



Prediction of sludge settleability, density and suspended solids of aerobic granular sludge in the presence of pharmaceutically active compounds by quantitative image analysis and chemometric tools

Cristiano Leal^a, Angeles Val del Río^b, Daniela P. Mesquita^{a,*}, António L. Amaral^{a,c,d}, Eugénio C. Ferreira^a

^a CEB - Centre of Biological Engineering, Universidade do Minho, Campus de Gualtar, 4710-057 Braga, Portugal

^b CRETUS Institute, Department of Chemical Engineering, Universidade de Santiago de Compostela, E-15705 Santiago de Compostela, Spain

^c Instituto Politécnico de Coimbra, ISEC, Rua Pedro Nunes, Quinta da Nora, 3030-199 Coimbra, Portugal

^d Instituto de Investigação Aplicada, Laboratório SiSus, Rua Pedro Nunes, Quinta da Nora, 3030-199 Coimbra, Portugal

ARTICLE INFO

Editor: Yang Liu

Keywords:

Granular, floccular, settleability and density prediction

Granules stability

Quantitative image analysis

Chemometric tools

Pharmaceutically active compounds

ABSTRACT

Steroid estrogens namely 17 β -estradiol (E2) and 17 α -ethinylestradiol (EE2) and antibiotics including sulfamethoxazole (SMX) are pharmaceutically active compounds (PhAC) of emerging concern due to their environmental and human health impacts even at ppb range concentrations. These compounds usually flow to wastewater treatment plants (WWTP) and are released to the aquatic systems due to inefficient removal in conventional biological systems. In this work, a sequencing batch reactor (SBR) with aerobic granular sludge (AGS) was operated in the presence of E2, EE2 and SMX. SVI₅, SVI₃₀/SVI₅ ratio, VSS, and TSS of mature AGS (in absence of PhAC), as well as in the presence of PhAC (0.221 mg L⁻¹ of E2, 0.278 mg L⁻¹ of EE2 and 0.290 mg L⁻¹ of SMX), were successfully predicted with multilinear regression (MLR) using morphological and structural parameters of floccular and granular fractions of AGS obtained from quantitative image analysis (QIA). Good prediction models were obtained for the SVI₅ (R² of 0.976), floccular VSS (R² of 0.949) and TSS (R² of 0.934), granular VSS (R² of 0.930) and TSS (R² of 0.916), SVI₃₀/SVI₅ ratio (R² of 0.917) and density (R² of 0.889). These results emphasize the usefulness of this methodology for monitoring dysfunctions in AGS in the presence of the studied PhAC.

1. Introduction

Concerns about pharmaceutically active compounds (PhAC), including 17 β -estradiol (E2), 17 α -ethinylestradiol (EE2), and sulfamethoxazole (SMX) have been increasing in the recent years, mainly due to their environmental impact and human health effects even in trace concentrations [1]. It is recognized that these compounds could be found covering a wide range of concentrations in different environmental compartments [2]. It is also known that these compounds inflow to wastewater treatment plants (WWTP) and are released into the environment as a result of inadequate removal during treatment. In fact, E2, EE2, and SMX can be found both in domestic and industrial wastewaters [1,3], with their impact in the biological treatment aggregates being far from negligible [4,5]. These three compounds were selected for this work based on the impacts that these substances can cause on the

environment, biological wastewater treatment systems and microbial aggregates structure.

Activated sludge (AS) is one of the most common biological wastewater treatments, but it is expected to be largely replaced by aerobic granular sludge (AGS) mainly due to the improved aggregated biomass structural features. Aerobic granules are a self-immobilized type of biofilm, much more compact and denser than AS, more resistant to shock loads and toxic compounds, and with improved settling properties [6]. Furthermore, the balance between the sludge floccular and granular fractions is considered critical for AGS stability [7], with the AGS density being used for the assessment and control of AGS systems [8,9]. It is known that steroid estrogens, namely E2 and EE2, can cause an increase in the sludge volume index (SVI) and a decrease in the SVI₃₀/SVI₅ ratio in AGS systems [10]. However, other reference found in literature reported no effects on the granules' size during the treatment of EE2 in

* Corresponding author.

E-mail address: daniela@deb.uminho.pt (D.P. Mesquita).

<https://doi.org/10.1016/j.jece.2022.107136>

Received 21 September 2021; Received in revised form 30 November 2021; Accepted 1 January 2022

Available online 5 January 2022

2213-3437/© 2022 Elsevier Ltd. All rights reserved.

AGS [11]. Regarding the treatment of antibiotics containing wastewaters, it has been referred that the presence of SMX and trimethoprim ($200 \mu\text{g L}^{-1}$ each) allowed maintaining good settling properties for the AGS, although the structural integrity of the aerobic granules was affected [12].

The use of quantitative image analysis (QIA) has been growing for the AS morphological and structural assessment, including the identification of different types of bulking [13]. QIA has already been used for anaerobic granular sludge monitoring and, more recently, in mature AGS assessment [14–16]. It is known that QIA requires proficiency in microscopy techniques, in predicting the AGS settleability, suspended solids and density by the use of structural and morphological parameters. In fact, the standard methods for these parameters' determination could be considered more time consuming when compared to the presented methodologies. The amount of data provided by QIA commonly requires the use of chemometric tools to organize the gathered data. Accordingly, QIA coupled to chemometric tools could be considered a complementary monitoring tool in wastewater treatment systems allowing to establish correlations between the aggregates' structure and the operational parameters [17,18]. In this regard, the use of principal component analysis (PCA), partial least squares (PLS), neural networks (NN), and multilinear regression (MLR), among others, are commonly reported in literature [19]. In this respect, MLR can provide for the most immediately perceptible information regarding the relationships between the dependent and independent variables when compared to the other techniques [20].

NN has been used to predict performance in AGS reactors treating industrial flowback water [21], while MLR predicted the biochemical oxygen demand (BOD) removal efficiency in different biological wastewater treatment systems [20,22] from the fruit and vegetable processing industry [23]. In constructed wetlands, MLR was previously used for the prediction of trace organic compounds' removal efficiency [24,25]. Moreover, MLR was also used to predict the suspended solids (SS) and SVI of mature and stable granules [16]. However, the prediction of such operational parameters in AGS in the presence of E2, EE2, and SMX remained unexplored so far.

Considering the above, the main objective of this research work was to predict the AGS density, stability, and settling ability, as well as SS in floccular and granular fractions, in the presence of environmentally relevant concentrations of PhAC, using MLR to analyze both AGS fractions morphological and structural data, obtained by QIA. In this work, innovative aspects include the use of QIA based morphological and structural parameters, for both the floccular and granular fractions of the sludge, for the prediction of AGS density, stability, and settling ability in the presence of PhAC. To the author's knowledge, this is also the first work presenting the prediction of the biomass density in the presence of E2, EE2 and SMX. Additionally, the successfulness of the proposed methodology, combining QIA and chemometric tools, in the present case study, opens the door for extending its usefulness in analog studies for other systems.

2. Materials and methods

2.1. Experimental survey: reactor set-up

In this work, a 5 L lab scale SBR inoculated with mature granules from a full-scale WWTP, was operated at room temperature ($18\text{--}23\text{ }^\circ\text{C}$), with AGS for the treatment of a synthetic effluent [26] containing 5.168 g L^{-1} of $\text{C}_2\text{H}_3\text{O}_2\text{Na}\cdot 3\text{H}_2\text{O}$ (as carbon source), 0.887 g L^{-1} of $\text{MgSO}_4\cdot 7\text{H}_2\text{O}$, 0.35 g L^{-1} of KCl, 0.596 g L^{-1} of Na_2HPO_4 , 0.286 g L^{-1} of KH_2PO_4 (as P sources) and 1.894 g L^{-1} of NH_4Cl . A volume of 10 mL L^{-1} of a trace elements solution was added containing 1.5 g L^{-1} of $\text{FeCl}_3\cdot 6\text{H}_2\text{O}$, 0.15 g L^{-1} of H_3BO_3 , 0.03 g L^{-1} of $\text{CuSO}_4\cdot 5\text{H}_2\text{O}$, 0.18 g L^{-1} of KI, 0.12 g L^{-1} of $\text{MnCl}_2\cdot 4\text{H}_2\text{O}$, 0.06 g L^{-1} of $\text{Na}_2\text{MoO}_4\cdot 0.2 \text{ H}_2\text{O}$; 0.12 g L^{-1} of $\text{ZnSO}_4\cdot 7\text{H}_2\text{O}$ and 0.15 g L^{-1} of $\text{CoCl}_2\cdot 6\text{H}_2\text{O}$. The composition of the synthetic wastewater was chosen to simulate a high strength

(industrial) wastewater [27]. The SBR operational cycles lasted for 6 h, encompassing 120 min of feeding, 232 min of aeration, 3 min of settling and 5 min of effluent withdrawal, and a hydraulic retention time of 12 h. The sludge retention time (SRT) was not controlled during the experiments. Air was supplied at 7.50 L min^{-1} resulting in a superficial air velocity above 1.8 cm s^{-1} during the monitoring period.

