

Optimal design and operation of activated sludge processes: State-of-the-art

Rainier Hreiz, M A Latifi, Nicolas Roche

▶ **To cite this version:**

Rainier Hreiz, M A Latifi, Nicolas Roche. Optimal design and operation of activated sludge processes: State-of-the-art. Chemical Engineering Journal, 2015, 281, pp.900 - 920. 10.1016/j.cej.2015.06.125 . hal-01458423

HAL Id: hal-01458423

https://amu.hal.science/hal-01458423

Submitted on 17 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Chemical Engineering Journal 281 (2015) 900–920

Contents lists available at ScienceDirect

Chemical Engineering Journal

journal homepage: www.elsevier.com/locate/cej

Review

Optimal design and operation of activated sludge processes: State-of-the-art

Rainier Hreiz [a,b,*], M.A. Latifi [a], Nicolas Roche [b]

[a] Laboratoire Réactions et Génie des Procédés, CNRS-ENSIC, Université de Lorraine, BP-20451, 1 rue Grandville, 54001 Nancy Cedex, France
[b] Aix-Marseille University – CNRS – Centrale Marseille, Laboratoire de Mécanique, Modélisation et Procédés Propres – M2P2 – UMR7340, Europôle de l'Arbois, 13545 Aix en Provence Cedex 4, France

HIGHLIGHTS

- The paper deals with the optimal control and design of activated sludge processes.
- Studies on the topic are critically reviewed.
- Focus is set on the problem formulation that allows one getting reliable solutions.
- The selection of appropriate cost functions and constraints is discussed.
- Aspects related to dealing with model mismatch are presented.

ARTICLE INFO

Article history:
Received 3 April 2015
Received in revised form 23 June 2015
Accepted 25 June 2015
Available online 9 July 2015

Keywords:
Wastewater treatment
Activated sludge
Dynamic optimization
Multi-objective optimization
Pareto front
Modeling

ABSTRACT

The activated sludge process (ASP) is the most commonly used process for wastewater treatment. Improving its performance is necessary from economic and environmental point of views. In this context, dynamic optimization is a powerful tool for assisting engineers in determining optimal operations and designs for ASPs.

However, the real optimality of the solution strongly depends on the optimization problem statement, for which unfortunately, there is no standard or commonly accepted formulation. In a hopeful attempt to provide a guideline for future works on the topic, this paper reviews the literature devoted to optimal control and design of ASPs. The main issues to be addressed in order to get reliable solutions are discussed, among which: (1) Managing the inevitable mismatch between the model predictions and the real ASP operation. (2) Dealing with the unpredictable variations in the wastewater characteristics. (3) Accounting for the slowest dynamic processes occurring in ASPs. (4) Appropriately selecting the decision variables and the flowsheet structure in order to simplify the problem formulation from a mathematical perspective. (5) Conveniently choosing the cost functions expressions/correlations. (6) Successfully selecting the mathematical constraints in order to guarantee physically relevant operations.

© 2015 Elsevier B.V. All rights reserved.

Contents

1. Introduction . . . 901
2. Conventional ASPs configurations . . . 904
3. Optimization . . . 904
   3.1. Single-objective optimization . . . 904
   3.2. Optimization algorithms . . . 907
   3.3. Multiple objective optimization . . . 908
      3.3.1. The Pareto front . . . 908
      3.3.2. Methods for computing the Pareto front . . . 908

* Corresponding author at: Laboratoire Réactions et Génie des Procédés, CNRS-ENSIC, Université de Lorraine, 1 rue Grandville, 54001 Nancy Cedex, France. Tel.: +33 (0) 383 175 179.
E-mail address: Hreiz.Rainier@gmail.com (R. Hreiz).

http://dx.doi.org/10.1016/j.cej.2015.06.125
1385-8947/© 2015 Elsevier B.V. All rights reserved.

4. Optimal control/design of ASPs: problem formulation . . . 908
4.1. Process model . . . 908
4.1.1. Biochemical reactions modeling . . . 909
4.1.2. Settler modeling . . . 909
4.2. Dealing with the mismatch between the model predictions and the real process behavior . . . 910
4.2.1. Process model mismatch . . . 910
4.2.2. Influent characteristics . . . 910
4.2.3. ASP controllability . . . 911
4.2.4. Long-term ASP performance . . . 911
4.3. Effectiveness of various optimization techniques . . . 911
5. Optimal control/design of ASPs: objective functions, flowsheets structures and DVs . . . 912
5.1. Objective function(s) . . . 912
5.1.1. Economic criterion . . . 912
5.1.2. Other criteria . . . 912
5.2. Flowsheet structure/process configuration . . . 912
5.3. Decision variables . . . 913
5.3.1. Continuous time-independent decision variables . . . 913
5.3.2. Continuous time-varying decision variables . . . 913
5.3.3. Integer and binary decision variables . . . 913
6. Cost functions formulation . . . 913
6.1. Exploitation costs . . . 914
6.1.1. Aerators power consumption . . . 914
6.1.2. Pumps power consumption . . . 915
6.1.3. Mixing power consumption . . . 915
6.1.4. Additives costs . . . 915
6.1.5. Sludge disposal/valorization . . . 916
6.1.6. Effluent fines . . . 916
6.2. Investment/capital costs . . . 916
6.3. Maintenance costs . . . 917
7. Constraints . . . 917
7.1. Discharge constraints . . . 917
7.2. Concentration of suspended solids in the bioreactors . . . 917
7.3. Duration of the aeration-on/aeration-off periods in sequentially aerated bioreactors . . . 918
7.4. Dissolved oxygen concentration in the bioreactors . . . 918
7.5. Alkalinity . . . 918
7.6. Sludge retention time . . . 918
7.7. Reactors size . . . 918
8. Conclusion . . . 919
Acknowledgments . . . 919
References . . . 919

1. Introduction

A wastewater treatment plant, WWTP, is an industrial facility where a combination of mechanical, physical, chemical and biological processes is used to achieve pollutants removal from the incoming wastewater. These pollutants are various, and in the case of some industrial wastewaters, they may require specific treatment strategies/techniques: this review paper will focus only on biological nutrient removal from sewage water, i.e. domestic and municipal wastewaters. In such WWTPs, a primary treatment ensures the separation of particulate pollutants (debris, sand, oils, grease and particulate wastes), in particular, thanks to gravitational separation in a large settling tank (referred to as primary settler). However, colloidal and dissolved wastes are not sufficiently removed by these separation processes, and therefore, must be eliminated by a secondary treatment stage. In the case of sewage, characterized by a low COD (chemical oxygen demand) to BOD_5 (five-day biological oxygen demand) ratio, biological secondary treatment is the most cost-effective process. Nowadays, given its efficiency, relatively simple operation and low cost, the activated sludge process (ASP) is by far the most widely used for sewage biological secondary treatment.

The infrastructure of a basic ASP (Fig. 1) consists of: (1) a single bioreactor operated continuously where suspended microorganisms consume the colloidal and dissolved organic matter. The reactor is aerated to provide dissolved oxygen (DO) for aerobic biodegradation. Bacteria consume one part of the colloidal and dissolved carbonaceous compounds to satisfy their energetic needs (catabolism), and synthesize another part – along with a small proportion of ammonium and phosphorus – into new cellular tissues (anabolism). (2) A settling tank (referred to as secondary settler or clarifier) where activated sludge (flocculated biomass) is gravitationally separated from the treated wastewater. The effluent overflows into the receiving waterbody, but in some WWTPs, it may undergo additional treatments (e.g. filtration and disinfection) before being discharged. A minimum solids retention time (SRT, also referred as 'sludge age') of about 3 days is required for bacteria to aggregate into flocs (bio-flocculation) so as to enable their gravitational separation. (3) A sludge recycle line returning the major proportion of the settled sludge to the bioreactor, thus allowing

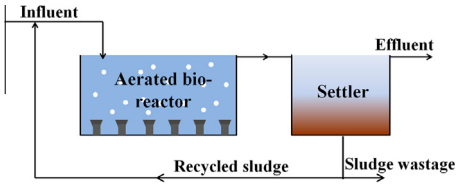


Fig. 1. The basic ASP.

Table 1
Summary of studies dealing with the SOO of ASPs operation. 'Discharge', 'TSS', Sequences durations', 'DO', 'Alkalinity', 'SRT', and 'Reactors volumes' refer to constraints discussed in Sections 7.1–7.7 respectively.

	[1]	[2]	[3]	[4]	[6]	[7]	[9]	[10]	[12]	[14]	[26]
Process configuration	Fed-batch SBR	AAS	Fed-batch SBR	4 STRs in series	AAS	Pre-denitrification	AAS	A^2O	Superstructure (5 STRs in series)	Superstructure (5 STRs in series)	Pre-denitrification
Biological model	ASM1	Linearized ASM1 (calibrated)	Simplified ASM1 (calibrated)	ASM1 (calibrated)	ASM1	ASM1	ASM3	ASM2d	ASM1	ASM1	ASM1
Settler model	–	–	–	Not specified	Takács [31]	Takács [31]	Takács [31]	Takács [31]	Takács [31]	Takács [31]	Takács [31]
Open-/Closed-loop problem	Open-loop	Closed-loop (RTO)	Open-loop	Closed-loop	Open-loop	Open-loop	Open-loop	Closed-loop	Closed-loop	Closed-loop	Closed-loop
Steady/Dynamic influent	Dynamic feed	Dynamic influent	Dynamic feed	Dynamic influent	Dynamic influent	Steady influent / Dynamic influent	Steady influent	Dynamic influent	Dynamic influent	Dynamic influent	Dynamic influent
Decision variables type	Includes time-varying DVs	Includes time-varying DVs	Includes time-varying DVs	All DVs are continuous and time-independent	Includes time-varying and discrete DVs	All DVs are continuous and time-independent / Includes time-varying DVs	Includes time-varying DVs	Includes time-varying DVs	Includes discrete DVs	Includes discrete DVs	All DVs are continuous and time-independent
Optimization algorithm	Sequential Quadratic Programming (SQP)	Not specified	SQP	Rosenbrock algorithm	SQP	Interior-point method	Genetic algorithm	Pattern search	Several algorithms were tested	Several algorithms were tested	Several algorithms were tested
Objective function	Min. batch time	Min. aeration-on duration	Min. batch time	Min. air supply to the reactors	Min. aeration-on duration	Min. aeration energy / Min. a weighted sum of aeration energy and pollutants discharge	Min. aeration-on duration	Min. operating costs	Min. a weighted sum of operating costs and PI controllers' ISE	Min. a weighted sum of operating costs and PI controllers' ISE	Min. a weighted sum of the PI controllers' ISE
Technical and operational constraints (related to the ASP)	Discharge	Discharge, sequences durations	Min. batch time	Discharge, sequences durations	Discharge, sequences durations	Discharge	Discharge, DO, sequences durations	–	–	–	–

