Qualitative trend analysis for fault detection in batch processes

Watermatex2015 – Gold Coast – 17.06.2015

Kris Villez
Jonathan Habermacher

Eawag: Swiss Federal Institute of Aquatic Science and Technology
Promise of qualitative methods
Setup

Activated sludge reactor (type SBR)

Influent wastewater

Sludge recirculation

Effluent

Excess sludge removal

Side-stream reactor (SSR)
Operator’s opinion

Visual inspection of 410 batch cycles

Initial independent analysis
- Normal: 350
- Abnormal: 41
- Different opinion: 19

After consensus search
- Normal: 348
- Abnormal: 46
- Different opinion: 16
Shape Constrained Spline (SCS) functions

Example

![Shape constraints - After optimization](chart)

SSR = 424.67

- $y$
- $\hat{y}$ (off-set)
Shape Constrained Spline (SCS) functions

For anomaly detection
Principal Component Analysis
Principal Component Analysis

Identification

Relative Variance [%]

Cumulative Relative Variance [%]

#PC
Principal Component Analysis

For anomaly detection
Performance evaluation

Anomalies identified positively by one method only
Performance evaluation
Receiver Operating Characteristic (ROC)
Qualitative trend analysis

• Summary of results
  • Performance
    • Better data approximation
    • Better fault detection
    • Intuitive (operator / manager trust)
  • Requirements
    • Limited prior knowledge
    • No calibration
Qualitative trend analysis

• What is new?
  • Application to experimental data
  • Benchmarking against operators and PCA
  • Method for discontinuous trends
• Use for
  • Anomaly detection (this presentation)
  • Automation (IT&Water)
This presentation

• What is next?
  • Complexity
    • Multivariate time series
      ↔ Univariate
    • Computer-based qualitative simulation
      ↔ Operators’ qualitative description
  • Pattern discovery (unsupervised)
    ↔ Pattern recognition (supervised)
• What is in it for you?
  If you know a shape, …
  And you want a computer to recognize it, …
Check out the toolbox (Matlab/Octave):
  http://homepages.eawag.ch/~villezkr/spike/
• Now - SCS toolbox, v1.0
• Soon - SCS toolbox, v2.0 + data

Email: kris.villez@eawag.ch
This presentation

Thanks

• Ing. Jonathan Habermacher
• Dr. Nicolas Derlon
Need for anomaly detection

- Increasing intensity of data collection
  - Monitoring / Diagnosis
  - Optimization
  - Automation
- Potential achieved with
  - Good quality data
  - Predictable processes
- Safe-guards are needed
  - Anomalous data
  - Anomalous process conditions
Principal Component Analysis

Geometry 2D

1. Calibration
2. Projection
3. Detection
Principal Component Analysis
Geometry 3D

1. Calibration
2. Projection
3. Detection
Shape Constrained Spline (SCS) functions

Concept

• Fit a flexible function with many parameters
  • Spline function with spline coefficients

• Constrain the parameters so that a desired shape results
  • Linear decrease – convex decrease – concave increase

• Find the locations (in time) where the change of shape is most likely
  • Changepoint / Transition
Shape Constrained Spline (SCS) functions

Execution

• Fit a flexible function with many parameters
  • quadratic problem (analytic solution)

• Constrain the parameters so that a desired shape results
  • second-order cone program (interior-point)

• Find the locations (in time) where the change of shape is most likely
  • Nonlinear program (branch-and-bound)
Shape Constrained Spline (SCS) functions

For anomaly detection

- Use SSR as test statistic
- Apply a suitable upper control limit (UCL) for detection

- Reference signal
  - Anomaly present / absent
  - Consensus between two process experts
    - 394 batches (16 without consensus)
Principal Component Analysis

For anomaly detection

- Use SSR = SPE = Q ~ DmodX as test statistic
- Apply a suitable upper control limit (UCL) for detection
  - Approximating distribution available under a number of assumptions
Anomaly detection

Basic flow

- Prior knowledge
- Historical data
- Model
- Information
- Program
- Online data
Anomaly detection

Traditional quantitative methods

**White box**
- Prior knowledge
- Model

**Methods**
- Balance equations
- Mechanistic models

**Challenges**
- Extrapolation
- Knowledge limited
- Time-consuming

**Black box**
- Historical data

**Methods**
- Artificial Intelligence
- Machine Learning
- Statistics

**Challenges**
- Data graveyards
- Data silos
- Interpretation
Anomaly detection
New(er) qualitative methods

White box
Prior knowledge

Model

Black box

Methods
Graph theory
Qualitative process theory

Benefits
Limited extrapolation
Only basic knowledge

Methods
Qualitative trend analysis

Benefits
Low need for data
Intuitive

Historical data
This contribution

- Development of new method (SCS)

- Application to experimental data
  ↦ benchmark simulation

- Use for anomaly/fault detection
  ↦ diagnosis/ fault identification

- Comparison with principal component analysis
  ↦ other qualitative trend analysis method