Water Practice & Technology

IWA PUBLISHING

Evaluation and optimization of ASM1 parameters using large-scale WWTP monitoring data from a subtropical climate region in Brazil

A. C. O. Martins Jr. (iD)*, M. C. A. Silva (iD) and A. D. Benetti IWA (iD)

Postgraduate Program in Water Resources and Environmental Sanitation, Hydraulic Research Institute, Federal University of Rio Grande do Sul, 9500 Bento Gonçalves Ave., Building 44302, Agronomia, Porto Alegre, Rio Grande do Sul, Brazil
*Corresponding author. E-mail: antonio.martins@marquette.edu

(iD) ACOM, 0000-0002-1609-0587; MCAS, 0000-0002-1104-8355; ADB, 0000-0002-5940-8866

ABSTRACT

This study aimed at providing a set of optimal kinetic and stoichiometric parameters of ASM1 representative of wastewater from a subtropical climate region in Brazil. ASM1 was applied on the STOAT program, and the model parameters were evaluated and optimized with sensitivity analysis and Response Surface Methodology (RSM) to reach minimum prediction errors of effluent TSS, COD, and NH_3. Six sensitive parameters were identified: Y_H, Y_A, μ_A, K_{NH}, b_A, and k_{OA}. Predictions of RSM regression models were strongly correlated to the STOAT predictions. Y_H mainly affected TSS and COD, and the other parameters affected NH_3. ASM1 calibration with estimated optimal values of sensitive parameters resulted in approximately null prediction errors for modeling state variables. NH_3 presented similar results in the ASM1 validation; meanwhile, TSS and COD presented high errors related to the increase in Y_H due to the RSM optimization. The optimal parameters, mainly Y_A, μ_A, K_{NH}, b_A, and k_{OA}, constitute references for other studies on ASM1 modeling using wastewater data from a subtropical climate region. Y_H optimal value should be evaluated as well as the effect of sludge wastage methods and the simulation periods.

Key words: mathematical modeling, response surface methodology, sensitivity analysis, systematic model calibration

HIGHLIGHTS

- A large-scale WWTP was modeled using standard monitoring data, ASM1, Version 3, and STOAT.
- Sensitivity analysis and Response Surface Methodology improved ASM1 calibration.
- Estimated optimal kinetic and stoichiometric parameters of ASM1 represented wastewater from a subtropical climate region in Brazil.
- The high experimental range deserves attention in the optimization of the heterotrophic yield (Y_H) parameter.

GRAPHICAL ABSTRACT
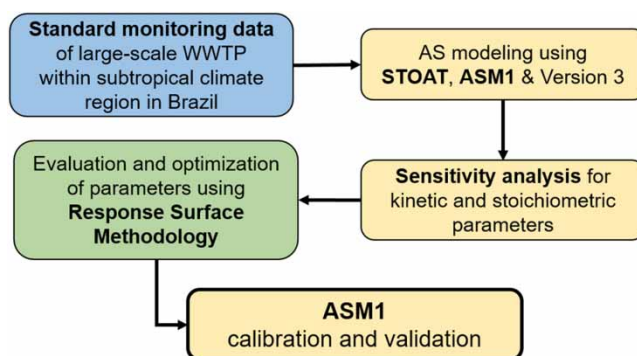
INTRODUCTION

The Sewage Treatment Operation and Analysis over Time (STOAT©) (WRC Plc 1994) consists of a free simulation program in which one can select the broadly known Activated Sludge Model N°1 (ASM1)

(Henze *et al.* 1987) to model an activated sludge (AS) process. STOAT and ASM1 may be used with scarce data sets available from standard monitoring of large-scale Wastewater Treatment Plants (WWTP) for setting up a facility and modeling the performance of AS under steady-state conditions (Andraka *et al.* 2018).

The calibration is one of the most critical steps in the modeling procedure to provide reliable predictions accordingly with the specific conditions of AS process (Rieger *et al.* 2013). ASM1 modeling and calibration depend on wastewater composition fractionation and model parameters determination, respectively. Such fractionation and determination may need extra laboratory analysis, which is not commonly viable for some plants (Borzooei *et al.* 2019). In some cases, typical ratios of municipal wastewater may be used for fractionation (Henze & Comeau 2008, p. 36), and kinetic and stoichiometric ASM1 parameters calibration on trial and error procedure may be adopted. However, this calibration procedure is not appropriate due to the importance of prioritizing methods that guarantee as much information as possible to form a suitable parameter combination (Petersen *et al.* 2003).

Alternatively, some authors have proposed a systematic calibration approach to find optimal parameters for ASMs. The strategy comprises modeling in association with sensitivity analysis for parameter selection, and design of experiments and Response Surface Methodology (RSM) for parameter evaluation and optimization (Kim *et al.* 2009; Lim *et al.* 2012; Ahn *et al.* 2014).

Most studies on the evaluation and optimization of ASM1 parameters were developed in Europe, North America, and Asia (Hauduc *et al.* 2011). Furthermore, ASM1 presents default values of parameters for European wastewater characteristics (Ahn *et al.* 2014). Thus, applications in other regions are relevant due to the variation of model parameters correspondingly to wastewater composition through space and time (Von Sperling *et al.* 2020).

The geographic expansion may also contribute to evaluating ASM1 modeling using STOAT in other climate regions. That program is broadly used in the temperate climate of the United Kingdom and was validated in tropical climate conditions in India, where the microorganisms' growth rate is higher than the temperate climate conditions (Sarkar *et al.* 2010).

Particularly in Brazil, AS modeling is essential to improve wastewater treatment. The country presents 354 WWTPs with such treatment process (10% of the total) (ANA 2020), and more than half of the population lack wastewater treatment (Brasil 2019). Some authors have used ASM1 and STOAT to model a WWTP using monitoring data of domestic wastewater treatment from a subtropical climate region in Brazil (Pistorello 2018; Baptista 2020). However, to the best of our knowledge, no studies have performed an estimation of optimal ASM1 parameters associating the use of monitoring data of a large-scale WWTP within a subtropical climate region with sensitivity analysis and RSM.

Therefore, this study aimed at providing a set of optimal kinetic and stoichiometric parameters of ASM1 representative of wastewater from a subtropical climate region in Brazil. Thus, these optimal values may compose the references of ASM1 parameters and be effective on AS modeling using the STOAT program and monitoring data of large-scale WWTP in the same climate conditions.

METHODS

WWTP of study

The research was conducted on domestic wastewater generated from São João Navegantes WWTP in Porto Alegre, Brazil, located in a subtropical climate region (29°59′29″S; 51°11′43.5″W). The average raw-water flow is 0.44 m^3/s, corresponding to 150 thousand inhabitants. The wastewater system comprises extended aeration activated sludge (Figure 1).

Data collection and reconciliation

The following wastewater characteristics were determined: flow, temperature, total suspended solids (TSS), chemical oxygen demand (COD), ammonia (NH$_3$), and pH. Influent and effluent means of these variables were considered after treating missing and censored data, as well as outliers of the WWTP monitoring data set, as described in von Sperling *et al.* (2020). Annual means of those variables of 2018 and 2019 were used for ASM1 calibration and validation, respectively (Table 1).

Activated sludge modeling

The ASM1 and Version 3 models were used, respectively, for the aeration and secondary sedimentation tanks. Both models are available on the STOAT$^{©}$ simulation program and the guidelines of Rieger *et al.* (2013) and WRC PLC
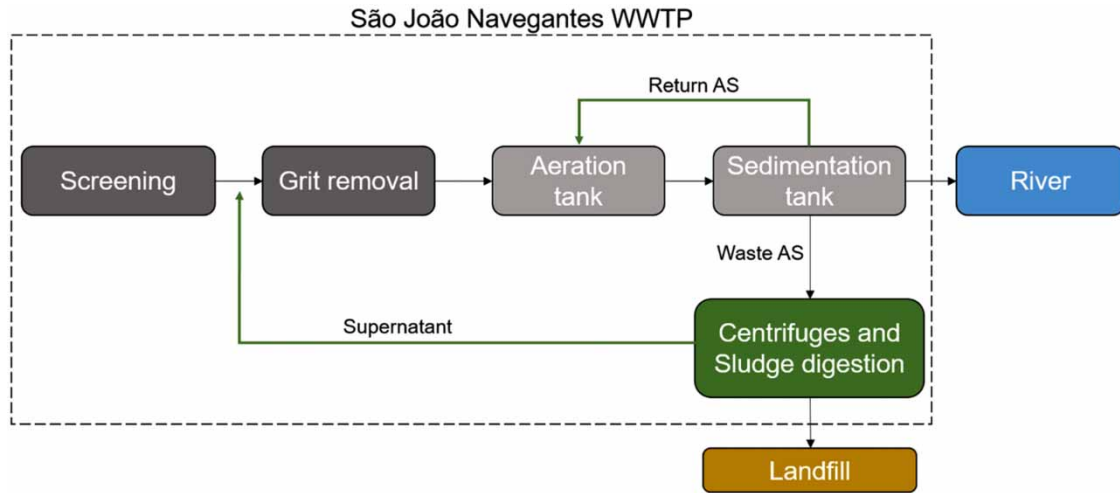
Figure 1 | The treatment system of the São João Navegantes WWTP.

