Environmental Science and Ecotechnology 24 (2025) 100527

Contents lists available at ScienceDirect

Environmental Science and Ecotechnology

journal homepage: www.journals.elsevier.com/environmental-science-and-ecotechnology/

Original Research

Real-time quantification of activated sludge concentration and viscosity through deep learning of microscopic images

Hewen Li [a, 1], Yu Tao [a, 1], Tiefu Xu [b, *], Hongcheng Wang [a], Min Yang [a], Ying Chen [a], Aijie Wang [a, **]

[a] School of Eco-Environment, Harbin Institute of Technology, Shenzhen, 518055, China
[b] School of Civil Engineering, Heilongjiang University, Harbin, 150080, China

ARTICLE INFO

Article history:
Received 28 March 2024
Received in revised form
14 January 2025
Accepted 15 January 2025

Keywords:
Activated sludge
MLSS
Viscosity
Image
Deep learning

ABSTRACT

The parameters of activatedg sludge are crucial for the daily operation of wastewater treatment plants (WWTPs). In particular, mixed liquor suspended solids (MLSS) and apparent viscosity provide metrics for the biomass and rheological properties of activated sludge. Traditional methods for determining these parameters are time-consuming, require separate measurements for each index, and fail to provide real-time data for future 'smart' WWTPs. Here we show a real-time online microscopic image data analysis system that quantitatively identifies MLSS and apparent viscosity. Microscopic videos of activated sludge are captured in lab-scale sequencing batch reactors under chemical oxygen demand shock, yielding 41482 high-quality images. The Xception convolutional neural network architecture is used to establish both qualitative and quantitative correlations between these microscopic images and MLSS/apparent viscosity. The accuracies of qualitative identification for MLSS and apparent viscosity are both higher than 97%, and the quantitative correlation coefficients are 0.95 and 0.96, respectively. This quantitative correlation between microscopic images of activated sludge and its physical parameters, specifically MLSS and apparent viscosity, provides a basis for real-time online measurements of activated sludge parameters in WWTPs.

1. Introduction

Wastewater treatment plants (WWTPs) are crucial components in the development of a digital water infrastructure [1]. Industry 4.0 integrates technologies such as the Internet of Things, artificial intelligence, and big data with industrial operations, collectively supporting the development of digital systems for water networks. This integration facilitates improvements in the operational efficiency of WWTPs through real-time monitoring, predictive analysis, and automated control [2]. However, in practical operations, WWTPs face challenges associated with outdated infrastructure, limited regulatory environments, and cybersecurity vulnerabilities [3]. These challenges impede the provision of the real-time foundational data necessary for digital water networks. Mixed liquor suspended solids (MLSS) and apparent viscosity are key parameters for assessing activated sludge's physical and rheological properties in WWTPs. Any sudden variations in these parameters could signal a potential malfunction in a WWTP.

By observing and controlling MLSS, it is possible to regulate the amount of sludge reflux and excess sludge discharge in real-time. However, the conventional method of measuring MLSS involves a gravimetric approach that requires filtration, drying, and weighing, with at least 2 h required for a single reading. This makes it impractical for online monitoring purposes. Although MLSS sensing probes have been developed, they are expensive and require specific rheological conditions for installation. As a result, the data from these probes can only be used as a reference in practical applications, and WWTPs still rely on manual measurements.

The apparent viscosity in aeration tanks reflects sludge characteristics such as the floc size, density, and the biological activity of bacteria within the flocs [4,5]. The floc size distribution provides a direct representation of the apparent viscosity [6,7]. However, different methods of measuring the apparent viscosity yield

* Corresponding author.
** Corresponding author.
E-mail addresses: xutiefu@hlju.edu.cn (T. Xu), wangaijie@hit.edu.cn (A. Wang).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ese.2025.100527
2666-4984/

varying results. Therefore, rheological models have been established to deduce the apparent viscosity from MLSS [8–12], aiming to eliminate the viscosity measurement step. This approach, however, has proven unreliable [13]. Currently, viscometers and rheometers are used for batch measurements of viscosity [13], requiring approximately 3 min per sample. Although this delay is relatively short, it still poses challenges when rapid decision-making is required.

Real-time, online quantitative measurements of MLSS and apparent viscosity remain a scientific and technological challenge for WWTPs. This highlights the need for a new, rapid quantitative method for assessing these activated sludge parameters.

Microscopic inspection is a vital empirical method for assessing the state of activated sludge in WWTPs [14]. Traditional optical microscopy only yields disconnected information about biomass, and struggles to achieve real-time continuous observations. Thus, it cannot be used to generate high-frequency data, which are the key to deciphering massive, in-depth information [15]. Obtaining the necessary data requires the automatic, real-time, online capture of microscopic images of activated sludge [16]. Certain characteristics of these microscopic images have been found to correlate with sludge parameters [17,18]. In 1997, research on microscopic images of activated sludge indicated that online monitoring of the average equivalent circular diameter and average shape factor of activated sludge flocs could be used to estimate MLSS [19]. Subsequently, Leal et al. used stereo microscopes to capture segmented microscopic images of aerobic granular sludge, and extracted specific features that allowed the suspended solids content to be identified [20]. Khan et al. established models using only floc image characteristics to calculate MLSS [21], while Campbell et al. identified qualitative correlations between microscopic images of activated sludge and apparent viscosity by extracting the filamentous bacteria length and floc morphology characteristics, noting a significant impact of filament length on viscosity [22]. Research has also explored the relationship between changes in floc structure and rheological properties [23]. Currently, MLSS can be quantified based on microscopic image features, but there is a low degree of standardization because image acquisition and feature extraction and selection are subject to many human influences. The analysis of apparent viscosity in relation to microscopic images of activated sludge remains largely qualitative.

Quantitative image analysis is the process of quantifying and analyzing data within images based on image processing and analysis algorithms, allowing for a quantitative expression of images [24]. This technology primarily encompasses image acquisition, image processing, and image analysis [25]. Image analysis methods are mainly divided into radiomics and deep learning. Radiomics involves manually designing and extracting features, whereas deep learning leverages convolutional neural networks (CNNs) to extract features and build models. Deep learning minimizes subjective interference, simplifies the feature extraction process, and enables the learning of more advanced feature representations [24]. Thus, this paper explores a deep learning-based method for the microscopic quantitative identification of MLSS and apparent viscosity in activated sludge.

In this study, we induced activated sludge bulking in a sequencing batch reactor (SBR) using stepwise chemical oxygen demand (COD) shock in the influent and monitored the sludge performance daily. Using the proposed real-time online microscopic image data analysis system (ROMIDAS), we captured microscopic videos of the activated sludge, resulting in 41,482 effective images. In-depth analysis of these images enabled quantitative correlations between microscopic images of activated sludge and the MLSS and apparent viscosity, providing a foundation for real-time online measurements of these parameters in WWTPs.