Before the experiments with PhAC, the reactor was operated in absence of PhAC acting as a control (CONT). The experiments with PhAC were conducted operating the SBR in the presence of environmentally relevant concentrations of E2, EE2, and SMX, namely 0.221 mg L^{-1} , 0.278 mg L^{-1} and 0.290 mg L^{-1} , in average, respectively [28], added every week to the system during the experiments. Care was taken to allow the AGS attaining somewhat similar characteristics, mainly in terms of the overall removal performance and granular biomass (fraction and large granules size), in the beginning of the monitoring period for all experiments (CONT, E2, EE2, and SMX). For that purpose, the AGS was allowed to recover to the initial steady-state conditions, between experiments and for a period of roughly one month without any PhAC addition.

2.2. Analytical methods

Chemical oxygen demand (COD), ammonium nitrogen ($\text{NH}_4^+\text{-N}$), nitrite ($\text{NO}_2^-\text{-N}$), nitrate ($\text{NO}_3^-\text{-N}$), and phosphorus (P) concentrations were determined by Hach Lange cell tests. The total inorganic nitrogen (TIN) concentration in the influent was assumed to be equal to the ammonium concentration, while the TIN in the effluent was determined as the sum of $\text{NH}_4^+\text{-N}$, $\text{NO}_2^-\text{-N}$ and $\text{NO}_3^-\text{-N}$ concentrations. COD, nitrogen species and P concentrations were determined both in the influent and in the effluent of the SBR.

The total and volatile suspended solids for the overall ($\text{TSS}_{\text{total}}$ and $\text{VSS}_{\text{total}}$), floccular (TSS_{floc} , VSS_{floc}), and granular (TSS_{gran} , VSS_{gran}) biomass fractions, as well as the sludge volume index at 5 min (SVI_5) and at 30 min (SVI_{30}), were determined according to standard methods [29]. The aggregates density was determined with Blue dextran by the method described by [30]. E2, EE2, and SMX concentrations were assayed using a Shimadzu Corporation apparatus (Tokyo, Japan) consisting of an UHPLC (Nexera) with a multi-channel pump (LC-30 CE), an autosampler (SIL-30AC), an oven (CTO-20AC), a diode array detector (M-20A), and a system controller (CBM-20A) with built-in software (LabSolutions), according to [31,32].

2.3. Sludge sampling methodology

A volume of 600 mL of sludge was collected at mid-point depth in the reactor, to obtain homogeneous and representative biomass samples, at the beginning of the aeration phase. This volume sample was kept under moderate agitation conditions to avoid settling and promote the mixture between the solid and liquid phases to collect the corresponding aliquots of 35 mL for the CONT and E2 experiment and 10 mL for the EE2 and SMX experiments. After this aliquots collection the remaining amount of sludge was returned to the AGS-SBR. A $500 \mu\text{m}$ sieve was next employed to separate the granular and suspended (floccular) fractions of these aliquots according to [16]. For this purpose, care was taken to avoid the formation of a "filtration cake", as much as possible, so that the smaller aggregates passage though the sieve would not be impaired by larger ones. Furthermore, the retained granules were carefully picked up by rising distilled water [33].

Images of the entire set of granules present in the sieved volumes were acquired with a Petri dish in an Olympus SZ 40 (Olympus, Shinjuku, Japan) and in a Leica S8AP0 (Leica, Wetzlar, Germany) stereomicroscope at a total magnification of $15\times$ for the CONT, E2 and EE2 experiments and of $16\times$ for the SMX experiment. Regarding the floccular fraction, images were acquired in an Olympus BX51 (Olympus, Shinjuku, Japan) and in a Nikon Eclipse Ci-L (Nikon Corporation, Tokyo, Japan) in bright field at a total magnification of $40\times$ for the CONT and

the E2 experiments and $100 \times$ for the EE2 and SMX experiments. A total of 150 images (50 images per 10 μL of sample by triplicate) were acquired according to [13].

2.4. QIA and chemometric analysis

Previously in house developed image processing programs were used to determine the QIA based data, both for the floccular and granular biomass [34]. The floccular (F) and granular (G) fractions data was further divided into 3 classes considering its equivalent diameter (Deq): F1 (<25 μm), F2 (25–250 μm), and F3 (>250 μm); and G1 (<0.25 mm), G2 (0.25–2.5 mm), and G3 (>2.5 mm). It is worth mentioning though, that the floccular biomass fraction passing through the 500 μm sieve overwhelmingly (>99%) presented aggregates below 250 μm in diameter [33]. On the other hand, with respect to the granular fraction, in some cases a significant percentage (around 15.5%) consisted on aggregates with an equivalent diameter below 250 μm trapped in the sieve [33]. This was mainly due to two reasons: i) smaller aggregates impaired in their passage by larger ones; and ii) granules eccentricity meaning that the larger dimension (length) of these granules could impair their passage through the sieve.

The F1-F3 fractions of sludge have been previously used for activated sludge classifications purposes [13], while the G1-G3 size classifications were previously used for anaerobic granular sludge systems' monitoring purposes [16,28]. Detailed information on the employed morphological and structural parameters for the sludge floccular and granular fractions can be found elsewhere [33,34].

MLR is a multilinear technique suited to determine the most important predictors for a given response variable, among a number of potential explanatory variables [35]. In the current study, MLR was employed to obtain linear model fits for TSS and VSS (both for the floccular and granular biomass), SVI_5 , $\text{SVI}_{30}/\text{SVI}_5$ ratio and biomass density (Y dataset) from the QIA based data (X dataset). Due to the inexistence of density data regarding the mature AGS experiment, the parameters prediction was obtained both using the mature AGS dataset (excluding density) and the biomass density dataset (excluding the mature AGS). Thus, it should be stressed, that the mature AGS dataset (MAGS) includes the CONT experiment data and the PhAC experiments data excluding the density parameter. On the other hand, the biomass density dataset (BD) only includes the PhAC experiments data and includes all of the studied parameters (and, thus, the density parameter as well). In all cases two thirds of the collected data were used for training purposes and one third for validation purposes.

Raw data was fed into the multilinear regression (MLR), meaning that the absolute values of each studied dependent (Y) and independent (X) variables varied significantly. Therefore, the value of the X variables coefficients (and, hence, their exponents) differed accordingly, and are not directly correlated with their importance, but with their absolute scale. With respect to the most appropriate model Y variables selection, an iterative process was employed for the prediction of each studied X variable, with a stepwise MLR running on all remaining parameters at each step. This technique allowed for the sequentially determination of each statistically significant ($p < 0.01$) parameter at each step, until no further significant variable could be added.

The obtained prediction abilities were evaluated considering the ensemble (training and validation) data, by the coefficient of determination (R^2), p-value (p), root mean square error (RMSE), and residual predictive deviation (RPD, reflecting the ratio between the population standard deviation – SD, and the prediction standard error of cross validation – SECV). An RPD value above 3 is recommended for screening purposes [36].

All QIA routines and MLR analyses were performed in Matlab 7.3 (The Mathworks, Inc. Natick, USA).

3. Results and discussion

3.1. Overall biomass structure evolution

The overall biomass structure regarding the floccular and granular sludge fractions, assessed by QIA, was found to be crucial for the prediction of the sludge settleability, density, and SS. Considering the floccular fraction, the evolution of the main overall parameters is presented in Fig. 1.

The overall flocs contents (Fig. 1(a)) ranged between 298 and 17,253 $\text{mm}^2 \text{mL}^{-1}$ during the monitoring period, presenting the largest values during the E2 experiment, and showing a similar trend to the floccular fraction TSS and VSS evolution (presented in the next section). With respect to the overall flocs morphology (in terms of robustness and convexity, see Fig. 1(b) and (d)), the E2 experiment also led (from day 82 onwards) to the most constant and regular floccular structures, in accordance with the larger TSS and VSS. On the other hand, the overall flocs size (in terms of diameter) ranged from an average 14–42 μm , with the largest values (except for day 61) found at the beginning of the CONT experiment, relating to the largest SVI_5 and $\text{SVI}_{30}/\text{SVI}_5$ values (lower sludge settling ability and stability).

In a similar way, the granular fraction structure was also assessed by QIA, with the evolution of the main overall parameters presented in Fig. 2.

The overall granules contents (Fig. 2(a)) ranged between 64 and 487 $\text{mm}^3 \text{mL}^{-1}$ during the monitoring period, presenting the largest values during the EE2 and SMX experiments, the two experiments presenting the larger granular fraction TSS and VSS (presented in the next section). The granular fraction presented a relatively stable overall biomass in terms of its morphology (assessed by the robustness and convexity presented in Fig. 2(b) and (d)), though a degrading trend is apparent throughout the SMX experiment.

Regarding the granules size (in terms of diameter in Fig. 2(c)) the overall granular biomass ranged from an average 0.77–2.69 mm during the monitoring period, with the largest values obtained at the end of the CONT and throughout the EE2 experiments. On the other hand, the E2 and SMX experiments led to a decrease on the overall granules' average diameter values.

