# Real-Time Outlier Detection with Dynamic Process Limits

Process Control 2023

Marek Wadinger[1], Michal Kvasnica[1]

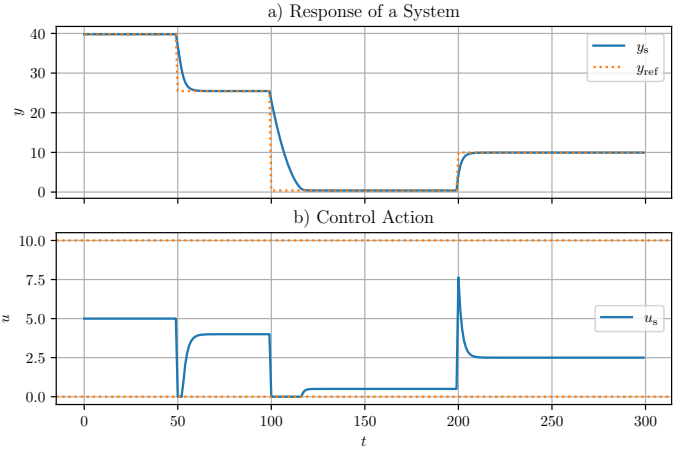
[1]Institute of Information Engineering, Automation, and Mathematics
*marek.wadinger@stuba.sk*