2. Materials and methods

2.1. Experimental setup

This study used a cylindrical SBR made of organic glass with an effective radius of 10 cm and an effective height of 40 cm. The influent and effluent pipes connect to the SBR at heights of 35 and 20 cm, respectively (Fig. 1). Two peristaltic pumps (Langer BT100-2J, China) were used to control the inflow and outflow. The SBR bottom was equipped with a microporous aeration disc, which was controlled using an ACO-006-type aeration pump and an LZB-4WB-type gas flowmeter, ensuring that the dissolved oxygen stabilized at 6.0–7.0 mg L^{-1} during the latter stages of aeration. An upflow mixer was used to keep the sludge adequately suspended. The experimental procedure was adjusted to effectively observe the sludge bulking phenomenon in the later stages of the experiment. The reactor effluent was first collected in an effluent bucket, allowed to settle for 1 h, and then discharged through a valve located at a height of 15 cm. All processes were controlled by timers, following a predetermined schedule.

In this setup, each SBR discharged 380–400 mL of sludge daily, with a sludge retention time of 25–27 days. The specific operational parameters and modes of the reactor are detailed in Table 1. The total cycle duration of the reactor was 4 h, beginning with stirring at the same time as the inflow, thus entering the anoxic phase. The reactor operated for six cycles per day, with no idle periods.

One SBR was designated as the control (0#) and the other was used as the experimental reactor (1#). For the control, the influent COD was maintained at 600 mg L^{-1} with a C:N:P ratio of 100:5:1. In the experimental reactor, the influent COD was incrementally increased from 600 to 900, 1200, 1500, and finally 1800 mg L^{-1}. Prior to reaching the 1800 mg L^{-1} stage, the C:N:P ratio was maintained at 100:5:1; after the influent COD had reached 1800 mg L^{-1}, the C:N:P ratio was adjusted to 100:5:3 (Table 2) to enhance the possibility of bulking.

2.2. Analytical methods

The MLSS was determined using the gravimetric method. A specific volume of activated sludge suspension was filtered through a quantitative filter paper, dried at 105 °C, and then weighed. Apparent viscosity was measured using a digital rotary viscometer (NDJ-8S, Shanghai Fangrui, China), with measurements taken using the number 0 rotor at 60 rpm for 1 min before reading the values. The sludge-settling velocity 5 or 30 min (SV_5 or SV_{30}) denotes the volume percentage occupied by sludge after the mixed liquor from

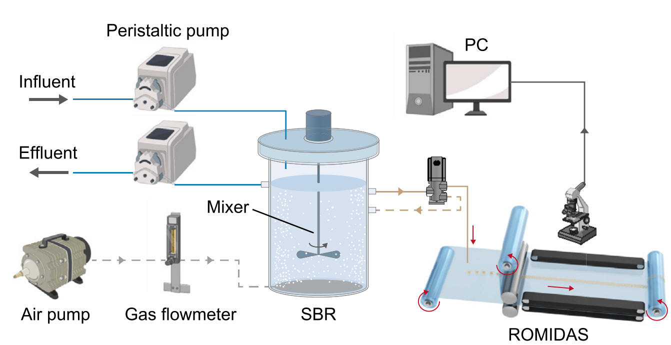

Fig. 1. Diagram of the experimental system setup. The real-time online microscopic image data analysis system (ROMIDAS) continuously captures video images of activated sludge, which is transported from a sequencing batch reactor (SBR). PC refers to a personal computer.

Table 1
SBR parameters.

Parameters	Index	Period	Time (h)
Sludge age	25–27 days	Inflow	0.25
Effective volume	10 L	Hypoxia	1
Volume exchange rate	40%	Aerobic	2
Dissolved oxygen	6.0–7.0 mg L^{-1}	Precipitate	0.75
Temperature	20–25 °C	Effluent	0.25

Note: Hypoxia: This phase involves low oxygen levels, promoting denitrification to remove nitrogen from the wastewater; Aerobic: During this phase, the reactor is aerated to support aerobic bacteria in breaking down organic matter and oxidizing ammonia to nitrates; Precipitate: In this stage, the mixed liquor is allowed to settle, separating solids from the clear supernatant, which can then be decanted.

the aeration tank was left to settle in a graduated cylinder for 5 or 30 min. The sludge volume index was determined using the ratio method (SV_{30}/MLSS), while the specific oxygen uptake rate was measured using the dissolved oxygen meter method [26]. The total COD, soluble COD, ammonium nitrogen (NH_4^+-N), and orthophosphate (PO_4^{3-}-P) of the influent and effluent were measured using the previously mentioned methodology [27].

2.3. ROMIDAS

ROMIDAS consists of a sample reflux system, a pressing system, a motion system, and a microscopic image acquisition and analysis system. The sample reflux system includes an inlet tube, a stepper motor-driven single-channel peristaltic pump, a three-channel mini electromagnetic valve, and a reflux tube. The peristaltic pump has a flow rate of 400 mL min^{-1}, ensuring that the sludge suspension is transferred from the reactor to the pressing component in less than 3 s in a uniformly mixed state. The electromagnetic valve dispenses 28 μL per drop, with an interval of 300 s between drops.

The pressing system comprises upper and lower transparent pressing conveyor belts (TCPBs) and rubber rollers, producing a TCPB with the activated sludge sample. The motion system comprises the produced TPCB, synchronous wheels, rubber belts, and a stepper motor. The stepper motor operates with a speed of 25 r min^{-1}, driving the drive wheel to rotate at an extremely low speed, thus pulling the TPCB forward at a speed of 1.8 cm min^{-1}. This ensures that the field of view inside the microscope moves forward at an appropriate speed to maintain image clarity.

The microscopic image acquisition and analysis system comprises fixed rollers, a microscope, and an electronic eyepiece. The fixed roller secures the TPCB on a focal plane, ensuring the microscope can obtain clear microscopic images of the activated sludge in real-time. The microscope has an actual magnification of 100 × and uses a halogen lamp as a light source. The electronic eyepiece has a resolution of 1824 × 1216. As the intensity of the halogen lamp changes during use, the exposure time for image capture can be manually adjusted from 20 to 40 ms, depending on the light intensity of the day. The white balance is set with fixed parameters (red: −71; green: −47; blue: 0) to achieve the optimal image capture effect.

2.4. Dataset

2.4.1. Total dataset

During the experiment, a total of 3.19 TB of microscopic videos of activated sludge was acquired from the experimental and control groups, with a frame rate of 30 fps and a resolution of 1824 × 1216 pixels. A timed frame extraction algorithm (extracting one image every 50 frames) was used to create the image dataset. After manually filtering out blank and blurred frames, a standardized collection of 41,482 images was compiled. This comprehensive dataset formed the foundation of the experiment. The distribution of MLSS and apparent viscosity in this complete dataset is depicted in Fig. 2 and Table 3.

