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

# Hybrid models integrating microbial images of activated sludge

L. Verhaeghe*,**,***, J. Verwaeren**, E. Torfs*,***

* model*EAU*, Université Laval, 1065 avenue de la Médecine, Québec G1V 0A6, QC, Canada, loes.verhaeghe.1@ulaval.ca, elena.torfs@gci.ulaval.ca

** BIOVISM, Department of Data Analysis and Mathematical Modelling, Faculty of Bioscience Engineering, Ghent University, Coupure links 653, 9000 Gent, Belgium, jan.verwaeren@ugent.be

*** BIOMATH, Department of Data Analysis and Mathematical Modelling, Faculty of Bioscience Engineering, Ghent University, Coupure links 653, 9000 Gent, Belgium

**Abstract:** Wastewater Resource Recovery Facilities (WRRFs) generate large volumes of data, much of which remains underused in current modeling practices. This study investigates the integration of microbial images into hybrid modeling (HM) structures to enhance mechanistic model predictions. Two HM configurations are explored: a parallel model, where microbial images are used by a data-driven component to predict the residual error of a mechanistic model, and a serial model, where images inform the dynamic estimation of the oxygen transfer coefficient ($K_{La}$). In the parallel model, improvements are targeted for effluent total suspended solids (TSS) and airflow rate predictions. Image features are extracted using two different approaches: a pretrained Convolutional Neural Network (CNN) and a convolutional autoencoder, allowing for a comparison between supervised feature extraction based on pretraining and unsupervised feature learning. Results from a 1.5-year pilot-scale WRRF dataset show that the CNN-enhanced hybrid models modestly improve predictions, particularly for airflow rate and $K_{La}$ dynamics, even under highly variable conditions. These findings highlight the potential of microbial images as a valuable, underutilized data source in WRRF modeling and motivate future testing on more stable full-scale systems.

**Keywords**: Wastewater treatment, hybrid modeling, settling, aeration, microbial imaging

## INTRODUCTION

Wastewater Resource Recovery Facilities (WRRFs) generate vast amounts of data from a variety of sources, much of which are currently either not used or underused in traditional models of water systems. Examples include microbial images and molecular data, as well as raw sensor signals (such as spectrograms from spectrolyzers) that are typically reduced to a few predefined variables, yet they still retain a wealth of additional information that remains untapped. This data has the potential to provide rich insights into processes and operating conditions. Currently, WRRFs are primarily modeled using mechanistic approaches based on system knowledge, with inputs limited to sensor data and lab measurements of water quality and flow variables. While interpretable, these models often oversimplify complex processes, resulting in limited predictive accuracy. Hybrid Models (HMs) merge mechanistic and data-driven modeling techniques and combine in this way the advantages of the two approaches. Within the HM structure, data-driven models can leverage underused data sources to complement mechanistic models. This study investigates how microbial images, routinely collected to assess activated sludge properties, can be integrated into HM structures. Previous research has shown that these images contain valuable information on microbial communities and floc characteristics, which influence key processes such as settling and aeration (Arelli et al., 2009; Boztoprak et al., 2016; Campbell et al., 2019; Koivuranta et al., 2014; Wilén et al., 2004). The developed HMs aim to support improved predictions, real-time monitoring, and the development of digital twins to drive automation and optimization in WRRFs.

## METHODOLOGY

A 1.5-year measurement campaign was conducted at a pilot-scale WRRF, pil*EAU*te (Québec, Canada), which treats domestic wastewater and rainwater. The plant configuration includes a pumping station, a storage tank, a primary settler, and a biological reactor with one anoxic basin, two aerobic basins, and a secondary clarifier. The measurement campaign collected sensor data and lab measurements from the influent, biological reactor, recycle streams, and effluent to gain a

general process understanding. Additionally, weekly microbial images of activated sludge were captured using a phase contrast microscope with a 10x objective lens, with at least 40 images taken each week. The data are collected from a pilot-scale WRRF, which is subject to frequent operational changes, testing scenarios, and rapidly shifting conditions that do not reflect stable full-scale operation. This resulted in a highly variable dataset, with limited consistency over time and some unrealistic process conditions (e.g. MLSS<0.4g/L). Additionally, data quality and quantity are limited for certain variables, e.g. effluent TSS was only measured through weekly grab samples, reducing the temporal resolution and reliability of key target variables.