	[30]	[33]	[49]	[50]	[51]	[54]	[56]	[59]	[60]	[74]	[80]
Process configuration	Superstructure (11 STRs in series)	AAS	AAS	AAS	AAS	A^2O	Superstructure (5 STRs in series) + Primary settler + Sludge treatment facility	Pre-denitrification	Superstructure (5 STRs in series)	Pre-denitrification	AAS
Biological model	ASM2d (calibrated)	ASM1	ASM1	ASM3	ASM1 (calibrated)	Modified ASM2d	PW-AS	ASM1	ASM3	ASM1	ASM1
Settler model	1D reactive settler model	Perfect point-settler	Perfect point-settler	Takács [31]	Point-settler	Takács [31]	Takács [31] for the secondary settler. Point-settler for the primary settler	Takács [31]	Takács [31]	Not specified	Takács [31]
Open-/Closed-loop problem	Open-loop	Open-loop	Open-loop	Open-loop	Open-loop	Closed-loop	Open-loop	Open-loop	Open-loop	Open-loop	Open-loop
Steady/Dynamic influent	Dynamic influent	Dynamic influent	Dynamic influent	Steady influent	Dynamic influent	Dynamic influent	Steady influent	Dynamic influent	Steady influent	Steady influent	Dynamic influent
Decision variables type	Includes discrete DVs	Includes time-varying and discrete DVs	Includes time-varying DVs	Includes time-varying DVs	Includes time-varying and discrete DVs	All DVs are continuous and time-independent	All DVs are continuous and time-independent	Includes time-varying DVs	All DVs are continuous and time-independent	All DVs are continuous and time-independent	Includes time-varying and discrete DVs

Table 1 (*continued*)

	[30]	[33]	[49]	[50]	[51]		[54]	[56]	[59]	[60]	[74]	[80]
Optimization algorithm	- (trial and error)	SQP	DYNO code	SQP	SQP		Pattern search	SQP	SQP	DICOPT algorithm	- (trial and error)	Genetic algorithm
Objective function	Min. operating costs	Min. aeration-on duration	Min. aeration-on duration	Min. aeration-on duration	Min. aeration-on duration	Min. nitrogen discharge	Min. operating costs	Max. the WWTP electrical autonomy	Min. Aeration costs	Min. operating costs	Min. operating costs	Min. a weighted sum of pollutants discharge
Technical and operational constraints (related to the ASP)	–	Discharge, sequences durations	Discharge, sequences durations	Discharge, DO, sequences durations	Discharge, sequences durations	Sequences durations	–	Discharge, TSS, SRT	Discharge	Discharge	–	Sequences durations

to maintain a high bacterial concentration in the reactor so as to intensify the biological nutrient removal. A first benefit of sludge recycling is decoupling the hydraulic retention time (HRT) and SRT: hence, an efficient treatment can be achieved for an HRT of the order of ten hours only. (4) A sludge wastage line at the bottom of the clarifier, from where a small fraction of sludge is withdrawn in order to stabilize the biomass concentration in the bioreactor and to fix an adequate SRT. The excess sludge withdrawn is then treated separately.

Given the large involved expenses in wastewater treatment, a lot of research contributions have aimed at improving the ASP efficiency. As a result, numerous sophisticated processes have been proposed as alternatives to the basic ASP (see Section 2 for details). These modern ASPs include combinations of aerated and non-aerated reactors, recirculation lines between these reactors, and sludge recycle lines between the settler and the bioreactors, etc. Compared to the basic ASP, these advanced treatment processes allow enhancing the biological nutrient removal, especially with respect to phosphorus and nitrogen removal, increasing the process flexibility, and reducing the cost and footprint of the plant.

Nowadays, because of the strengthening of the environmental regulations, it is necessary to improve ASPs operation and design, but also from an economic point of view, to reduce their operating costs. In addition, most of the exiting WWTPs have to undertake major retrofitting to comply with the new regulations. However, even for expert engineers, determining the optimal operating strategy for ASPs remains quite difficult and laborious due to the high non linearity and complexity of the underlying biochemical phenomena and their potential interaction, the large number of operating and design parameters to determine (e.g. ASP configuration, aeration rates and sludge recycle flow rate), and the variety of objectives to deal with (e.g. enhancing the treatment efficiency and minimizing the operating costs).

In this context, dynamic optimization reveals to be a powerful tool for assisting and supporting designers in determining the optimal operating conditions for existing WWTPs, or simultaneously predicting the optimal design (units' sizing) and operation for future plants (i.e. no infrastructure built yet). Optimization refers to the method of identifying operating conditions and/or design parameters that make a system as effective or functional as possible with respect to a given criterion, while satisfying specified constraints. Optimization is formulated as follows: (1) A mathematical model describing the process behavior is provided, e.g. a model linking the degradation rate of organic matter in an ASP to its operating strategy. (2) The designer selects a – or many – criterion expressing the process's feature(s) to be optimized, e.g. minimizing the operating costs of an ASP. This criterion is referred as 'objective function'. (3) The designer specifies the parameters that one can control in order to optimize the objective function, e.g. sludge recycle flow rate, aeration rate and/or reactor volume in an ASP. These parameters are referred as decision variables, DVs. (4) The designer chooses the constraints that should be satisfied by the optimal solution, e.g. effluent quality must comply with regulations. (5) Provided the optimization problem formulation explained above, an optimization algorithm is used to systematically search for the solution(s) (i.e. the DVs values) best meeting the designer's criteria while respecting the predefined constraints.

Numerous studies dealing with optimal control and design of ASPs have been published in the last decade, most of them aiming at determining the most lucrative pollutants abatement strategy. Studies dealing with the optimal operation of existing ASPs are summarized in Table 1, those addressing the optimal control and design of future ASPs are summarized in Table 2, while Table 3 compiles papers devoted to the multi-objective optimization of ASPs. These papers reveal that optimization allows finding operating conditions/designs that greatly enhance the ASP performance, a

fact that has been confirmed experimentally in some of these studies [1–4]. Thus, the optimal control/design of ASPs addresses an important academic-industrial research issue.

However, one should be aware that the real optimality of the computed solution strongly depends on the optimization problem formulation. For example, the preference between alternative operating strategies is closely related to the objective functions used like cost functions, for which there is unfortunately no commonly approved expressions. Moreover, there is no model that can exactly predict the complex biological phenomena occurring in ASPs. Therefore, engineers may lack confidence regarding the reliability of the computed solution and even the true potential of optimization.

In an attempt to settle these misgivings and to provide a general guideline to help researchers choosing an appropriate and reliable formulation for their optimization problem, this paper reviews the literature dealing with the optimal control and design of ASPs. Section 2 overviews the most common ASPs' configurations and their operating principles, and Section 3 presents some general basics of optimization: this should help researchers in the field of optimization to better understand issues related to wastewater treatment and vice versa. Section 4 considers the formulation of problems addressing the optimal control/design of ASPs, and discusses some important issues that should be considered in the problem formulation in order to reduce the inevitable mismatch between the model prediction and the real process behavior, so as to get solutions of industrial relevance, i.e. applicable in practice to real-scale ASPs. Section 5 reviews general aspects relative to the choice of objective functions, flowsheets structures and DVs. It is discussed how a convenient selection of the DVs allows to greatly simplify the optimization problem. Section 6 critically reviews the principal cost functions (i.e. the formulations of economic objective functions) that have been used so far in the literature. Their limits and potentials are discussed. Finally, Section 7 summarizes the major technical and operational constraints employed in the literature. This review paper is hopefully expected to provide a guideline for future works on optimal control and design of ASPs.

2. Conventional ASPs configurations

The basic ASP, described above, allows an efficient removal of carbonaceous organics and ammonium/ammonia (S_{NH}) from wastewater. However, due to the increasing concerns about eutrophication, stricter regulations forced its upgrading to improve nitrogen and phosphorus removal. For this purpose, and thanks to the advances in microbiology, the classical ASP has evolved into several more sophisticated processes of which some are described below.

Concerning the enhanced nitrogen removal, the two-steps nitrification-denitrification process revealed to be the most economical abatement strategy. Nitrification is carried out by nitrifying autotrophic bacteria under aerobic conditions, and consists in the oxidation of S_{NH} into nitrate/nitrite (S_{NO}). Since the growth rates of nitrifying bacteria are low, they require a minimum SRT of about 10 days to establish themselves. Denitrification is the process of S_{NO} reduction into nitrogen gas, and is achieved by some heterotrophic bacteria under anoxic conditions, i.e. DO depletion but S_{NO} presence as a substitute terminal electron acceptor. Carbonaceous matter is consumed by bacteria under both aerobic and anoxic conditions.

In large-size WWTPs, anoxic and aerobic conditions are carried out in separate reactors. The pre-denitrification (Fig. 2a) and the post-denitrification (Fig. 2b) configurations are the most widespread. In the pre-denitrification design, the anoxic tank(s) is placed upstream of the aerobic reactor(s). Therefore, internal recirculation of the nitrified liquor to the anoxic reactor is needed, which involves additional pumping costs. In the post-denitrification configuration, the anoxic tank(s) is placed downstream of the aerobic reactor(s). This configuration allows a better S_{NO} removal efficiency than the pre-denitrification process. However, it is possible that denitrification continues to occur within the settler, and hinders sludge settling due to the generation of nitrogen bubbles. Moreover, since carbonaceous substrates are consumed in the aerated tank(s), they may become limiting in the anoxic reactor(s). Therefore, the addition of carbonaceous matter (methanol, acetate, and ethanol are the most commonly used) may be required in order to achieve an efficient denitrification.

In most small-size WWTPS, nitrification and denitrification are carried out in the same bioreactor, i.e. are shifted in time by running the aerators sequentially to create periodically alternating aerobic and anoxic conditions. This Alternating Activated Sludge (AAS) process is the most widespread in small-size WWTPs. Indeed, since it requests the same infrastructure than the basic ASP, it allows a cost-efficient retrofitting of existing WWTPs: only a modest addition of equipment such as controllers may be needed.

Concerning phosphorous, the conventional ASP leads to a removal efficiency around 10–30%. Enhanced phosphorus removal can be achieved by chemical precipitation, or by exposing the biomass to alternating anaerobic and aerobic conditions, which intensifies phosphorus uptake by some microorganisms. The most common ASP configuration enabling enhanced nitrogen and phosphorus removal is the A^2O process, which consists of an Anaerobic tank followed by an Anoxic and an Oxic (i.e. aerobic) reactor.

Finally, it is noteworthy that although most current ASPs are run continuously, some are operated in batch mode or semi-continuously as the Sequencing Batch Reactor (SBR) process. Such discontinuous processes (as well as the AAS) are inherently dynamic and require specific formulations of the optimization problem as explained later. For details about modern ASPs configurations, the reader is invited to consult specialized text books, e.g. [5].

3. Optimization

3.1. Single-objective optimization

A general single-objective optimization problem, SOOP, can be stated as follows:

$$\underset{\boldsymbol{u},t_f}{\text{Min}} J(\boldsymbol{x},\boldsymbol{u},t_f) \tag{1}$$

Subject to:

$$\text{Process model equations}: \boldsymbol{f}(\boldsymbol{x},\dot{\boldsymbol{x}},\boldsymbol{u},\boldsymbol{t}) = \boldsymbol{0} \tag{2}$$

$$\text{Initial conditions}: \boldsymbol{x}(t_0) = \boldsymbol{x}_0 \tag{3}$$

$$\text{Equality constraints}: \boldsymbol{h}(\boldsymbol{x},\boldsymbol{u}) = \boldsymbol{0} \tag{4}$$

$$\text{Inequality constraints}: \boldsymbol{g}(\boldsymbol{x},\boldsymbol{u}) \leqslant \boldsymbol{0} \tag{5}$$

$$\text{Decision variables bounds}: \boldsymbol{u}_L \leqslant \boldsymbol{u} \leqslant \boldsymbol{u}_U \tag{6}$$

where t is time and t_f the time horizon. $\boldsymbol{x}$ represents the set of process state variables (and $\dot{\boldsymbol{x}}$ its time derivative). Although the above formulation considers the state variables as differential, it is important to notice that algebraic and integer (and binary) variables can be handled as well. However, discrete variables give rise to differentiability issues and thus, problems involving such variables are difficult to solve using deterministic techniques (Section 3.2). $\boldsymbol{u}$ is the

Table 2
Summary of studies dealing with the SOO of ASPs operation and design.

