ICA2025OSLO
14th IWA Specialised Conference on
Instrumentation, Control and Automation

# Causal vs. Computer Vision-Based Risk Models for Assessing Microbiology-Driven Solids Separation Problems in WRRFs

Sina Borzooei*, Leonardo Scabini**, Jun-Jie Zhu***, Saba Daneshgar****, *****, Lukas Deblieck******, Elias Van Den Broeck******, Odemir Bruno**, Ingmar Nopens****, Elena Torfs******

*IVL Swedish Environmental Research Institute, P.O. Box 210 60, Stockholm SE-100 31, Sweden
**São Carlos Institute of Physics, University of São Paulo, 13560-970, São Carlos - SP, Brazil
***Department of Civil and Environmental Engineering and Andlinger Center for Energy and the Environment, Princeton University, Princeton, NJ 08544, USA
****BIOMATH, Department of Data Analysis and Mathematical Modelling, Ghent University, Coupure links 653, 9000 Ghent, Belgium/Centre for Advanced Process Technology for Urban Resource Recovery (CAPTURE), Frieda Saeysstraat 1, 9000 Gent, Belgium
***** VITO, Water and Energy Transition Unit, Boeretang, 200, 2400, Mol, Belgium
******Pantarein Water, Egide Walschaertsstraat 22L, 2800 Mechelen, Belgium
*******Université Laval, 1065 avenue de la Médecine, Québec G1V 0A6, QC, Canada

**Abstract:** This study compares two paradigms for assessing filamentous bulking (FB) risks in an industrial Water Resource Recovery facility (WRRF) employing a sequencing batch reactor (SBR). A refined causal-based Risk Assessment Model (RAM) was developed to predict FB risk using fuzzy logic and operational triggers, such as low organic loading, nutrient deficiency, and insufficient dissolved oxygen (DO). However, variations in FB causes, significantly alter the morphology of filamentous bacteria and flocs, which operational triggers fail to capture. To address this limitation, computer vision (CV) models, including convolutional neural networks (CNNs) and vision transformers (ViTs) were implemented to analyse microscopy images (MIs) for FB risk assessment. Results show MI-based models outperformed RAM, with ViTs excelling in recognizing complex spatial patterns. These findings highlight the importance of sludge morphology in FB risk assessment and advocate for hybrid approaches that integrate operational triggers with CV-based morphological insights, enabling proactive sludge management strategies.

**Keywords**: Risk Assessment Model; Computer vision; Filamentous bulking; Sequencing batch reactor; Wastewater treatment

## 1. INTRODUCTION AND BACKGROUND

Filamentous bulking (FB) poses a critical challenge in biological nutrient removal-activated sludge (BNRAS) systems, leading to poor settleability, operational inefficiencies, and environmental issues in Water Resource Recovery Facilities (WRRFs). The consequences include increased chemical use, energy consumption, maintenance costs, and investments in advanced treatment technologies. Poor effluent quality also threatens aquatic ecosystems and risks regulatory penalties for permit violations.

Traditional FB risk assessment relies on metrics like the sludge volume index (SVI), which reactively indicates problems after they occur. Proactive Risk Assessment Models (RAMs), based on heuristic and empirical knowledge, identify risks by evaluating operational triggers such as low organic loading, nutrient deficiencies, and insufficient dissolved oxygen (DO) (Dalmau et al. 2008). These causal-based models, integrated with mechanistic process simulations, offer insights into control strategies (Comas et al. 2008); however, available RAMs face several limitations, including dependence on process variables often influenced by local operational practices and design. They lack microbial insights, failing to account for the complex interactions within the microbial community that drive FB. Their knowledge formation steps have been based on conventional activated sludge systems of municipal WRRFs, limiting their applicability to emerging treatment technologies and industrial wastewater.

In this study, we use microscopy images (MIs) to address these limitations and capture microbial changes associated with FB in an industrial WRRF employing a sequencing batch reactor (SBR). Building on our previous work (Borzooei et al. 2024), where convolutional neural networks (CNNs) predicted SVI values from MIs, we advance the approach by directly classifying images as "bulking" or "non-bulking" for precise FB identification. We incorporate two computer vision

models, Vision Transformers (ViTs) and CNNs, to analyse sludge morphology and provide microbial-level FB risk insights. These data-driven methods are compared with an improved RAM, highlighting their strengths and limitations.