2.4.2. Classification dataset

Based on the quartiles of the total dataset, MLSS and apparent viscosity were divided into four levels to establish an image classification dataset. This approach was intended to identify the qualitative correlation between the microscopic images of activated sludge and the MLSS or apparent viscosity, as detailed in Table 3. MLSS was classified of 'less than 5 g L^{-1}', '5–7 g L^{-1}', '7–9 g L^{-1}', and 'greater than 9 g L^{-1}'. Apparent viscosity was classified as 'less than 2 cp', '2–3 cp', '3–4 cp', or 'greater than 4 cp'.

The duration of model training is directly related to data volume. To accelerate the determination of the correlation between the microscopic images and activated sludge parameters, the random.sample() function was used to randomly extract images from the total classification dataset. Two data subsets, labeled as simple and formal, were constructed (Table 4). The simple dataset was employed to explore the correlation between the microscopic images and activated sludge parameters, whereas the formal dataset was used to validate the findings from the previous step and establish a qualitative correlation between the microscopic images and activated sludge parameters.

2.4.3. Regression dataset

Considering the impact of COD step shocks on the sludge status, an image regression dataset was established based on the time points of the overall experiment for model training. Fifty images per day were extracted to form a test dataset; the remaining 36,682

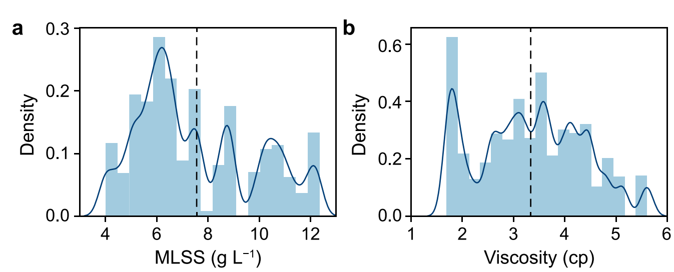

Fig. 2. Distribution of the total dataset. **a**, Mixed liquor suspended solids (MLSS) ranges from 3.999 to 12.353 g L^{-1}; **b**, Viscosity ranges from 1.70 to 5.60 cp. The line represents the estimate of the probability density function, and the bars represent the distribution density of each data point.

Table 2
Actual influent parameters.

Parameters	COD 600 (Days 1–31)	COD 900 (Days 32–35)	COD 1200 (Days 36–40)	COD 1500 (Days 41–43)	COD 1800 (Days 44–70)
COD (mg L^{-1})	487–899	799–983	1142–1178	1468–1574	1729–1937
NH_4^+-N (mg L^{-1})	29–48	36–41	53–67	75–106	102–133
PO_4^{3-}-P (mg L^{-1})	5–9	6–10	11–14	14–26	55–63

Table 3
Distribution of total dataset.

Parameters	Min	25%	50%	75%	Max
MLSS (g L^{-1})	3.999	5.762	6.803	8.883	12.353
Apparent viscosity (cp)	1.70	2.56	3.32	4.15	5.60

images were divided into three groups according to the stage of the experimental process and used to create training datasets. During training, the ratio of training to validation images was maintained at 9:1, as detailed in Table 5. The test set was divided into Test-R-1, Test-R-2, and Test-R-3, corresponding to the same experimental stages, and the images were categorized as either 0#SBR or 1#SBR for identification purposes (Table 6).

2.5. Modeling

2.5.1. Classification and regression model

Since the advent of deep learning, various CNN architectures have emerged. From the classic VGG16, the field has progressed to developing lightweight neural networks such as MobileNet V1, introduced residual blocks in ResNet50, employed depthwise separable convolutions in Xception, and incorporated dense layers in DenseNet121. This study explored the performance of eight deep learning architectures: VGG16, MobileNet V1, Inception V3, ResNet50, ResNet101, InceptionResNet V2, Xception, and DenseNet121 [28].

The classification model determines the qualitative correlation between images and parameters, while the regression model establishes the quantitative correlation. The classification model uses image features extracted through CNN architectures. These features are then passed through a GlobalAveragePooling2D module to flatten the data into a one-dimensional format. Subsequently, a dense layer with four units and the Softmax activation function is added to achieve multi-classification. For the regression model, a terminal dense layer with one unit and the ReLU activation function is used to fulfill the objective of regression. The model architecture is shown in Fig. 3.

2.5.2. Model training

Considering computational limitations, the images were batch-normalized from their original dimensions of 1824 × 1216 to 228 × 152 for model training, with the batch size set to 50. The training methods were divided into transfer learning and traditional training.

Transfer learning uses pre-trained weights from past datasets as the initial weights and combines this with retraining on the current dataset to reduce the overall training duration. In this study, the pre-trained weights of the eight CNN architectures specified in section 2.5.1 were directly employed; these were pre-trained on ImageNet. Transfer learning was employed to train models for each architecture. During the training of each epoch, the learning rate decayed according to:

$$\text{Learning rate} = \text{Initial learning rate} \times e^{-0.1\times \text{epoch}} \quad (1)$$

Traditional learning was also employed, with no pre-set initial weights in the CNNs. This was because most images in the ImageNet collection are of natural scenery, people, animals, food, and movie posters, which differ significantly from sludge microscopic images. The learning rate decay function in equation (1) was again employed during the training process.

Table 4
Classification dataset.

Parameters	Dataset-C-Simple	Dataset-C-Formal
Number per level	1100	3000
Number per index	4400	12000
Train:Val:Test	9:1:1	4:1:1

Table 5
Regression training dataset.

Name	Train-R-1	Train-R-2	Train-R-3
Number of images	10498	8908	17276
Stage	COD 600	COD 600–COD 1500	COD 1800
MLSS (g L^{-1})	5.743–8.883	4.522–8.623	3.999–12.353
Viscosity (cp)	2.65–5.60	1.82–3.16	1.70–4.86

Table 6
Regression test dataset.

Reactor	Test Name	MLSS (g L^{-1})	Viscosity (cp)
0#SBR	Test-R-0-1	5.743–8.570	2.65–4.86
	Test-R-0-2	4.522–6.533	1.82–2.62
	Test-R-0-3	3.999–6.238	1.70–2.65
1#SBR	Test-R-1-1	5.843–8.883	3.13–5.60
	Test-R-1-2	5.036–8.623	2.60–3.16
	Test-R-1-3	8.369–12.353	3.02–4.86

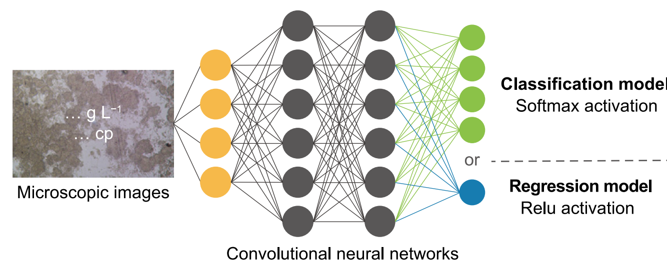

Fig. 3. Classification and regression model.