	[8]	[11]	[13]	[21]	[27]	[28]	[29]
Process configuration	Superstructure (2 STRs and 2 settlers)	Superstructure (up to 3 plug-flow reactors in series)	Superstructure (2 STRs and 2 settlers)	A^2O	A^2O	Primary settler + Basic ASP	Basic ASP
Biological model	A simple model	ASM1	A simple model	ASM2d (calibrated) simplified using a MLA	ASM3 + EAWAG Bio-P (calibrated) simplified using a MLA	ASM3	ASM1
Settler model	Simple 1D model	Perfect point-settler	Simple 1D model	Not specified	Not specified	Point-settler model with variable separation efficiency	ATV
Open-/Closed-loop problem	Closed-loop	Open-loop	Closed-loop	Open-loop	Open-loop	Open-loop	Open-loop
Steady/Dynamic influent	Dynamic influent	Steady influent	Dynamic influent	Dynamic influent	Dynamic influent	Steady influent	Steady influent
Decision variables type	All DVs are continuous and time-independent	Includes discrete DVs	All DVs are continuous and time-independent	All DVs are continuous and time-independent	All DVs are continuous and time-independent	All DVs are continuous and time-independent	All DVs are continuous and time-independent
Optimization algorithm	Several algorithms were tested	Simulated annealing	Several algorithms were tested	Accelerating genetic algorithm	Accelerating genetic algorithm	GAMS software	Interior-point method
Objective function	Min. a weighted sum of: operating costs, capital costs and PI controller's ISE	Min. pollution discharge	Min. a weighted sum of: operating costs, capital costs and PI controller's ISE	Min. Pollutants discharge	Min. Pollutants discharge	Min. Capital + operating + maintenance costs	Min. Capital + operating + maintenance costs
Technical and operational constraints (related to the ASP)	Numerous constraints were used	–	Numerous constraints were used	–	–	Reactor volume, discharge	Discharge, DO, TSS

	[34]	[35]	[36]	[38]	[39]	[40]	[41]	[57]
Process configuration	Primary settler + A^2O	Basic ASP	Basic ASP	Superstructure (up to 5 STRs in series)	Superstructure (up to 7 STRs in series)	Alpha process	RDN and DRDN	Regional wastewater system: 3 cities, a sewer system and 3 decentralized WWTPs
Biological model	ASM3 + EAWAG Bio-P module	ASM1	ASM1	ASM3	ASM3 + EAWAG Bio-P module	ASM1	ASM1 (calibrated)	–
Settler model	Point-settler model with variable separation efficiency	Takács [31] + ATV	Takács [31] + ATV	Takács [31]	Takács [31]	Perfect point-settler	Point-settler (separation efficiency was not reported)	–
Open-/Closed-loop problem	Open-loop	Open-loop	Open-loop	Open-loop	Open-loop	Open-loop	Open-loop	Open-loop
Steady/Dynamic influent	Steady influent	Steady influent	Steady influent	Steady influent	Steady influent	Steady influent / Dynamic influent	Steady influent	Steady influent
Decision variables type	All DVs are continuous and time-independent	All DVs are continuous and time-independent	All DVs are continuous and time-independent	All DVs are continuous and time-independent	All DVs are continuous and time-independent	All DVs are continuous and time-independent	All DVs are continuous and time-independent	All DVs are continuous and time-independent
Optimization algorithm	GAMS software	SQP	Interior-point method	Generalized Reduced Gradient (GRG)	GRG	GRG2	GRG2	Genetic algorithm
Objective function	Min. Capital + operating + maintenance costs	Min. Capital + operating + maintenance costs	Min. Capital + operating + maintenance costs	Min. Capital + Operating costs / Min. Operating costs	Min. Capital + Operating costs / Min. Operating costs	Min. Total reactors volume	Min. total reactors volume / Min. Capital + operating costs	Min. Capital + operating costs
Technical and operational constraints (related to the ASP)	Reactor volume, discharge	Discharge, DO	Reactor volume, Discharge, DO	Discharge	Discharge	Discharge, DO, TSS	Discharge, DO, TSS	–

Table 3
Summary of studies dealing with the MOO of ASPs operation and/or design.

	[18]	[19]	[20]	[42]	[43]	[44]
Process configuration	5 STRs in series	5 STRs in series	Prim. Settler +7 STRs in series + a sludge treatment facility	Basic ASP	Basic ASP	Pre-denitrification
Biological model	ASM3	ASM3	ASM3	Simple model (calibrated)	ASM1	ASM1
Settler model	Takács [31]	Takács [31]	Takács [31]	Point-settler	Takács [31] + ATV	Takács [31]
Open-/Closed-loop problem	Closed-loop	Closed-loop	Open-loop	Open-loop	Open-loop	Closed-loop
Steady/Dynamic influent	Steady influent	Steady influent	Steady influent	Steady influent	Steady influent	Dynamic influent
Decision variables type	All DVs are continuous and time-independent	All DVs are continuous and time-independent	All DVs are continuous and time-independent	All DVs are continuous and time-independent	All DVs are continuous and time-independent	All DVs are continuous and time-independent
Optimization algorithm	Controlled Random Search algorithm	Controlled Random Search algorithm	A global optimizer (no details are reported)			
Method for computing the Pareto front	– (NIMBUS interactive method)	– (NIMBUS interactive method)	– (NIMBUS interactive method)	Multi-objective genetic algorithm: NSGA II	Multiple-objective genetic algorithm: MEGA	Multi-objective genetic algorithm: NSGA II
Objective functions	*J*1: Min. S_{NH} discharge *J*2: Min. aeration power consumption *J*3: Min. sodium bicarbonate addition	*J*1: Min. S_{NH} discharge *J*2: Min. aeration power consumption *J*3: Min. sodium bicarbonate addition	*J*1: Max Nitrogen removal rate *J*2: Min. Aeration energy *J*3: Min. External carbon addition *J*4: Min. Sludge production *J*5: Max. Gas production of the anaerobic digester	*J1*: Max. the influent wastewater throughput *J2*: Min. BOD_5 discharge *J3*: Min. Operating costs	*J1*: Min. Operation + Investment + Maintenance costs *J2*: Min. Pollutants discharge	*J1*: Min. Operation costs *J2*: Min. Pollutants discharge
Technical and operational constraints (related to the ASP)	Alkalinity	Alkalinity	Discharge, TSS	–	Discharge, DO	–

	[47]	[48]	[52]	[53]	[55]	[58]	[81]
Process configuration	Superstructure (1 STR)	Superstructure (4 STRs in series)	Biodenipho process	5 STRs in series	A^2O	Storm tank + basic ASP + Sewer system + River	A^2O
Biological model	ASM1	ASM1	Model from the literature	ASM1	Extended ASM2d	ASM1	ASM2d
Settler model	Takács [31]	Takács [31]	Not specified	Takács [31]	Takács [31]	Takács [31]	Takács [31]
Open-/Closed-loop problem	Open-loop	Open-loop	Open-loop	Open-loop	Closed-loop	Closed-loop	Closed-loop
Steady/Dynamic influent	Dynamic influent	Dynamic influent	Steady influent	Steady Dynamic influent influent	Dynamic influent	Dynamic influent	Steady influent
Decision variables type	Includes time-varying DVs	All DVs are continuous and time-independent	Includes discrete DVs	All DVs are continuous and time-independent	Includes time-varying DVs	All DVs are continuous and time-independent	All DVs are continuous and time-independent
Optimization algorithm			Single-objective genetic algorithm	Single-objective scatter search algorithm			
Method for computing the Pareto front	A multi-objective genetic algorithm	A multi-objective genetic algorithm	ε-constraint	ε-constraint	– (Random points generation)	Multi-objective genetic algorithm: NSGA II	Multi-objective genetic algorithm: NSGA II
Objective functions	*J*1: Min. operation cost *J*2: Min. Nitrogen discharge	*J*1: Min. operation cost *J*2: Min. Nitrogen discharge	*J*1: Min. Operation + Investment + Maintenance costs *J*2: Min. Nitrogen discharge	*J*1: Min. Operation costs *J*2: Min. Pollutants discharge	*J*1 : Min. Operating costs *J*2: Min. Pollutants discharge *J*3: Min. Risk of microbiology-related solids separation problems	Several objectives related to the water quality in the river	*J*1: Min. operating costs *J*2: Min. Pollutants discharge
Technical and operational constraints (related to the ASP)	Discharge, SRT, DO, Reactor volume	Discharge, SRT, DO, Reactor volume	Discharge	–	–	–	–

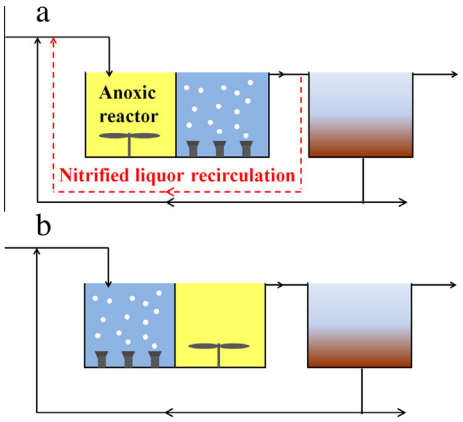


Fig. 2. Two conventional ASP configurations for enhanced nitrogen removal. (a) The pre-denitrification configuration. (b) The post-denitrification configuration.

set of DVs, and $\boldsymbol{u}^L$ and $\boldsymbol{u}^U$ are its lower and upper bounds respectively. J is the objective function to be minimized (or maximized; in that case, negative J is minimized). $\boldsymbol{h}$ and $\boldsymbol{g}$ represent respectively the set of equality and inequality constraints that should be fulfilled by the optimal solution.

The problem is qualified as nonlinear if the objective function, and/or the process model and/or at least one of the constraints is/are nonlinear. Obviously, non-linear problems are much harder to solve than linear ones, and they require the use of specific algorithms. Biochemical phenomena involved in ASPs (and liquid–solid separation in the settler) are highly non-linear.

An optimization problem is considered as dynamic if time is involved in the model, i.e. the system presents a temporal evolution. Otherwise, the problem is said to be static. Dynamic problems are more complicated and time-consuming to solve, since the model is to be integrated over the entire time horizon at each iteration of the optimizer used. ASPs operation is naturally dynamic since the wastewater flow rate and composition vary with time (due to the variation of human/industrial activities during the day, rain events, etc.): indeed, these temporal variations constitute the main source of disturbances of ASPs operation.

In dynamic optimization problems, DVs may be considered as constant over the time horizon (e.g. a constant aeration rate in a reactor), or may be allowed to vary with time (e.g. aeration rate in an AAS). The use of time-varying DVs is mandatory for handling inherently dynamic processes, e.g. AAS or semi-continuous and batch ASPs. However, new considerations arise from such a formulation, since the dimension of the DVs vector becomes infinite (indeed, the optimal values of the DVs are required at every moment). To be solved, such problems must be broken into simpler sub-problems of finite dimension, then, classical optimization principles are applied. In studies dealing with the optimal control of ASPs, the control vector parameterization (CVP) approach has been the most used method to handle such problems (e.g. [6]; only Luo and Biegler [7] have employed another method, the simultaneous collocation approach). The CVP approach consists in dividing the time horizon into a finite number of intervals, and then using simple functions (e.g. piece-wise constant functions) to approximate the DVs temporal variations within these intervals. For illustration, if we consider a problem aiming at determining the optimal aeration rate over a day, the time horizon may be split for example into 24 intervals of 1 hour each. Using piece-wise constant functions to approximate the DVs converts the original complex problem into a classic optimization problem, involving 24 time-independent DVs, i.e. the aeration rates at each hour.