3.2. Prediction of the AGS floccular and granular fractions contents

The prediction of the floccular and granular fractions for VSS and TSS was performed by MLR considering the biomass (filamentous, floccular, and granular) contents and structure assessed by QIA. Given that the stability of AGS systems is dependent on the balance between the suspended and granular fractions, the prediction of VSS and TSS was performed separately for both fractions. Moreover, two different models were studied in each case, including either the mature aerobic granular sludge (MAGS) or the biomass density (BD) datasets. Considering that no biomass density could be obtained for the mature biomass (CONT experiment), the use of one of these two datasets led to the exclusion of the other.

The prediction of the floccular fraction VSS, using the MAGS dataset, led to a R^2 value of 0.919 ($p < 0.01$, RMSE of 0.216 g L^{-1} (8.6% of the studied range), and RPD of 3.50), was obtained for the prediction model shown in Eq. (1), based on the overall and small flocs contents, intermediate flocs fraction, intermediate and large flocs size, and overall flocs morphology. Furthermore, the predictions obtained in the presence of PhAC were also compared with previous models for MAGS under PhAC free conditions (Leal et al., 2020b), to bring new insights regarding the influence of the studied PhAC. Thus, while in MAGS under PhAC free conditions, the most important parameters for the prediction of the floccular fraction VSS were related to the intermediate and large flocs morphology and overall flocs contents, the presence of PhAC led to the inclusion of a more class stratified size and contents parameters. These results point out how the resulting differences in the flocs classes, from

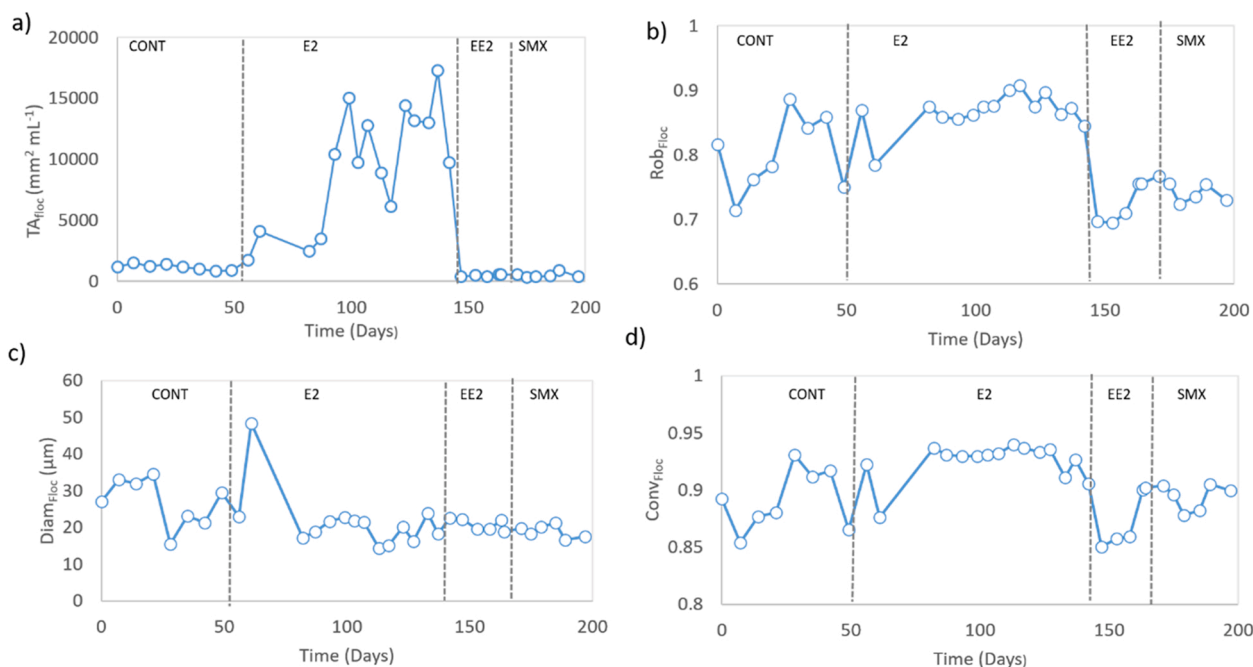


Fig. 1. Evolution of main QIA parameters regarding the overall floccular fraction: a) Total flocs area per sample volume (TA_{floc}), b) Robustness (Rob), c) Diameter (Diam), and d) Convexity (Conv). The dashed lines separate the different operational periods (CONT, E2, EE2, and SMX).

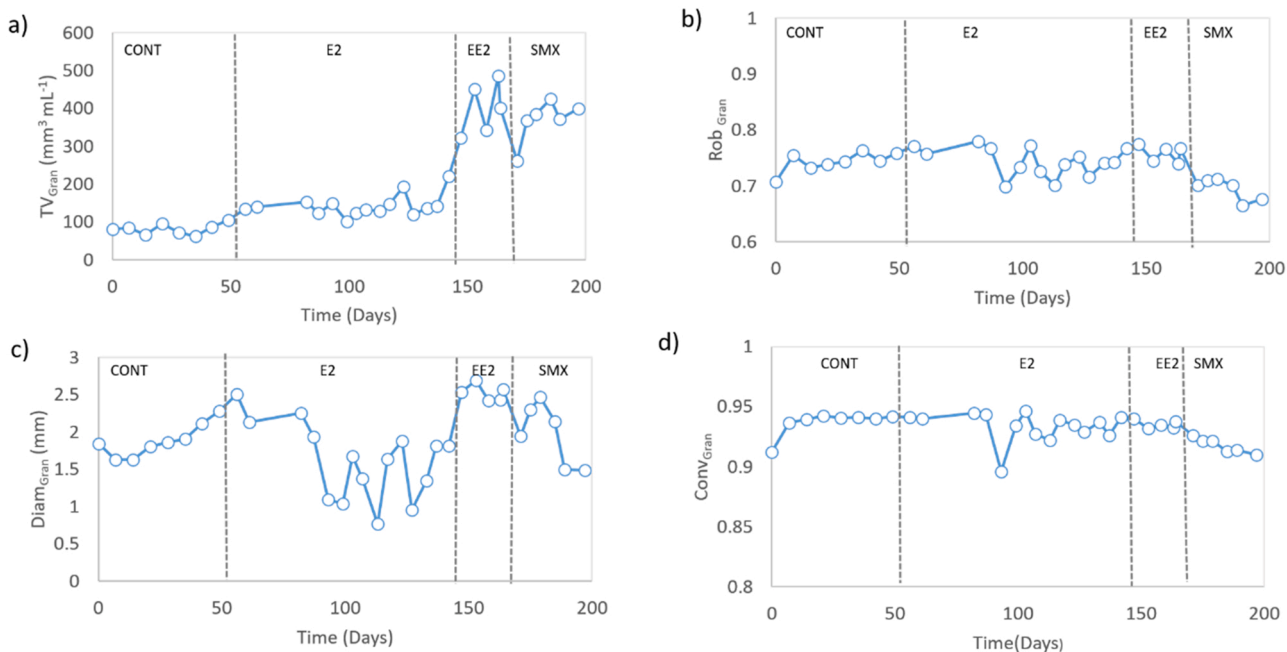


Fig. 2. Evolution of main QIA parameters regarding the overall granular fraction: a) Total granules volume per sample volume (TV_{gran}), b) Robustness (Rob), c) Diameter (Diam), and d) Convexity (Conv). The dashed lines separate the different operational periods (CONT, E2, EE2, and SMX).

the presence of PhAC, have impacted on the floccular biomass VSS. Taking also in consideration the higher R^2 (0.986), and lower RMSE value ($0.06 \text{ g L}^{-1} - 4.7\%$ of the studied range), obtained for the MAGS under PhAC free conditions, it can be inferred that the prediction of floccular VSS in the presence of PhAC becomes more challenging.

$$\begin{aligned}
 VSS_{flocMAGS} = & -3.09 + 8.35 \times 10^{-5} (TA_{floc}) + 6.40 \times 10^{-4} (\%Nb_{F2})^2 \\
 & + 6.48 (Rob_{floc}) - 8.56 \times 10^{-8} (Nb_{F1}) - 5.84 \times 10^{-8} (Area_{F2})^2 \\
 & + 1.98 \times 10^{-11} (Area_{F3})^2 + 1.34 \times 10^{-3} (Diam_{F2})^2 + 9.29 \\
 & \times 10^{-2} (Diam_{F2})
 \end{aligned}
 \tag{1}$$

On the other hand, when using the BD dataset, shown in Eq. (2), the most important parameters were found to be related with the overall and large flocs contents, large flocs fraction, overall, intermediate and large

flocs size, and overall flocs morphology. For this prediction a R^2 value of 0.940 ($p < 0.01$, RMSE of 0.205 g L^{-1} (8.2% of the studied range), and RPD of 3.93) was obtained.