2.5.3. Evaluation metrics

Model training attempts to minimize the loss function and identify the optimal model parameters. In this study, the classification model employed the categorical_crossentropy as the loss function, using accuracy as the performance metric. The results are presented using a confusion matrix.

For the regression model, the outputs are continuous and likely to deviate from the true values. Thus, accuracy is not a suitable metric. Instead, the Huber loss and mean absolute error (*MAE*) were used to evaluate the training process. In this paper, the identification results are visualized using violin plots, and the coefficient of determination (R^2) is used to assess the fit between the median of the identification results and the true values.

3. Results and discussion

3.1. MLSS

3.1.1. Qualitative identification

The MLSS-C model was constructed using Xception, with the architecture detailed in section 2.5.1. The initial weights were set to those obtained using ImageNet, and the initial learning rate was set to 0.1%. The training process and identification results for MLSS-C based on Dataset-C-Simple and Dataset-C-Formal are depicted in Fig. 4. MLSS-C trained on Dataset-C-Simple achieved an identification accuracy of 99.50% on the Dataset-C-Simple test set (Fig. 4b),

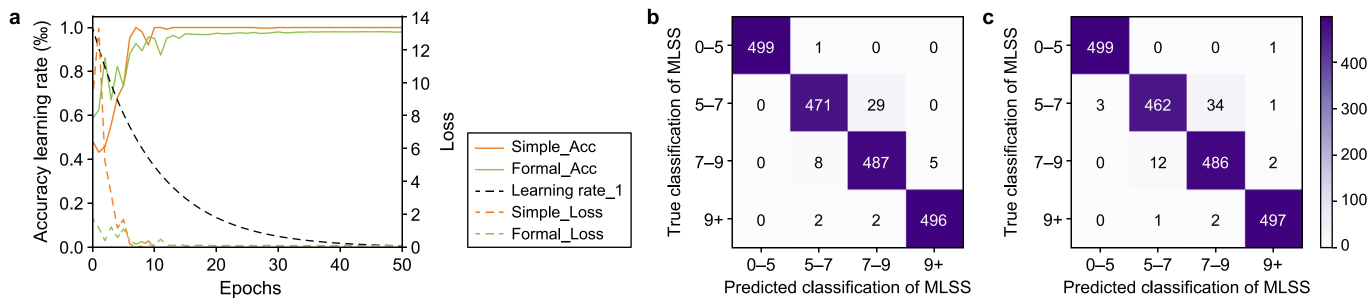

Fig. 4. Training and testing of MLSS-C with Dataset-C-Simple/Formal. **a**, Training of MLSS-C; **b**, Testing of MLSS-C for Dataset-C-Simple; **c**, Testing of MLSS-C for Dataset-C-Formal. Accuracy of 'less than 5 g L^{-1}', '5–7 g L^{-1}', '7–9 g L^{-1}', and 'greater than 9 g L^{-1}' is 99.9%, 92.4%, 87.2%, and 99.4%, respectively; misrecognition focuses on the '5–7 g L^{-1}' and '7–9 g L^{-1}' groups. MLSS: mixed liquor suspended solids. Acc: accuracy.

Table 7
Training parameters for MLSS-R.

Parameters	MLSS-R-1	MLSS-R-2	MLSS-R-3
Applied Dataset	Train-R-1	Train-R-1 Train-R-2	Train-R-1 Train-R-2 Train-R-3
Initial weight	ImageNet	MLSS-R-1	MLSS-R-2
Initial learning rate (‰)	1	0.3	0.05

preliminarily confirming a qualitative correlation between microscopic images of activated sludge and MLSS. MLSS-C trained on Dataset-C-Formal achieved an identification accuracy of 97.2% on the Dataset-C-Formal test set, further establishing this qualitative correlation. Model selection and parameter adjustments are detailed in the Supplementary Materials.

3.1.2. Quantitative identification

An image regression model was constructed using Xception; the specific architecture is detailed in section 2.5.1. MLSS-R-1, MLSS-R-2, and MLSS-R-3 models were trained using the Train-R dataset (Table 7).

The pre-trained weights from ImageNet were used as the initial weights in the training of MLSS-R-1 based on the Train-R-1 dataset and an initial learning rate of 0.1%. Subsequently, the training datasets Train-R-2 and Train-R-3 were incrementally added, and the weights from models MLSS-R-1 and MLSS-R-2 were used as the initial weights for each subsequent training stage to reduce the training duration. MLSS-R-2 and MLSS-R-3 were then trained under initial learning rates of 0.3‰ and 0.05‰, respectively, with the decay process given by equation (1). The model training process is illustrated in Fig. 5.

MLSS-R-1, MLSS-R-2, and MLSS-R-3 were used to recognize the test sets Test-R-0 and Test-R-1. The identification results of each model are presented in Fig. 6. In the identification of Test-R-0-1 and Test-R-1-1, model MLSS-R-1 achieved R^2 values of 0.98 and 0.89, respectively, between the median of the identification results and the true values. For Test-R-0-2 and Test-R-1-2, however, the R^2 values between the median of the identification values and the true values dropped to −4.83 and −0.06, respectively. In both cases, there was a relatively good front-end fit but only a similar variation trend for the back end. The discrepancy in R^2 values is hypothesized to be caused by the MLSS fluctuation range in the Test-R-1-2 images being greater within the range of Train-R-1 than in Test-R-0-2. This indicates that the model's ability to recognize unknown ranges is lower than for known ranges. Hence, there may be a quantitative correlation between microscopic images of activated sludge and MLSS, but the richness of the dataset limits the model's performance.

After progressively enriching the training dataset, models MLSS-R-2 and MLSS-R-3 were developed. MLSS-R-3 achieved R^2 values of 0.95 and 0.97 for Test-R-0 and Test-R-1, respectively, indicating a strong quantitative correlation between microscopic images of activated sludge and MLSS. Thus, for a fixed reactor or WWTP, the theoretical basis for quantitative identification of MLSS based on video segments has been established.