Table 1 | Means of wastewater composition for modeling periods

Variable/Year	Influent		Effluent	
	2018	2019	2018	2019
TSS (mg/l)	119	170	30	21
COD (mg/l)	297	359	52	45
NH₃ (mg/l)	45.5	39.5	4.4	3.7
pH	7.2	7.2	6.9	7
Temperature (°C)	22.6	23.8	22.9	24.1
¼ Treated flow (m³/s)	–	–	0.077	0.076

(1994) were considered for modeling conduction. Modeling was developed for one of the four parallels flows of the treatment plant. Information on tank sizes and operational data were provided by the WWTP (Table 2).

On Version 3, the wastage method selected was fixed-rate (0.00153 m³/s) over variable time (0–8 h) to maintain a specific Mixed-Liquor Suspended Solids (MLSS) set-point (3,000 mg/l). By assuming extended aeration correlation, sewage calibration data for the same model considered Stirred Sludge Volume Index (SSVI) (3.5 g/l) equal to 100 ml/g (average for fair sludge sedimentation) (von Sperling 1994).

AS Solids Retention Time (SRT) was estimated using Equation (1) (Tchobanoglous *et al.* 2014), according to the WWTP monitoring data (Table 2). V is the volume of the aeration tank, X is the MLSS concentration, Q_W is the sludge wastage flow (0.00153 m³/s considering a pump run time of 8 h), and X_R is the sludge TSS. The first simulation period was assumed as 3 SRT (60 d) (Rieger *et al.* 2013).

$$SRT \approx \frac{VX}{Q_W X_R} \tag{1}$$

The influent profile assumed on STOAT was the sinusoidal pattern. Thus, the following conditions were selected for most wastewater variables (Table 1): 0 h for phase, 30% for amplitude, and frequency equal to 0.261799 (representing daily fluctuations). Specifically for temperature, the frequency was 7.27E-4 (annual fluctuations), and amplitude was 0% for the same variable and pH. Typical ratios of municipal wastewater were applied at the fractionation of TSS, COD, and N contents (Table 3).

The modeling state variables were the mean effluent concentrations of TSS, COD, and NH₃ (Table 1). The following runs started from the end of the first simulation.

ASM1 parameters and respective default values are presented in Table 4. Some kinetic parameters display different units on STOAT; that is, h⁻¹ for 15 °C ($P_{15°C}$). By using the temperature coefficients (θ) provided on

Table 2 | Characteristics of the WWTP AS process for ¼ flow of the treatment

Aeration tank	
Volume (m³) (V)	2,070
Length (m)	30
Width (m)	15
Depth (m)	4.6
MLSS (mg/l) (X)	4,000
Secondary sedimentation tank	
Surface area (m²)	720
Depth (m)	3.5
Depth of feed (m)	1.75
AS	
Return AS (RAS) flow (m³/s)	0.15
Sludge wastage flow (m³/s) (Q_W)	0.00153
Wastage pump run time (h)	8
Wastage cycle time (h)	24
MLSS set-point (mg/l)	3,000
Sludge TSS (mg/l) (X_R)	10,000
SRT (d)	20

Table 3 | Fractionation of wastewater composition for modeling periods

COD[a]	Typical Ratio	2018	2019
Soluble undegradable (S_U)	$0.04*COD_{total}$	11.9	14.4
Soluble biodegradable (S_B)	$0.36*COD_{total}$	106.9	129.2
Particulate biodegradable (X_B)	$0.4*COD_{total}$	118.8	143.6
Particulate undegradable (X_U)	$0.2*COD_{total}$	59.4	71.8
TSS[a]			
Volatile solids	$0.7*TSS$	83.3	119
Non-volatile solids	$0.3*TSS$	35.7	51
Nitrogen (N)[b]			
Total N (N_{total})[c]	$1.61*NH_3$	73.24	63.51
Soluble organic N	$0.06*N_{total}+0.03*N_{total}$	6.59	5.72
Particulate organic N	$0.06*N_{total}+0.1*N_{total}$	11.72	10.16

Sources: (a) Henze & Comeau (2008, p. 34); (b) Rössle & Pretorius (2001); (c) Barroso Júnior (2020, p. 94–95).

STOAT for each of those parameters, Equation (2) (Tchobanoglous *et al.* 2014) was applied to convert the values onto the default unit; that is, d⁻¹ for 20 °C ($P_{20°C}$). STOAT uses such coefficients to model different wastewater temperatures.

$$P_{20°C} = P_{15°C} \exp^{\theta(20-15)} \tag{2}$$

Sensitivity analysis

A one-way sensitivity analysis was performed to indicate the critical parameters of ASM1 (Table 4). This analysis consisted of sequentially varying each parameter while keeping the others constant. A 10% increase was applied for each parameter to calculate its sensitivity coefficient (S_j^i) (Equation (3)) related to each state variable; where Y

Table 4 | Kinetic and stoichiometric parameters of ASM1

Stoichiometric	Symbol	Unit	Default (20 °C)[a]	Literature
Heterotrophic yield	Y_H	g cell COD formed (g COD oxidized)$^{-1}$	0.67	0.62–0.67[b]
Autotrophic yield	Y_A	g cell COD formed (g N oxidized)$^{-1}$	0.24	0.07–0.28[c]
Fraction of biomass yielding particulate products	f_P	Dimensionless	0.08	–
Mass N/Mass COD in biomass	i_{XB}	g N (g COD)$^{-1}$ in biomass	0.086	0.079–0.086[b]
Mass N/Mass COD in products from biomass	i_{XP}	g N (g COD)$^{-1}$ in endogenous biomass	0.06	–
Kinetic				
Heterotrophic maximum specific growth rate	μ_H	d^{-1}	6 (0.1768 h^{-1})[d]	5.7–6[b]
Half-saturation coefficient (hsc) for heterotrophs	K_S	g COD.m^{-3}	20	5–225[c]
Oxygen hsc for heterotrophs	K_{OH}	g O$_2$.m^{-3}	0.2	0.01–0.2[c]
Nitrate hsc for denitrifying heterotrophs	K_{NO}	g NO$_3$-N.m^{-3}	0.5	0.1–0.5[c]
Heterotrophic decay rate	b_H	d^{-1}	0.62 (0.0147 h^{-1})[d]	0.05–1.6[c]
Autotrophic maximum specific growth rate	μ_A	d^{-1}	0.8 (0.0207 h^{-1})[d]	0.2–1[c]
Ammonia hsc for autotrophs	K_{NH}	g NH$_3$-N.m^{-3}	1	0.75–1[b]
Oxygen hsc for autotrophs	K_{OA}	g O$_2$.m^{-3}	0.4	0.4–2[c]
Autotrophic decay rate	b_A	d^{-1}	0.15 (0.0036 h^{-1})[d]	0.05–0.15[a]
Correction factor for anoxic growth of heterotrophs	η_g	Dimensionless	0.8	0.6–1[c]
Ammonification rate	k_A	m^3(g COD d)$^{-1}$	0.08 (0.0024 h^{-1})[d]	0.07–0.08[b]
Maximum specific hydrolysis rate	K_h	g slowly biodeg.(g cell COD d)$^{-1}$	3 (0.0722 h^{-1})[d]	2.2–3[b]
Hsc for hydrolysis of slowly biodegradable substrate	K_X	g slowly biodeg.(g cell COD d)$^{-1}$	0.03 (0.0173 h^{-1})[d]	–
Correction factor for anoxic hydrolysis	η_h	Dimensionless	0.4	0.4–0.5[b]

Sources: (a) Henze *et al.* (1987); (b) Hauduc *et al.* (2011); (c) Ahn *et al.* (2014); (d) value displayed on STOAT for 15 °C.

is the state variable output, P is the ASM1 parameter value, and the subscripts *0* e *1* stand for the default and the changed values, respectively. A parameter is considered sensitive if $S_j^i \geq 0.25$ (Liwarska-Bizukojc *et al.* 2011; Andraka *et al.* 2018; Chen *et al.* 2020). Herein, the period of each simulation was equivalent to the estimated SRT (Table 2).

$$S_j^i = \frac{\dfrac{Y_{i1} - Y_{i0}}{Y_{i0}}}{\dfrac{P_{j1} - P_{j0}}{P_{j0}}} \tag{3}$$

Parameters optimization

Sensitive kinetic and stoichiometric parameters of ASM1 were optimized aiming at minimum prediction errors for state variables on simulations, according to the WWTP monitoring data (Lim *et al.* 2012). Response Surface Methodology (RSM) and Central Composite Design (CCD) were conducted using Minitab Statistical Software (Version 20.2), whose input consisted of the difference between the observed data recorded by the WWTP (Table 1) and the result of simulation.