$$\begin{aligned} VSS_{floc\ BD} = & 0.578 + 1.02 \times 10^{-4} (TA_{floc}) + 2.84 \times 10^{-7} (Area_{floc})^2 \\ & - 6.10 \times 10^{-9} (Area_{F2})^2 - 2.11 \times 10^{-10} (Nb_{F3}) \\ & - 1.16 \times 10^{-3} (Conv_{floc})^2 \end{aligned} \quad (2)$$

Regarding the prediction of the floccular fraction TSS, using the MAGS dataset, a R^2 value of 0.908 ($p < 0.01$, RMSE of 0.243 g L^{-1} (9.0% of the studied range) and RPD of 3.28) was obtained for the prediction model presented in Eq. (3), based on the overall and small flocs contents, small, intermediate and large flocs size, and overall and large flocs morphology. On a similar way to the VSS, the flocs TSS have also been assessed previously for MAGS in PhAC free conditions [16]. Again, some changes in the model could be inferred by the presence of PhAC. While, for MAGS under PhAC free conditions, the intermediate flocs contents, and large and overall flocs morphology, were found to be crucial for TSS prediction, in the presence of PhAC the most important variables again included a larger set of class stratified size and contents parameters. These findings reveal, once again, how the presence of the studied PhAC have impacted the different flocs classes and the floccular biomass TSS. Again, taking also in consideration the higher R^2 (0.971), and lower RMSE value (0.10 g L^{-1} – 7.2% of the studied range), obtained for the MAGS under PhAC free conditions, it can be inferred that the prediction of floccular TSS in the presence of PhAC also becomes more challenging.

$$\begin{aligned} TSS_{floc\ MAGS} = & -11.9 + 1.09 \times 10^{-4} (TA_{floc}) + 6.98 \times 10^{-5} (Area_{F1})^2 \\ & + 7.36 (Rob_{floc}) - 1.05 \times 10^{-3} (TA_{F1}) - 1.03 \times 10^{-3} (Area_{F2}) \\ & + 1.53 \times 10^{-1} (Diam_{F2}) + 6.08 \times 10^{-6} (Area_{F3}) + 1.57 (Ecc_{F3})^2 \end{aligned} \quad (3)$$

Again, the floccular fraction of TSS was also predicted using the BD dataset, with the most important parameters being related to the overall and large flocs contents, overall and intermediate flocs size, and overall flocs morphology. The obtained results are presented in Eq. (4) and revealed a R^2 value of 0.934 ($p < 0.01$, RMSE of 0.219 g L^{-1} (8.2% of the studied range) and RPD of 3.89).

$$\begin{aligned} TSS_{floc\ BD} = & 1.13 + 1.20 \times 10^{-4} (TA_{floc}) + 3.25 \times 10^{-7} (Area_{floc})^2 \\ & - 9.21 \times 10^{-9} (Area_{F2})^2 - 3.53 \times 10^{-10} (Nb_{F3}) - 1.61 \times 10^{-3} (Conv_{floc})^2 \end{aligned} \quad (4)$$

As expected, the floccular TSS and VSS models show a high similitude, and a direct dependency on the overall flocs' contents, size and morphology. Indeed, it could be inferred that higher floccular VSS and TSS values were (mostly) positively correlated with larger and more regular flocs. It could also be inferred that the use of the biomass density data (and exclusion of the MAGS data), in both VSS and TSS predictions for the floccular fraction, resulted in an increase in the models quality, taking in consideration the higher R^2 and RPD values, and lower RMSE. The obtained RPD values, being above 3 in all cases, confirm the models' adequateness for screening purposes, with error (RMSE) values around or below 9.0% (below 8.2% for the inclusion of the biomass density data). No significant differences were found with respect to the accuracy of the VSS regarding to the TSS prediction by the use of the QIA data.

The evolution of the predicted and experimentally determined suspended solids (SS) values, regarding the AGS floccular fraction, in the monitoring period, is presented in Fig. 3.

Regarding the VSS prediction for the granular fraction using the MAGS dataset, a R^2 value of 0.930 ($p < 0.01$, RMSE of 1.99 g L^{-1} (8.9% of the studied range) and RPD of 3.61) was obtained for the prediction model shown in Eq. (5), based on the small and large granules contents, intermediate granules size and overall and small granules morphology. With respect to the granular VSS prediction in the presence of PhAC, compared to MAGS under PhAC free conditions based on the

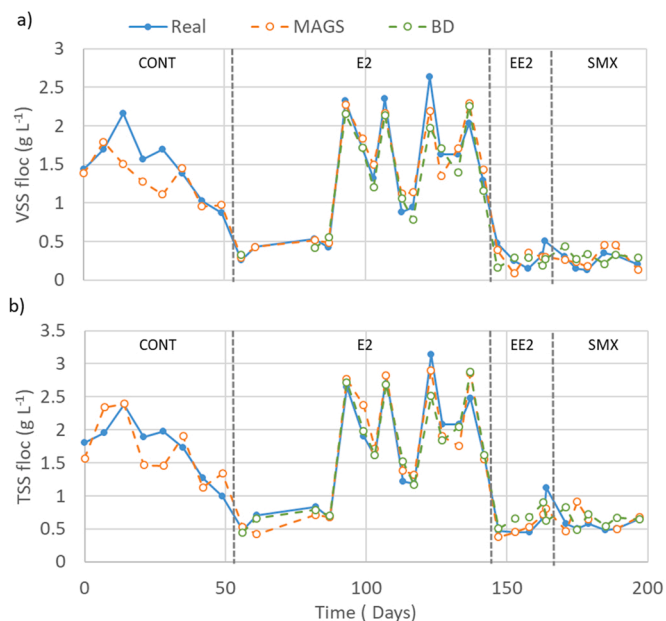


Fig. 3. Evolution of the predicted and experimental (real) SS values, with the MAGS and BD datasets, for the floccular fraction. a) VSS b) TSS. The dashed lines separate the different operational periods (CONT, E2, EE2, and SMX).

intermediate and large granules contents and intermediate granules morphology, the model, although recognizing the importance of the large granules contents, shifted towards the inclusion of size, morphology and contents of a larger set of granules classes. These results point out how the resulting differences in the granules classes, from the presence of PhAC, have impacted on the granular biomass VSS. Taking also in consideration the higher R^2 (0.984), and lower RMSE value (0.123 g L^{-1} – 6.5% of the studied range), obtained for the MAGS under PhAC free conditions, it can be inferred that the prediction of granular VSS in the presence of PhAC becomes more challenging. This is in accordance with the increased challenges found in estimating also the floccular fraction VSS in the presence of PhAC.

$$\begin{aligned} VSS_{gran\ MAGS} = & -42.9 + 4.48 \times 10^{-2} (TV_{G3}) - 8.20 \times 10^{-3} (Diam_{G2}) \\ & + 7.13 \times 10^1 (Rob_{G1}) + 4.98 \times 10^{-4} (Diam_{G2})^2 + 8.93 \times 10^4 (TV_{G1}) \\ & - 3.77 \times 10^1 (Ecc_{gran})^2 \end{aligned} \quad (5)$$

In a similar way to the floccular fraction predictions, the granular fraction VSS was also predicted using the BD dataset, presented in Eq. (6), leading to a smaller R^2 value of 0.874 ($p < 0.01$ RMSE of 1.59 g L^{-1} (8.3% of the studied range) and RPD of 2.84). The biomass density, overall and large granules contents, and intermediate and small granules size were found to be crucial for this prediction.

$$\begin{aligned} VSS_{gran\ BD} = & -53.3 + 0.137 (Nb_{gran}) + 5.06 \times 10^{-2} (Dens) \\ & + 2.92 \times 10^{-1} (Diam_{G1}) + 3.85 \times 10^{-1} (Nb_{G3}) - 2.34 \times 10^{-9} (Vol_{G2}) \end{aligned} \quad (6)$$

With respect to the granular TSS, using the MAGS dataset, a R^2 value of 0.916 ($p < 0.01$, RMSE of 2.18 g L^{-1} (9.6% of the studied range) and RPD of 3.44) was obtained for the prediction model presented in Eq. (7), based on the overall and large granules contents, large granules fraction and morphology and small granules size. Similar to VSS, the most relevant parameters for the prediction of granular TSS in MAGS under PhAC conditions, were related to intermediate and large granules contents and intermediate granules morphology. However, in the presence of PhAC the model prediction again shifted towards the inclusion of size, morphology, and contents of a different set of granules classes (namely

small and large granules). These findings reveal, once again, how the presence of the studied PhAC have impacted the different granules classes and the granular biomass TSS. Again, taking also in consideration the higher R^2 (0.985), and lower RMSE value (0.13 g L^{-1} – 6.6% of the studied range), obtained for the MAGS under PhAC free conditions, it can be inferred that the prediction of granular TSS in the presence of PhAC also becomes more challenging. Again, this is in accordance with the increased challenges found in estimating also the suspended fraction TSS in the presence of PhAC.

