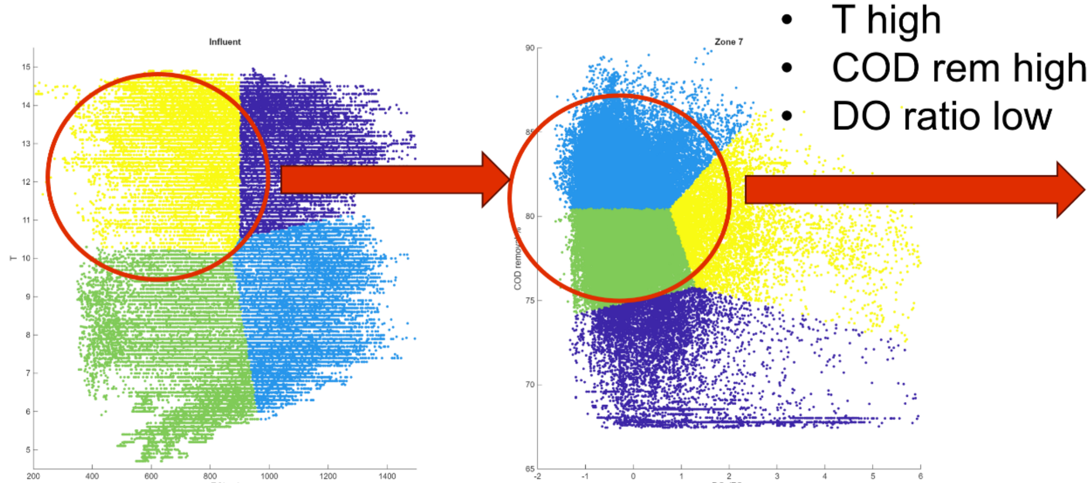
Clustering and optimization of nutrient removal in biological wastewater treatment

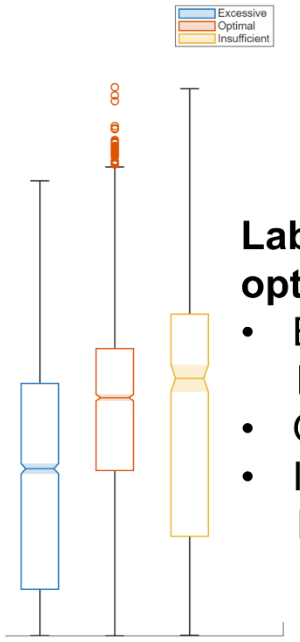
Professor Tiina Komulainen*,
Ranya M.S.H.A. Fahmy* Jehan Kanaganathan*,
Arvind Keprate* Katrine.Marsteng.Jansen**
***Oslo Metropolitan University**
** Hias IKS, Norway

Contents

- Background
- Challenge and Goal
- Data preprocessing and labeling
- Clustering and Box-chart analysis
- Conclusions and further work