## 2. METHODOLOGY

### 2.1 Data Collection and Acquisition

This study was conducted at an industrial WRRF in Antwerp, Belgium, operated by Pantarein Water. The facility processes, on average, 2300 $m^3$/day of malting wastewater using a SBR for biological and chemical phosphorus and nitrogen removal, followed by a membrane bioreactor (MBR) for solids separation. Weekly samples were collected over 21 months (November 2020–July 2022), including operational data and MIs. Measured parameters included influent and effluent characteristics (COD, TN, TP), SRT, DO, pH, and mixed liquor-suspended solids (MLSS). Settling characteristics were assessed via SVI using APHA protocols, with total suspended solids (TSS) determined through filtration and drying. MIs were captured using a phase-contrast microscope (Olympus BX43) at 4× magnification. Images were pre-processed (512×384 pixels) and normalized with ImageNet standards, resulting in 494 MIs representing microbial morphology. Each image is labelled as bulking (279 samples) or non-bulking (215 samples).

### 2.2 A Fuzzy-Logic-Based Risk Assessment Model

This study builds on the foundational work of Dalmau et al. (2008), adapting their rule-based RAM to the dynamic conditions of an industrial WRRF employing an SBR. The original model, designed for CAS processes in municipal systems, required modifications to address site-specific conditions. The RAM evaluates FB risk using three primary rules: (1) low DO during high food-to-microorganism ($F/M_{Removed}$) and high SRT, (2) nutrient deficiencies indicated by high $BOD_5$/TN or $BOD_5$/TP ratios, and (3) substrate-limiting conditions characterized by low $F/M_{Fed}$ and low readily biodegradable COD (rbCOD). Seven key variables derived from laboratory, sensor, and operator data informed these rules. To account for industrial wastewater, rbCOD ranges were adjusted from 0–145 mg/L (municipal) to 1689 ± 353 mg/L, and DO thresholds were refined to reflect cyclic SBR operations, requiring >1.5 mg/L for at least 45% of the cycle. With triangular and pseudo-trapezoidal membership functions, the Mamdani fuzzy inference approach was used for fuzzification, translating numerical data into qualitative descriptors. A Centre of Gravity method was applied for defuzzification, producing a crisp risk value (0–100%). The model was implemented using Python 3.8.2 for data processing and MATLAB's Fuzzy Logic Toolbox for modelling.

### 2.3 MI-Based Risk Assessment Models

Building on prior work (Borzooei et al., 2024), we applied transfer learning and fine-tuned two neural network architectures for MI analysis: the ConvNeXt-nano CNN (Liu et al., 2022) and the multi-resolution transformer CrossViT-18 (Chen et al., 2021). ConvNeXt-nano, pre-trained on ImageNet-21k, demonstrated superior performance for SVI prediction, while CrossViT-18, pre-trained on ImageNet-1k, was chosen for its ability to handle varying input resolutions and aspect ratios, aligning with our rectangular images. The models were fine-tuned by replacing the final layer with binary classification (bulking vs. non-bulking) and optimizing a binary cross-entropy loss function using the AdamW optimizer.

Data augmentation techniques were employed to enhance robustness and reduce overfitting while preserving the morphological properties of flocs in MIs. Transformations, including grayscale conversion, flips, translation, rotation, brightness adjustments, and random region erasing, were applied iteratively during training. We used a weight decay of $10^{-8}$ and a drop path rate of 0.1.

Each CV model was evaluated using a stratified 10-fold cross-validation protocol. Models were trained on nine folds and tested on the remaining fold, repeating this process ten times. The training was conducted with a learning rate of $10^{-4}$, a batch size of 16, and over 50 epochs with cosine-

annealing learning rate decay. The best-performing models, determined by the lowest test loss, computed *SoftMax* probabilities across two output neurons. Probabilities for multiple images per date were averaged, producing robust and reliable performance metrics while minimizing overfitting.