$$\begin{aligned} TSS_{granMAGS} = & 1500 + 1.95 \times 10^{-1} (Nb_{gran}) + 1.73 \times 10^{-3} (\%Vol_{G3}) \\ & - 5.33 \times 10^{-6} (Vol_{G1}) + 8.82 \times 10^{-4} (Diam_{G1})^2 \\ & + 3.76 \times 10^{-5} (TV_{G3}) - 4.51 \times 10^3 (Ecc_{G3}) + 3.25 \times 10^3 (Ecc_{G3})^2 \end{aligned} \quad (7)$$

Regarding the TSS prediction for the AGS granular fraction, when the BD dataset was used (Eq. (8)), a decrease on the model prediction ability was perceived (R^2 value of 0.874, p -value < 0.01, RMSE of 1.63 g L^{-1} (8.7% of the studied range) and RPD of 2.75). It was found that the biomass density, overall, large, intermediate, and small granules contents, small granules size and large granules morphology were crucial for the prediction.

$$\begin{aligned} TSS_{granBD} = & -121 + 0.990 (Nb_{gran}) + 5.79 \times 10^{-2} (Dens) \\ & + 2.27 \times 10^{-1} (Diam_{G1}) - 2.54 \times 10^3 (TV_{G1})^2 + 9.35 \times 10^1 (Rob_{G3}) \\ & - 7.86 \times 10^{-1} (Nb_{G2}) - 1.56 \times 10^{-2} (Nb_{G3})^2 \end{aligned} \quad (8)$$

As expected, the granular VSS and TSS models show a high similitude, and a direct dependency on the overall or (dominant) large granules contents and, again, the VSS could be predicted slightly more accurately than the TSS. Furthermore, it could be inferred that higher granular VSS and TSS values were positively correlated with the biomass density and that the exclusion of the MAGS data decreased the prediction ability for both granular VSS and TSS. The obtained RPD values, being above 3 for the use of the mature AGS dataset, confirm these models' adequateness for screening purposes, with error (RMSE) values around or below 9.6%. However, when the MAGS data was excluded, the RPD values dropped slightly below 3 (around 2.8), although the error (RMSE) values did not surpass a maximum of 8.7%.

The evolution of the predicted and experimentally determined SS values, regarding the AGS granular fraction, in the monitoring period, is presented in Fig. 4.

3.3. Prediction of the AGS density and settleability

Opposite to the previous analysis, the biomass density, SVI_5 and SVI_{30}/SVI_5 ratio were predicted including the ensemble floccular and granular fractions. Regarding the biomass density prediction (based on TSS) a R^2 value of 0.889 ($p < 0.01$, RMSE of $11.6 \text{ g TSS L}^{-1}$ biomass (8.8% of the studied range) and RPD of 3.02) was obtained for the prediction model shown in Eq. (9), based on the granular biomass and intermediate granules contents, large granules fraction and both large granules and flocs morphology.

$$\begin{aligned} \text{Density (TSS)} = & 659 + 0.164 (TSS_{gran})^2 - 3.29 (Nb_{G2}) + 1.34 \times 10^3 (Conv_{F3})^2 \\ & - 1.23 \times 10^3 (Conv_{F3}) - 1.50 \times 10^{-2} (\%Nb_{G3})^2 - 4.59 \times 10^2 (Rob_{G3})^2 \end{aligned} \quad (9)$$

Similar results were obtained for the density prediction when based on VSS, presenting a R^2 value of 0.882 ($p < 0.01$, RMSE of $11.6 \text{ g VSS L}^{-1}$ biomass (8.7% of the studied range) and RPD of 2.95) for the prediction model shown in Eq. (10), based on the granular biomass and intermediate granules contents, small granules fraction and large flocs morphology.

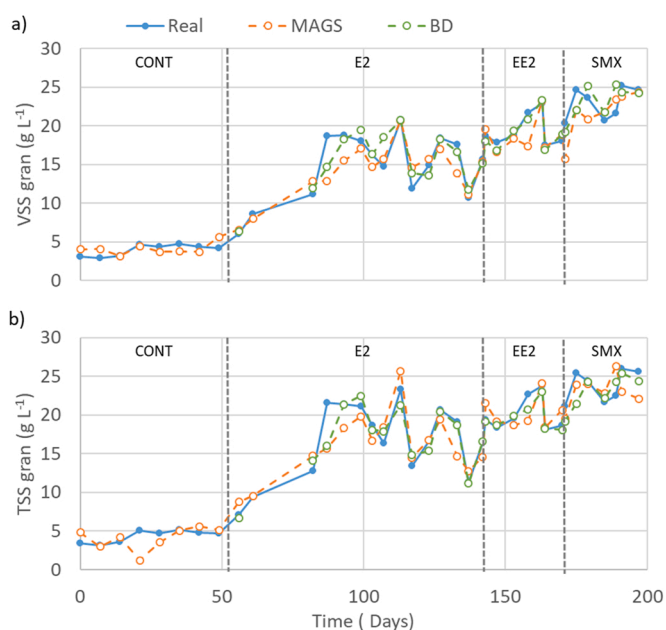


Fig. 4. Evolution of the predicted and experimental (real) SS values, with the MAGS and BD datasets, for the granular fraction. a) VSS b) TSS. The dashed lines separate the different operational periods (CONT, E2, EE2, and SMX).

$$\begin{aligned} \text{Density(VSS)} = & 337 + 1.28 \times 10^{-1} (VSS_{gran})^2 - 6.59 \times 10^{-2} (Nb_{G2})^2 \\ & + 1.32 \times 10^3 (Conv_{F3})^2 - 1.23 \times 10^3 (Conv_{F3}) + 2.67 \times 10^{-2} (\%Nb_{G1})^2 \end{aligned} \quad (10)$$

As expected, the density prediction model based on either the VSS or TSS show a high similitude. Given the above results it can be inferred that the biomass density can be predicted based upon the granular biomass (directly dependent) and intermediate granules (inversely dependent) contents. Furthermore, the large flocs morphology was also found to be a key factor on the biomass density prediction (with convexity values around 0.45 leading to lower density values). The obtained RPD values, around 3 in both cases, are on the threshold of confirming the models' adequateness for screening purposes, with error (RMSE) values around or below 8.8%.

These results could be inferred also by the evolution of the predicted and experimentally determined density values in the monitoring period, presented in Fig. 5.

The use of the MAGS dataset the SVI_5 prediction ability decreased to a R^2 value of 0.965 ($p < 0.01$, RMSE of $2.37 \text{ mL g}^{-1} \text{ VSS}$ (4.6% of the studied range) and RPD of 5.38) for the prediction model shown in Eq. (11). The main parameters that allowed for this prediction ability included the granular and intermediate granules contents, apparent

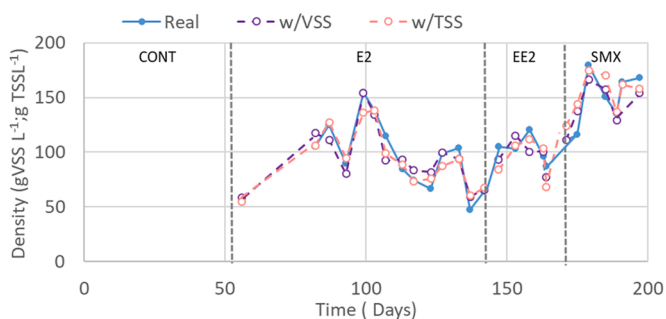


Fig. 5. Evolution of the predicted and experimental (real) density values, predicted by VSS and by TSS. The dashed lines separate the different operational periods (CONT, E2, EE2, and SMX). No biomass density values were determined for the CONT experiment.

floccular density, intermediate flocs fraction, overall flocs size, large granules and intermediate flocs morphology.

$$\begin{aligned} SVI_5(VSS)_{MAGS} = & 149 - 2.15(VSS_{gran}) + 2.00 \times 10^{-2} (\%Nb_{F2})^2 \\ & - 1.32 \times 10^7 (VSS_{floc}/TA_{floc}) - 3.00 \times 10^2 (Conv_{G3}) \\ & - 2.31 \times 10^{-2} (Area_{floc})^2 + 2.61 \times 10^{-2} (Nb_{G2})^2 \\ & + 4.81 \times 10^2 (Conv_{F2}) - 2.84 \times 10^2 (Conv_{F2})^2 \end{aligned} \quad (11)$$

The AGS settling properties, in terms of the SVI_5 , could also be successfully predicted, by using either VSS or TSS. In the former case, the use of the biomass density dataset allowed for a R^2 value of 0.976 ($p < 0.01$, RMSE of 1.219 mL g⁻¹ VSS (5.1% of the studied range) and RPD of 5.26) for the prediction model shown in Eq. (12). The main parameters that allowed for the SVI_5 prediction were related to the overall and granular biomass contents, overall biomass density, overall flocs size and small granules fraction.