3.2. Optimization algorithms

Optimization algorithms used to solve nonlinear problems can be broadly classified into three categories: local-deterministic methods, global-deterministic techniques, and stochastic/meta-heuristics approaches. Contrary to stochastic algorithms, starting from a particular initial point, deterministic approaches always yield the same solution, i.e. they produce the same series of iterates.

Local-deterministic techniques are iterative methods that proceed by successive improvements of an initial – arbitrarily guessed – solution. They define the search direction based on information about the neighborhood of the current solution, i.e. the evaluation of the gradient. These methods suffer from several drawbacks: (1) Most of them cannot efficiently handle models involving non-differentiability. (2) When dealing with multimodal problems (i.e. when numerous local optima exist, which is the case in most chemical engineering problems), the provided solution will depend on the initially guessed one. To overcome this issue, the so-called multi-start strategy is often used: the optimization procedure is carried out repeatedly, starting from different initial points, hence increasing the probability of finding a better solution. However, mathematically speaking, this procedure does not guarantee reaching the global optimum.

Global-deterministic methods are designed to escape local minima and to converge to the global optimum within a desired tolerance. However, as discussed in [8], to guarantee global optimality, most of these methods have a number of requirements (e.g. twice differentiability of the model) that are not met in many realistic dynamic problems. Moreover, these techniques are relatively hard to implement, and the associated computational effort increases very rapidly with the problem's size.

Stochastic methods are optimization techniques that introduce randomness into the search procedure, and attempt to connect the DVs to the objective function in a black box manner. They incorporate artificial intelligence: they mimic physical or biological phenomena to guide the search into promising zones in the solutions' space. A primary advantage of stochastic techniques is that, unlike most deterministic methods, they do not have much mathematical requirements. Moreover, they are easy to implement. Stochastic methods are regarded as global optimization techniques (since the solution procedure is not 'trapped' in local optima), although they do not guarantee global optimality in finite time. Nevertheless, they usually provide good solutions in moderate computation times. Stochastic algorithms may be combined with local-deterministic techniques: these two-stage optimization methods are known as hybrid algorithms. First, a global search is conducted by the stochastic algorithm to guide the solution to the vicinity of the global optima, i.e. its attraction basin. In a second step, the local search routine is activated to precisely locate that

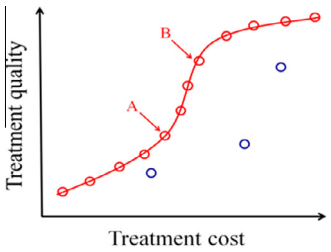


Fig. 3. Schematic representation of a Pareto front: red circles represent Pareto optimal solutions, while blues ones represent dominated solutions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

optimum. Several stochastic techniques were used in studies dealing with optimal operation/design of ASPs: genetic algorithms (e.g. [9]), pattern search (e.g. [10]), simulated annealing [11], ant colony [12], scatter search [13], tabu search [14], etc.

3.3. Multiple objective optimization

Most real-world processes require the simultaneous optimization of several objectives, e.g. minimizing the exploitation costs and maximizing the treatment efficiency of an ASP. These criteria are often conflicting and incommensurable (cannot be expressed in the same units), therefore, they cannot be 'instinctively' combined in a single meaningful objective function. Accordingly, handling such problems using the traditional SOO approach leads to unnecessary simplifications. Indeed, it hides the real nature of the problem since the conflicting aspects between the different objectives are not clearly recognized. Instead, multi-objective optimization (MOO) allows accounting simultaneously and comprehensively for several criteria, and thus provides the designer with information about the trade-offs involved between the different objectives that he/she would like to optimize.

Multi-objective optimization problems (MOOPs) are formulated similarly as SOOPs, but Eq. (1) where the objective function is a scalar is replaced with Eq. (7) where the objective function is a vector:

$$\underset{u,t_f}{\mathrm{Min}}\boldsymbol{J}(\boldsymbol{x},\boldsymbol{u},t_f) = [J_1, J_2, \ldots, J_n]^{\mathrm{T}} \tag{7}$$

where T stands for the transpose, and n for is the number of objectives.

3.3.1. The Pareto front

Since conflicting criteria are tackled together in MOO, there is no unique solution that can simultaneously optimize the set of defined objectives. Therefore, the notion of optimality is redefined and the 'Pareto dominance' concept is used to compare solutions: a feasible solution dominates another one only if it leads to a superior or equal performance with respect to all objective functions. Thus, a solution is said to be – Pareto – optimal if it is not dominated by any other feasible solution. Using the Pareto dominance concept, MOO leads to a set of optimal solutions known as the Pareto front, PF (Fig. 3). A Pareto optimal solution cannot be improved with respect to any objective without impairing at least another one. In other words, for a problem aiming at optimizing the treatment quality and cost of an ASP for example, the PF represents the minimal cost enabling a given treatment efficiency.

From a mathematical perspective, all Pareto optimal solutions are equally optimal, i.e. solution vectors cannot be ordered mathematically. Therefore, a decision maker (DM) should examine these solutions, and based on his/her 'subjective' preferences, selects which one to be implemented on the real-scale ASP (this final solution is referred to as the most preferred solution). The DM is generally the project manager or an expert of the topic. Note that he/she may rely on decision making tools that would help him/her for ranking and comparing alternative solutions. Compared to SOO, a primary advantage of MOO is that it allows benefiting from the DM experience. Indeed, given his/her expertise, he/she will avoid for example solutions that may lead to subnormal operating conditions (e.g. poor sludge settleability or overgrowth of filamentous bacteria). Moreover, it allows the DM to assess the true optimality of each solution. If we consider for example the PF illustrated in Fig. 3, while both points A and B represent optimal solutions, it is clear that the DM would have preference for solution B, since for a small increase in the operating cost, it allows an important improvement of the treatment quality.

3.3.2. Methods for computing the Pareto front

Techniques allowing the computation of the PF can be broadly classified into two classes, scalarization and vectorization approaches. Scalarization techniques consist in converting the original MOOP into a series of parametric SOOPs. The most widespread approach is the weighted sum method. It consists in aggregating all the objective functions within a single objective, using a weighted sum. Hence, Eq. (7) is replaced by:

$$\underset{\boldsymbol{u},t_f}{\mathrm{Min}}(\omega_1 J_1, +\omega_2 J_2 + \cdots + \omega_n J_n) \tag{8}$$

$$\sum_{i=1}^{i=n}\omega_i = 1 \tag{9}$$

$$\omega_i \geqslant 0 \tag{10}$$

Thus, the MOOP is converted into a parametric SOOP: by parametrically varying the weights values, a set of optimal solutions is produced, i.e. the PF can be computed. However, this approach suffers from several drawbacks: it does not allow exploring the non-convex parts of the PF, and moreover, uniformly varying the weights does not necessarily result in equally spaced solutions over the PF.

Contrary to scalarization approaches, vectorization techniques directly tackle the MOOP. The most widespread vectorization methods are based upon genetic algorithms, given that the concept of Pareto optimality can be easily integrated within such algorithms. Since they are a population-based metaheuristic, genetic algorithms are well adapted for handling MOOPs, and moreover, they are not – or little – susceptible to the shape and convexity of the PF. Several variants of MOO genetic algorithms exist. They use variant techniques (e.g. niching, fitness sharing) in order to generate uniformly distributed solutions over the entire PF. The interested reader may refer to [15–17] among others for details about these algorithms.

Finally, the existence of interactive approaches is to be mentioned. These methods allow solving MOOPs without computing the PF. They are computationally efficient since they focus only on the part of the PF that interests the DM. They are iterative techniques where the DM continuously interacts with the solution procedure. At each step, the DM expresses his/her opinion on how the current solution should be improved, and then, a new solution is computed based on his/her preferences. The procedure continues until the DM is satisfied and believes that no better solution can be achieved. Additional details about these methods can be found in [18–20].

4. Optimal control/design of ASPs: problem formulation

This section surveys the formulations of problems addressing the optimal control/design of ASPs. Mathematical models describing the ASPs dynamic behavior are first presented in Section 4.1. Section 4.2 discusses some primary issues that should be addressed in order to get a realistic numerical solution, applicable in practice to the ASP. Finally, Section 4.3 reviews the effectiveness of the various classes of optimization algorithms when applied to such complex problems.

4.1. Process model

In WWTPs, the secondary treatment stage is the most important process considering both performance and involved costs. Accordingly, most studies dealing with optimal control/design of WWTPs have restricted their flowsheet to the ASP solely.

The main phenomena that should be accounted for in a detailed mathematical modeling of ASPs operation are: hydrodynamics phenomena, biochemical reactions and their kinetics, and liquid-solid separation in the settler. However, all papers dealing with optimal control/design of ASPs have treated the bioreactors (and the settler) as stirred tank reactors (STRs), i.e. hydrodynamics effects were neglected. Plug-flow reactors, e.g. carrousel type ditches, were modeled as a series arrangement of continuous STRs, e.g. [21]. Indeed, using such an assumption allows avoiding partial differential equations (PDEs) in this process model: by neglecting spatial gradients, the process behavior will be described by a system of ordinary differential equations, ODEs. In fact, although from a theoretical perspective, models involving PDEs can be handled by optimization, however, they require unaffordable computation times if industrial-scale ASPs are to be dealt with.

4.1.1. Biochemical reactions modeling

Given their good predictive results, ASM models, developed by the International Water Association, are nowadays established as standards by the scientific community to model the dynamic biochemical reactions occurring in ASP reactors. ASM1 [22] and ASM3 [23] address carbon and nitrogen degradation under aerobic and anoxic conditions, while ASM2 [24] and ASM2d [25] account also for phosphorus removal. ASMs track the temporal evolution of the ASP key state variables, such as the concentrations of DO, heterotrophic and autotrophic bacteria, soluble and particulate nitrogen and carbonaceous substrates, etc. Using the STR assumption, mass balances in the reactor for each of these variables give rise to a system of highly non-linear ODEs. ASMs model the conversion processes of each variable, which represent source and sink terms in the dynamic mass balance equations. In the case of the biomass for example, these terms correspond to the bacterial growth and death respectively.

However, ASMs lead to a very complex process model, making the optimization procedure highly time-consuming. Therefore, some authors, e.g. [13], have opted for simpler models, which are however probably less accurate than ASMs. Nonetheless, the use of simple models may be necessary when computation time is an issue, i.e. in the case of real-time optimization (RTO) (see Section 4.2.3). Therefore, Kim and coworkers [2] who considered the RTO of an AAS have used a linearized version of ASM1. Indeed, faster calculations of the optimal operating strategy enable more frequent updates of the optimal trajectory.

To avoid ASMs complexity, some authors [21,26,27] have used surrogate based models: (1) Simulations are conducted with the original complex process model, for different sets of DVs values. (2) The simulations inputs (i.e. DVs values) and outputs (i.e. the objective function value and the fulfillment or not of the specified constraints) are supplied to a Machine Learning Algorithms (MLA) (e.g. artificial neural network). Based on these data, the MLA derives a new model approximating the original one, while being, computationally speaking, much cheaper to evaluate. (3) Optimization is performed using the derived surrogate model, hence allowing a substantial reduction in computation time.