## 3. Results and Discussion

Identifying bulking periods was essential for evaluating the models' performance. This process relied on four key data sources: (1) online turbidity data, (2) offline water and sludge quality measurements (MLSS and SVI), (3) offline microbiological data, and (4) in-situ operator observations. Operator input served as the primary signal, supported by trends in MLSS and SVI values. The findings are summarized in Fig. 1.

Fig. 1(a) illustrates the individual risks contributing to FB, with Low DO emerging as the most critical factor, followed by nutrient deficiency and low organic loading. Fig. 1(b) shows the aggregated daily risks calculated by RAM, with red-shaded regions indicating bulking periods. Fig. 1(c) and Fig. 1(d) display predictions from MI-based models, highlighting average daily risks (black dashed lines) and variability (shaded regions). Evaluation criteria flagged bulking days as high risk (≥80%) and non-bulking days as low risk (<80%). MI-based models achieved 81–91% accuracy in flagging bulking days and >90% accuracy during non-bulking periods, with ViTs outperforming CNNs due to their superior ability to capture spatial patterns. In contrast, RAM flagged fewer than 30% of bulking days as high risk and achieved only 43% accuracy during non-bulking periods.

This performance difference shows the fundamental contrast between causality-driven and morphology-based approaches. MI-based models consistently tracked FB events and recovery periods. In contrast, RAM relies on external operational triggers, often influenced by interventions, resulting in inconsistent predictions. For instance, during the final FB period, RAM flagged high risk only toward the end and continued signalling risk even after the period ended. This likely reflected operational adjustments (e.g., low DO) rather than true microbiological recovery. Conversely, MI-based models accurately identified the resolution of bulking events, reflecting actual sludge recovery from a microbiological perspective.

## 4. Conclusions and Future Directions

This study demonstrates the superiority of MI-based models in predicting FB events, emphasizing the critical role of sludge morphology in achieving accurate risk assessment. The CrossViT-18 outperformed CNNs by effectively capturing complex spatial patterns, reflecting advancements in computer vision research. In contrast, RAM's reliance on operational triggers rather than direct biological indicators resulted in inconsistent predictions, especially during prolonged bulking events.

Different FB causes lead to the dominance of specific filamentous bacteria, each exhibiting unique morphological characteristics. For example, oxygen-limited conditions favor *Sphaerotilus natans*, nutrient deficiencies encourage *Microthrix parvicella*, and low organic loading promotes the growth of organisms like *Type 021N*. These bacteria influence the density, length, and branching patterns of sludge flocs, impacting overall floc architecture and settleability. The interactions between filamentous and floc-forming bacteria further modify sludge morphology, driven by the environmental triggers of FB. Hence, future research should focus on hybrid models that combine the interpretability of RAM with the precision of MI-based approaches. By integrating operational triggers with morphology-based insights, such models could provide robust, interpretable, and practical solutions for proactive sludge management in wastewater treatment plants.