3.2. Apparent viscosity

3.2.1. Qualitative identification

The Viscosity-C model was constructed in the same way as MLSS-R. The initial weights were set to those obtained from ImageNet, and the initial learning rate was set to 0.1% (Fig. 7a). When trained on Dataset-C-Simple, Viscosity-C achieved an

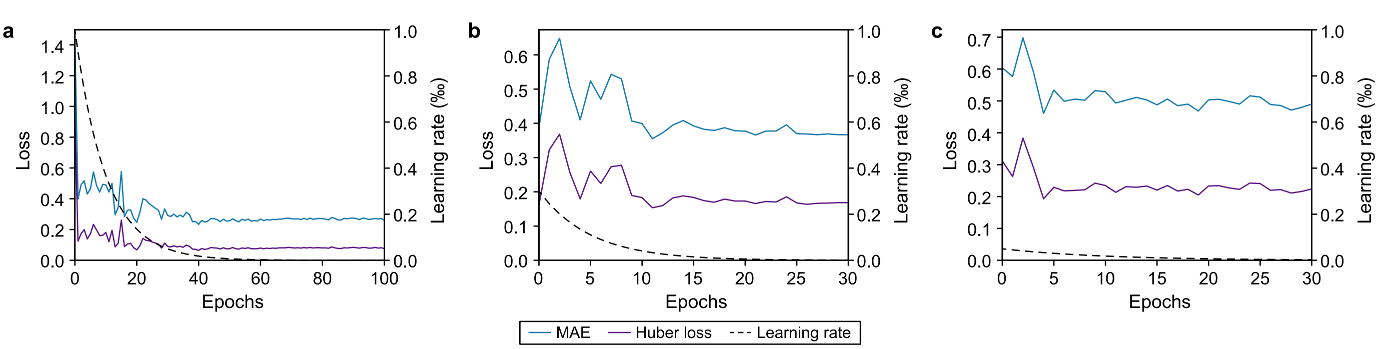

Fig. 5. Training of MLSS-R. **a**, Training of MLSS-R-1, requiring 50 epochs to stabilize; **b**, Training of MLSS-R-2, requiring 15 epochs to stabilize; **c**, Training of MLSS-R-3, requiring 10 epochs to stabilize. MAE: mean absolute error.

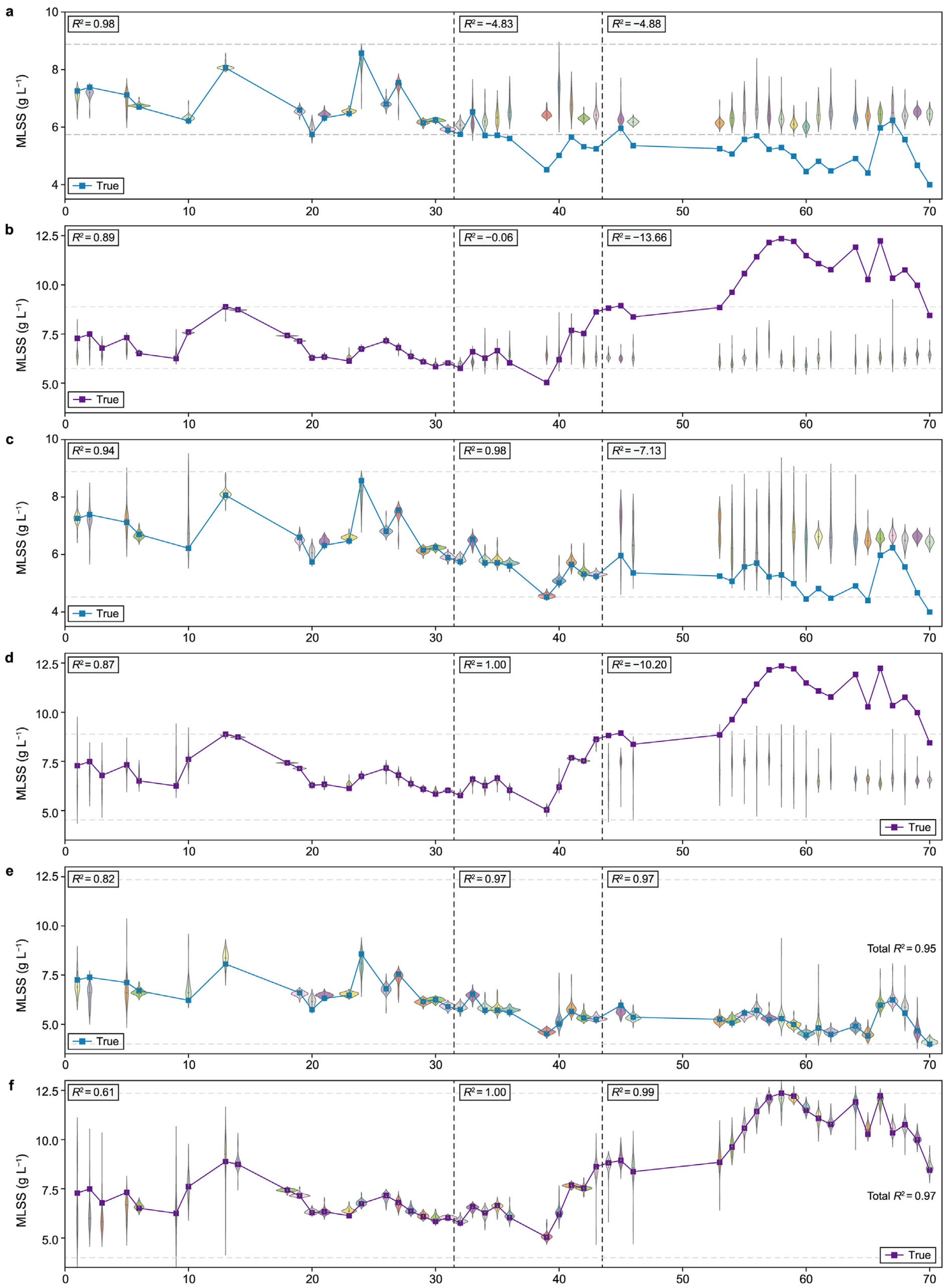

Fig. 6. Identification results of MLSS-R. **a**, Identification results of MLSS-R-1 for Test-R-0; **b**, Identification results of MLSS-R-1 for Test-R-1; **c**, Identification results of MLSS-R-2 for Test-R-0; **d**, Identification results of MLSS-R-2 for Test-R-1; **e**, Identification results of MLSS-R-3 for Test-R-0, total $R^2 = 0.95$; **f**, Identification results of MLSS-R-3 for Test-R-1, total $R^2 = 0.97$. As the dataset becomes richer, the performance of MLSS-R-2 and MLSS-R-3 gradually improves compared with MLSS-R-1. The blue and purple lines indicate a designation of 'true' for 0#SBR and 1#SBR, respectively. The single violin plot represents the test results of 50 samples from a single day. The model-building and testing process utilized three distinct periods. As only the final two sub-figures employed the complete period, they are the sole instances for which the total R^2 was computed.