Finally, it is noteworthy that the kinetics of the biochemical reactions depend on the operating temperature. To the authors' knowledge, no study has considered the temporal variability of the ambient temperature (e.g. due to daily and seasonal cycles) in their model. However, El-Shorbagy and coworkers [28] have investigated the optimal design of a basic ASP at different operating temperatures, 10, 20, 30 and 40 °C. Temperature effects were quantified by varying a number of kinetic parameters according to the Arrhenius type of equation, while other reactions rates were supposed to remain unchanged. Results revealed that the optimal designs of the primary and secondary settlers were unaffected by the operating temperature, while the optimal reactor volume was deeply modified. The most economic ASP plant was obtained for an operating temperature of 20 °C. Indeed, according to [28], at low temperatures, since bacterial growth is slow, a big reactor and a large SRT are required to achieve the desired effluent quality. On the other hand, at high temperatures, the increase in the microorganisms' death rate leads to a high concentration of inert particles in the sludge. Hence, greater HRT, SRT and aeration rates are needed to reach the intended treatment efficiency.

4.1.2. Settler modeling

As discussed in [29], the secondary settler plays a crucial role in ASPs: (1) As a thickener, it concentrates the sludge to be recycled to the reactors. (2) As a clarifier, it separates sludge to produce a good quality effluent. (3) As a storage tank, it allows conserving sludge during peakflows. Therefore, an accurate modeling of the settler dynamics is needed for reliable optimal results.

Most settler models have focused only on the liquid-solid separation phenomena, and neglected the biochemical reactions that would occur, especially in the sludge layer at the bottom of the settler. However, these models may be easily coupled with a biological model to enable characterizing bacterial activity within the settler. Nonetheless, for simplification purposes, apart Li and coworkers [30], all studies concerning optimal control/design of ASPs have opted for a non-reactive settler unit.

Nowadays, the model proposed by Takács and coworkers [31] is the most commonly employed for describing the clarification and thickening processes in the secondary clarifier. It assumes a 1D settler consisting of 10 STR layers in series. Mass balances of the state variables are performed in each layer. Sludge settling velocity is assumed to depend only on the local concentration of total suspended solids (TSS) (Kynch theory). It is calculated through a double exponential function, a formulation that accounts for both hindered and flocculated settling. Clarification is considered above the feed layer, and thickening below the inlet level and in the sludge blanket. In layers where thickening occurs, solids sedimentation flux is limited to the flux that can be handled by the layer below (limiting flux concept).

The Takács model is generally regarded as accurate but computationally expensive since it gives rise to numerous ODEs. Moreover, the limited flux concept is modeled using the non-differentiable 'Minimum' function, which in turn can raise issues for deterministic optimization algorithms. Therefore, many studies have opted for simpler models, of which the point-settler approach has been the most widely used. It considers a constant separation efficiency regardless of the separator operating conditions, e.g. incoming flow rate and TSS concentration. Concentrations of soluble components in the effluent are generally assumed to be the same as at the settler inlet, i.e. a nil HRT in the settler is considered. However, delay functions (i.e. plug-flow approximation) may be used to allow better comparison with experimental data [32]. Obviously, the point-settler approach underestimates the settler role and disregards the strong coupling between the settler efficiency and the reactors operation. Nevertheless, as pointed in [33], the ideal point-settler approximation (100% separation efficiency) may reveal sufficient in the case of an over-designed settler.

El-Shorbagy and coworkers [24,28] have used a more sophisticated point-settler model: separation efficiency was not predefined as in the classic approach, but was calculated depending on the incoming flow rate, TSS concentration, and the settler cross-section and height.

Espírito-Santo and coworkers [35,36] have followed the ATV (for 'Abwasser technik verband' in German) design guideline [37]. ATV is a design procedure for the secondary settler, relying on empirical engineering knowledge, and accounting explicitly

for the influence of the settler cross-section and depth. It contemplates rain and storm events which may lead to sludge washing due to the sharp increase in the incoming wastewater flow rate: hence, a portion of the settler volume is dedicated to sludge storage purposes, so as to preserve a sufficient amount of sludge further to a peakflow event. Espírito-Santo and coworkers [35] investigated the influence of three settler modeling approach on the optimal design of a basic ASP: the ATV procedure, the Takács model and a coupled ATV–Takács approach. Then, in order to evaluate the robustness of the three achieved designs, i.e. if they achieve or not an efficient treatment under disturbances, simulations were carried out applying a flow rate five times greater than the one used in the design step. The design achieved using the coupled ATV–Takács approach was shown to handle the best such adverse flow rate conditions given that its effluent quality was the less compromised.

Finally, it is noteworthy that, although newer and more accurate models for sludge separation have been developed, they have not been used yet in studies dealing with optimal operation and design of ASPs. The relatively old Takács model remains the most used (see Tables 1–3) as it seems to provide a good trade-off between the model's complexity and accuracy. Indeed, given the numerous involved uncertainties (e.g. influent, rain events) and the inevitable mismatch between the model and practice (see Section 4.2), the Takács model is probably sufficient for modeling liquid-solid separation in the settler. Moreover, the real issue when modeling the settler operation is the evolution of the kinetic parameters of flocs' settling depending on the operating conditions of the ASP and on the wastewater type. However, to the authors' knowledge, no current model is able to predict such phenomena.

4.2. Dealing with the mismatch between the model predictions and the real process behavior

Problems dealing with optimal control/design of ASPs are quite interesting to address from an academic and a theoretical perspectives given the complexity of the process model. Some authors [12–14] have even considered such challenging problems with the aim of testing the effectiveness of the optimization algorithms that they developed. However, to be of industrial relevance, several aspects relative to the problem formulation should be considered, so as to achieve an optimal solution that is applicable to the real-scale ASP. These issues are discussed in the following paragraphs.

4.2.1. Process model mismatch

As with any process involving biological reactions, mismatches between the model predictions and the real ASP performance are unavoidable. They are mainly due to parameter uncertainties, neglected model dynamics and disturbances. This disparity should be as small as possible so to get a reliable and realistic numerical solution. Therefore, the use of advanced models such as the ASMs and the Takács model [31] should be promoted, since they can follow the process behavior with a relatively good accuracy. However, these models contain a large number of kinetic, stoichiometric and physical parameters (e.g. bacteria growth rates) whose values depend on the wastewater strength and type, as well as on the operating conditions, etc. Therefore, if the default parameters values are used, the optimal solution will reflect a trend rather than a precise prediction of the optimal design/operating strategy. Unfortunately, given the complexity of the involved biochemical processes, there are no reliable correlations to estimate these parameters values based on given operating conditions.

Therefore, if the problem deals with a WWTP to be designed, due to the lack of alternatives, the parameters values should be selected conservatively. Default parameters values may be used, but safety factors should be applied then to the selected design. On the other hand, if the problem copes with the upgrade of an existing facility, the process model should be calibrated from input/output experimental data, i.e. measurement of the influent and effluent compositions. Calibration, i.e. fitting or regression analysis, allows estimating the actual parameters values for a given ASP: it is an optimization procedure for determining the set of parameters values minimizing the difference between numerical results and experimental data. However, for practical identification reasons, based on a sensitivity analysis, only the most sensitive parameters are generally estimated (the reader may refer to [33] among others); default literature values are attributed to the remaining parameters.

Finally, it is noteworthy that once the optimal strategy is applied to the ASP, RTO may be used to deal efficiently with disparities between the model results and the actual system behavior (see Section 4.2.3).

4.2.2. Influent characteristics

To predict an ASP behavior, in addition to the process model, a description of the system inputs, i.e. the influent characteristics, is required. Indeed, the computed optimal solution closely depends on the assumed wastewater flow rate and composition [28,34,38,39], as well as on the temporal variations of these characteristics [40], since they constitute the main source of disturbances of the ASP operation. While daily/weekly/seasonally variations of the wastewater characteristics due to human activity (the result of households' life-cycles and/or industrial activities) are quasi-periodic and can be reasonably estimated over the short and middle term, weather-related perturbations such as precipitation are unpredictable.

Optimal solutions based on time-varying wastewater characteristics (referred to as dynamic influent) are more conservative (and presumably more realistic) than those based on a steady influent assumption [40]. Therefore, typical patterns of disturbances should be accounted for whenever available. Dry weather influent conditions or standard influent characteristics following a rain/storm event may be used, e.g. [10]. It is noteworthy that the temporal variations of the influent characteristics may be approximated using Fourier series (e.g. [32]): such a simplification allows reducing the computational time necessary for integrating the model over the time horizon.

The model integration time may be further reduced if steady influent conditions are assumed: therefore, this simplification has been used in numerous studies (see Tables 1–3). Moreover, if all DVs are time-independent, with the use of such an assumption, it becomes possible to convert the original dynamic problem into a static optimization problem as follows: all time derivatives are considered nil, so the process model becomes a set of equality constraints, and the state variables become a set of additional DVs which values are to be determined by the optimization algorithm [29,35,36,40–43]. Although the static optimization problem involves additional DVs, however, it does not require model integration. Its solution corresponds to the steady-state part of the original dynamic problem's solution.

As noted in [40], if steady influent conditions are adopted, in order to get a realistic solution, critical conditions as a high pollutants load or a low operating temperature should be considered. Otherwise, safety factors should be applied to the optimal solution. Moreover, an a posteriori robustness analysis should be performed: simulations using dynamic influent conditions should be conducted to check whether or not the achieved solution ensures a reasonable plant operation under disturbances [35,44]. In this case, the use of MOO presents an additional benefit, since the DM will take into account the robustness criterion when choosing the most preferred solution.

4.2.3. ASP controllability

Finding the real optimal operation strategy or design of an ASP requires the perfect knowledge of the influent characteristics over a large time horizon. Therefore, given the unpredictability of the influent variations, running the ASP following pre-defined operating conditions does not ensure neither optimal nor adequate functioning. In practice, apart in some small-size WWTPs, a control system is used to counteract disturbances effects.

Control systems require monitoring some process key-variables, e.g. redox potential, DO or S_{NH} concentration in the reactor(s). In the case of AASs for example, a widespread and simple feedback control of aeration consists in turning off the aerators when nitrification is achieved, i.e. when the concentration of S_{NH} or DO in the reactor reaches a predefined threshold value.

Optimization studies dealing with ASPs can be broadly classified into open-loop and closed-loop optimization (Tables 1–3). Open-loop problems focus on determining the ASP optimal operating strategy and/or design only: the tracking of the computed optimal trajectory and disturbance rejection are not considered. Such problems regard the process control as an independent second step, and any interaction between control and optimal operation is ignored. On the other hand, closed-loop problems tackle simultaneously the ASP optimal operation/design and the controllability issues. In other words, they recognize the practical need to control the process: they aim at obtaining a cost-efficient and operable process systematically, i.e. finding the best performance that a given control structure can achieve. For example, in the case of an AAS process, while open-loop problems aim at selecting the optimal aeration profile, closed-loop problems will look at determining the optimal DO concentration setpoint that triggers the aerators turning off. It is noteworthy that solutions achieved using a closed-loop formulation are expected to be more robust than those obtained from an open-loop optimization; in fact, they are likely to be less sensitive to disturbances effects which are partially rejected by the control system.

RTO is a particularly interesting case of closed-loop optimization. It allows dealing efficiently with the model inaccuracies as well as with disturbances, especially if coupled with a feed-forward control strategy. RTO works as follows: (1) An optimizer is regularly supplied with measurements of the monitored variables. (2) The optimizer recalculates then the optimal operating conditions, taking into consideration the mismatch between the experimental measurements (i.e. the system actual state) and the model predictions. Therefore, on-line monitoring of the process variables enables more frequent updates of the optimal solution, and hence, a more efficient and flexible optimal control of the ASP. (3) The controlled variables are manipulated according to the current optimal solution.