First, microbial images are integrated into a parallel hybrid model architecture, where they serve as input to a data-driven component designed to predict the residual error of a mechanistic model. This approach is applied to a mechanistic settler model, with the data-driven model predicting the residual in effluent Total Suspended Solids (TSS). In addition, the same methodology is tested for improving predictions of the airflow rate supplied to the biological reactor to maintain the dissolved oxygen setpoint. The predicted residuals are added to the output of the mechanistic model, correcting its predictions (Figure 1, left). This setup leverages the hypothesis that residuals may capture effects related to microbial communities and floc structure, factors not explicitly accounted for in the mechanistic model. By incorporating information from microbial images, the hybrid model could enhance prediction accuracy while preserving the interpretability and structure of the mechanistic framework. Secondly, a serial hybrid model is explored to enhance airflow rate predictions. In this configuration, microbial images are used as input to a data-driven model that estimates the volumetric oxygen mass transfer coefficient ($K_{La}$), a key parameter governing oxygen transfer from the gas to the liquid phase. $K_{La}$ is influenced by the properties of the activated sludge, such as floc structure and microbial composition. The predicted $K_{La}$ values can then be used as dynamic inputs to the mechanistic model, allowing it to adjust in real-time based on microbial data (Figure 2, right). This enables more accurate and responsive modeling and control of aeration by capturing changes in microbial communities and floc characteristics.

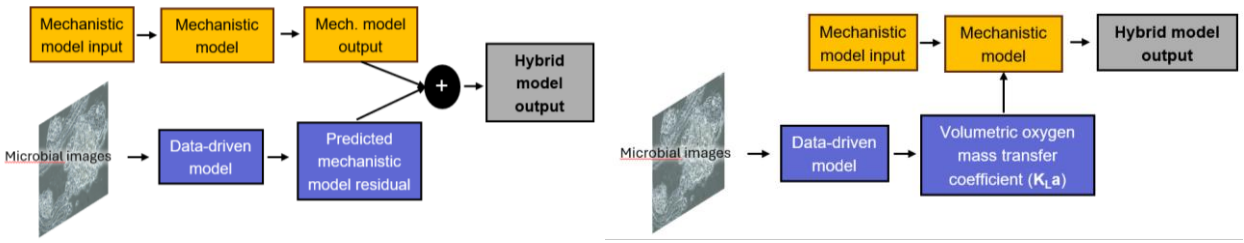


**Figure 1** Scheme of the parallel HM structure (left) and serial HM structure (right).

Two data-driven approaches are explored within the hybrid model structures. The first follows a common image-based prediction strategy using ConvNeXt-nano, a Convolutional Neural Network (CNN) pretrained on the large-scale ImageNet-21k dataset (>14 million non-microbial images). CNNs are deep learning models that automatically learn to extract meaningful features from image data. Pretraining is essential, as it enables the network to learn general feature representations, which are then fine-tuned on the microbial image dataset from pil*EAU*te. The final layers are adapted to perform regression and predict the target value. This approach is based on the best-performing model from Borzooei et al. (2024), where further details on optimization, training and model setup are provided. Secondly, a convolutional autoencoder is used for unsupervised feature extraction from microbial images. An autoencoder is a type of neural network that learns to compress input data into a lower-dimensional representation (latent space) and then reconstruct it, capturing the most relevant features without needing labeled output data (Xu & Duraisamy, 2020). In this case, the autoencoder compresses high-dimensional image data into latent features that retain essential visual patterns. These features are used as input to a separate Random Forest regression

model that predicts the target variable. This approach is valuable because it allows feature extraction independent of the prediction task, potentially capturing broader and more generalizable information from the images. Details on the optimization strategy and final data-driven model architectures are provided in Appendix 1.