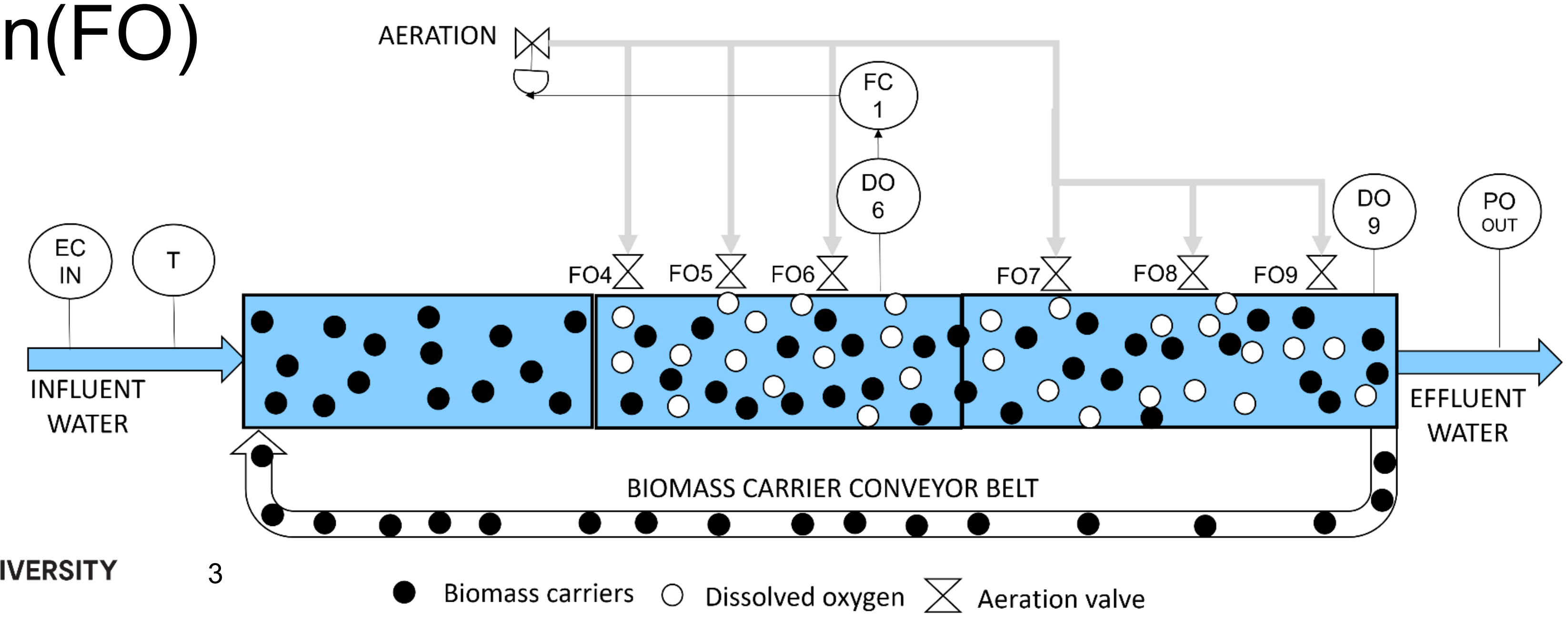
Cluster 8
- ECload low
- T high
- COD rem high
- DO ratio low



Labeled groups optimization potential
- Excessive: savings
  Decrease DO setpoint
- Optimal: no actions
- Insufficient: trouble-shoot
  Increase DO setpoint