Despite the benefits it can achieve, to the authors' knowledge, only Kim and coworkers [2] have investigated the RTO of ASPs. They focused on a bench-scale AAS process, with the aim at minimizing the aeration duration, while respecting the S_{NH} discharge limit. Measurements of the influent composition and S_{NH} concentration in the effluent were supplied to the optimizer once each day. The new optimal strategy was computed after adapting the current S_{NH} discharge constraint, according to the difference between the calculated and measured effluent S_{NH} concentration, i.e. taking into consideration the model inaccuracy. The experimental results revealed that RTO allowed achieving good COD and nitrogen removal efficiencies, even under shock loads (i.e. steep increase in the influent COD or/and S_{NH} concentration).

4.2.4. Long-term ASP performance

The computed optimal operating strategy must guarantee a reliable and sustainable plant operation, i.e. that the ASP functioning remains cost-effective over a large time horizon. Therefore, the problem formulation should ensure that the optimal solution does not depend on the chosen time horizon, neither on the plant initial state since its effects should vanish with time. Otherwise, the problem initial conditions, which are often arbitrary guessed, will promote/disadvantage some operation strategies/designs over others. Two main approaches are used to address these issues.

The first method consists in considering a large time horizon so as to reach the steady-state if a steady influent is assumed and if the problem involves time-independent DVs only, and a cyclic steady-state otherwise. The considered time horizon should be several times larger than the SRT, i.e. the characteristic time of the slowest dynamic processes occurring in ASPs. In fact, as shown in [33], operating strategies determined from a short time horizon optimization (1 day in their case), if applied repeatedly over a long time horizon (60 days in their case), reveal to be suboptimal and may even lead to biomass washout.

When this first approach is employed, the quantities of interest (e.g. electrical consumption, effluent quality) should be computed during the steady-state only, i.e. after a sufficient time delay: hence, the effects of the ASP initial state on the computed optimal solution are mitigated. This formulation is commonly used for computing the objective function since it has been popularized thanks to [45,46]. Recently, it has been applied in [47,48] to constraints also: path constraints expressing discharge limits were taken into consideration only after a large time delay. Indeed, activating a constraint during the transient regime makes the computed solution depending on the arbitrary guessed start-up conditions.

Although quite simple and intuitive, this first approach is computationally expensive since the process model is to be integrated over a large time horizon. An alternative approach consists in reformulating the problem, so as to restrict its feasible solutions space to the system possible steady-states only. If a steady influent is assumed and all the DVs are time-independent, then all the time derivatives should be considered nil, which leads to a static optimization problem as explained in Section 4.2.2. Otherwise, if the permanent regime corresponds to a cyclic steady-state, a more complicated methodology is applied [40,49]. The model is integrated over a single cycle only, and equality constraints are added in order to achieve time-periodicity of the system state:

$$\boldsymbol{x}^{Initial} = \boldsymbol{x}^{End} \tag{11}$$

where $\boldsymbol{x}^{Initial}$ and $\boldsymbol{x}^{End}$ are the state variables vectors at the beginning and end of the cycle respectively. They are treated as a vector of additional DVs. Although very attractive, this approach generally leads to considerably multimodal problems [32], necessitating the use of advanced optimization algorithms.

A third method allowing to account for the ASP slowest dynamic processes, while using a short time horizon, was proposed in [50]: biomass concentration in the reactor was allowed to vary with an amount of 0.01% only over the time horizon. Hence, this formulation guarantees that the computed solution does not lead to biomass washout over the long term. However, this method makes the solution very sensitive to the initial bacteria concentration. Thus, doubts may be raised concerning the true optimality of the achieved solution.

4.3. Effectiveness of various optimization techniques

As discussed in Section 4.1, mathematical models describing ASPs operation are highly non-linear and often non-differentiable. Therefore, when using local-deterministic techniques, a multi-start procedure is required in order to achieve a satisfying solution. However, some studies [6,28,51] reported that, starting from different sets of initial conditions, their

local-deterministic algorithm returned the same solution, whence the possibility that it might correspond to a global optimum. On the other hand, numerous studies adopting a multi-start procedure reported that the achieved results corresponded to a large number of local solutions [8,12,38,39,52,53].

Several papers compared the performance of stochastic/hybrid algorithms to local-deterministic techniques coupled with a multi-start procedure [8,12,52,53]. Their results indicated that, although local-deterministic techniques may achieve satisfying solutions, they are outperformed by stochastic/hybrid algorithms.

Some studies [8,13,14] compared the efficiency of stochastic/hybrid and global-deterministic algorithms. Their results revealed that stochastic techniques generally outperform global-deterministic solvers in terms of both objective function value and computation time requirement. In fact, as discussed in [8], to guarantee global optimality, global-deterministic methods have a number of requirements that are not met in such complex problems. On the other hand, with a moderate computation time, stochastic/hybrid algorithms can identify sufficiently good solutions that satisfy the designer in practice.

5. Optimal control/design of ASPs: objective functions, flowsheets structures and DVs

This section reviews general aspects relative to the choice of objective functions (Section 5.1), flowsheets structures (Section 5.2) and DVs (Section 5.3). It is discussed how a convenient selection of the DVs allows simplifying the optimization procedure.

5.1. Objective function(s)

5.1.1. Economic criterion

When dealing with ASPs, focus is generally put on the cost-effectiveness of the treatment. Thereby, in SOOP, a convenient objective function must express an economic criterion: pollutants discharge are dealt with through inequality constraints (see Section 7.1), or are incorporated as fines into the objective function to be minimized (see Section 6.1.6).

For existing WWTPs, the objective function should involve operating costs, e.g. power consumption by aerators and pumps. Some exploitation costs are independent of the operating strategy, thus, it is not necessary to include them in the objective function (e.g. wastewater pumping cost since it is dictated by the influent flow rate). Equipment cost should be accounted for unless the required material is available: in this case, the DVs bounds should represent the equipment characteristics, e.g. the maximum flow rate that can be delivered by the available pumps. Among papers dealing with existing ASPs upgrading, only Ostace and coworkers [54] have incorporated equipment related costs in the objective function. Their study aimed at comparing the optimal performance of several control systems, suggesting that the necessary equipment was not purchased yet. Therefore, in addition to the exploitation cost, their objective function integrated the necessary equipment price (sensors, controllers and filtration units), maintenance and installation charges.

For WWTPs to be built, the objective function should represent all the costs related to the ASP operation and design, at least exploitation and construction costs. Some authors have also accounted for some equipment costs [19,28,34–36,38,39,41] and for the maintenance of the constructions and equipment [28,29,35,36]. El-Shorbagy and coworkers [28] have also considered the salaries of the staff maintaining and operating the WWTP.

As stated above, the objective functions representing an economic criterion involve several contributions (e.g. aerators power consumption, reactors construction costs) that must be expressed in the same units, generally monetary units, in order to be commensurable. This is achieved using cost functions which are expressions/correlations linking the costs of each equipment/unit to its operating parameters, capacity or size. The solution of an optimization problem is closely dependent on the used cost functions since they determine the preference between alternative designs or operating strategies. Therefore, in order to get realistic and relevant solutions, reliable cost functions are required. Unfortunately, as detailed in Section 6, for the same operating conditions, cost functions used in the literature often provide very different estimates of the costs involved. Indeed, costs may depend on the executing engineering firm and its standards in term of design and equipment. Therefore, instead of relying on literature data, it is advisable to collaborate with local wastewater engineering companies in order to formulate reliable cost functions.

5.1.2. Other criteria

Obviously, cost is not the unique worthy criterion to be improved and optimized when dealing with ASPs. However, since economic aspects may not be neglected, non-economic criteria should not be considered alone, but simultaneously with the ASP operating/total costs: both objectives can be gathered using a weighted sum, or handled together in a MOOP.

The most used non-economic objective is the minimization of pollutants concentration in the effluent (e.g. [51]). Studies [1] and [3] who dealt with a semi-continuous SBR considered minimizing the batch time, i.e. maximizing the ASP productivity. Iqbal and Guria [42] investigated the maximum wastewater flow rate that the ASP can treat efficiently, i.e. the maximum plant throughout/capacity. Guerrero and coworkers [55] considered minimizing the risk of microbiology-related solids separation problems, i.e. bulking, foaming and rising sludge. Among papers addressing the closed-loop optimal control of ASPs, several ones (e.g. [8]) aimed at maximizing the process controllability. This objective is generally stated as the minimization of the controllers' integral square error (ISE).

5.2. Flowsheet structure/process configuration

Numerous ASP configurations have been addressed in the literature: semi-continuous SBR (e.g. [1]), AAS (e.g. [2]), A^2O (e.g. [10]), alpha-process [40], RDN and DRDN [41], etc. Some studies have included the primary clarifier within their flowsheet [20,28,34,52,56].

Conventional flowsheets represent a single ASP configuration whose operation/design is to be optimized. However, for the considered influent characteristics, it is very likely that a different ASP configuration allows a more cost-efficient treatment. Using a conventional flowsheet, finding the best process structure requires: (1) Determining which ASP configurations are interesting to investigate. (2) Computing the optimal solution for each of these options. (3) Comparing the optimal solutions to identify the best one. However, this procedure may be simplified by the use of a superstructure flowsheet.

A superstructure is a general flowsheet embedding all the interesting process configurations. The most used superstructures consist of several reactors in series. However, contrary to conventional flowsheets, the environmental conditions in the reactors are not specified in advance: aeration is allowed in all reactors, and the optimal solution determines if a reactor should operate under aerobic, anoxic or anaerobic conditions. In some superstructures, even the number of reactors is considered as a DV [11]. Hence, optimizing a superstructure allows determining simultaneously the optimal operation/design and the most appropriate ASP configuration.

Given their generality, superstructures have more degrees of freedom than conventional flowsheets; thereby, the corresponding optimization problems are more complex to solve. For example, in a conventional flowsheet, aeration is allowed only in aerobic reactors, while in a superstructure, aeration must be allowed in all reactors, which increases the number of DVs. Therefore, the size of a superstructure should be chosen carefully so as to ensure a good trade-off between the generality of the flowsheet and the required computational effort.

Finally, it is noteworthy that some authors have considered flowsheets comprising other processes in addition to the ASP. Descoins and coworkers [56] and Hakanen and coworkers [20] integrated a sludge treatment facility comprising an anaerobic digester (for biogas production) in their flowsheet. Using such a general formulation, the entire WWTP may be considered as a single unit: compared to the optimization of the ASP and the sludge treatment facility separately, this plant-wide control approach allows identifying better operating policies for the WWTP since the interactions between the two processes are accounted for.

A few works have even considered a flowsheet larger than the WWTP. Indeed, Brand and Ostfeld [57] incorporated in a single flowsheet, 2 cities (i.e. wastewater sources), 3 decentralized WWTPs and a pipeline network. Fu and coworkers [58] considered a flowsheet covering the urban wastewater system: it comprises a sewer network, a WWTP consisting of a basic ASP and a storm tank, and a river, i.e. a receiving waterbody.

5.3. Decision variables

5.3.1. Continuous time-independent decision variables

In open-loop optimization problems, the most often selected DVs are the aeration rate (generally expressed in terms of achieved k_La, DO concentration or injected air flow rate) and the flow rate of the external recycle, internal recirculation and sludge wastage. Indeed, these operating parameters have a great impact on both treatment cost and quality. Other commonly used continuous DVs include the external carbon dosage and the influent flow fractioning among the reactors in the case of a step-feed. When dealing with ASPs to be built, the reactors' volumes and the settler design are added to the DVs set. In closed-loop optimization problems, the setpoints of the manipulated variables and the controllers tuning parameters are often taken as additional DVs.