2023-05-24

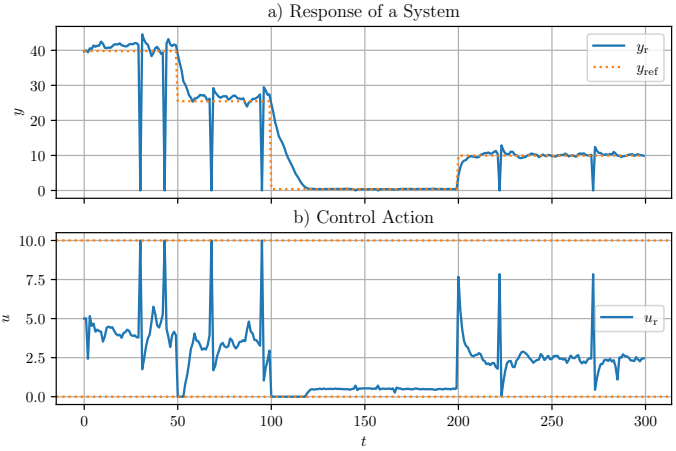


SLOVAK UNIVERSITY OF
TECHNOLOGY IN BRATISLAVA
FACULTY OF CHEMICAL
AND FOOD TECHNOLOGY

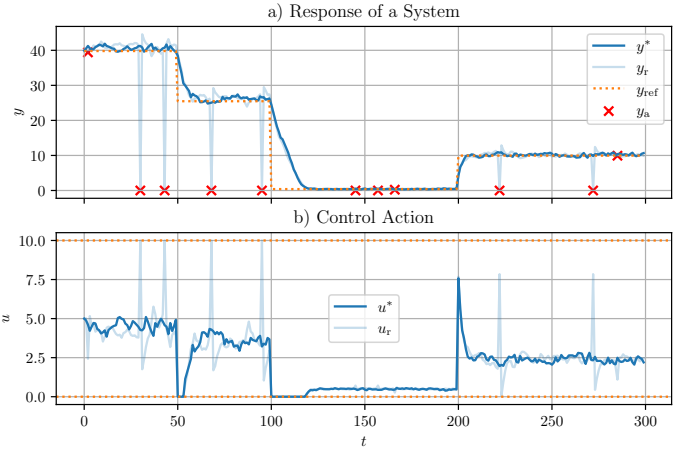
# Simulation Results



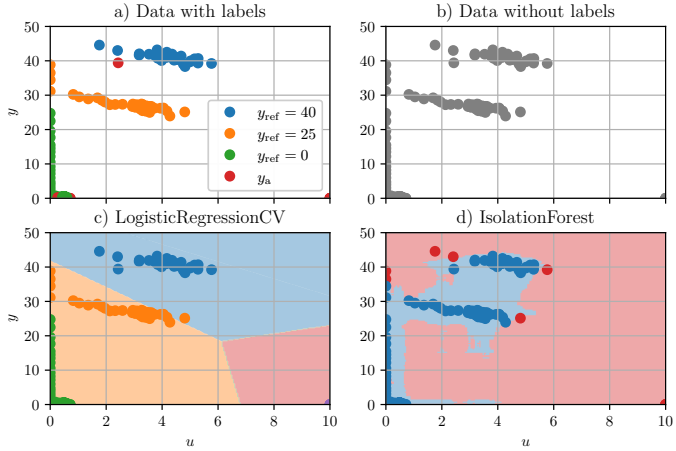
# Practical Scenario



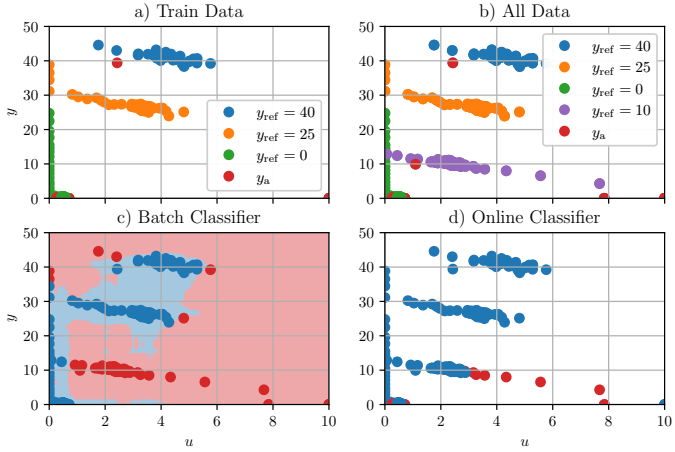
# Control Engineering meets Artificial Intelligence



# Lack of Labels



# Changes in Distribution



## Goals

We need to make detector that:

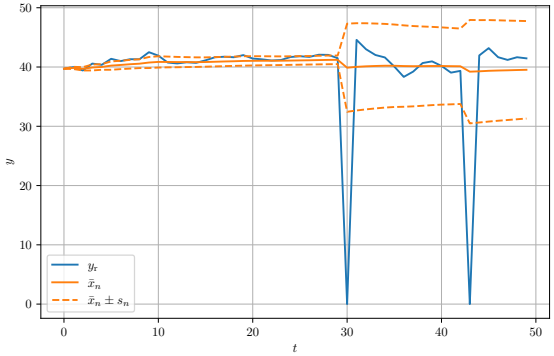
- Does not require huge amount of data
- Adapts to unseen operation
- Offers credible decision boudary
- Does not alter operation of existing systems

## Proposed Solution

Real-Time Outlier Detection with Dynamic Process Limits combining:

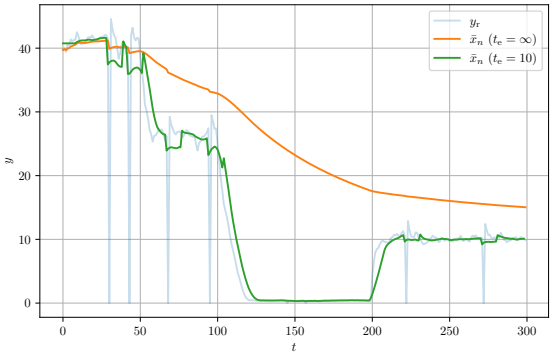
- Online Learning
- Outlier Detection
- Self-supervised Learning
- Soft Real-Time System
- Invertible Probabilistic Model