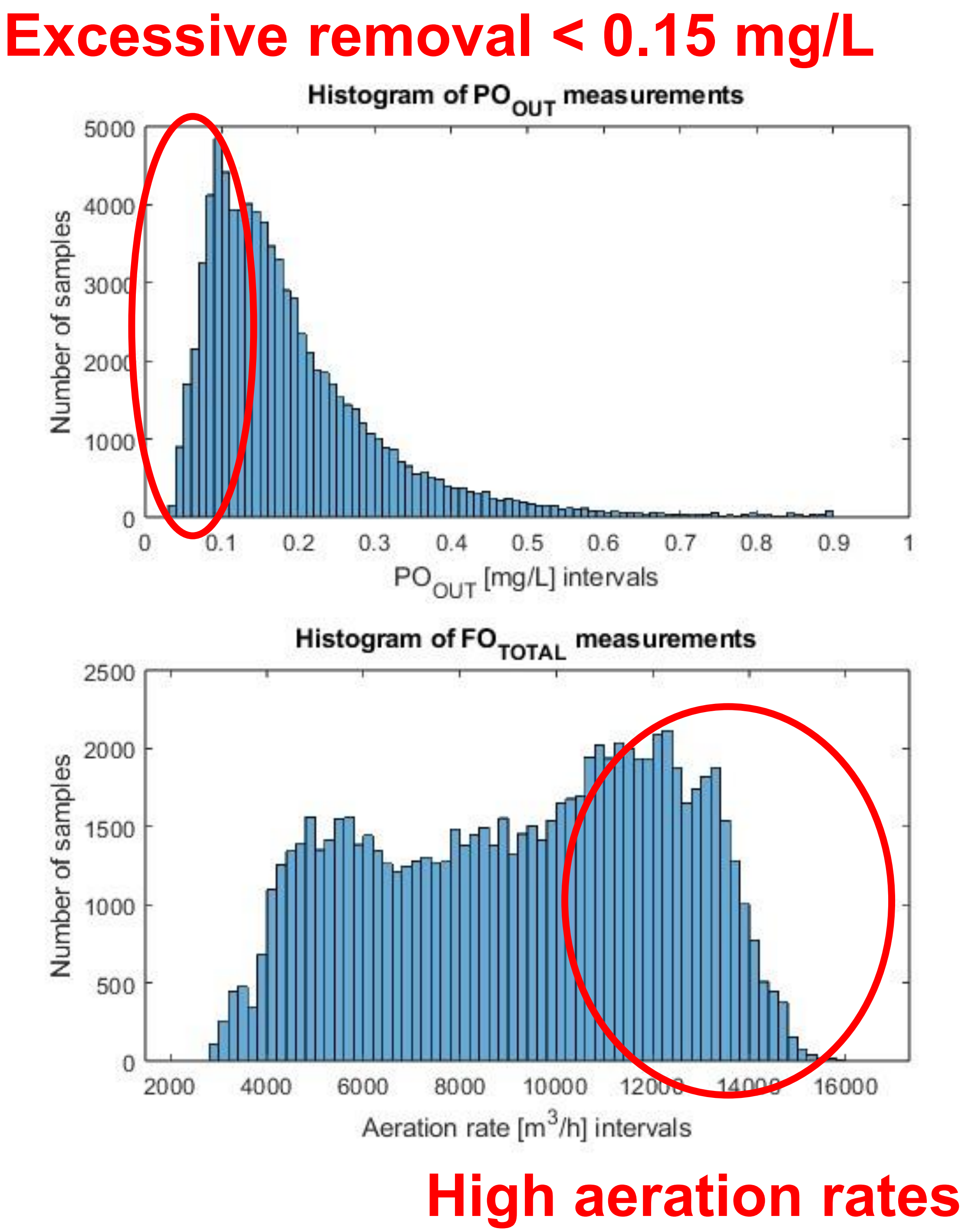
Background

- The Hias Process – moving bed bio-reactor with enhanced biological phosphorus removal and simultaneous nitrification and denitrification
- Municipal WRRF in Hamar, Norway with 192 000 PE
- Phosphorus: influent ~5 mg/L, effluent < 0.4 mg/L
- Main cost: aeration (FO)
- Feedback control loops between dissolved oxygen (DO)-aeration(FO)



Challenge

- Challenge: Most of the effluent phosphorus analyzer data show excessive removal <0.15mg/L.
- This can indicate excessive use of aeration and possibility for cost savings.
- RQ: Can clustering indicate operational strategies leading to excessive removal? Can clustering help identifying control strategies optimizing removal performance?