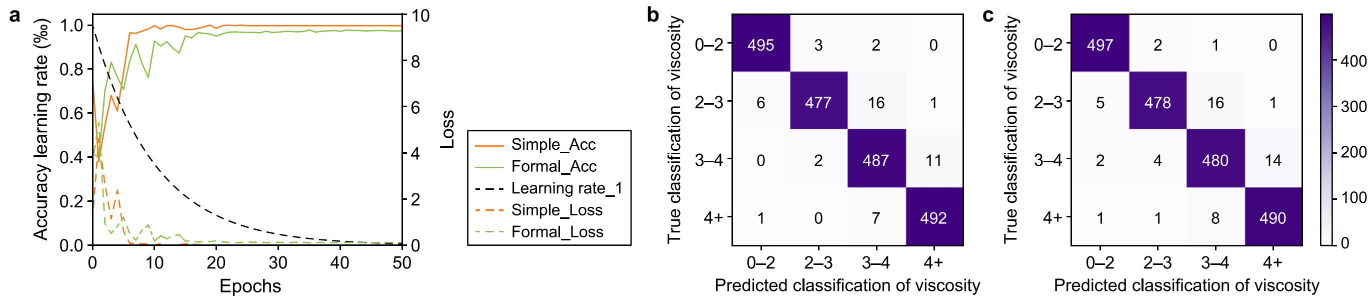

Fig. 7. Training process and identification results for the apparent viscosity-based Dataset-C-Simple and Dataset-C-Formal. **a**, Training of Viscosity-C; **b**, Testing of Viscosity-C for Dataset-C-Simple; **c**, Testing of Viscosity-C for Dataset-C-Formal. The accuracy of 'less than 2 cp', '2—3 cp', '3—4 cp', and 'greater than 4 cp' is 99.4%, 95.6%, 96%, and 98%, respectively. Acc: accuracy.

Table 8
Training parameters for Viscosity-R.

Parameters	Viscosity-R-1	Viscosity-R-2	Viscosity-R-3
Applied Dataset	Train-R-1	Train-R-1 Train-R-2	Train-R-1 Train-R-2 Train-R-3
Initial weight	ImageNet	Viscosity-R-1	Viscosity-R-2
Initial learning rate (‰)	1	0.5	0.1

identification accuracy of 99.25% with the Dataset-C-Simple test set, preliminarily confirming a qualitative correlation between microscopic images of activated sludge and apparent viscosity (Fig. 7b). Viscosity-C trained on Dataset-C-Formal achieved an identification accuracy of 97.2% with the Dataset-C-Formal test set, further establishing the qualitative correlation between microscopic images of activated sludge and apparent viscosity.

3.2.2. Quantitative identification

Viscosity-R was constructed in the same way as MLSS-R. The Train-R-1, Train-R-2, and Train-R-3 datasets were successively added, and the weights from ImageNet, Viscosity-R-1, and Viscosity-R-2 were used as initial weights for each training stage (Table 8). Initial learning rates of 1‰, 0.5‰, and 0.1‰ were used to train Viscosity-R-1, Viscosity-R-2, and Viscosity-R-3, respectively (Fig. 8). The specific training process and the model identification results are illustrated in Figs. 8 and 9.

Similar to the performance of MLSS-R on the test set scenarios, Viscosity-R produced a higher degree of fit when the actual values were within the training set parameter range (Fig. 9). This is particularly evident in the identification results of Viscosity-R-1, where the R^2 values were −12.48 for Test-R-0-2 and −1.18 for Test-R-1-2. In the MLSS case, the parameter fluctuations in the test set were entirely outside the training set range, resulting in no correlation between the identification and actual values. In contrast, Test-R-1-2 contained parameter fluctuations (2.60—3.16 cp) that were closer to those of the training set (2.65—5.60 cp) than Test-R-0-2 (1.82—2.62 cp), leading to identification values that were better aligned with the actual values. Additionally, Viscosity-R-3 demonstrated strong performance, with R^2 values of 0.99 and 0.96 for Test-R-0 and Test-R-1, respectively, further establishing a quantitative correlation between microscopic images of activated sludge and apparent viscosity. This aligns with the hypothesis derived from the performance analysis of the initial two models.

4. Conclusion

This study has enhanced the acquisition of foundational data for WWTPs by developing the ROMIDAS system, which automates and standardizes the real-time capture of microscopic images of activated sludge. Models for both qualitative and quantitative identification of MLSS were established, achieving a qualitative identification accuracy of 97.2% and a quantitative fitting with an R^2 value of 0.95. Similarly, models for the qualitative and quantitative identification of apparent viscosity were developed, with a qualitative identification accuracy of 97.2% and a quantitative fitting with an R^2 value of 0.96. These models demonstrate the strong quantitative correlation between microscopic images of activated sludge and MLSS or apparent viscosity, providing a foundation for real-time online measurement of these parameters in WWTPs.

The data used for model training were exclusively derived from microscopic videos of activated sludge suspension during the normal operation of an SBR as captured by ROMIDAS. As a result, the models struggled to identify the MLSS and apparent viscosity of

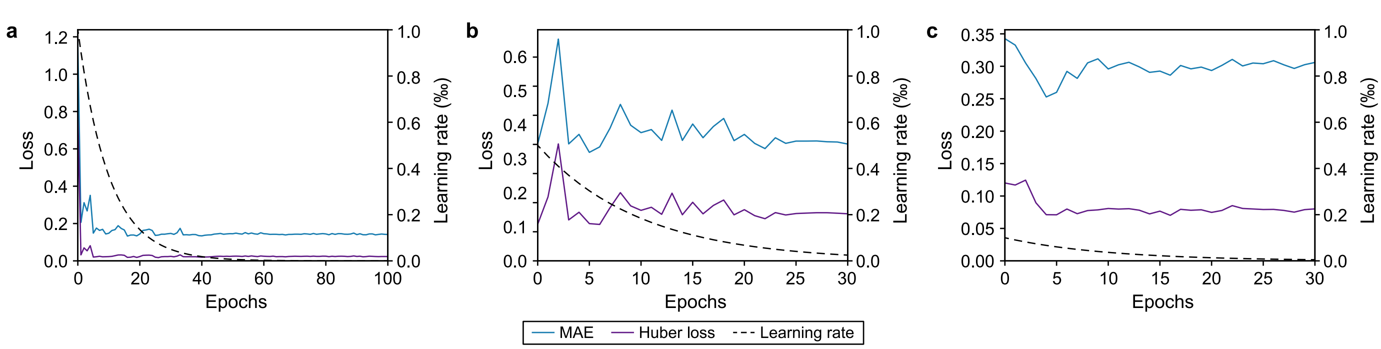

Fig. 8. Training of Viscosity-R. **a**, Training of Viscosity-R-1, requiring 40 epochs to stabilize; **b**, Training of Viscosity-R-2, requiring 25 epochs to stabilize; **c**, Training of Viscosity-R-3, requiring 10 epochs to iteratively stabilize. MAE: mean absolute error.