$$\begin{aligned} SVI_5(VSS)_{BD} = & 81.8 - 5.21(VSS_{gran}) + 1.22 \times 10^{-1} (VSS_{total})^2 \\ & - 1.22 \times 10^{-6} (Area_{floc})^2 - 5.65 \times 10^{-2} (Dens) - 2.40 \times 10^3 (\%Vol_{G1})^2 \end{aligned} \quad (12)$$

The SVI_5 prediction, based on TSS, was performed using the MAGS dataset, shown in Eq. (13), with the main parameters found to be the granular biomass contents, overall flocs and filamentous bacteria contents, overall granules, overall and large flocs size. This model presented a R^2 value of 0.938 (p -value < 0.01 , RMSE of 3.32 mL g TSS⁻¹ (6.5% of the studied range) and RPD of 3.84).

$$\begin{aligned} SVI_5(TSS)_{MAGS} = & 41.9 - 2.17(TSS_{gran}) + 1.22(TL)^2 \\ & + 1.85 \times 10^{-10} (Area_{F3})^2 - 2.22(TL/TSS_{floc}) + 9.61 \times 10^{-10} (Vol_{gran}) \\ & + 6.01 \times 10^{-4} (TA_{floc}) + 3.49 \times 10^{-2} (TSS_{gran})^2 - 1.32 \times 10^{-6} (Area_{floc})^2 \end{aligned} \quad (13)$$

Regarding the SVI_5 predictions based on TSS, the use of the biomass density dataset, shown in Eq. (14), presented a R^2 value of 0.971 ($p < 0.01$, RMSE of 1.14 mL g TSS⁻¹ (4.7% of the studied range) and RPD of 5.63). The most important parameters for the prediction were found to be related with the overall and granular biomass contents, overall biomass density, overall flocs size and small granules fraction.

$$\begin{aligned} SVI_5(TSS)_{BD} = & 85.1 - 5.22(TSS_{gran}) + 1.09 \times 10^{-1} (TSS_{total})^2 \\ & - 1.29 \times 10^{-6} (Area_{floc})^2 - 4.10 \times 10^{-2} (Dens) - 2.13 \times 10^3 (\%Vol_{G1})^2 \end{aligned} \quad (14)$$

The obtained models were able to predict the SVI_5 , based upon the overall biomass density, granular contents and flocs size (all inversely dependent) and overall biomass contents (directly dependent). The obtained models also revealed that the use of the biomass density data (excluding the MAGS data) led to better SVI_5 prediction abilities, taking in consideration the obtained higher R^2 and RPD values, and lower RMSE. Furthermore, no significant differences were found with relation to the use of VSS or TSS data. The obtained RPD values, largely above the value 3, confirm the models' adequateness for screening purposes, with error (RMSE) values around or below 6.5% (below 5.1% for the inclusion of the biomass density data).

The evolution of the predicted and experimentally determined SVI_5 values in the monitoring period is presented in Fig. 6.

No SVI_5 values were predicted in the MAGS under PhAC free conditions study, however the SVI_{30} was predicted resulting in a R^2 of 0.975, and a RMSE value of 2.24 mL g TSS⁻¹ (6.7% of the studied range). Considering the equivalent model obtained for the presence of PhAC (data not shown) an increased model complexity (presenting six more variables) was obtained with somewhat comparable R^2 (0.953) and RMSE value (3.06 mL g TSS⁻¹ - 5.9% of the studied range). Thus, and although also requiring an increased complexity of the prediction model, the prediction ability of the SVI seems to not be as strongly affected as for the SS.

In any given AGS system, the sludge stability can be inferred by the ratio between the SVI_{30} and SVI_5 [8]. Considering the above, the SVI_{30}/SVI_5 ratio was also predicted by using either the biomass density

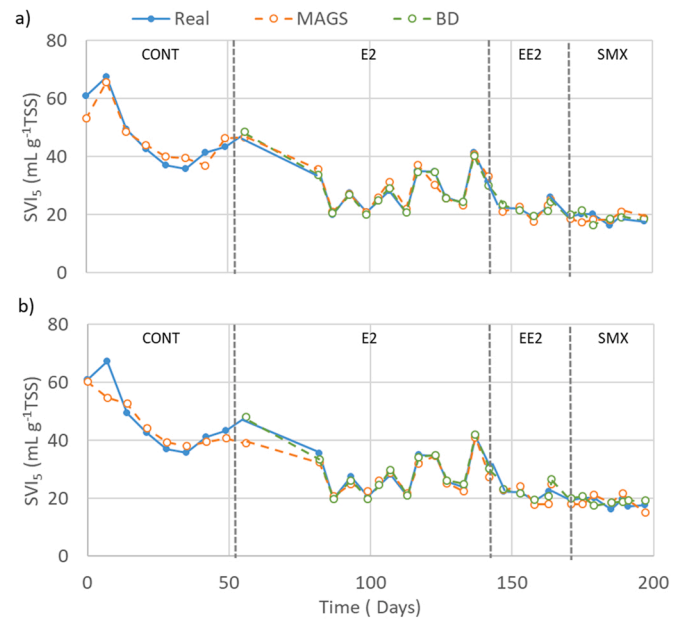


Fig. 6. Evolution of the predicted and experimental (real) SVI_5 values, with the MAGS and BD datasets. a) based on VSS b) based on TSS. The dashed lines separate the different operational periods (CONT, E2, EE2, and SMX).

dataset or the mature granules dataset, and VSS or TSS. The SVI_{30}/SVI_5 ratio based on VSS data was predicted using the mature AGS dataset, presented in Eq. (15), leading to a smaller R^2 value of 0.737 ($p < 0.01$ RMSE of 0.0242 (12.1% of the studied range) and RPD of 1.85). The floccular biomass contents, filamentous bacteria contents, small flocs fraction, large granules and small flocs size, and large flocs morphology were found to be crucial for this prediction.

$$\begin{aligned} SVI_{30}/SVI_5(VSS)_{MAGS} = & 0.709 + 3.39 \times 10^{-2} (TL)^2 - 5.16 \times 10^{-5} (Area_{F1})^2 \\ & + 1.16 \times 10^{-2} (Diam_{F1})^2 - 6.31 \times 10^{-3} (TL/VSS_{floc}) \\ & - 0.358 (Ecc_{F3}) - 8.88 \times 10^{-5} (Diam_{G3}) \\ & - 1.07 \times 10^{-2} (VSS_{floc})^2 - 2.35 \times 10^{-3} (\%Area_{F1}) \end{aligned} \quad (15)$$

The biomass density dataset allowed for a R^2 value of 0.917 ($p < 0.01$, RMSE of 0.0045 (8.7% of the studied range) and RPD of 3.52) for the prediction model using VSS data shown in Eq. (16). The main parameters that allowed for the SVI_{30}/SVI_5 prediction were related to the floccular and large flocs contents, overall flocs size and morphology.

$$\begin{aligned} SVI_{30}/SVI_5(VSS)_{BD} = & 1.04 - 3.40 \times 10^{-3} (VSS_{floc})^2 \\ & - 2.17 \times 10^{-8} (Area_{floc})^2 + 4.00 \times 10^{-5} (Area_{floc}) \\ & + 3.45 \times 10^{-9} (TA_{F3})^2 - 1.86 \times 10^{-5} (TA_{F3}) - 8.15 \times 10^{-2} (Ecc_{floc}) \end{aligned} \quad (16)$$

The SVI_{30}/SVI_5 prediction, based on TSS, was performed using the MAGS dataset, and shown in Eq. (17), with the main parameters found to be the granular biomass contents, overall flocs and filamentous bacteria contents, overall granules, overall and large flocs size. This model presented a R^2 value of 0.701 ($p < 0.01$, RMSE of 0.0248 (12.4% of the studied range) and RPD of 1.81).

$$\begin{aligned} SVI_{30}/SVI_5(TSS)_{MAGS} = & 0.501 + 4.19 \times 10^{-2} (TL)^2 \\ & - 5.61 \times 10^{-5} (Area_{F1})^2 + 1.28 \times 10^{-2} (Diam_{F1})^2 \\ & - 1.88 \times 10^{-2} (TL/VSS_{floc}) - 0.303 (Ecc_{F3}) - 7.01 \times 10^{-5} (Diam_{G3}) \\ & - 8.33 \times 10^{-3} (VSS_{floc})^2 - 2.07 \times 10^{-3} (\%Area_{F1}) \end{aligned} \quad (17)$$

Regarding the SVI_{30}/SVI_5 ratio predictions based on TSS, the use of the BD, shown in Eq. (18), presented a R^2 value of 0.874 ($p < 0.01$, RMSE of 0.0033 (6.4% of the studied range) and RPD of 4.78). The most

important parameters for the prediction were found to be related with the overall floccular biomass contents, large granules and flocs contents, apparent floccular density, small and large flocs fraction and overall flocs size.

$$\begin{aligned} SVI_{30}/SVI_5(TSS)_{BD} = & 0.869 - 5.29 \times 10^{-3} (TSS_{floc})^2 - 5.39 \times 10^{-8} (Area_{floc})^2 \\ & + 1.51 \times 10^{-9} (TA_{F3}) + 1.67 \times 10^{-4} (Area_{floc}) - 5.54 \times 10^{-2} (\%Nb_{F3})^2 \\ & + 2.02 \times 10^{-3} (\%Area_{F1}) + 1.43 \times 10^4 (TSS_{floc}/TA_{floc}) - 3.93 \times 10^{-5} (Nb_{G3})^2 \end{aligned} \quad (18)$$

The models obtained using the BD allowed to predict, at some extent, the SVI_{30}/SVI_5 values, based upon mainly the overall floccular biomass contents (inversely dependent), the flocs size (with area values between 1000 and 1500 μm^3 leading to higher SVI_{30}/SVI_5 values) and overall flocs morphology. Besides the overall flocs, the large flocs morphology seemed also to affect the SVI_{30}/SVI_5 values. Furthermore, the use of VSS instead of TSS data, led to slightly better prediction abilities. The obtained RPD values were above the value 3 for the models encompassing the biomass density dataset, confirming these models' adequateness for screening purposes, presenting error (RMSE) values below of 8.7% (6.4% for the use of the TSS data).