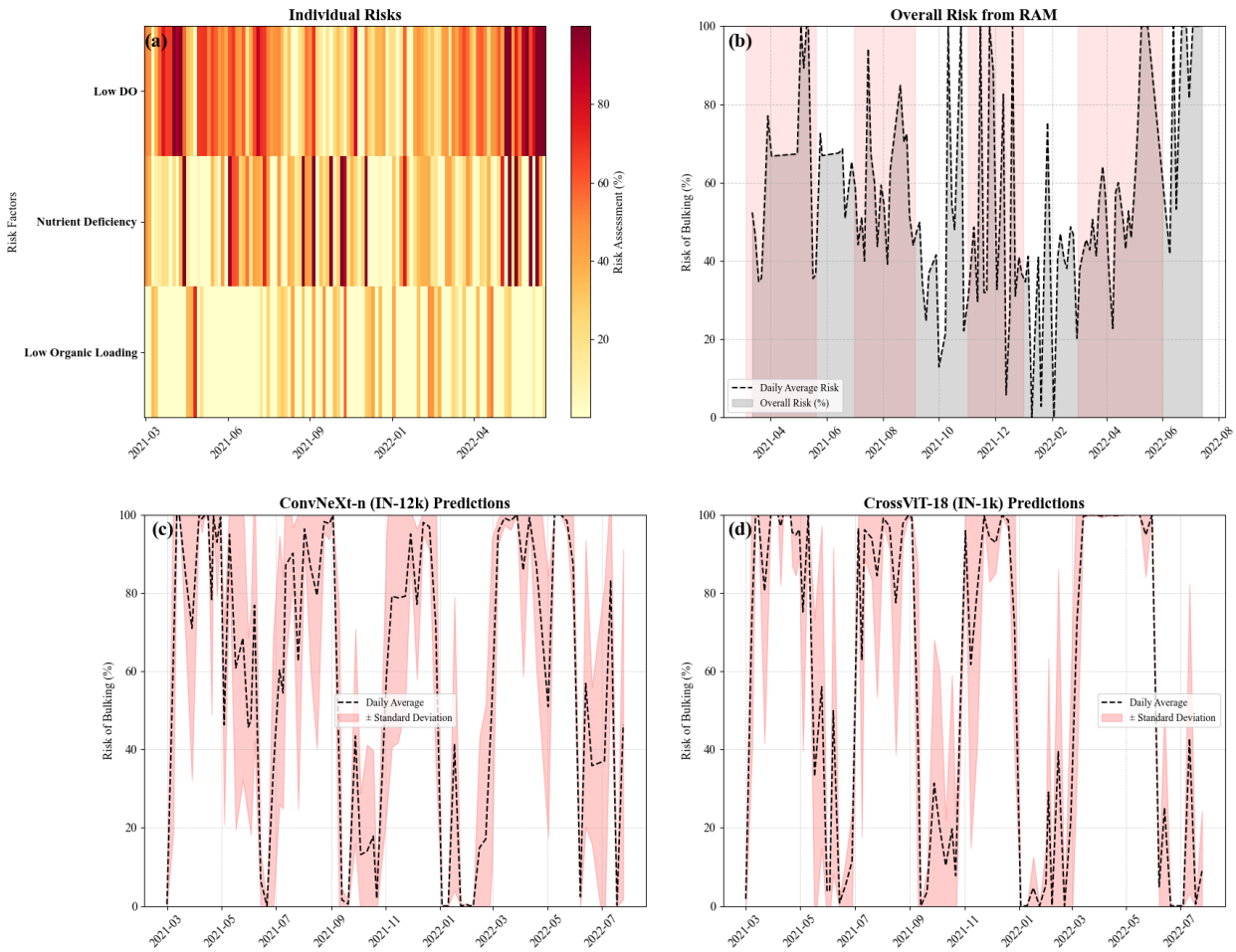


Fig. 1. Comparative Performance of Risk Assessment Models: Visualization of Individual Risk Factors (a) Overall Risk from RAM (b), MI-Based Model Predictions (ConvNeXt-n and CrossViT-18) (c, d)

## REFERENCES

Borzooei, S., L. Scabini, G. Miranda, S. Daneshgar, L. Deblieck, O. Bruno, P. De Langhe, B. De Baets, I. Nopens, and E. Torfs. 2024. "Evaluation of activated sludge settling characteristics from microscopy images with deep convolutional neural networks and transfer learning." *Journal of Water Process Engineering*, 64: 105692. https://doi.org/10.1016/j.jwpe.2024.105692.

Comas, J., I. Rodríguez-Roda, K. V. Gernaey, C. Rosen, U. Jeppsson, and M. Poch. 2008. "Risk assessment modelling of microbiology-related solids separation problems in activated sludge systems." *Environmental Modelling & Software*, 23 (10–11): 1250–1261. Elsevier.

Dalmau, J., I. Rodríguez-Roda, and J. Comas. 2008. "Risk model for microbiology-related settling problems implementation."

Chen, C. F. R., Q. Fan, R. Panda. 2021. "Crossvit: Cross-attention multi-scale vision transformer for image classification." Proceedings of the IEEE/CVF international conference on computer vision.