5.3.2. Continuous time-varying decision variables

The use of time-varying DVs is necessary to perform open-loop optimizations of discontinuous processes such as AASs and SBRs. In the case of reactors in series, aeration rates and other DVs are generally treated as time-independent. Nonetheless, time-varying operating parameters allow to better handle large load variations. For example, aeration can be increased under high loads to handle the incoming pollution, and decreased during low loads to save energy.

Some studies dealing with large-size WWTPs have considered time-varying aeration rates [7,53,59]. Guerrero and coworkers [10,55] used time-varying setpoints for the manipulated variables, to allow identifying different optimal setpoints during weekdays and weekends. Hreiz and coworkers [47] considered the sludge wastage flow rate as a time-varying DV. Hence, the optimizer was allowed to adapt the TSS concentration in the reactors so as to avoid the settler overload under peakflows for example. Egea and Gracia [53] treated the internal recycle flow rate as a time-varying DV. On the other hand, Hreiz and coworkers [47,48] considered the external recycle and internal recirculation flow rates as time-independent DVs. They argued that, in practice, recycle flows are lifted by screw pumps which are generally operated at constant capacity.

5.3.3. Integer and binary decision variables

Integer variables used in studies addressing optimal control/design of ASPs include the position of the feed layer among the 10 layers of the Takács [31] model (e.g. [12,14]), the number of daily aeration cycles of an AAS (e.g. [6]) and the number of reactors in a superstructure [11]. Doby and coworkers [52] employed a binary DV to choose whether or not a primary clarifier should be used. Mussati and coworkers [60] used binary DVs to select the process configuration: the pre-, post- or pre-post-denitrification ASP.

Problems involving discrete and binary DVs are known to be extremely difficult to solve using deterministic techniques because of differentiability issues. In that case, the use of such DVs should be avoided as discussed in the following paragraphs.

5.3.3.1. Optimal number of reactors in a superstructure. To determine the optimal number of reactors in a superstructure while avoiding the use of integer DVs, Alasino and coworkers [38,39] treated the reactors volumes as continuous DVs: if the optimal volume of a reactor is almost nil, this means that the reactor in question can be deleted from the optimal design.

It is noteworthy that, as discussed in [61], reactors staging allows better treatment efficiencies. Therefore, if the reactors construction costs are calculated according to Eq. (19) instead of Eq. (18) (see Section 6.2 for details), then the optimal solution will probably correspond to an infinite number of reactors in series.

5.3.3.2. Optimal number of daily aeration cycles in a AAS. Studies dealing with AASs (see Table 1) aimed at determining the optimal durations of the aeration-on and aeration-off sequences, allowing to minimize the total aeration time, while fulfilling the discharge limits as well as other technical constraints. In all these papers, it has been assumed that aeration is on/off controlled, i.e. that the aerator operates at a constant capacity only (fixed rotation speed for a surface aerator or fixed air flow rate for a diffused aerator).

In their general form, such problems involve an integer DV, N_c, the number of aerations cycles per day, and $2 \times N_c$ continuous DVs, the durations of the aeration-on and aeration-off periods (a cycle refers to the duration between two consecutive starts of the aerators, i.e. consecutive aerator-on and aerator-off periods). To avoid using discrete DVs, Chachuat and coworkers [6,33,51] determined the optimal N_c value by enumeration: (1) The problem was solved for different predefined values of N_c: hence, only $2 \times N_c$ continuous DVs were involved. (2) The achieved solutions were compared so as to identify the optimal one.

It is noteworthy that Chachuat and coworkers [6,33,51] have limited the maximum duration of aeration-on and aeration-off periods to 2 hours (see Section 7.3). Without the use of such constraints, the optimal solution would correspond to an infinite value of N_c. Indeed, mathematically speaking, increasing the number of cycles per day would obviously result in better treatment efficiencies.

6. Cost functions formulation

As mentioned earlier, cost functions are expressions/correlations allowing to calculate the costs related to the construction/operation/maintenance of a unit, according to its operating parameters, capacity or size. The solution of an optimization problem is closely dependent on the used cost functions since they determine the preference between alternative ASP designs or operating strategies. Unfortunately, there are no commonly accepted formulations for the cost functions, and the expressions proposed in the literature often lead to diverging results, sometimes different of several order of magnitude.

In this section, cost functions provided by the literature are critically reviewed: their limits and potentials are discussed, hoping to provide a general guideline to assist designers selecting appropriate and reliable formulations.

6.1. Exploitation costs

6.1.1. Aerators power consumption

According to [56,62], the aerators power consumption accounts for about 60–80% of the WWTP operating costs. Aeration is achieved by circulating the mixed liquor using a rotating turbine, or by injecting air bubbles through diffusers placed near the bottom of the reactor. Oxygen, the species of interest in the context of ASPs, is sparsely soluble in water: resistance to mass transfer is much more important in the liquid than in the gas side of the interface. Therefore, according to the double film model, the oxygen mass flow rate transferred from the gas to the liquid phase, F [Mt^{-1}], is given by:

$$F = k_L a V_L (DO^* - DO) \tag{12}$$

where DO^* [ML^{-3}] is the oxygen saturation concentration in the liquid phase, V_L [L^3] the aerated liquid volume, k_L [LT^{-1}] the liquid-side oxygen mass transfer coefficient, and a [L^{-1}] the specific interfacial area. The mass transfer rate is generally characterized through the k_La factor, since k_L and a are very difficult to access individually in complex flows such as those occurring in activated sludge reactors.

Wastewater and mixed liquors are more complex than pure air–water systems. Indeed, they contain many pollutants having an important impact on k_La [63]. Moreover, activated sludge exhibit a complex non-Newtonian behavior depending mainly on their TSS concentration, wastewater strength and type, as well as on the reactors performance [64]. The activated sludge viscosity significantly influences k_La since it impacts the DO convective transport rate. All the effects of this complex environment on the mass transfer rate are gathered together into an α factor, generally lower than one in the case of wastewater and mixed liquor, and that is expressed as:

$$\alpha = \frac{k_L a_{Mixed\ Liquor}}{k_L a_{Pure\ Water}} \tag{13}$$

In addition to the numerous factors mentioned before, α also depends on the aerator type (which influences the interfacial areas). α may be estimated using correlations, however, since it varies from an ASP to another, it is preferable to determine it experimentally when dealing with existing WWTPs.

If the aeration system is operated at a constant capacity only, which is generally the case in AASs, then the aerator power consumption is proportional to the aeration time. Indeed, the electrical consumption of diffused aerators is almost independent of the fluid rheological behavior. On the other hand, in the case of surface aerators, as shown in [32], for the commonly encountered designs and operational speeds and typical TSS concentrations in the mixed liquor, the flow around the impeller is fully turbulent; hence, the aerator energy input is independent of the TSS concentration in the medium. Therefore, when dealing with AASs, if the aerator energy input is to be minimized, it is unnecessary to calculate its equivalent monetary costs since the problem reduces to minimize the duration of the aeration-on periods [2,6,9,32,33,49,50]. However, it is noteworthy that fully turbulent flow conditions do not imply that the achieved k_La is independent of the TSS concentration, although this assumption was implicitly made in all these studies.

Except the specific case mentioned above, aeration power consumption needs to be estimated using cost functions. Aeration charges can then be calculated simply by multiplying the power input by the actual cost of a kWh. In the case of surface aerators, some studies Dey and coworkers [65] and Li and coworkers [30] adopted a constant specific power per kg of DO equal to 0.7 kg O_2/kWh [66] and 1.48 kg O_2/kWh [67] respectively. However, this approach assumes that aeration costs are proportional to F (Eq. (12)), i.e. the aerator efficiency depends on the DO concentration in the reactor, which is obviously incorrect.

For the case of diffused aerators, Copp [45] considered aeration costs in a 1333 m^3 aerobic reactor to be a quadratic function of k_La. However, such an expression assumes aeration charges to be independent of the aerated liquid volume. Therefore, it should absolutely not be used when dealing with reactors of different volumes, and especially when investigating the optimal design of ASPs since it will mistakenly promote large aerobic reactors. Alex and coworkers [46] considered the aeration requirement of a fine bubbles diffuser to be a linear function of ($k_La \times V_L$), which is valid only over a restricted range of air flow rates. In fact, it is well known that the mass transfer efficiency decreases at increasing air flow rates. It is noteworthy that for the same ASP and operating conditions considered in their papers, the aeration power as calculated in [45] is about two times greater than that calculated in [46] (N.B.: these two formula are the most commonly used in literature). Stare and coworkers [68] and Alasino and coworkers [38,39] expressed the aerator power consumption as a quadratic function of ($k_La \times V_L$): such a formulation combines the merits of both expressions used in [45,46].

Instead of estimating the aeration power requirement based on the achieved k_La, some authors used a more rigorous approach: they expressed the aeration costs (as well as the achieved k_La) as a function of the injected air flow rate, Q_{Air}, i.e. the variable actually manipulated by the ASP control system. The blower electrical consumption can be calculated then knowing the device efficiency (generally assumed constant for simplifying purposes), the diffuser immersion depth, and assuming an adiabatic compression for example [69]. Espírito-Santo and coworkers [29,35,36] considered a linear relationship between k_La and Q_{Air}, and a 7% oxygen transfer efficiency. El-Shorbagy and coworkers [28,34] calculated the DO demand of the ASP as the sum of the DO required for the removal of carbonaceous matter, plus the DO required for nitrification [66]. Q_{Air} was calculated then assuming 10% oxygen transfer efficiency. Hreiz and coworkers [47,48] used the empirical formula derived by Gillot and Héduit [70] from on site measurements in ASPs:

$$k_L a = 1.477 h^{-0.136} S^{-1.175} S_p^{0.042} S_a^{0.145} Q_{Air}^{1.037} \tag{14}$$

where h is the water column height in m, S_a the surface covered by the diffuser modules in m^2, S_p the total membrane surface in m^2, k_La is in h^{-1} and Q_{Air} in $Nm^3\ h^{-1}$. This formula has been derived from measurements in aerobic reactors, and therefore, is valid for important aeration rates only (whence the quasi-linear relation between k_La and Q_{Air}). In [47,48], thanks to the used DO-related constraints (see Section 7.4), Eq. (14) was expected to provide realistic results over the feasible solutions space.

To account for the effects of the mixed liquor characteristics on the mass transfer rate, Hreiz and coworkers [47,48] corrected Eq. (14) using the α factor. α was supposed to depend mainly on the TSS concentration in the mixed liquor, and the corrective formula was deduced from experimental results [71] (valid for $0\ g\ L^{-1} \leqslant TSS \leqslant 10\ g\ L^{-1}$; TSS concentration in ASPs typically lies in the range 3–5 $g\ L^{-1}$):

$$\alpha = -0.072 \times TSS + 1 \tag{15}$$

It is noteworthy that α is independent of the air flow rate [71]. Their results (Fig. 4) proved the usefulness of accounting for the detrimental effect of the TSS concentration on oxygen transfer rate: neglecting this phenomenon leads to a significant underestimation of the ASP exploitation costs.

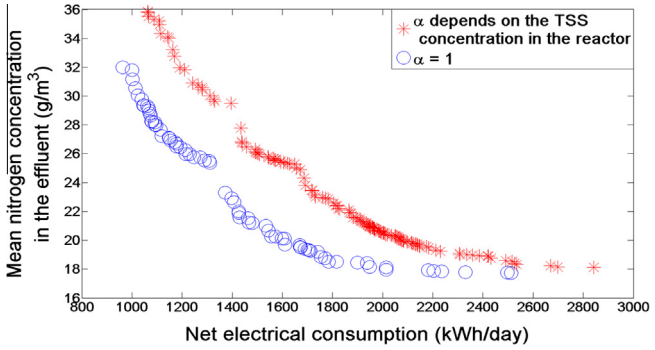


Fig. 4. Effect of the reduction of the oxygen transfer rate on the Pareto front [47].

