concept for water distribution networks based on Principal Components Analysis (PCA)
- Only one fitting parameter ( $N\text{-}\sigma$  environment for alarm threshold)
- Implemented as real-time capable software
- Tested by means of historical data of Berlin, Paris, Strasbourg and online at a laboratory plant

## Future research

- Incremental learning (consideration e.g. of long term drifts)
- Plausibility check of events by hydraulic and quality model
- Use information about water source and operational information



SPONSORED BY THE



Federal Ministry  
of Education  
and Research

The project "SMaRT-Online WDN" has been supported by the German Federal Ministry of Education and Research (BMBF; reference project: 13N12180) and by the French Agence Nationale de la Recherche (ANR; reference project: ANR-11-SECU-006)

---

# Thanks for your attention!

Project website: [www.smart-onlinewdn.eu](http://www.smart-onlinewdn.eu)

Thomas.Bernard@iosb.fraunhofer.de