Regression models were created for each modeling state variable, and ANOVA and standardized effects were used to evaluate the influence of changes in sensitive parameters on prediction errors of effluent TSS, COD, and NH$_3$. Regression models' adequacy was verified using the Lack-of-fit test. Also, the correlation between results of simulations on STOAT and predictions of the regression models was analyzed, as well as the *RMSE* (Equation (4)) of each regression model (Kim Rao & Yoo 2009; Lim *et al.* 2012; Ahn *et al.* 2014).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_i - Y_i)^2}{n-1}} \tag{4}$$

X is the simulation prediction using STOAT, Y is the regression prediction via RSM, and n is the number of simulations designed by CCD.

In the multiple optimization, the desirability function based on a target value was used to estimate optimal values for sensitive parameters, which would lead to minimum prediction errors of effluent TSS, COD, and NH$_3$ (Kim Rao & Yoo 2009; Lim *et al.* 2012). According to the modeling objective, a range of minimum prediction error for each state variable was assumed: ± 5 mg/l for TSS and COD, and ± 1 mg/l for NH$_3$ (Rieger *et al.* 2013).

Finally, optimized parameters of ASM1 were applied on simulations on STOAT for model calibration and validation. It was verified whether or not the results complied with the previous range of prediction errors determined for each state variable. The period of each run for parameter optimization was also the estimated SRT (Table 2).

RESULTS AND DISCUSSION

Sensitive parameters of ASM1

The sensitivity analysis required 20 runs on STOAT, which is considered a fast way to evaluate sensitive model parameters. The present study identified six sensitive parameters of ASM1 on modeling using domestic wastewater data from the São João Navegantes WWTP (Figure 2). COD presented no sensitive parameters; meanwhile, NH$_3$ was the most affected state variable due to changes in the ASM1 kinetic and stoichiometric parameters.
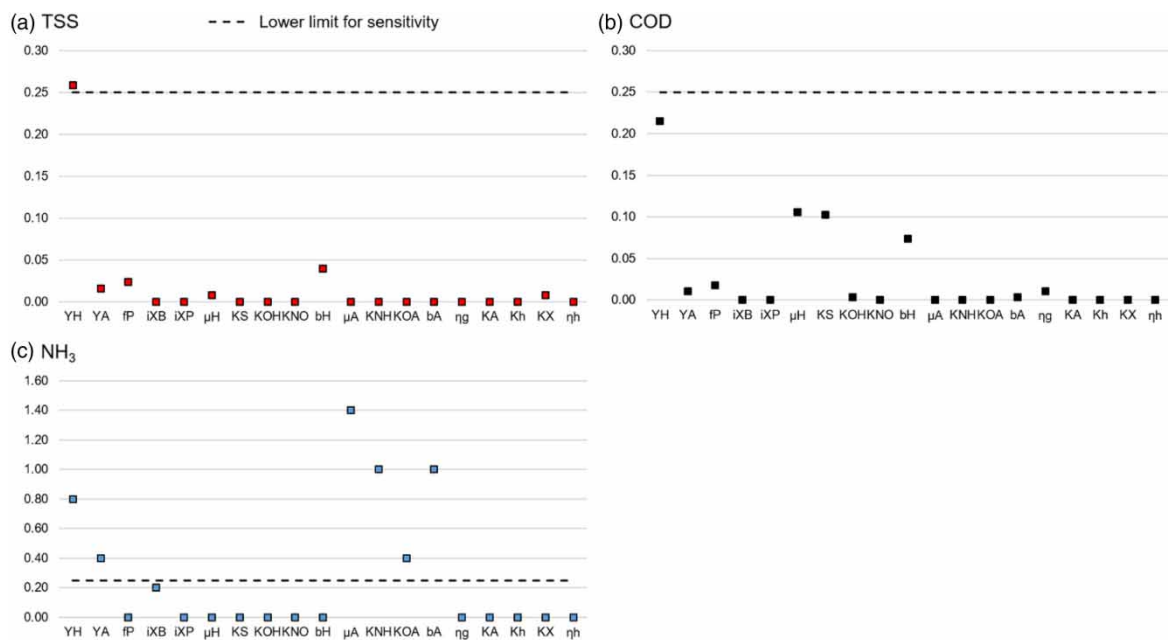

Figure 2 | Sensitivity coefficients calculated for kinetic and stoichiometric parameters of ASM1 for each modeling state variable.

Y_H was sensitive for the TSS and NH_3 outputs. Although the same parameter was not sensitive for COD, it presented a coefficient of 0.22, indicating an effect on that state variable. Lim *et al.* (2012) observed similar results and commented that the relationship between Y_H and the state variables TSS and COD are related to sludge production. The heterotrophic yield also depends upon the nature of the substrate and the population of microorganisms performing degradation (Henze *et al.* 1987).

Other sensitive parameters only affected NH_3 predictions. In particular, the sensitivity coefficients of μ_A, K_{NH}, and b_A presented high values (≥ 1), characterizing a substantial effect on the NH_3 outputs of the WWTP AS process modeling. Contrarily, μ_A also affected effluent TSS in another study (Lim *et al.* 2012). Nevertheless, the high sensitivity of μ_A may be related to elevated ammonia loads and to a lack of complete nitrification (Levy 2007). This parameter is the most critical one for characterizing the growth of the autotrophic biomass and is influenced by many environmental factors, such as pH and temperature (Henze *et al.* 1987). On the other hand, the value of b_A is considered difficult to measure and may be assumed (Levy 2007), as well as the values of Y_A and K_{OA} that also affected effluent NH_3 (Figure 2(c)). K_{NH} and Y_H are recommended to be evaluated for each type of wastewater (Henze *et al.* 1987).

It is noteworthy that 13 parameters of ASM1 were not sensitive for the predictions of effluent TSS, COD, and NH_3 of the São João Navegantes WWTP, regarding the specific modeling conditions. It meant that fewer parameters need to be evaluated, which sped up the ASM1 calibration. However, even though the parameters I_{XB}, μ_H, k_S, and b_H were not considered sensitive, they deserved attention as long as they affected COD and NH_3 predictions (Figure 2(b) and 2(c)).

Half of the six sensitive parameters identified (Y_H, μ_A, and b_A) were also subject to change in another study that used ASM1 modeling results of 18 WWTPs in Europe, three in Asia, and one in North America (Hauduc *et al.* 2011). Other authors have confirmed the relevance of Y_H, μ_A, K_{NH}, and K_{OA} for either steady-state or dynamic AS simulation conditions (Petersen *et al.* 2003). Furthermore, sensitive parameters determined for steady-state conditions can significantly support the dynamic calibration (Liwarska-Bizukojc *et al.* 2011). Therefore, the results of the present study indicate a set of sensitive parameters of ASM1, which besides being representative of a subtropical climate region, may be used in dynamic modeling practices.

Central composite design

CCD resulted in 90 cases randomly designed to find optimal values for sensitive parameters that minimize the prediction errors of effluent TSS, COD, and NH_3. The levels of the test were coded and the ranges are shown in Table 5.

Table 5 | Experimental design for parameter optimization

Sensitive parameters	Range of test		
	−1	0	+1
Y_H	0.98	1	1.02
Y_A	0.2	0.24	0.28
μ_A^a	0.0101	0.0126	0.0151
K_{NH}	1	1.25	1.5
b_A^a	0.0048	0.006	0.0072
K_{OA}^a	0.4	0.5	0.6

Note: (a) Value for 15 °C (h^{-1}) as displayed on STOAT.

The experimental ranges of Y_A, K_{NH}, and K_{OA} were close to the default values of these parameters (Table 4). It corroborates other findings, as changes in the Y_A aren't usual, and half-saturation coefficients (K_{NH} and K_{OA}) depend on environmental conditions and don't require significant changes (Hauduc *et al.* 2011). However, Y_H, μ_A, and b_A were tested in different ranges (Table 5), which may be related to the variability of such parameters (Hauduc *et al.* 2011).

ASM1 is a deterministic model and repeated runs result in the same outputs. Since RSM is a statistical model, it requires variability. Thus, random changes were performed on the other parameters which affected the state

variables but were not considered sensitive by the 0.25 limit (Figure 2) to provide such variability. These changes were made in 13 out of 14 repeated cases designed by CCD, and the parameters changes were i_{XB}, b_H, μ_H, and K_S (Table 6).

Table 6 | Alternate parameters changes to provide variability in repeated cases of CCD

Case (run)	i_{XB}	b_H	μ_H	K_S
78	0.0825	Default	Default	40
79	0.075	Default	Default	Default
80	0.075	Default	0.1179	40
81	0.09	Default	0.1179	Default
82	0.09	Default	Default	60
83	0.085	Default	0.1179	Default
84	0.0825	0.0237	Default	Default
85	0.075	0.0237	Default	Default
86	0.08	Default	Default	Default
87	0.08	Default	0.1179	40
88	0.0825	Default	0.1179	60
89	0.09	Default	0.1179	60
90	0.085	Default	0.1179	40

Note: default values are presented in Table 4.