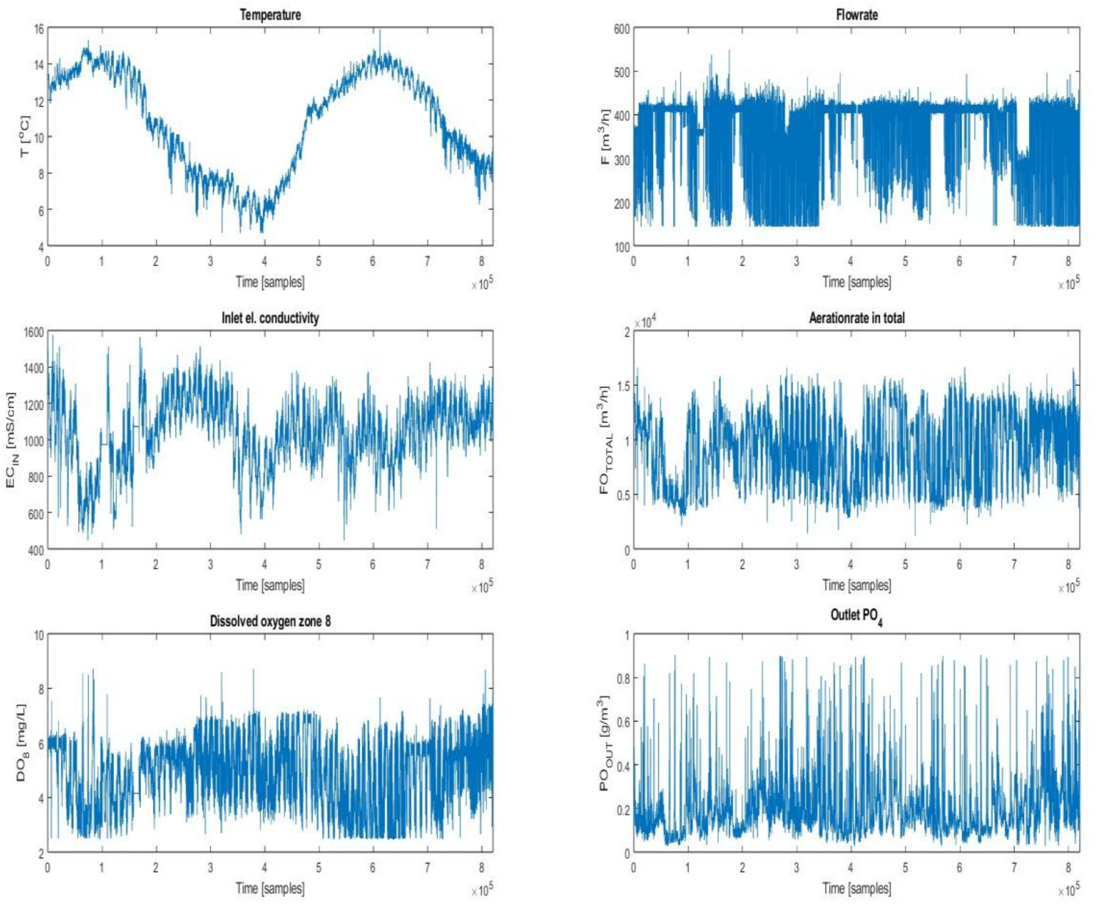


Goal and methods

- Goal is to find operational strategies that lead to excessive nutrient removal and could provide insights in process optimization.
- Use data exploration techniques to identify variables, variable transformations and variable combinations that relate to nutrient removal efficiency
- Label nutrient removal efficiency subgroups excessive – optimal - insufficient
- Apply clustering methods to find groups of inlet conditions and aeration
- Apply box-charts to visualize the subgroups and analyze which conditions lead to excessive nutrient removal

Data pre-processing

- Forward filling of N/A with "previous"
- Backward filling of N/A with "next"
- Outlier removal ±3std from mean
- Downsampling from 1 min to 10 min using mean value
- Dynamic time-lagging using hydraulic retention time for variables at inlet and intermediate zones
- Normalize prior to clustering



Variable selection and labeling

- Variables: 5 influent, 14 intermediate, 2 effluent
- Different variable transformations to improve clustering
  - Organic load ECIN·F, dissolved oxygen to aeration ratio DO/FO

Influent variables
T - temperature
EC_{IN} – electic conductivity
F – wastewater flowrate
COD_{IN} organic matter
NO_{3IN} nitrates