## RESULTS AND DISCUSSION

Figure 3 presents the results of the mechanistic model and parallel HM for predicting effluent TSS concentrations and air flowrate with the pretrained CNN as data-driven component. During the training period, the HM demonstrates accurate predictions, showcasing its strong ability to capture the underlying patterns in the data. During the testing periods, the HMs demonstrate some capability to capture system dynamics (peaks and trends), particularly in predicting airflow rates, although a significant bias persists. These results indicate that microbial image data contain exploitable information on settling and aeration processes, but substantial enhancement of the mechanistic model has yet to be achieved. It should be noted that the last months of data in Figure 3 represent the most rigorous measurement campaign. When the test and training periods are rearranged, e.g. using the first months as a test period, the model's performance declined notably.

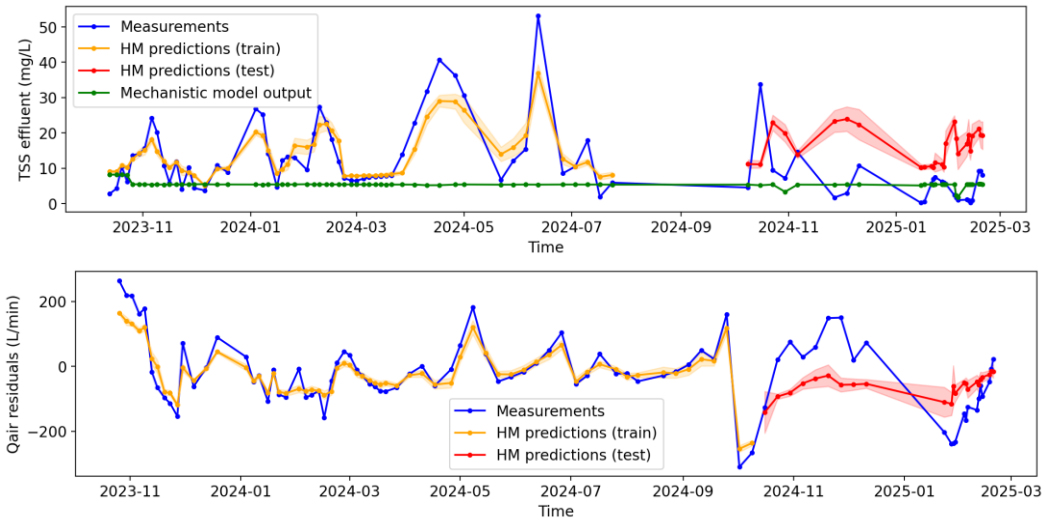


**Figure 3** Results of the parallel hybrid model during the training (orange) and testing (red) periods of the data-driven CNN model. The top figure shows the results for TSS effluent, and the bottom figure for the air flowrate residuals.

Figure 4 presents the results of the $K_{La}$ prediction by directly using the pretrained CNN (no hybrid model). The model performs well during the training period, potentially even showing signs of overfitting. Nonetheless, during the testing period, it is still able to capture certain dynamic patterns in $K_{La}$, indicating that the data-driven component can extract meaningful information from microbial images alone. The Kla predictions could subsequently be used as input into a serial hybrid model.

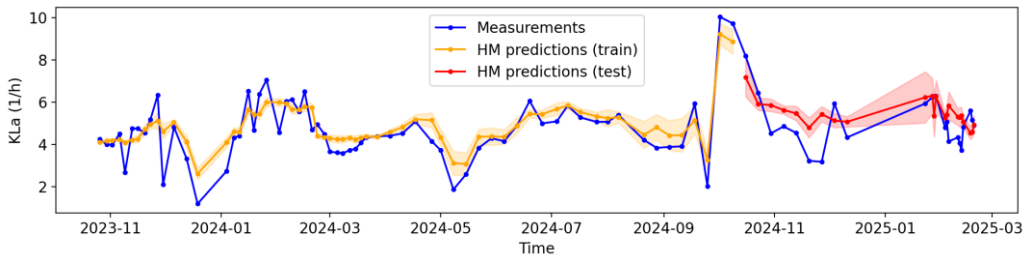


**Figure 4** Results of the $K_{La}$ prediction during the training (orange) and testing (red) periods of the data-driven CNN model.