Regression models

The prediction errors of effluent TSS, COD, and NH_3 were calculated from the results of runs designed by CCD, which means the difference between the observed data recorded by the WWTP and the STOAT simulation's prediction. Such prediction errors were input onto RSM to generate the regression models for each state variable. The regression coefficients estimated by RSM are presented in Table 7. The respective non-

Table 7 | Regression models coefficients of modeling state variables generated by RSM

Term	TSS	COD	NH_3
Constant	3.270	6.085	−0.047
Y_H	10.062	11.973	−0.1021
Y_A	0.703	0.835	0.5201
μ_A	0.342	0.405	−3.1580
K_{NH}	−0.055	−0.065	0.4933
b_A	−0.256	−0.303	2.4448
K_{OA}	−0.061	−0.073	0.5749
Y_H*Y_H	1.3441	0.9401	–
Y_A*Y_A	0.3216	−0.3249	–
$\mu_A*\mu_A$	0.1966	−0.4742	1.1151
$K_{NH}*K_{NH}$	0.3278	−0.3180	–
b_A*b_A	0.2816	−0.3724	0.3826
$K_{OA}*K_{OA}$	0.3241	−0.3224	–
$Y_A*\mu_A$	–	–	−0.2970
μ_A*b_A	–	–	−1.2083
μ_A*K_{OA}	–	–	−0.2789
b_A*K_{OA}	–	–	0.2458

significant quadratic relationships and interactions between parameters weren't considered to simplify the regression models.

ANOVA showed similar results for TSS and COD. Changes in most sensitive parameters affected TSS and COD predictions ($p < 0.001$). Only K_{NH} e k_{OA} didn't affect those state variables (TSS: $F_{(1,89)} = 0.22$; $p = 0.639$ and $F_{(1,89)} = 0.28$; $p = 0.598$, respectively; COD: $F_{(1,89)} = 0.36$; $p = 0.552$ and $F_{(1,89)} = 0.45$; $p = 0.505$, respectively). All quadratic relationships also influenced TSS and COD, especially Y_H*Y_H for both variables ($F_{(1,89)} = 197.60$ and 109.12, respectively; $p < 0.001$). However, interactions between sensitive parameters weren't statistically significant for TSS and COD and were disregarded in their regression models (Table 7).

The ANOVA of the NH_3 regression model showed that Y_H was the only sensitive parameter that did not affect the predictions of such state variable ($F_{(1,89)} = 1.69$; $p = 0.198$). Although only quadratic relationships of μ_A and b_A were statistically significant in that regression model ($F_{(1,89)} = 303.27$ and 35.70, respectively; $p < 0.001$), four interactions between parameters affected effluent NH_3 prediction errors: $Y_A*\mu_A$; μ_A*b_A; μ_A*K_{OA} and b_A*K_{OA} (Table 7). Response surfaces of these significant interactions are shown in Figure 3.
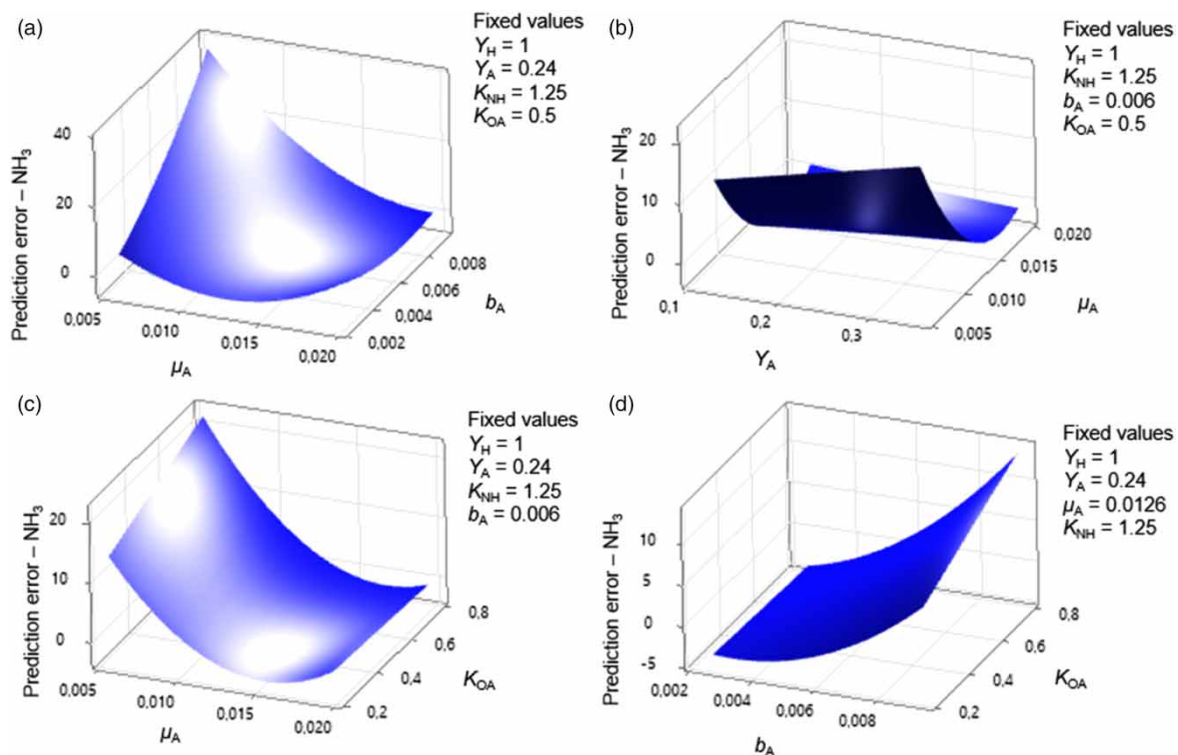

Figure 3 | Response surfaces of significant interactions between sensitive parameters in the regression model for NH_3 prediction errors: (a) μ_A*b_A; (b) $Y_A*\mu_A$; (c) μ_A*K_{OA}; and (d) b_A*K_{OA}.

The relevant impact of μ_A*b_A interaction ($F_{(1,89)} = 189.17$; $p < 0.001$) indicates the many combinations that may be assumed for such parameters to estimate minimum prediction errors for effluent NH_3 (Figure 3(a)). Hauduc *et al.* (2011) highlighted the variability of values for those parameters. Such a relationship is usually inversely proportional, as reduction of μ_A and b_A increases and decreases, respectively, ammonia effluent concentration. Also, an increase in b_A may inhibit cell growth and replace nutrients in the system, elevating NH_3 concentration. However, the effect of b_A is not the most important for effluent NH_3 control (Levy 2007).

In the interactions of $Y_A*\mu_A$ ($F_{(1,89)} = 11.43$; $p = 0.001$) and μ_A*K_{OA} ($F_{(1,89)} = 10.08$; $p = 0.002$), the prediction errors of effluent NH_3 generated by the regression models were null for higher tested values of μ_A (Figure 3(b) and 3(c)). Changes in Y_A and K_{OA} on such interactions didn't show a distinct effect on prediction errors, as observed by other authors (Petersen *et al.* 2002).

Finally, the interaction between b_A*K_{OA} ($F_{(1,89)} = 7.83$; $p = 0.006$) was the less significant in the NH_3 regression model. In that relationship, the whole experimental range tested for K_{OA} was suitable to obtain minimum effluent NH_3 prediction errors (Figure 3(d)), though only lower values adopted for b_A worked appropriately.