The evolution of the predicted and experimentally determined SVI_{30}/SVI_5 values in the monitoring period is presented in Fig. 7.

Taking all the above results in consideration, it can be inferred that the floccular SS were affected mainly by the overall flocs contents, size and morphology, rising with the increase of these parameters, whereas the granular SS rise mainly by the increase on the overall granules contents and overall AGS density. Indeed, a direct dependency was found between the overall AGS density and the granular SS, which could be expected in AGS systems presenting a large granular fraction. Both parameters, as well as the overall flocs size, were negatively correlated with the sludge settling ability, with their increase leading to a decrease on the SVI_5 , and opposite to the effect of the overall SS. Indeed, it could be expected that the increase of the floccular fraction in AGS systems, and mainly of small flocs, could lead to worsen settling abilities. The SVI_{30}/SVI_5 ratio was affected mainly by the floccular biomass contents and size, though its relationship needs further enlightenment.

4. Applicability and practical relevance of this work

In this work, morphological and structural parameters of floccular and granular fractions biomass, obtained by QIA, combined with chemometric tools, were used to successfully predict the density, settleability and SS of AGS in both absence and presence of PhAC (E2, EE2 and SMX). QIA application relies on sample preparation, survey and acquisition, requiring basic equipment and knowledge of microscopy and image processing techniques. Approximately 4 h of man-work (on average) were needed to evaluate each sample set, including sample collection, sieving and mounting, microscopy survey and image acquisition and QIA data retrieval. Taking into consideration that standard approaches for VSS, TSS, and SVI determination (the later one relying on SS determination), are far more time consuming, the use of the presented methodology can be considered of major interest in daily based operations for the above purposes.

The obtained results allowed for a new insight, on an operational point of view, in AGS monitoring and contributed to process optimization and granules stability in the presence of the studied PhAC. It is, thus, expected that this methodology may be used for monitoring full-scale AGS systems in the presence of PhAC in the near future, contributing to timely identification and control of dysfunctions in biomass morphology, structure and settleability in the presence of environmentally relevant concentrations of E2, EE2, and SMX.

5. Conclusions

AGS processes stability is quite dependent on the balance between

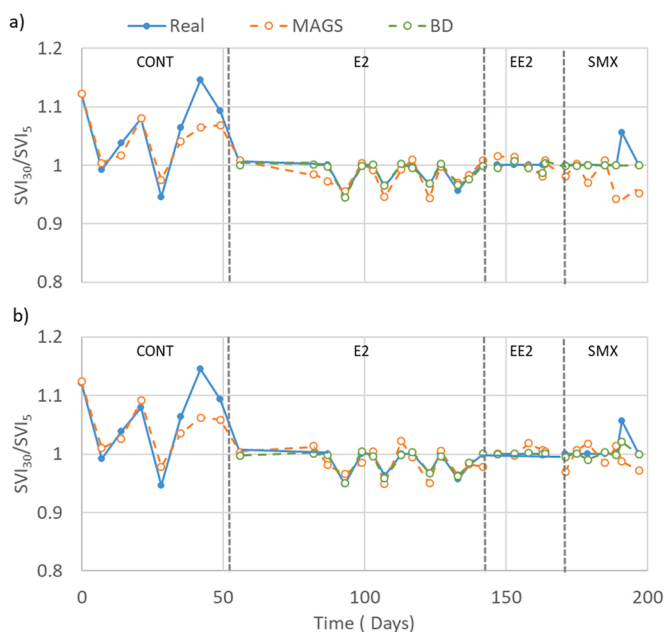


Fig. 7. Evolution of the predicted and experimental (real) SVI_{30}/SVI_5 values, with the MAGS and BD datasets. a) based on VSS b) based on TSS. The dashed lines separate the different operational periods (CONT, E2, EE2, and SMX).

floccular and granular biomass contents, and the assessment of their contents, settling and density properties is of major interest. In accordance, the prediction of parameters such as AGS density, SVI_5 , SVI_{30}/SVI_5 ratio, VSS, and TSS in absence as well as in the presence of PhACs (E2, EE2, and SMX), was successfully studied by MLR techniques using data obtained by QIA. The use of the BD dataset increased the prediction abilities of the floccular fraction suspended solids concentrations, and the AGS settling ability and stability. On the other hand, the predictions regarding the granular fraction suspended solids, revealed an opposite behavior with a decrease in the prediction ability by the exclusion of the MAGS data.

As expected, the overall flocs contents, size and morphology played a major role in predicting the floccular biomass, and the overall granules contents in predicting the granular biomass contents. Also, a clear dependency between the sludge density and the granular biomass contents was found. The sludge settling ability was found to depend on the AGS density, granular biomass contents and overall flocs size, whereas the stability was related to the floccular biomass contents and size, though its relationship needs further enlightenment.

CRedit authorship contribution statement

Cristiano Leal: Investigation, Conceptualization, Writing - Original Draft
Angeles Val del Río: Investigation, Writing - Review & Editing
Daniela P. Mesquita: Supervision, Writing - Review & Editing
António L. Amaral: Supervision, Methodology, Writing - Review & Editing
Eugénio C. Ferreira: Funding acquisition, Writing - Review & Editing.

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.

Acknowledgments

The authors thank the Portuguese Foundation for Science and Technology (FCT) under the scope of the strategic funding of UIDB/

04469/2020 unit and the project AGE_{NT} - PTDC/BTA-BTA/31264/2017 (POCI-01-0145-FEDER-031264). The authors wish to thank the company Águas do Tejo Atlântico, S.A. for supplying the granules. Cristiano Leal is recipient of a fellowship supported by a doctoral advanced training (call NORTE-69-2015-15) funded by the European Social Fund under the scope of Norte2020 - Programa Operacional Regional do Norte. A. Val del Río is supported by Xunta de Galicia (ED418B 2017/075) and program Iacobus (2018/2019). Cristiano Leal also thanks René Benevides for all the support during the experimental activities.