The results of the hybrid models using the autoencoder combined with a random forest, presented in Appendix 2, indicate that this data-driven approach performs less effectively than the CNN-based model. Although the autoencoder performed well in reconstructing the images after compression to a latent space, it did not sufficiently extract meaningful information related to settling or aeration into its latent space (despite different preprocessing techniques as well as targeted loss function configurations, see Table 1). This highlights the advantage of leveraging pretrained networks, which benefit from prior training on large-scale non-microbial image datasets. The overall performance of the hybrid models, while not outstanding, must be interpreted in the context of the challenging nature of the dataset. Despite the extremely variable dataset, it is noteworthy that the data-driven models were still able to capture meaningful dynamic patterns, particularly peaks and trends. This suggests that microbial images, even under highly variable conditions, contain valuable information that can enhance mechanistic models when properly leveraged. To further evaluate the potential of these methods, future work will focus on applying them to datasets from full-scale WRRFs. Such systems typically exhibit more stable operating conditions and offer higher-resolution data, which may enable more reliable training and improved model performance.

## REFERENCES

Arelli, A., Luccarini, L., and Madoni, P. 2009. Application of image analysis in activated sludge to evaluate correlations between settleability and features of flocs and filamentous species. *Water Science and Technology*, **59**(10), 2029 – 2036.

Borzooei, S., Scabini, L., Miranda, G., Daneshgar, S., Deblieck, L., Bruno, O., De Langhe, P., De Baets, B., Nopens, I., and Torfs, E. 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.

Boztoprak, H., Özbay, Y., Güçlü, D., and Küçükhemek, M. 2016. Prediction of sludge volume index bulking using image analysis and neural network at a full-scale activated sludge plant. *Desalination and Water Treatment*, **57**(37), 17195 – 17205.

Campbell, K., Wang, J., and Daniels, M. 2019. Assessing activated sludge morphology and oxygen transfer performance using image analysis. *Chemosphere*, **223**, 694 – 703.

Koivuranta, E., Keskitalo, J., Stoor, T., Hattuniemi, J., Sarén, M., and Niinimäki, J. 2014. A comparison between floc morphology and the effluent clarity at a full-scale activated sludge plant using optical monitoring. *Environmental Technology*, **35**(13), 1605 – 1610.

Wilén, B. M., Gapes, D., Blackall, L. L., and Keller, J. 2004. Structure and microbial composition of nitrifying microbial aggregates and their relation to internal mass transfer effects. *Water Science and Technology*, **50**(10), 213-220.

Xu, J. and Duraisamy, K. 2020. Multi-level convolutional autoencoder networks for parametric prediction of spatio-temporal dynamics. *Computer Methods in Applied Mechanics and Engineering*, **372**, 113379.

## APPENDIX 1

**Table 1** Overview of optimization strategies explored for the data-driven components.

| *Aspect* | *CNN-Based Model* | *Autoencoder-Based Model* |
|---|---|---|
| ***Image Preprocessing*** | Normalization, edge detection | Resize, grayscale conversion, histogram stretching, edge detection, Contrast Limited Adaptive Histogram Equalization (CLAHE) |
| ***Architecture Tuning*** | Pretrained ConvNeXt-nano, following Borzooei et al. (2024) | Variation of latent space size, convolutional layers/channels, batch normalisation, activation functions, regularization |
| ***Loss Function*** | Mean Squared Error (MSE), following Borzooei et al. (2024) | MSE, Structural similarity index measure (SSIM), combined MSE+SSIM, perceptual loss, joint loss formulation |

| | | |
|---|---|---|
| ***Downstream Model*** | Linear regression layer, following Borzooei et al. (2024) | Random Forest (on flattened latent features), CNN, fully connected neural network (on flattened latent features) |
| ***Additional Inputs*** | Inclusion of MLSS as extra feature in regression layer | Inclusion of MLSS as extra feature in downstream model |