The standardized effects of the regression models pointed out the influence of Y_H on TSS (87%) and COD (109%) prediction errors (Figure 4(a) and 4(b)). Linear and quadratic relationships of other parameters presented effects lower than 15%, even though they were significant in the ANOVA results.
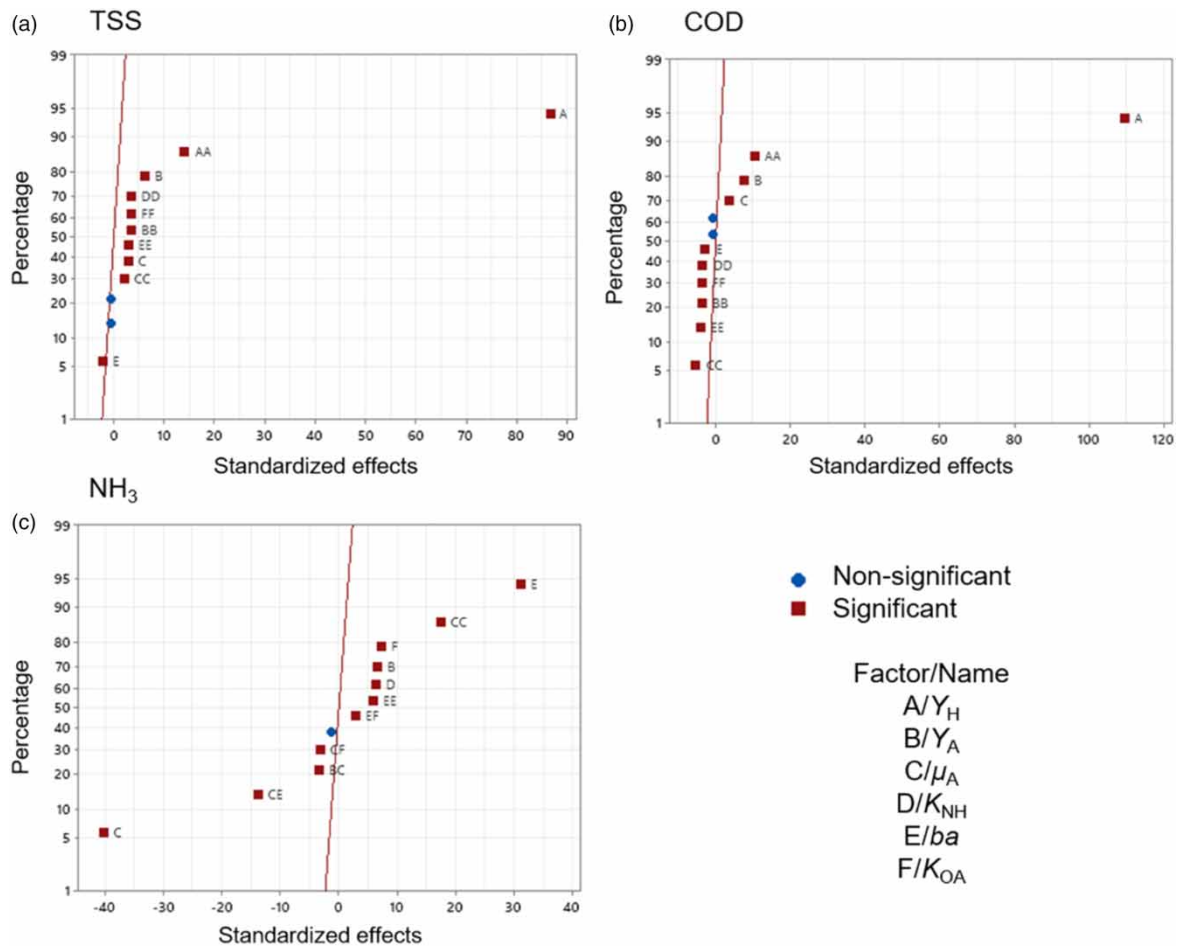

Figure 4 | Standardized effects of RSM regression models for modeling state variables (α=0.05).

For NH$_3$ prediction errors, μ_A and b_A stood out and indicated contrary effects on responses: -40 and 31%, respectively (Figure 4(c)). This situation is also shown in Figure 3(a).

The lack-of-fit tests showed that p-values were not statistically significant for regression models of TSS and COD, which means that the alternative hypothesis that indicates the lack of fit of those regression models wasn't selected. On the other hand, the NH$_3$ regression model presented opposite results, revealing that such a model didn't show the goodness of fit (Table 8).

Table 8 | Measures of the regression model's adequacy

	TSS	COD	NH$_3$
Lack-of-fit test	$F_{(64,89)}=0.06$; $p=1$	$F_{(64,89)}=0.11$; $p=1$	$F_{(64,89)}=15.51$; $p<0.001$
R²	0.9902	0.9938	0.977
RMSE	0.97	0.91	0.65

For that specific case, to correct the lack of fit of the regression model, the original variable predicted on simulation using STOAT was used (prediction of NH$_3$ effluent concentration instead of the prediction error). Then, it

was expected to obtain only positive values, which could be transformed to improve the fit of the regression model. However, it didn't show different results, and the first regression model was maintained (Table 7).

The estimated prediction errors of all state variables provided by RSM were compared with those predicted using STOAT. Such relationships presented a strong correlation ($R^2 > 0.97$), as shown in Figure 5. Thus, the regression models explain at least 97% of the variations in prediction errors of effluent TSS, COD, and NH_3. The statistical error via *RMSE* presented lower values; that is, close to 0, indicating that there wasn't a discrepancy between prediction errors estimated by RSM and predicted via STOAT. Those results are similar to the findings of other authors (Kim *et al.* 2009; Ahn *et al.* 2014).
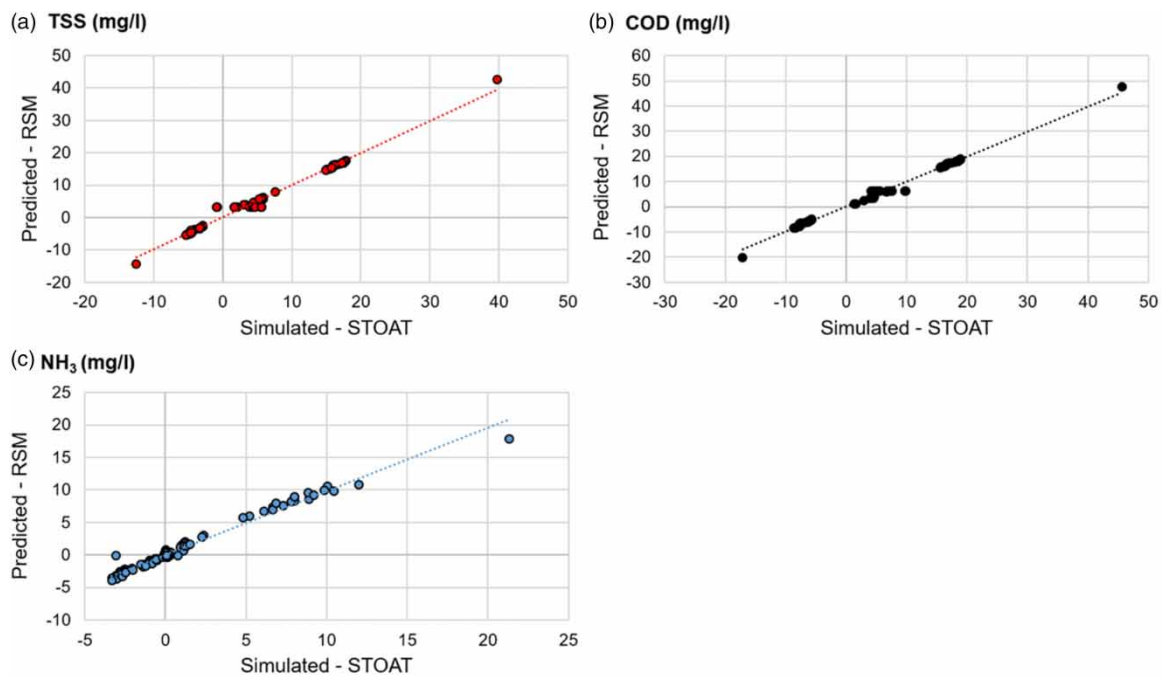

Figure 5 | Correlation between prediction errors via RSM and prediction errors via STOAT for modeling state variables.

Parameter optimization

The generated regression models of RSM were used to estimate optimal values for the sensitive parameters to be applied on the calibration and validation of ASM1. The target value onto the desirability function in the optimization procedure was zero to estimate parameters to reach minimum prediction errors of effluent TSS, COD, and NH_3.

RSM optimization generated the output shown in Figure 6. The optimal values for the sensitive parameters are displayed in red within a specific range. The optimal prediction errors for each state variable are in blue. The vertical lines in red show the optimal parameters within the experimental range of CCD (Table 5).

Figure 6 also presents the relationships between each parameter and the modeling state variables. Changes in Y_H mainly affected prediction errors of effluent TSS and COD. Variations in the values of the other sensitive parameters affected NH_3 prediction errors. It's important to emphasize the effect of μ_A and b_A, whose interaction was the most significant in the ANOVA results of the NH_3 regression model. Nonetheless, μ_A showed two optimal points, on the central and upper regions of the parameter experimental range. The first one was more prominent and consisted of the optimal value determined for the parameter. That situation may be related to the great importance of μ_A for the NH_3 predictions (Petersen *et al.* 2002), as shown in Figure 3(a)–3(c).