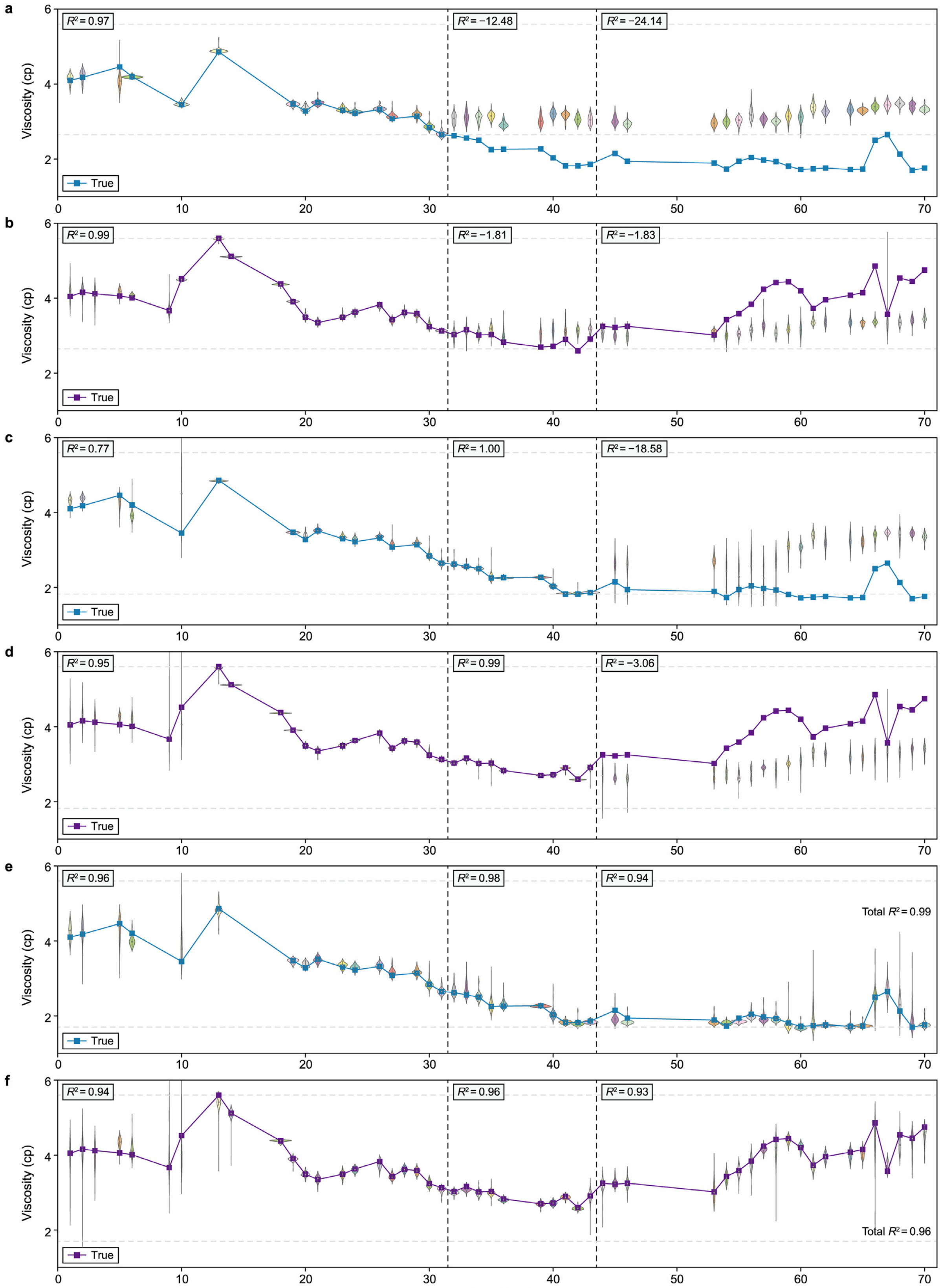

Fig. 9. Identification results of Viscosity-R. **a**, Identification results of Viscosity-R-1 for Test-R-0; **b**, Identification results of Viscosity-R-1 for Test-R-1; **c**, Identification results of Viscosity-R-2 for Test-R-0; **d**, Identification results of Viscosity-R-2 for Test-R-1; **e**, Identification results of Viscosity-R-3 for Test-R-0, total $R^2 = 0.99$; **f**, Identification results of Viscosity-R-3 for Test-R-1, total $R^2 = 0.96$. Similar viscosities can be achieved under varying reactor conditions. When the parameter range of the training set adequately encompasses that of the test set, the general trends of the test set parameters can be identified. However, accurate parameter identification requires a sufficiently diverse range of scenarios to provide comprehensive coverage. The blue and purple lines indicate a designation of 'true' for 0#SBR and 1#SBR, respectively. The single violin plot represents the test results of 50 samples from a single day. The model-building and testing process utilized three distinct periods. As only the final two sub-figures employed the complete period, they are the sole instances for which the total R^2 was computed.

artificially diluted or concentrated activated sludge, where the microscopic structure of the flocs does not change significantly. Future work will attempt to enrich the training dataset and explore the correlation between microscopic images of activated sludge and other parameters (such as the sludge volume index and specific oxygen uptake rate) under various impact factors.

CRediT authorship contribution statement

Hewen Li: Writing - Original Draft, Visualization, Investigation, Formal Analysis. **Yu Tao:** Writing - Review & Editing, Supervision, Funding Acquisition, Conceptualization. **Tiefu Xu:** Supervision, Conceptualization. **Hongcheng Wang:** Methodology. **Min Yang:** Writing - Review & Editing. **Ying Chen:** Writing - Review & Editing. **Aijie Wang:** Supervision, Visualization, Funding acquisition, Project Administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Dr. Yu Tao, the Managing Editor, and Dr. Aijie Wang, the Executive Editor of Environmental Science and Ecotechnology, were not involved in the editorial review or the decision to publish this article.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos. 52321005, 52293443, and 52230004), Shenzhen Science and Technology Program (Grant No. KQTD20190929172630447), Shenzhen Key Research Project (Grant No. GXWD20220817145054002), and Talent Recruitment Project of Guangdong (Grant No. 2021QN020106).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ese.2025.100527.