**Table 2**: Final model configurations of the data-driven components.

| Feature Extractor | Pretraining | Image preprocessing | Optimizer / Training Strategy | Loss function | # Trainable Parameters | Validation |
|---|---|---|---|---|---|---|
| **CNN-BASED MODEL** | ImageNet-21k (14M images) | Resizing, normalization, data augmentation | AdamW, LR=0.0001 (adaptable), early stopping based on validation loss | Loss based on target prediction: MSE | ~15M | 10-fold cross-validation |
| **AUTO-ENCODER** | None | Resizing, CLAHE, no data augmentation | Adam, LR=0.0001, early stopping based on validation loss | Loss based on image reconstruction: MSE+10*SSIM | ~3.5M | 10% of training data used for validation |

## APPENDIX 2

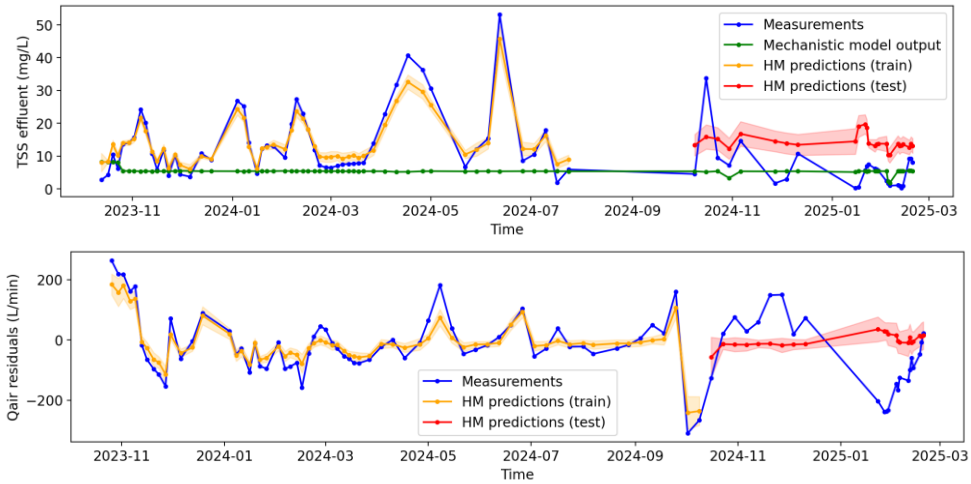


**Figure 5** Results of the parallel hybrid model during the training (orange) and testing (red) periods of the data-driven autoencoder model. The top figure shows the results for TSS effluent, and the bottom figure for the air flowrate residuals.

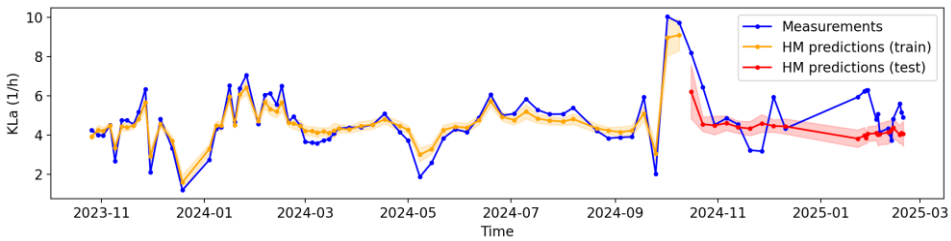


**Figure 6** Results of the $K_{La}$ prediction during the training (orange) and testing (red) periods of the data-driven autoencoder model.