The single value of D, which lies between zero and one, is the overall desirability of the combined response levels. D close to one means that the balance of the properties becomes more feasible. On the other hand, D close to zero indicates that one of the response variables is unacceptable (Lim *et al.* 2012). The composite desirability showed $D = 0.9181$, pointing out a balance among the responses regarding the optimal values estimated for
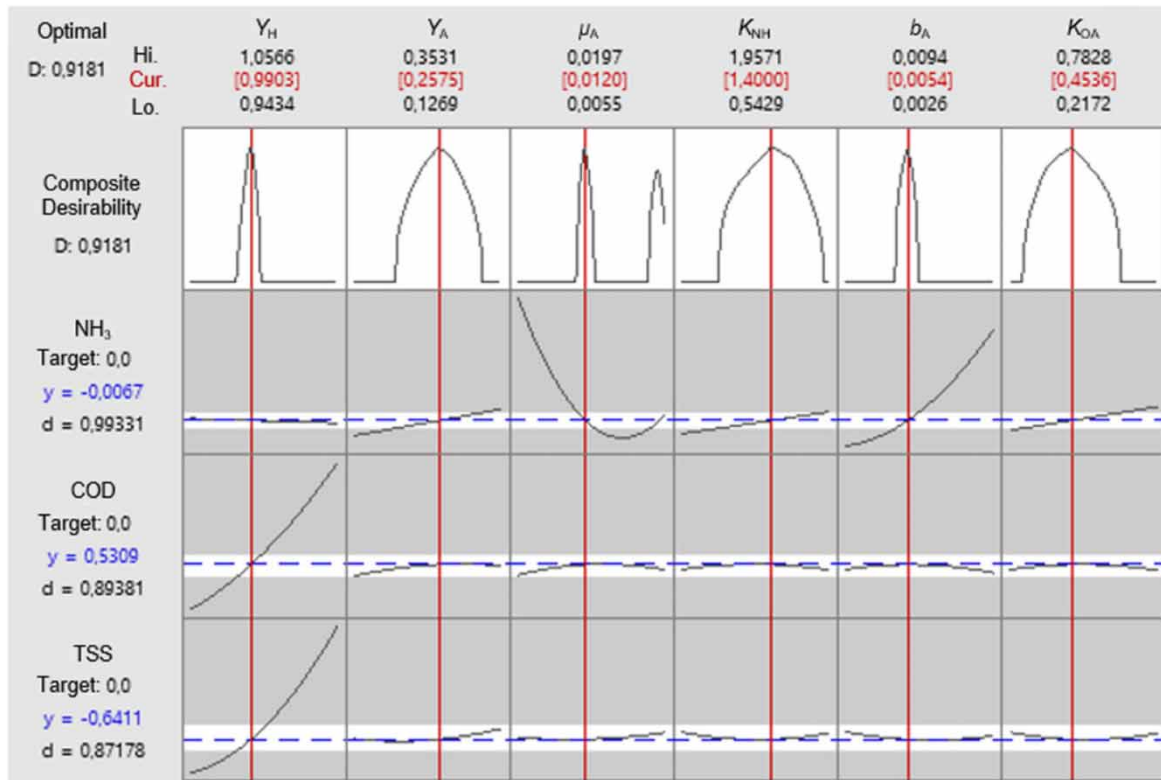
Figure 6 | The composite desirability (*D*) and optimal values estimated for the sensitive parameters of ASM1.

the sensitive parameters (Kim *et al.* 2009; Ahn *et al.* 2014). Each state variable also presented an acceptable value of *d* (>0.87). Those *d*s may also be related to the lower and upper limits of prediction error selected for the state variables in the optimization procedure (± 5 mg/l for TSS and COD, and ± 1 mg/l for NH_3). Then, the higher the limit (i.e., for TSS and COD), the higher the estimated prediction error, even though all three were within ± 1 mg/l (in blue in Figure 6).

Moreover, the NH_3 desirability function (*d*=0.99331) was the highest among all state variables, besides presenting the lowest *RMSE* (Table 8). Despite the lack of fit of the NH_3 regression model, *d* and *RMSE* values are strong reasons to use this regression model in the parameter optimization (Kim *et al.* 2009; Ahn *et al.* 2014).

The optimal value estimated for Y_H was 0.9903, placed between the minimum and central levels of the experimental design (Table 5). That value is relatively higher than the default one (0.67) (Henze *et al.* 1987). Also, other authors optimized the same parameter and had a maximum value of 0.67 (Hauduc *et al.* 2011). Baptista (2020) used Y_H 0.45 on AS modeling of the same WWTP of study. However, that author considered a period of monitoring data between 2006 and 2010 which presented lower values of effluent COD (43 mg/l) and TSS (15 mg/l), compared to the means obtained in the present study for 2018 (COD: 52 mg/l; TSS: 30 mg/l) (Table 1). Such difference may explain the lower value of heterotrophic yield used on ASM1 calibration by that author.

Furthermore, the optimal value calculated in the present study for Y_H is close to 0.9380, defined in the optimization performed by Lim *et al.* (2012). The authors also worked on AS modeling using ASM1, even though the physical conditions (flow, sludge wastage, and RAS) were lower than the records of the São João Navegantes WWTP. This could have accounted for the difference in the Y_H value. In addition, the experimental range tested by those authors was also lower, and the optimal value estimated in that study comprised the higher level of test; that is, close to the one found in the present study. Those results highlight the variability of Y_H and its relationship with the influent wastewater composition (Henze *et al.* 1987; Hauduc *et al.* 2011).

The other sensitive parameters of ASM1 evaluated via RSM are related to NH_3 predictions, mainly μ_A and b_A, whose changes had a considerable effect on that state variable (Figure 6). The optimal values estimated for those parameters were 0.012 h^{-1} (15° C) and 0.0054 h^{-1} (15° C), respectively. These values were converted to the default unit of ASM1 using Equation (2) and resulted in 0.5 d^{-1} (20 °C) and 0.22 d^{-1} (20 °C). The optimal

value estimated for μ_A remained in the usual range presented in the literature (0.2–1) (Ahn *et al.* 2014). For b_A, the optimal estimative was higher than the maximum value observed (0.05–0.15) (Henze *et al.* 1987).

The optimal values of Y_A 0.2575 and K_{OA} 0.4536 (Figure 6) were similar to the default values of such parameters: 0.24 and 0.4, respectively (Ahn *et al.* 2014). On the other hand, K_{NH} 1.4 also showed a higher value than the usual range of 0.75–1 (Hauduc *et al.* 2011).

Table 9 shows the comparison of observed data, default results, and verification of parameter optimization results. Prediction error ranges and prediction errors estimated by RSM and via simulations on STOAT are also presented.

Table 9 | Observed and predicted means of effluent concentrations, and prediction errors of modeling state variables

	TSS	COD	NH$_3$
Observed in 2018 (WWTP)[a]	30	52	4.4
Prediction error range[b]	±5	±5	±1
Prediction error via RSM[c]	−0.64	0.53	0
Using default parameters[d]			
Predicted mean	13	28	0.5
Prediction error	−17	−24	−3.89
Using optimal parameters[c]			
ASM1 calibration	31	51	4.25
Prediction error	+1	−1	−0.14
Observed in 2019 (WWTP)[a]	21	45	3.7
ASM1 validation	114	141	3.8
Prediction error	+93	+96	+0.1

Units in mg/l.
Notes: (a) Table 1; (b) Rieger *et al.* (2013); (c) Figure 6; (d) Table 4.

The predictions of the simulation using default parameters resulted in high prediction errors for all state variables (>45%) compared to the observed data of 2018. Effluent NH$_3$ prediction of that simulation presented the highest variation (90%). Such difference was also found by Baptista (2020); meanwhile, for TSS and COD, the author obtained higher prediction errors compared to the observed data of that study. Therefore, the simulation results using default parameters reinforce the importance of ASM1 calibration for every type of wastewater. Most parameters of ASM1, especially kinetic ones, vary accordingly to the wastewater characteristics throughout space and time (Von Sperling *et al.* 2020).

The ASM1 calibration used the optimal parameters, resulting in approximately null prediction errors. These errors were different from the ones estimated via RSM, though they were within the range selected for each state variable. Thus, ASM1 was calibrated by applying the optimal parameters estimated via RSM.

The ASM1 validation was performed to verify the applicability of the estimated optimal parameters in simulation using data of a different period; that is, 2019 (Table 1). The results of validation presented around null prediction error for effluent NH$_3$ (Table 9). However, high differences were obtained for TSS and COD.

Such differences are related to the increase in Y_H compared to the parameter default value (Table 4); that is, 0.67 to 0.99. Even though the regression models of TSS and COD presented goodness of fit, the values tested for Y_H (Table 5) may not have considered the whole experimental domain of the parameter. Thus, a new experimental range may be evaluated in association with other ranges determined for the other sensitive parameters identified in the present study.

Another explanation for the discrepancy observed consists in the differences in effluent concentrations of TSS and COD between 2018 and 2019, used on ASM1 calibration and validation, respectively. The calibration data presented higher values of TSS and COD compared to 2019 (9 and 7 mg/l higher, respectively), requiring an increase in the heterotrophic yield to reduce the prediction errors for both state variables. The effluent concentration of NH$_3$ was also higher in 2018, though it presented a slight difference related to 2019 (0.7 mg/l higher).

The increase in Y_H raises the COD transformation efficiency, determining the biomass concentration in the reactor. Hence, it has a high effect on the microbial growth rates and on the sludge production (Petersen *et al.* 2002; Levy 2007; Chen *et al.* 2020). Therefore, adjustments in the sludge wastage method may lead to obtaining a suitable Y_H value for ASM1 calibration and validation, aiming at minimum prediction errors of TSS and COD for the specific modeling conditions of the present study.