# Welford Algorithm



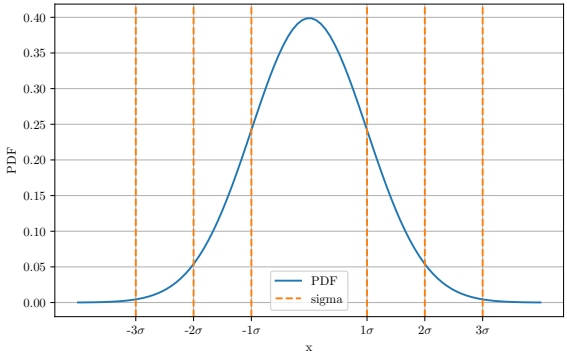

\+ One-Pass Algorithm | − Adaptation Slows Down

# Inverse Welford Algorithm



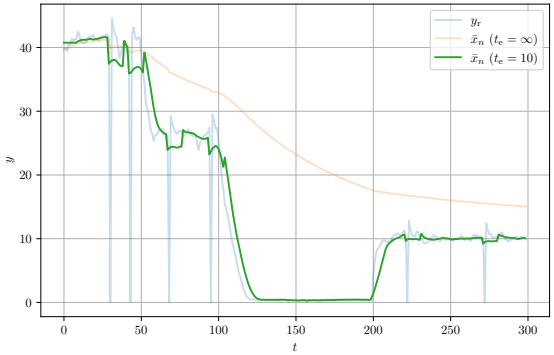

\+ Constant Adaptation | − Memorizes Data Window

# Distance-based Outlier Detection



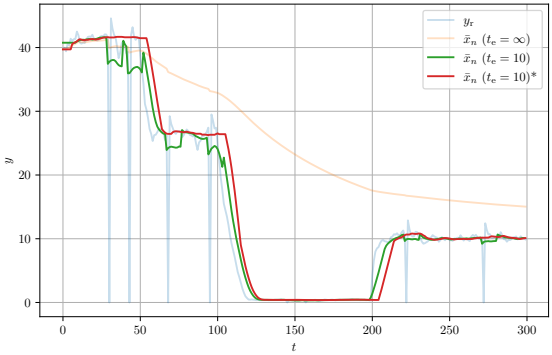

$$y_i = \begin{cases} 0 & \text{if } q \leq F_X(x_i; \bar{x}_n, s_n) \quad (1a) \\ 1 & \text{if } q > F_X(x_i; \bar{x}_n, s_n) \quad (1b) \end{cases}$$

# Self-Supervised Learning



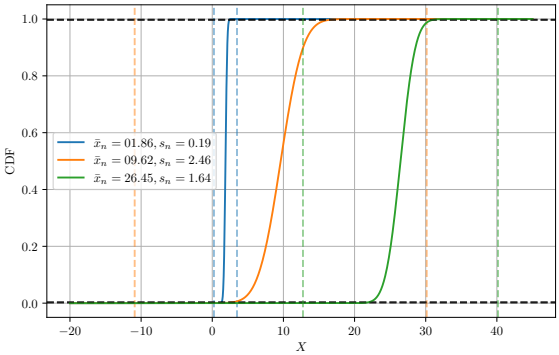

$$y_i = \begin{cases} 0 & \text{if } q \leq F_X(x_i; \bar{x}_n, s_n) \quad (1a) \\ 1 & \text{if } q > F_X(x_i; \bar{x}_n, s_n) \quad (1b) \end{cases}$$