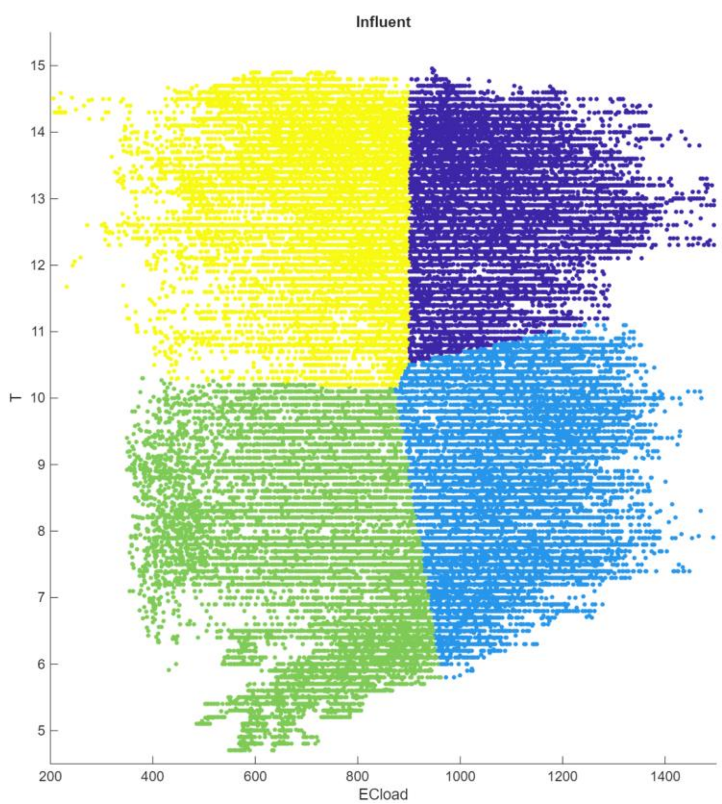
Intermediate variables
DO_4-DO_{10} aeration rates
DO_4-DO_9 dissolved oxygen
COD_7 organic matter
NOX_7 nitrites, nitrates

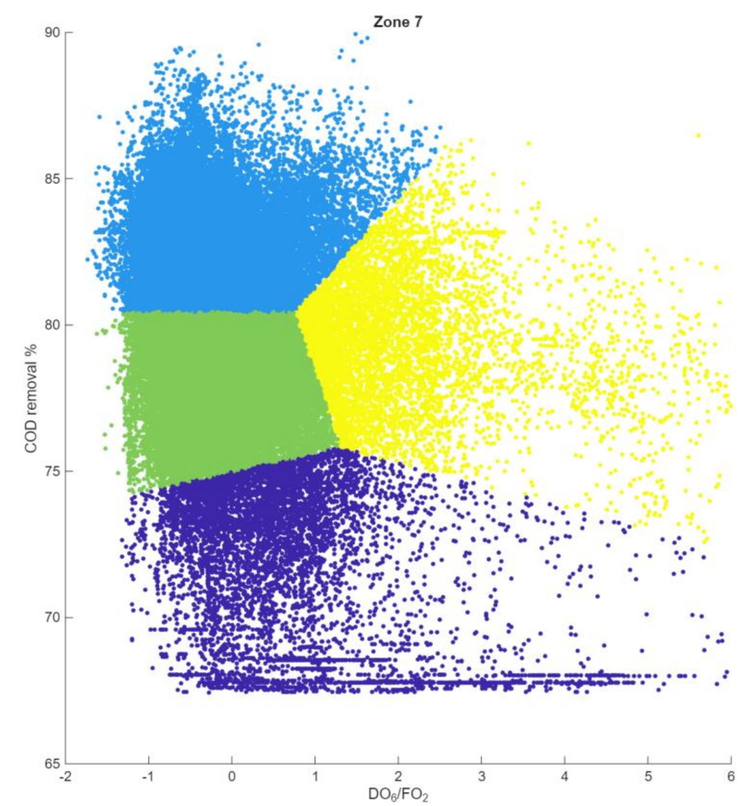
Effluent variables
PO – phosphorus
SS – suspended solids

- Labeling of effluent PO removal performance subgroups
  - Excessive PO<0.15mg/L or PO removal>97%
  - Optimal 0.15mg/L <PO<0.25mg/L or 95%<PO removal<97%
  - Insufficient PO>0.25mg/L or PO removal< 95%