References

[1] M. Garrido-Baserba, L. Corominas, U. Cortes, D. Rosso, M. Poch, The fourth-revolution in the water sector encounters the digital revolution, Environ. Sci. Technol. 54 (8) (2020) 4698–4705.
[2] D. Savic, Digital water developments and lessons learned from automation in the car and aircraft industries, Engineering 9 (2022) 35–41.
[3] C.E. Richards, A. Tzachor, S. Avin, R. Fenner, Rewards, risks and responsible deployment of artificial intelligence in water systems, Nature Water 1 (5) (2023) 422–432.
[4] J. Perez, C. Picioreanu, M. van Loosdrecht, Modeling biofilm and floc diffusion processes based on analytical solution of reaction-diffusion equations, Water Res. 39 (7) (2005) 1311–1323.
[5] K.H. Chu, H.M. van Veldhulzen, M.C.M. van Loosdrecht, Respirometric measurement of kinetic parameters: effect of activated sludge floc size, Water Sci. Technol. 48 (8) (2003) 61–68.
[6] Y. Pechaud, S. Pageot, A. Goubet, C.D. Quintero, S. Gillot, Y. Fayolle, Size of biological flocs in activated sludge systems: influence of hydrodynamic parameters at different scales, J. Environ. Chem. Eng. 9 (4) (2021).
[7] M. Stricot, A. Filali, N. Lesage, M. Sperandio, C. Cabassud, Side-stream membrane bioreactors: influence of stress generated by hydrodynamics on floc structure, supernatant quality and fouling propensity, Water Res. 44 (7) (2010) 2113–2124.
[8] F.G. Meng, H.M. Zhang, F.L. Yang, S.T. Zhang, Y.S. Li, X.W. Zhang, Identification of activated sludge properties affecting membrane fouling in submerged membrane bioreactors, Sep. Purif. Technol. 51 (1) (2006) 95–103.
[9] A. Pollice, C. Giordano, G. Laera, D. Saturno, G. Mininni, Physical characteristics of the sludge in a complete retention membrane bioreactor, Water Res. 41 (8) (2007) 1832–1840.
[10] F. Yang, A. Bick, S. Shandalov, A. Brenner, G. Oron, Yield stress and rheological characteristics of activated sludge in an airlift membrane bioreactor, J. Membr. Sci. 334 (1–2) (2009) 83–90.
[11] G. Laera, C. Giordano, A. Pollice, D. Saturno, G. Mininni, Membrane bioreactor sludge rheology at different solid retention times, Water Res. 41 (18) (2007) 4197–4203.
[12] C. Giordano, A. Pollice, G. Laera, D. Saturno, G. Mininni, Influence of solid retention time on the rheology of MBR sludge, Water Sci. Technol. 56 (8) (2007) 151–159.
[13] N. Ratkovich, W. Horn, F.P. Helmus, S. Rosenberger, W. Naessens, I. Nopens, T.R. Bentzen, Activated sludge rheology: a critical review on data collection and modelling, Water Res. 47 (2) (2013) 463–482.
[14] D.H. Eikelboom, Process Control of Activated Sludge Plants by Microscopic Investigation, IWA publishing, 2000.
[15] R. Nathan, C.T. Monk, R. Arlinghaus, T. Adam, J. Alos, M. Assaf, H. Baktoft, C.E. Beardsworth, M.G. Bertram, A.I. Bijleveld, T. Brodin, J.L. Brooks, A. Campos-Candela, S.J. Cooke, K.O. Gjelland, P.R. Gupte, R. Harel, G. Hellstrom, F. Jeltsch, S.S. Killen, T. Klefoth, R. Langrock, R.J. Lennox, E. Lourie, J.R. Madden, Y. Orchan, I.S. Pauwels, M. Riha, M. Roeleke, U.E. Schlagel, D. Shohami, J. Signer, S. Toledo, O. Vilk, S. Westrelin, M.A. Whiteside, I. Jaric, Big-data approaches lead to an increased understanding of the ecology of animal movement, Science 375 (6582) (2022) 734.
[16] E. Koivuranta, T. Stoor, J. Hattuniemi, J. Niinimaki, On-line optical monitoring of activated sludge floc morphology, J. Water Process Eng. 5 (2015) 28–34.
[17] A. Jang, Y.H. Yoon, I.S. Kim, K.S. Kim, P.L. Bishop, Characterization and evaluation of aerobic granules in sequencing batch reactor, J. Biotechnol. 105 (1–2) (2003) 71–82.
[18] S. O'Meara, Small images, big picture: artificial intelligence to revolutionize microscopy, Science 372 (6540) (2021) 422–425.
[19] K. Grijspeerdt, W. Verstraete, Image analysis to estimate the settleability and concentration of activated sludge, Water Res. 31 (5) (1997) 1126–1134.
[20] C. Leal, A.V. del Rio, D.P. Mesquita, A.L. Amaral, P.M.L. Castro, E.C. Ferreira, Sludge volume index and suspended solids estimation of mature aerobic granular sludge by quantitative image analysis and chemometric tools, Sep. Purif. Technol. 234 (2020).
[21] M.B. Khan, H. Nisar, C.A. Ng, Generalized modeling of the sludge volume index and suspended solids to monitor activated sludge bioflocculation using image analysis, Environ. Protect. Eng. 46 (3) (2020) 17–37.
[22] K. Campbell, J. Wang, M. Daniels, Assessing activated sludge morphology and oxygen transfer performance using image analysis, Chemosphere 223 (2019) 694–703.
[23] V. Bakos, B. Gyarmati, P. Csizmadia, S. Till, L. Vachoud, P.N. Gode, G.M. Tardy, A. Szilagyi, A. Jobbagy, C. Wisniewski, Viscous and filamentous bulking in activated sludge: rheological and hydrodynamic modelling based on experimental data, Water Res. 214 (2022).
[24] A. Ibrahim, S. Primakov, M. Beuque, H.C. Woodruff, I. Halilaj, G. Wu, T. Refaee, R. Granzier, Y. Widaatalla, R. Hustinx, F.M. Mottaghy, P. Lambin, Radiomics for precision medicine: current challenges, future prospects, and the proposal of a new framework, Methods 188 (2021) 20–29.
[25] M. Oberholzer, M. Ostreicher, H. Christen, M. Bruhlmann, Methods in quantitative image analysis, Histochem. Cell Biol. 105 (5) (1996) 333–355.
[26] G.-H. Chen, W.-K. Yip, H.-K. Mo, Y. Liu, Effect of sludge fasting/feasting on growth of activated sludge cultures, Water Res. 35 (4) (2001) 1029–1037.
[27] H.-c. Wang, Y. Liu, Y.-m. Yang, Y.-k. Fang, S. Luo, H.-y. Cheng, A.-j. Wang, Element sulfur-based autotrophic denitrification constructed wetland as an efficient approach for nitrogen removal from low C/N wastewater, Water Res. 226 (2022) 119258.
[28] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.