# Self-Supervised Learning



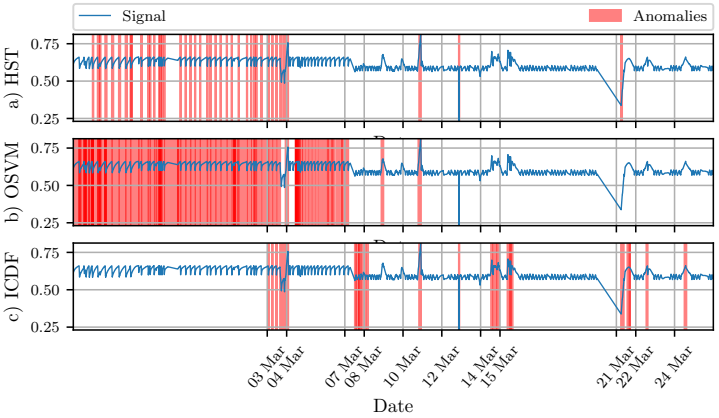

$$\frac{\sum_{y \in Y} y}{n(Y)} > q \tag{3}$$

# Inversion of CDF

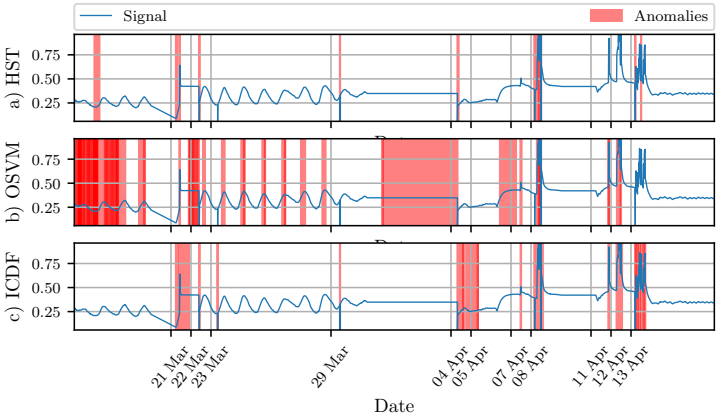



$$x_{\mathrm{l}} = F_X(1-q;\bar{x}_n,s_n)^{-1}$$
$$x_{\mathrm{u}} = F_X(q;\bar{x}_n,s_n)^{-1}$$

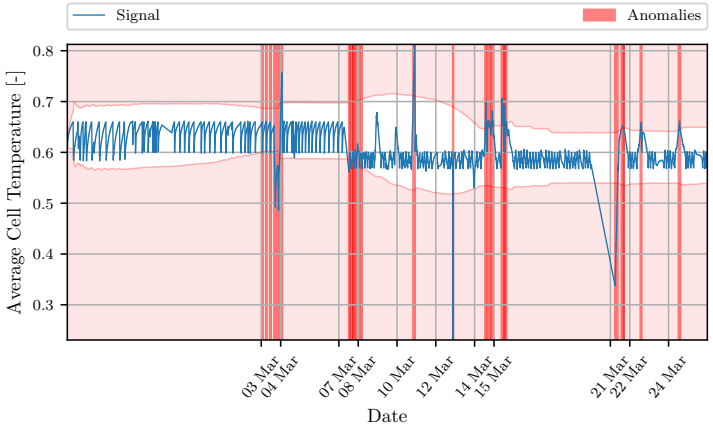
# ICDF-based Outlier Detection - BESS



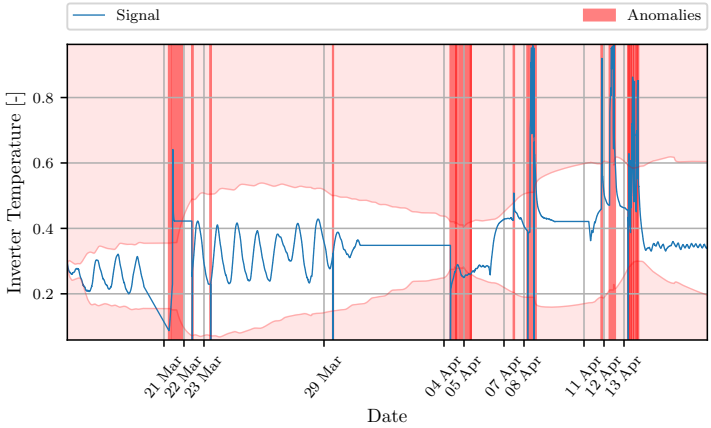
# ICDF-based Outlier Detection - Inverter