Clustering

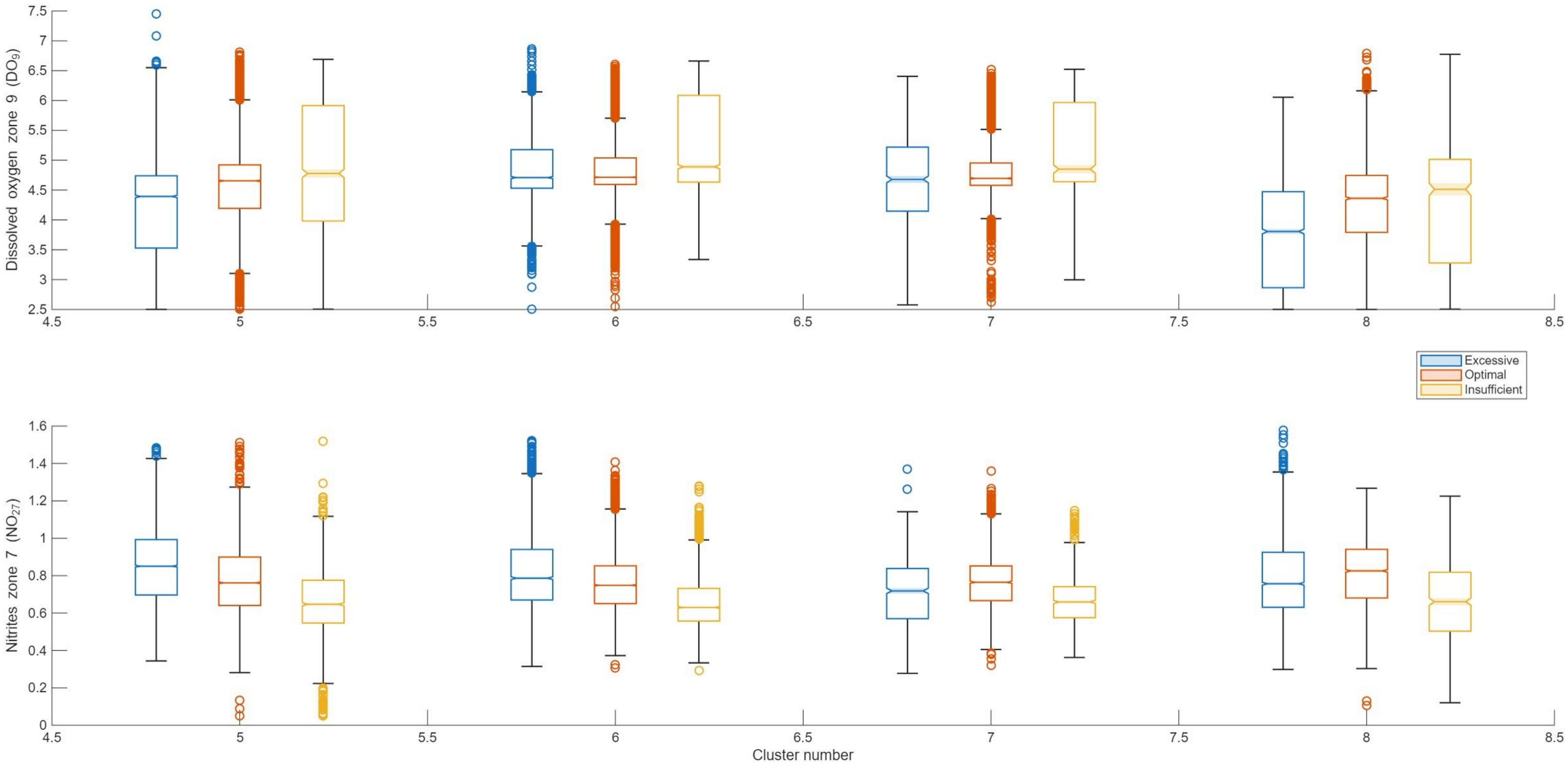
- Clustering methods tested in recent master theses: K-means, Gaussian Mixture Model, Self-Organizing Map with K-means clustering and hierarchical clustering.
  - Best cluster cohesion and separation using K-means
- Clusters for inlet variables clear separation on temperature and nutrients
- Clusters for intermediate variables with clear separation on COD removal performance and dissolved oxygen to aeration ration