After that, the sludge wastage method adopted in the present study; that is, fixed-rate over variable time, may also be the reason for the TSS and COD oscillation peaks observed in the ASM1 calibration and validation simulations (Figure 7).
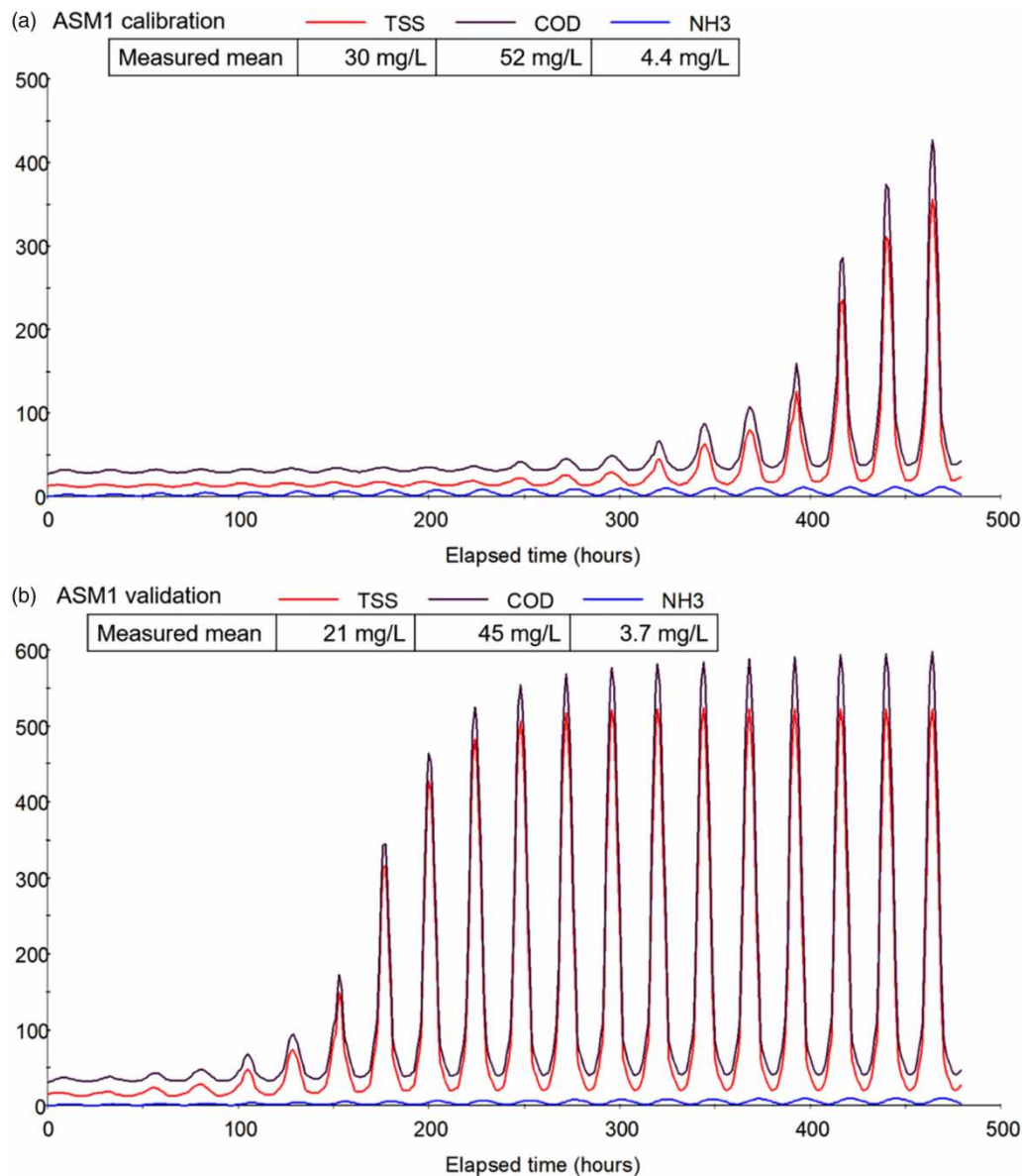

Figure 7 | Effluent concentration (mg/l) of the modeling state variables predicted on STOAT using estimated optimal parameters. Measured means are from the WWTP's monitoring dataset (Table 1).

As estimated values of the AS operation informed by the WWTP (Table 2) were used to determine the sludge wastage method, it may have affected the process performance. Furthermore, the SRT estimate was also based on WWTP data and is linked to the period determined for simulations (20 d=480 h). Baptista (2020) modeled the same WWTP in steady-state conditions and applied 10 days as the period of simulation. In the ASM1 calibration

of the present study, the oscillation peaks were at the end of the period; hence, after 10 days of simulation (Figure 7(a)). Although in the ASM1 validation the peaks started after 100 hours (Figure 7(b)); that is, less than 10 days of simulation, the SRT estimate may have been overrated. Therefore, accurate AS operation data would aid in improving process performance on modeling. Also, other methods of sludge wastage (constant or fixed time over variable rate), and changes in operational parameters of ASM1 and Version 3 may contribute to control predictions of effluent TSS and COD.

Moreover, such peaks always happened when the Y_H upper level of test was chosen in the simulations performed via CCD (Table 5) (see Results of CCD simulations in Supplementary Material). Then, those peaks are also related to the increase of Y_H and its influence on the biomass concentration in the treatment process (Levy 2007).

The heterotrophic yield sensitivity is highlighted in optimizing ASM1 parameters (Hauduc et al. 2011). It's important to determine the value of such parameter for every type of wastewater, as Y_H is commonly changed in ASM1 modeling (Henze et al. 1987) and each WWTP should be modeled assuming specific conditions related to its reality and challenges (Borzooei et al. 2019). Then, the wastewater physical and chemical composition, especially temperature, may have influenced Y_H hence the predictions of effluent TSS and COD on STOAT using the optimal value estimated for the parameter. Also, the results of the present study support that attention should be dedicated to the calibration of Y_H.

Otherwise, in addition to the high variation of NH_3 prediction in the simulation using default parameters (Table 9), all six sensitive parameters affected this state variable (Figure 2(c)). Five of these parameters are related to the autotrophic microorganisms (Y_A), N content, and specifically to ammonia (μ_A, K_{NH}, b_A, and k_{OA}). The optimal values estimated for all of them were appropriate for both calibration and validation of ASM1, and the NH_3 prediction errors were lower than 1 mg/l. Therefore, such parameters are related to wastewater concerning subtropical climate conditions being a reference for other applications.

CONCLUSIONS

AS process of a large-scale WWTP within a subtropical climate region in Brazil was modeled using standard monitoring data set, the ASM1 model, and the STOAT program.

With sensitive analysis, six parameters of ASM1 were identified as sensitive to effluent TSS, COD, and NH_3 predictions: two stoichiometric (Y_H and Y_A) and four kinetic (μ_A, K_{NH}, b_A, and k_{OA}). Sensitive parameters were evaluated and optimized via RSM aiming at minimum prediction errors for modeling state variables. Y_H mainly affected TSS and COD predictions; meanwhile, NH_3 predictions were influenced by changes in the other sensitive parameters.

Estimated optimal parameters were used on ASM1 calibration resulting in approximately null prediction errors for effluent TSS, COD, and NH_3. In the ASM1 validation, the same was obtained for NH_3, even though high prediction errors were observed for TSS and COD. Such errors are related to an increase in Y_H through parameter optimization that highlights the sensitivity of this parameter on ASM1 modeling.

The parameters optimized in the present study, mainly Y_A, μ_A, K_{NH}, b_A, and k_{OA}, whose optimal values were appropriate for effluent NH_3 predictions in the ASM1 calibration and validation, constitute references for other studies on modeling using wastewater data from a subtropical climate region. In particular, those optimal parameters may be effective on either steady-state or dynamic AS modeling using the STOAT program and monitoring data of large-scale WWTP in the mentioned climate conditions.

The quantity and quality of the WWTP monitoring data set used to characterize the AS process and wastewater, and the typical ratios adopted at influent fractionation restrict the application of the optimized parameters. Also, factors related to the sinusoidal simulation pattern and the sludge wastage method consist of specific modeling conditions and deserve attention.

The optimal value estimated for Y_H needs to be evaluated considering tests on a new experimental range and/or adjustments in the sludge wastage method.