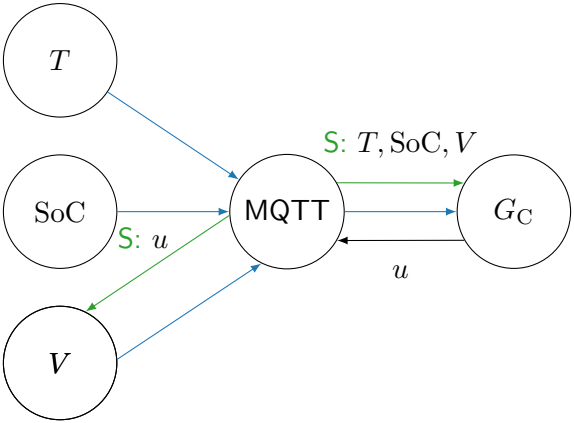
# Dynamic Process Limits



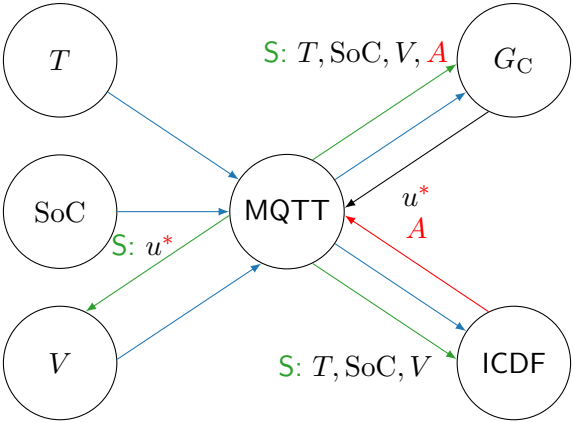
# Dynamic Process Limits



# Utilize Existing Infrastructure



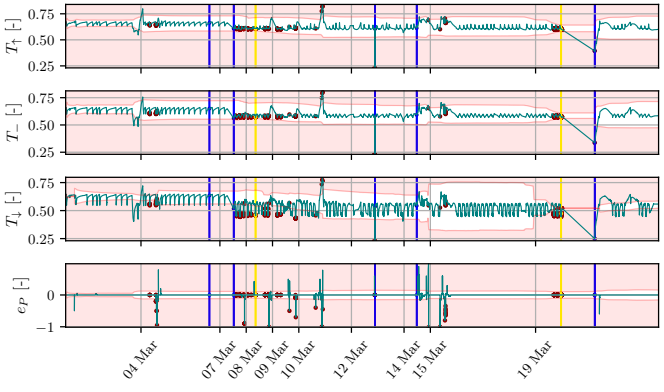
# Utilize Existing Infrastructure



## Summary

- Automated alerting without prior knowledge of process limits
- Assessment of environmental conditions and device aging
- Self-learning approach using streamed data
- Automated alerting threshold setup for multiple signals
- Seamless integration with existing IT infrastructure

# Follow-up research




[1]

[1] M. Wadinger and M. Kvasnica. Adaptable and interpretable framework for novelty detection in real-time iot systems. In Proceedings of the 62nd IEEE CDC, Singapore, 2023. under review.

# Online Anomaly Detection Workflow

**Input:** expiration period $t_e$, time constant $t_c$
**Output:** score $y_i$, threshold $x_{q,i}$

*Initialisation* :
$i \leftarrow 1;\ n \leftarrow 1;\ q \leftarrow 0.9973;\ \bar{x} \leftarrow x_0;\ s^2 \leftarrow 1;$
compute $F_X(x_0)$ ;
*LOOP Process*
**loop**
  $x_i \leftarrow$ RECEIVE();
  $y_i \leftarrow$ PREDICT($x_i$) ;
  $x_{q,i} \leftarrow$ GET($q, \bar{x}, s^2$);
  **if** (1a) **or** (3) **then**
    $\bar{x},\ s^2 \leftarrow$ UPDATE($x_i, \bar{x}, s^2, n$);
    $n \leftarrow n + 1$;
    **for** $x_{i-t_e}$ **do**
      $\bar{x},\ s^2 \leftarrow$ REVERT($x_{i-t_e}, \bar{x}, s^2, n$);
      $n \leftarrow n - 1$;
    **end for**
  **end if**
  $i \leftarrow i + 1$;
**end loop**