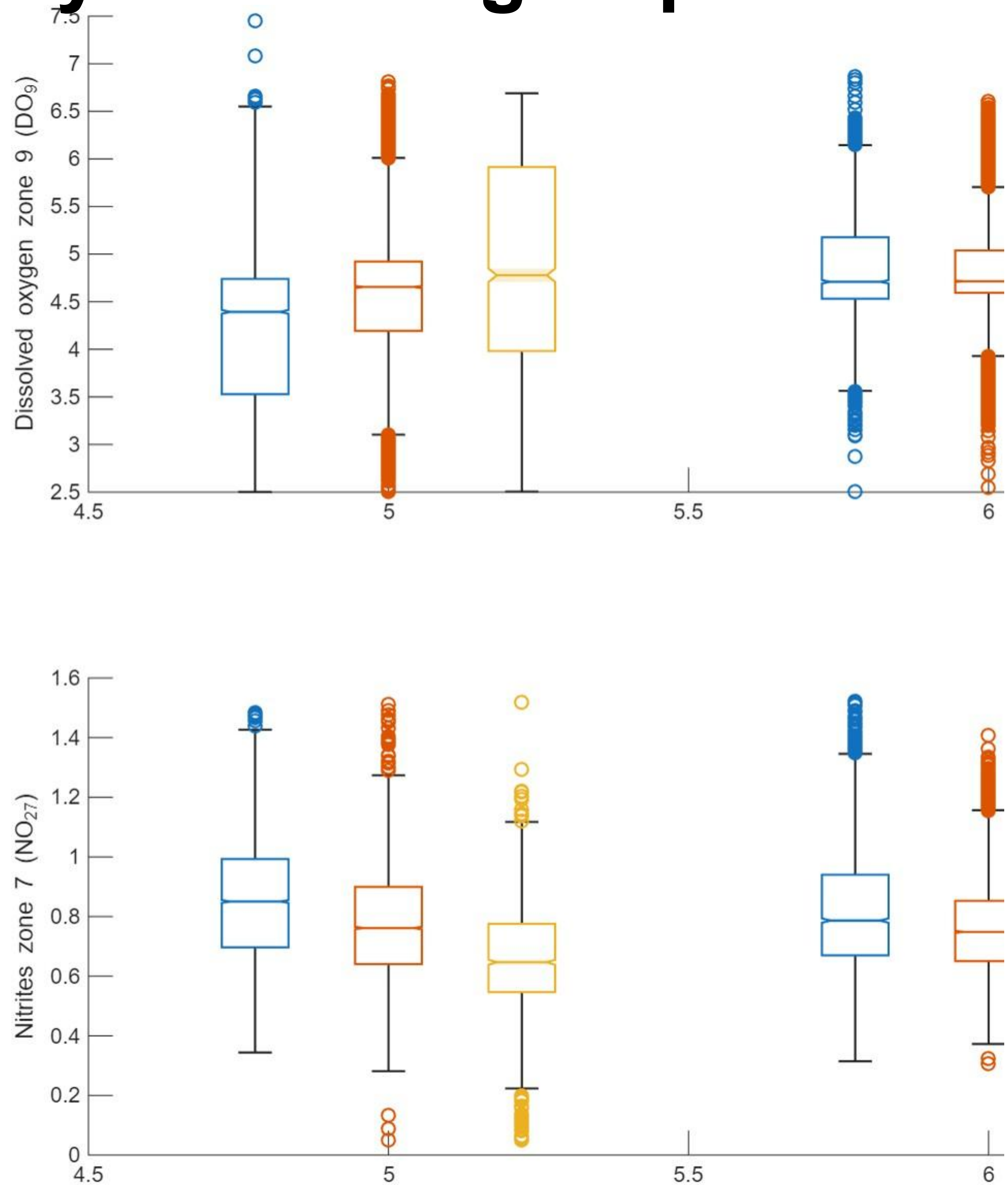
Box-chart analysis of subgroups

- Clusters and subgroups were illustrated with box-charts for all variables and variable combinations
- Analysis of box-charts with goal of finding consistent differences between subgroups within each cluster