References

- [1] C.S. Leal, D.P. Mesquita, A.L. Amaral, A.M. Amaral, E.C. Ferreira, Environmental impact and biological removal processes of pharmaceutically active compounds: the particular case of sulfonamides, anticonvulsants and steroid estrogens, *Crit. Rev. Environ. Sci. Technol.* 50 (2020) 698–742, <https://doi.org/10.1080/10643389.2019.1642831>.
- [2] M. Adeel, X. Song, Y. Wang, D. Francis, Y. Yang, Environmental impact of estrogens on human, animal and plant life: a critical review, *Environ. Int.* 99 (2017) 107–119, <https://doi.org/10.1016/j.envint.2016.12.010>.
- [3] C. Gadipelly, A. Pérez-González, G.D. Yadav, I. Ortiz, R. Ibáñez, V.K. Rathod, K. V. Marathe, Pharmaceutical industry wastewater: review of the technologies for water treatment and reuse, *Ind. Eng. Chem. Res.* 53 (2014) 11571–11592, <https://doi.org/10.1021/ie501210j>.
- [4] M.K. Pholchan, J.C. de Baptista, R.J. Davenport, W.T. Sloan, T.P. Curtis, Microbial community assembly, theory and rare functions, *Front. Microbiol.* 4 (2013) 1–9, <https://doi.org/10.3389/fmicb.2013.00068>.
- [5] Y. Zhu, Y. Wang, X. Jiang, S. Zhou, M. Wu, M. Pan, H. Chen, Microbial community compositional analysis for membrane bioreactor treating antibiotics containing wastewater, *Chem. Eng. J.* 325 (2017) 300–309, <https://doi.org/10.1016/j.cej.2017.05.073>.
- [6] S. Bengtsson, M. de Blois, B.-M. Wilén, D. Gustavsson, A comparison of aerobic granular sludge with conventional and compact biological treatment technologies, *Environ. Technol.* 40 (2019) 2769–2778, <https://doi.org/10.1080/09593330.2018.1452985>.
- [7] M. Hussain, Aqeel Weissbrodt, David Cerruti, S. Wolfaardt, M. Wilén Gidéon, Liss Britt-Marie, Drivers of bioaggregation from flocs to biofilms and granular sludge, *Environ. Sci. Water Res. Technol.* (2019) 20172, <https://doi.org/10.1039/c9ew00450e>.
- [8] S.F. Corsino, M. Capodici, M. Torregrossa, G. Viviani, Fate of aerobic granular sludge in the long-term: the role of EPSs on the clogging of granular sludge porosity, *J. Environ. Manag.* 183 (2016), <https://doi.org/10.1016/j.jenvman.2016.09.004>.
- [9] J. Czarnota, A. Masłó, M. Zdeb, G. Łagód, The impact of different powdered mineral materials on selected properties of aerobic granular sludge, *Molecules* 25 (2) (2020) 386, <https://doi.org/10.3390/molecules25020386>.
- [10] R.M. Castellanos, M. Dezotti, J.P. Bassin, COD, nitrogen and phosphorus removal from simulated sewage in an aerobic granular sludge in the absence and presence of natural and synthetic estrogens: performance and biomass physical properties assessment, *Biochem. Eng. J.* 176 (2021), 108221, <https://doi.org/10.1016/j.bej.2021.108221>.
- [11] J. Kent, J.H. Tay, Treatment of 17 α -ethinylestradiol, 4-nonylphenol, and carbamazepine in wastewater using an aerobic granular sludge sequencing batch reactor, *Sci. Total Environ.* 652 (2019) 1270–1278, <https://doi.org/10.1016/j.scitotenv.2018.10.301>.
- [12] A.R. Mendes Barros, T.S. Argenta, C. de Amorim de Carvalho, F. da Silva Oliveira, P.I. Milen Firmino, A. Bezerra dos Santos, Effects of the antibiotics trimethoprim (TMP) and sulfamethoxazole (SMX) on granulation, microbiology, and performance of aerobic granular sludge systems, *Chemosphere* 262 (2021), 129124, <https://doi.org/10.1016/j.chemosphere.2020.127840>.
- [13] D.P. Mesquita, A.L. Amaral, E.C. Ferreira, Identifying different types of bulking in an activated sludge system through quantitative image analysis, *Chemosphere* 85 (2011) 643–652, <https://doi.org/10.1016/j.chemosphere.2011.07.012>.
- [14] J.C. Costa, M.M. Alves, E.C. Ferreira, A chemometric tool to monitor high-rate anaerobic granular sludge reactors during load and toxic disturbances, *Biochem. Eng. J.* 53 (2010) 38–43, <https://doi.org/10.1016/j.bej.2008.12.006>.
- [15] J.C. Costa, I. Moita, E.C. Ferreira, M.M. Alves, Morphology and physiology of anaerobic granular sludge exposed to an organic solvent, *J. Hazard. Mater.* 167 (2009) 393–398, <https://doi.org/10.1016/j.jhazmat.2008.12.130>.
- [16] C. Leal, A. Val del Río, 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), 116049, <https://doi.org/10.1016/j.seppur.2019.116049>.
- [17] J.C. Costa, D.P. Mesquita, A.L. Amaral, M.M. Alves, E.C. Ferreira, Quantitative image analysis for the characterization of microbial aggregates in biological wastewater treatment: a review, *Environ. Sci. Pollut. Res.* 20 (2013) 5887–5912, <https://doi.org/10.1007/s11356-013-1824-5>.
- [18] D.P. Mesquita, A.L. Amaral, E.C. Ferreira, Activated sludge characterization through microscopy: a review on quantitative image analysis and chemometric techniques, *Anal. Chim. Acta* 802 (2013) 14–28, <https://doi.org/10.1016/j.aca.2013.09.016>.
- [19] A.L. Amaral, E.C. Ferreira, Activated sludge monitoring of a wastewater treatment plant using image analysis and partial least squares regression, *Anal. Chim. Acta* 544 (2005) 246–253, <https://doi.org/10.1016/j.aca.2004.12.061>.
- [20] H.Z. Abyaneh, Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters, *J. Environ. Heal. Sci. Eng.* 12 (2014), <https://doi.org/10.1186/2052-336X-12-40>.
- [21] J. Liang, Q. Wang, Q.X. Li, L. Jiang, J. Kong, M. Ke, M. Arslan, M. Gamal El-Din, C. Chen, Aerobic sludge granulation in shale gas flowback water treatment: assessment of the bacterial community dynamics and modeling of bioreactor performance using artificial neural network, *Bioresour. Technol.* 313 (2020), 123687, <https://doi.org/10.1016/j.biortech.2020.123687>.
- [22] P. Sharma, S. Sood, S.K. Mishra, Development of multiple linear regression model for biochemical oxygen demand (BOD) removal efficiency of different sewage treatment technologies in Delhi, India, *Sustain. Water Resour. Manag.* 6 (2020) 1–13, <https://doi.org/10.1007/s40899-020-00377-9>.
- [23] G. Mundi, R.G. Zytner, K. Warriner, H. Bonakdari, Machine Learning Models for Predicting Water Quality of Treated Fruit and Vegetable Wastewater, 2021 1–17.
- [24] H. Ilyas, I. Masih, E.D. Van Hullebusch, Prediction of the removal efficiency of emerging organic contaminants based on design and operational parameters of constructed wetlands, *J. Environ. Chem. Eng.* 9 (2021), 104592, <https://doi.org/10.1016/j.jece.2020.104592>.
- [25] H. Ilyas, I. Masih, E.D. van Hullebusch, Prediction of the removal efficiency of emerging organic contaminants in constructed wetlands based on their physicochemical properties, *J. Environ. Manag.* 294 (2021), 112916, <https://doi.org/10.1016/j.jenvman.2021.112916>.
- [26] M.K. De Kreuk, J.J. Heijnen, M.C.M. Van Loosdrecht, Simultaneous COD, nitrogen, and phosphate removal by aerobic granular sludge, *Biotechnol. Bioeng.* 90 (2005) 761–769, <https://doi.org/10.1002/bit.20470>.
- [27] R.S. Rana, P. Singh, V. Kandari, R. Singh, R. Dobhal, S. Gupta, A review on characterization and bioremediation of pharmaceutical industries' wastewater: an Indian perspective, *Appl. Water Sci.* 7 (2017) 1–12, <https://doi.org/10.1007/s13201-014-0225-3>.
- [28] C.S. Leal, M. Lopes, A. Val del Río, C. Quintelas, P.M.L. Castro, E.C. Ferreira, A. L. Amaral, D.P. Mesquita, Assessment of an aerobic granular sludge system in the presence of pharmaceutically active compounds by quantitative image analysis and chemometric techniques, *J. Environ. Manag.* 289 (2021), 112474, <https://doi.org/10.1016/j.jenvman.2021.112474>.
- [29] Standard Methods for the Examination of Water and Wastewater, 23rd Edition, American Public Health Association, American Water Works Association, Water Environment Federation, 2017.
- [30] J.J. Beun, M.C.M. Van Loosdrecht, J.J. Heijnen, Aerobic granulation in a sequencing batch airlift reactor, *Water Res.* 36 (2002) 702–712, [https://doi.org/10.1016/S0043-1354\(01\)00250-0](https://doi.org/10.1016/S0043-1354(01)00250-0).
- [31] Ana Paula Fonseca, Massano Cardoso, V. Esteves, Determination of estrogens in raw and treated wastewater by high-performance liquid chromatography-ultraviolet detection, *J. Environ. Anal. Toxicol.* 04 (2013) 1–5, <https://doi.org/10.4172/2161-0525.1000203>.
- [32] C. Quintelas, D.P. Mesquita, E.C. Ferreira, A.L. Amaral, Quantification of pharmaceutical compounds in wastewater samples by near infrared spectroscopy (NIR), *Talanta* 194 (2019) 507–513, <https://doi.org/10.1016/j.talanta.2018.10.076>.
- [33] C. Leal, A. Val del Río, E.C. Ferreira, D. Mesquita, A.L. Amaral, Validation of a quantitative image analysis methodology for the assessment of the morphology and structure of aerobic granular sludge in the presence of pharmaceutically active compounds, *Environ. Technol. Innov.* 23 (2021), 101639, <https://doi.org/10.1016/j.eti.2021.101639>.
- [34] A. Amaral, Image Analysis in Biotechnological Processes: Applications to Wastewater Treatment, University of Minho, 2003. (<http://hdl.handle.net/1822/4506>).
- [35] S.A. Hill, Statistics, in: *Found. Anesth.*, Elsevier, 2006, pp. 207–217, <https://doi.org/10.1016/B978-0-323-03707-5.50024-3>.
- [36] T. Fearn, Assessing calibrations: SEP, RPD, RER and R 2, *NIR News* 13 (2002) 12–13, <https://doi.org/10.1255/nirn.689>.