Towards a common benchmark for long-term process control and monitoring performance evaluation

Christian Rosen*, Ulf Jeppsson* and Peter Vanrolleghem*,**, *

* IEA, Dept. of Industrial Electrical Engineering and Automation, Lund University, Lund, Sweden

** BIOMATH, Dept. of Applied Mathematics, Biometrics and Process Control, Ghent University, Gent, Belgium
Outline

• Introduction
• Why benchmarking process monitoring
• Existing benchmark models – the IWA/COST benchmark
• Long-term control and monitoring benchmark – description of ideas
• Example
• Conclusions
Introduction

• Vision:
  – a realistic **dual** purpose simulation benchmark for **long-term control** and **process monitoring** evaluation - *the long-term benchmark simulation model no1* (BSM1_LT)

• Why do we present a draft and not a final definition?
  – we need input from users and researchers to ensure realism
Why a benchmark for monitoring?

- Process monitoring is relatively new
- Many different techniques are available
- Must be adapted to WWT
- Difficult to evaluate the efficiency from literature
- Authors often use plant specific data not available to everyone

- Tennessee Eastman process
What can be tested?

• Different monitoring methods’ applicability to WWT
  – What can be detected?
  – What can be isolated?

• Efficiency
  – Correct classifications
  – Misclassifications

• Sensitivity
  – Does the method handle process changes?

• Detection speed
  – How long does one have to wait?
Benchmark simulation model no 1 (BSM1)

- Developed by IWA/COST task groups
- Model, control system, benchmark procedure and evaluation criteria
- Five-reactor biological plant with secondary clarification
- Simulated 4 weeks with influent data according to file
- Performance of the last week evaluated
Ongoing extensions to the BSM1

- Benchmark simulation model no 2 (BSM2)
- Currently in preparation by the former COST benchmark group
- Plant-wide benchmark
- BSM1 + pre-treatment, sludge train, additional loads
- Long evaluation time (one year)
Shortcomings with BSM1 (and BSM2)

BSM1
• Too short evaluation period to be realistic
• No seasonal effects, failures and faults
• Focused on control

BSM2
• Not completed
• High complexity level
The BSM1_LT

The original BSM1 configuration with:

- 1 year evaluation period (+ 6 months initialisation)
- Influent characteristics model (BSM2)
- Time (temperature) varying parameters
- Sensor and actuator failures
- Process faults
- Additional actuators with long-term effects
Influent disturbances

- Coordinated with the work on BSM2
- Generate 1.5 year of treatment plant influent data
- Diurnal flow rate and concentration variations
- Weekend effects
- Seasonal flow rate and concentration variations
- Holiday effects
- Rain and storm events, and their effects on influent flow rate and concentration
- Varying temperature (seasonal effect)
Sensor/actuator faults

• Complete failure (break down)
  – No signal or error signal (sensors)
  – Zero capacity (actuators)

• Partial failure
  – Faulty signals (sensors)
  – Drift beyond expected (sensors)
  – Reduction in capacity (actuators)
  – Actuator drift “wear and tear” (actuators)
  – High variance (sensors and actuators)

• Calibration instances (sensors)
Process disturbances

*Defined as disturbances acting through model parameters*

- Inhibition of nitrification
- Toxicity strikes
- Varying settling properties
- Temperature
BSM1_LT benchmarking procedure

• Initialisation 6 months; the data can be used for data driven monitoring models
• One year simulation (summer to summer) with faults imposed
• Off-line or on-line evaluation - calculation of various performance indices
Evaluation criteria

• Detection performance
  – rate of correct detections
  – rate of missed detections
  – time to detection

• Isolation performance?
  – difficult to be objective
  – dependent of the method
  – often based on experience

• For control: as BSM1
Example - adaptive monitoring

![Graph showing influent flow rate and equipment failures over time.](image)

- **Influent flow rate [m³/d]:** The graph depicts the variation in influent flow rate over 120 days, with peaks and troughs indicating fluctuations in flow.
- **Equipment failures:** There are several instances marked where equipment failures occurred.
- **High flow rate:** The graph also indicates periods of high flow rate, with spikes above the typical range.

These data points illustrate the challenges and performance evaluation aspects of long-term process control and monitoring.
Example - continued

<table>
<thead>
<tr>
<th>Treatment method</th>
<th>Correct detections/true faults</th>
<th>Missed detections/true faults</th>
<th>False alarms/total detections</th>
<th>Average detection time (h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0: no pre-treatment</td>
<td>5/7</td>
<td>2/7</td>
<td>13/18</td>
<td>1.6</td>
</tr>
<tr>
<td>1: calibration instances removed</td>
<td>7/7</td>
<td>0/7</td>
<td>13/20</td>
<td>1.5</td>
</tr>
<tr>
<td>2: pre-treatment 1 + calculated fluxes</td>
<td>7/7 (6/7)*</td>
<td>0/7 (1/7)*</td>
<td>9/16 (10/16)*</td>
<td>8.7 (1.5)*</td>
</tr>
<tr>
<td>3: pre-treatment 2 + deviations from dry weather diurnal pattern</td>
<td>7/7</td>
<td>0/7</td>
<td>9/16</td>
<td>1.1</td>
</tr>
</tbody>
</table>

* One detection was very late (during the third fault event). The values within brackets are the result if this instance is classified as a missed detection.
Conclusions

• The BSM1_LT should provide more realistic benchmarking
• One year evaluation period together with seasonal and holiday effects
• Sensor/actuator as well as process faults
• Dual purpose: control and monitoring
• We need input to ensure realism!
Failure models

• Time for failure
  – Poisson distributed with specified rate depending on the sensor/actuator?
  – The number of failures during a certain time interval $t$ with $\lambda$ failure rate.
    \[ X \in Po(\lambda t) \]

• Time to repair
  – Lognormal distributed with a minimum repair time?
    \[ t_r - t_{\text{min}} \in \lognorm(\mu,\sigma) \]
Realisation of failures

Time for failures for different failure rates

\[ \lambda t = \frac{1}{7} \]

\[ \lambda t = \frac{1}{28} \]

\[ \lambda t = \frac{1}{182} \]

Time for repair

\[ T_{\text{min}} = 12 \text{ hours} \]

\[ \mu = \log(12) \]
Long-term control benchmark - motivation

• BSM1 - too short evaluation period to be realistic
• Slow actuators
  – e.g. WAS
• Actuators with long-term effects
  – e.g. equalisation tanks, sludge storage tanks
• Evaluation of robustness against sensor, actuator and process failures
• Slow processes, e.g. phosphorous (future)
Additional control handles

- Long evaluation period allows for additional control handles
- Wastage flow rate
- Equalisation tanks
- Sludge storage tanks