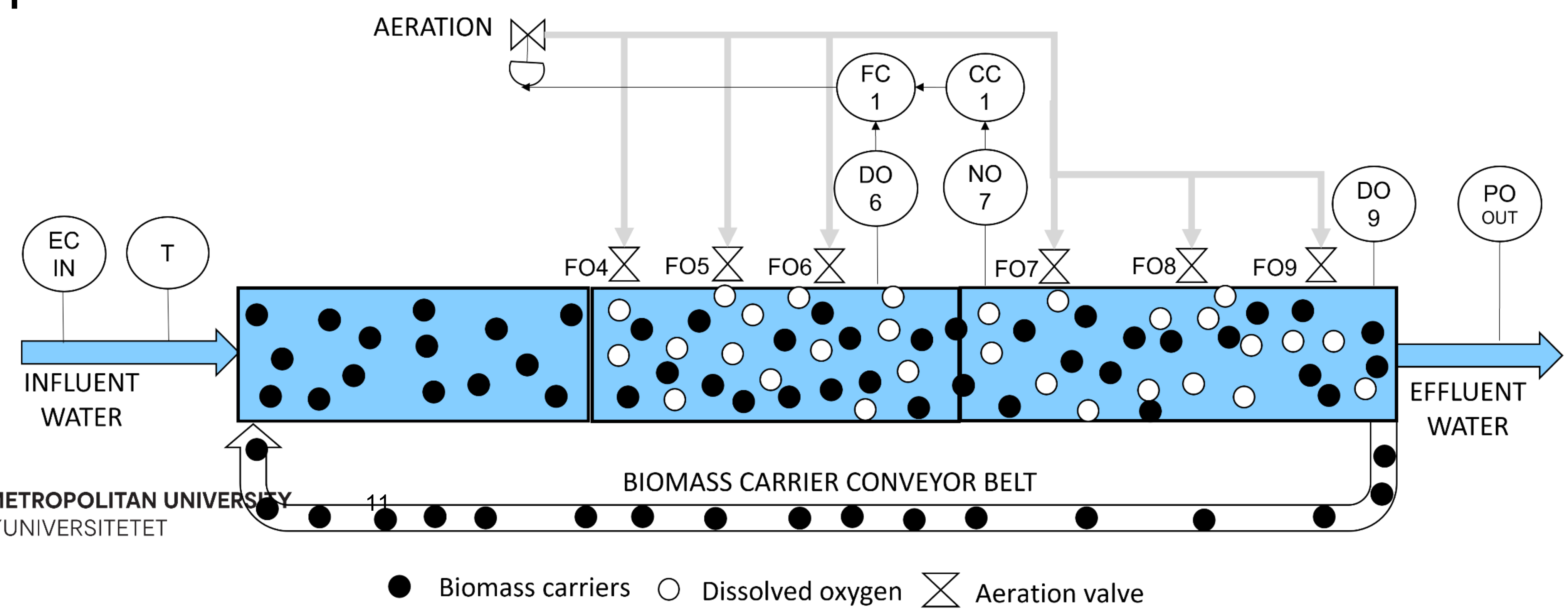
Box-chart analysis of subgroups

- Analysis of effluent dissolved oxygen (DO9) and intermediate nitrites (NO7) reveals significant differences between nutrient removal subgroups
- The excessive group has lowest DO and highest NO→ savings potential
- The insufficient group has highest DO and lowest NO → optimization necessary
- Cascade control strategy based on NO-DO-FO can optimize the aeration use



Conclusions and further work

- K-clustering gave coherent clusters
- Removal performance subgroup labels provided meaningful insights possible improvements
- Cascade control strategy based on NO-DO-FO can optimize the aeration use
- Refinement of clustering variables and labeling with additional process states: transition and disturbance
- Cascade control strategy development using cluster-based setpoints and linearized models



Thank you for your interest

Questions?

Tiina.Komulainen@oslomet.no