The optimal parameters may be tested on a ASM1 cross-calibration/validation; that is, data of validation (2019) for calibration and calibration (2018) for validation; on dynamic modeling studies; and applied at the modeling of similar WWTPs. It is recommended to use the optimal parameters for evaluating the simulation period and different sludge wastage methods on ASM1 and STOAT modeling. Furthermore, it's important to assess and optimize

Version 3 parameters and other operational parameters of AS process on modeling using the estimated optimal ASM1 parameters to improve the removal of TSS, COD, and NH₃.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

REFERENCES

Ahn, J. Y., Chu, K. H., Yoo, S. S., Mang, J. S., Sung, B. W. & Ko, K. B. 2014 Determination of optimal operating factors via modeling for livestock wastewater treatment: comparison of simulated and experimental data. *International Biodeterioration and Biodegradation* **95**(PA), 46–54. https://doi.org/10.1016/j.ibiod.2014.04.014.

ANA – AGÊNCIA NACIONAL DE ÁGUAS 2020 *Atlas Esgotos – Atualização da Base de Dados de Estações de Tratamento de Esgotos no Brasil (Wastewater Treatment Plants of Brazil Data Set Update)*. Brasília, p. 44. Available from: https://www.snirh.gov.br/portal/centrais-de-conteudos/central-de-publicacoes/encarteatlasesgotos_etes.pdf/view (accessed 31 May 2021).

Andraka, D., Piszczatowska, I. K., Dawidowicz, J. & Kruszynski, W. 2018 Calibration of activated sludge model with scarce data sets. *Journal of Ecological Engineering* **19**(6), 182–190. https://doi.org/10.12911/22998993/93793.

Baptista, M. T. 2020 *Avaliação dos modelos matemáticos ASM1 e ASM3 calibrados com dados de monitoramento padrão de um sistema de lodos ativados (Evaluation of the ASM1 and ASM3 Models Calibrated with Standard Monitoring Data of an Activated Sludge System)*. Thesis (Master in Water Resources and Environmental Sanitation), Hydraulic Research Institute, Federal University of Rio Grande do Sul, Porto Alegre, p. 96. Available from: https://lume.ufrgs.br/handle/10183/217074 (accessed 8 February 2021)

Barroso Júnior, J. C. A. 2020 *Avaliação de lagoas de tratamento com presença de macrófitas flutuantes e microalgas aplicadas ao pós-tratamento de esgoto sanitário em condições de clima subtropical (Evaluation of Lagoons Presenting Floating Macrophytes Applied at Domestic Wastewater Post-Treatment in Subtropical Climate Conditions)*. Thesis (PhD in Water Resources and Environmental Sanitation), Hydraulic Research Institute, Federal University of Rio Grande do Sul, Porto Alegre, p. 165. Available from: https://lume.ufrgs.br/handle/10183/211282 (accessed 14 April 2021)

Borzooei, S., Amerlinck, Y., Abolfathi, S., Panepinto, D., Nopens, I., Lorenzi, E., Meucci, L. & Zanetti, M. C. 2019 Data scarcity in modelling and simulation of a large-scale WWTP: stop sign or a challenge. *Journal of Water Process Engineering* **28**, 10–20. https://doi.org/10.1016/j.jwpe.2018.12.010.

Brasil. Ministério do Desenvolvimento Regional. Secretaria Nacional de Saneamento – SNS 2019 *Sistema Nacional de Informações sobre Saneamento (National System of Sanitation Information): 25° Diagnóstico dos Serviços de Água e Esgotos*. Brasília, p. 183. Available from: http://www.snis.gov.br/diagnosticos (accessed 30 July 2021)

Chen, W., Dai, H., Han, T., Wang, X., Lu, X. & Yao, C. 2020 Mathematical modeling and modification of a cycle operating activated sludge process via the multi-objective optimization method. *Journal of Environmental Chemical Engineering* **8**(6), 104470. https://doi.org/10.1016/j.jece.2020.104470.

Hauduc, H., Rieger, L., Ohtsuki, T., Shaw, A., Takács, I., Winkler, S., Héduit, A., Vanrolleghem, P. A. & Gillot, S. 2011 Activated sludge modelling: development and potential use of a practical applications database. *Water Science and Technology* **63**(10), 2164–2182. https://doi.org/10.2166/wst.2011.368.

Henze, M. & Comeau, Y. 2008 Wastewater Characterisation. In: *Biological Wastewater Treatment: Principles, Modelling and Design*, Vol. 7 (Henze, M., Van Loosdrecht, M. C. M., Ekama, G. A. & Brdjanovic, D., eds.). IWA Publishing, London, pp. 33–52.

Henze, C., Grady Jr, C. P. L., Gujer, W., Marais, G. V. R. & Matsuo, T. 1987 *Activated Sludge Model No1*. International Association on Water Pollution Research and Control, London, p. 33.

Kim, M., Rao, A. S. & Yoo, C. 2009 Dual optimization strategy for N and P removal in a biological wastewater treatment plant. *Industrial and Engineering Chemistry Research* **48**(13), 6363–6371. https://doi.org/10.1021/ie801689t.

Levy, A. L. L. 2007 *Modelagem e análise de sensibilidade do processo de tratamento de lodo ativado com reciclo (Modeling and Sensitivity Analysis of the Activated Sludge Process Treatment with Return Sludge)*. Thesis (Master in Science in Chemical Engineering), Federal University of Rio de Janeiro, Rio de Janeiro, p. 140. Available from: http://portal.peq.coppe.ufrj.br/index.php/dissertacoes-de-mestrado/2007-1/ (accessed 28 April 2021)

Lim, J. J., Kim, M. H., Kim, M. J., Oh, T. S., Kang, O. Y., Min, B., Rao, A. S. & Yoo, C. K. 2012 A systematic model calibration methodology based on multiple errors minimization method for the optimal parameter estimation of ASM1. *Korean Journal of Chemical Engineering* **29**(3), 291–303. https://doi.org/10.1007/s11814-011-0178-2.

Liwarska-Bizukojc, E., Olejnik, D., Biernacki, R. & Ledakowicz, S. 2011 Calibration of a complex activated sludge model for the full-scale wastewater treatment plant. *Bioprocess and Biosystems Engineering* **34**(6), 659–670. https://doi.org/10.1007/s00449-011-0515-1.

Petersen, B., Gernaey, K., Henze, M. & Vanrolleghem, P. A. 2002 Evaluation of an ASM1 model calibration procedure on a municipal-industrial wastewater treatment plant. *Journal of Hydroinformatics* **4**(1), 15–38. https://doi.org/10.2166/hydro. 2002.0003.

Petersen, B., Gernaey, K., Henze, M. & Vanrolleghem, P. A. 2003 Calibration of Activated Sludge Models: A Critical Review of Experimental Designs. In: *Biotechnology for the Environment: Wastewater Treatment and Modeling* (Agathos, S. N. & Reineke, W., eds.), Waste Gas Handling. 3C ed. Springer, Dordrecht, pp. 101–186. https://doi.org/10.1007/978-94-017-0932-3_5

Pistorello, J. 2018 *Simulação do co-tratamento de resíduo de tanque séptico em Estação de Tratamento de Esgoto Doméstico (Simulation of Septic Tank Waste co-Treatment in Domestic Wastewater Treatment Plant)*. Thesis (Master in Water Resources and Environmental Sanitation), Hydraulic Research Institute, Federa University of Rio Grande do Sul, Porto Alegre, p. 123. Available from: https://lume.ufrgs.br/handle/10183/174129 (accessed 8 February 2021)

Rieger, L., Gillot, S., Langergraber, G., Ohtsuki, T., Shaw, A., Takács, I. & Winkler, S. 2013 *Guidelines for Using Activated Sludge Models*. IWA Publishing, London, p. 281. Available from: https://www.iwapublishing.com/books/9781843391746/guidelines-using-activated-sludge-models (accessed 25 November 2020)

Rössle, W. H. & Pretorius, W. A. 2001 A review of characterisation requirements for in-line prefermenters paper 1: wastewater characterisation. *Water SA* **27**(3), 405–412. https://doi.org/10.4314/wsa.v27i3.4985.

Sarkar, U., Dasgupta, D., Bhattacharya, T., Pal, S. & Chakroborty, T. 2010 Dynamic simulation of activated sludge based wastewater treatment processes: case studies with Titagarh Sewage Treatment Plant, India. *Desalination* **252**(1–3), 120–126. https://doi.org/10.1016/j.desal.2009.10.014.

Tchobanoglous, G., Burton, F. L. & Stensel, H. D. 2014 *Wastewater Engineering: Treatment and Resource Recovery*, 5th edn. McGraw-Hill Higher Education, New York, p. 2018.

Von Sperling, M. 1994 A new unified solids flux-based approach for the design of final clarifiers: description and comparison with traditional criteria. *Water Science and Technology* **30**(4), 57–66. https://doi.org/10.2166/wst.1994.0157.

Von Sperling, M., Verbyla, M. E. & Oliveira, S. M. A. C. 2020 *Assessment of Treatment Plant Performance and Water Quality Data: A Guide for Students, Researchers and Practitioners*. IWA Publishing, London, UK, p. 644.

WRC PLC 1994 *WRC STOAT: Tutorials Guide*. Swindon, UK, p. 74. Available from: https://www.wrcplc.co.uk/ps-stoat (accessed 10 February 2021)