6.1.2. Pumps power consumption

Pumps account for about 25% of WWTPs total electrical consumption [56]. Pumping power consumption includes several contributions: influent pumping, sludge recycling, internal recirculation of the nitrified liquor and sludge wastage. Influent pumping costs are generally not included in the objective function since they are dictated by the incoming wastewater flow rate, i.e. they are independent of the ASP design and operation.

Pumping charges depend on the type of pumps used, the fluid rheological properties and the flow Reynolds number, etc. For simplification purposes, most authors dealing with optimal control/design have computed pumping costs according to Eq. (16), assuming a constant differential head:

$$\text{Pumping power} = \frac{\rho Q \Delta H}{\eta_P} \quad (16)$$

where ρ [ML^{-3}] is the fluid density, Q [L^3T^{-1}] the volume flow rate, g [LT^{-2}] the acceleration due to gravity, ΔH [L] the differential head and η_P the pump efficiency. Literature survey reveals a great discrepancy among the differential head values that have been employed in different studies (Table 4). However, it is not possible to judge if one is better than the other: each one may reveal more convenient for a given plant design and pump type. Nonetheless, the differential head of 10 bar used in [42] for the raw sewage pump seems unrealistic.

Concerning sludge wastage, most studies have associated it with an important differential head, probably to account for some further sludge treatment and pumping charges, e.g. sludge dewatering in a filter press. Hreiz and coworkers [47,48] considered sludge purge to be cost-free since it is gravity-driven. Sludge thickening and dewatering costs were accounted for separately than sludge wastage in their model.

6.1.3. Mixing power consumption

Activated sludge reactors are mechanically mixed to achieve homogenization, and most importantly, to maintain the biomass in suspension: otherwise, sludge would settle, which reduces the reactor capacity and may lead to undesired anaerobic conditions in the sludge blanket.

In aerobic reactors, for economic reasons, the aeration system is generally designed to satisfy mixing requirements, while in anoxic and anaerobic reactors, mixing is usually carried out by a slow-rotating submersible mixer. According to Rittman and McCarty [72] (cited in [65]), a minimum power input of 10 W/m^3 is necessary to ensure a completely mixed flow regime. Alex and coworkers [46] and Zakkour and coworkers [73] estimated the mixing energy requirements to 5 W/m^3 and 14 W/m^3 respectively.

In sequentially aerated reactors (e.g. AAS), mixing and aeration are decoupled: the mixer is operated during the aeration-off periods, while the aerator provides the necessary mixing during the aeration-on periods. However, the use of a mixer is not necessary (thus enabling equipment savings) provided that the duration of the aeration-off periods are not too large (see Section 6.3).

6.1.4. Additives costs

Chemicals addition is widespread in WWTPs. It aims at enhancing the treatment efficiency (e.g. chemicals for phosphorus precipitation) or ensuring a satisfying operation of the process (e.g. chemicals for pH control), etc. However, in studies dealing with optimal operation/design of ASPs, only the additions of carbonaceous matter and of chemicals for pH control have been considered.

The addition of external carbonaceous substrates aims at balancing the COD deficiency in the wastewater, i.e. enhancing bacterial growth/activity when carbon becomes limiting (which is often the case in post-denitrification ASPs as explained in Section 2). Alex and coworkers [46] used an external carbon source having a concentration of 400 kg-COD m^{-3}: one cubic meter of this additive (whose nature was not specified) was considered to cost as much as 1200 kWh of electrical energy. Mussati and coworkers [60] used an additive costing 0.3 € $kg\text{-}COD^{-1}$, but mentioned neither its nature nor its concentration. Ostace and coworkers [54] used acetate, having a price of 0.5 € kg^{-1} and a concentration of 400 kg-COD m^{-3}. Samuelsson and coworkers [74] used ethanol, and estimated its price and concentration to 550 € m^{-3} and 1200 kg-COD m^{-3} respectively. They have also investigated the situation where carbonaceous matter would be provided free of charge, for example if industrial by-products are available.

The biomass involved in ASPs requires a pH in the range of 6–8 to survive and proliferate. However, many industrial wastewaters have a high or a low pH (e.g. food process wastes), which necessitates chemicals additions (e.g. lime or acids) to adjust their pH. On the other hand, sewage water has generally a convenient pH, typically in the range of 7–8. Nevertheless, since some biochemical reactions occurring in ASPs consume alkalinity (especially nitrification), the addition of buffering agents for pH control may reveal necessary. However, among studies dealing with optimal control/design of ASPs, only [18,19] have accounted for such chemicals costs (probably because of the moderate involved costs). They considered sodium bicarbonate addition to maintain alkalinity within a suitable range, between 1.5 and 2 mol m^{-3} according to these

Table 4
Pumping differential head values according to literature studies.

	Raw sewage		Internal recirculation		Sludge recycling		Sludge wastage		Remarks
	ΔH(m H_2O)	η_P	ΔH(m H_2O)	η_P	ΔH(m H_2O)	η_P	ΔH(m H_2O)	η_P	
[45]	–		14.7	–	14.7	–	14.7	–	
[46]	–		1.5	–	2.9	–	18.4	–	
[42]	100	–	–	–	*		–	–	*A quadratic function is used to estimate the sludge recycling energy requirement
[28,34]	–		–		10	0.6	10	0.6	Wastage flow concerns both primary and secondary sludge
[47,48]	–	–	1	0.7	4	0.5	0	–	

authors. The sodium bicarbonate dosage was not expressed in monetary units, but considered as an objective function to be minimized in a MOOP.

6.1.5. Sludge disposal/valorization

Sludge produced by WWTPs contains pathogen microorganisms and may produce offensive odors. Therefore, for a long time, the wastewater industry has considered it as a hazardous waste that should be processed properly before being used as soil amendment or deposited in landfills, depending on its toxic content. Sludge treatment includes thickening and dewatering to dehydrate the watery sludge in order to facilitate its transport, and stabilization to reduce its evolution and the number of pathogen microorganisms. Sludge may be stabilized by anaerobic or aerobic digestion, and/or by chemical or thermal sterilization. Sludge treatment charges reported in the literature vary from about 0.2 € kg^{-1} of dry sludge [38,54,60] to around 0.5 € kg^{-1} [36,45,46,68].

However, nowadays, with the growing concerns about global warming and fossil fuel depletion, the actual trend is to consider sludge as a renewable raw material to be valorized through different energetic sectors. Some studies [41,47,48] have investigated sludge incineration for electricity production. Rivas and coworkers [41] did not report their cost functions. Hreiz and coworkers [47,48] estimated sludge thickening, dewatering and drying costs about 0.1 € kg^{-1}. The dry sludge calorific value was taken as 15 MJ kg^{-1}, and the efficiency of the in-site incineration engine was considered to be 30%. Sludge incineration was shown to be of a high energetic profit, especially when a high nitrogen discharge is tolerated (Fig. 5). Descoins and coworkers [56] and Hakanen and coworkers [20] have integrated a sludge treatment facility comprising an anaerobic digester within their flowsheets. Hakanen and coworkers [20] handled the biogas production as an objective function to be maximized in a MOOP. Descoins and coworkers [56] considered burning the produced biogas for electricity production. They found that, unless the discharge limits are very tight, anaerobic digestion of the sludge allows the WWTP to be electrically autonomous, and even to produce important surplus of electricity. However, such surprising result may be due to the problem formulation (e.g. the used cost functions), especially the steady influent assumption made in [56].

6.1.6. Effluent fines

In SOOPs aiming at minimizing ASPs costs, two main approaches are used to guarantee an efficient treatment strategy. The first method, discussed in Section 7.1, consists in imposing inequality constraints to the pollutants concentrations in the effluent, so as the feasible solutions achieve a treatment complying with the applicable regulations. The second approach, discussed in this paragraph, consists in penalizing pollutants discharges by applying fines relative to the effluent quality. The treatment level is hence 'converted' into monetary units. Thus, it can be combined with the treatment cost in a single economic objective function to be minimized.

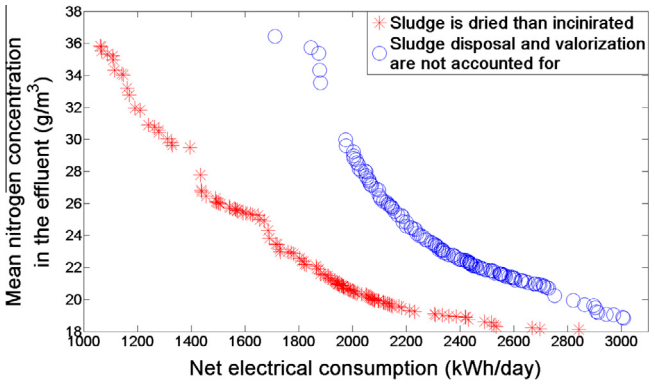


Fig. 5. Effect of sludge disposal and valorization on the Pareto front [47].

In some countries/federations, e.g. Denmark and Flanders, authorities charge according to the total amount of discharged pollution [75]. In such situations, converting the effluent quality into fines provides more realistic results than dealing with the treatment requirements through inequality constraints. Almost all studies adopting this approach (e.g. [10,55]) have used fines functions inspired from the Danish or the Flemish legislation as presented in [75]. However, it is noteworthy that the "converting into fines" approach emphasizes on economic interests rather than on environmental concerns. Indeed, it allows selecting the most lucrative abatement strategy among two alternatives: achieving a treatment complying with the specified discharge limits, or partially treating the wastewater and paying fines. In the latter case, a significant discharge of some species can be allowed. To remedy this issue, both above-mentioned approaches (i.e. using effluent fines and constraining pollutants concentrations) may be used simultaneously [38,39].

In some countries, e.g. France, the WWTP operator is taxed only when pollutants concentrations in the effluent exceed the specified limits. In situations where such a regulation applies, effluent fines should be calculated on the basis of the regulation violations rather than the amount of pollutants discharged. Such a formulation of the effluent fines has been used only in [30].

6.2. Investment/capital costs

When addressing ASPs optimal design, since design and operation are interlinked, the objective function should integrate investment (i.e. construction and equipment) and exploitation (i.e. operation and maintenance) costs so as to permit a rigorous economic evaluation of the facility over its life span. Purchase/construction costs can be generally expressed as a power law function of the prime attribute of the unit [76]:

$$Cost = aZ^n \tag{17}$$

where a and n are cost parameters, and Z the equipment cost attribute. Z is related to the unit capacity or size: for example, it corresponds to the volume of a reactor or to the maximum flow rate that a pump can deliver. n is generally less than one: the "larger" the equipment is, the lower its cost per unit capacity will be.

Numerous capital/investment cost functions have been reported in literature. The correlations presented in [77] (cited in [69]) were developed for municipal WWTPs in Flanders. Those presented in [28,34,78] are a compilation from literature. Espírito-Santo and coworkers [29] developed their cost functions based on data from a Portuguese company. Alasino and coworkers [38,39] did not mention their data source.

In the case of ASPs consisting of reactors in series, following Eq. (17), Alasino and coworkers [38,39] formulated the reactors construction costs as:

$$\text{Total construction cost} = a\sum_{i=1}^{m}(V_i)^n \tag{18}$$

where m is the total number of reactors and V_i the reactors volumes. On the other hand, Rivas and coworkers [40,41] considered minimizing the reactors total volume in order to reduce the construction charges. Hence, they implicitly assumed that constructions costs are expressed as:

$$\text{Total construction cost} = a\left(\sum_{i=1}^{m}V_i\right)^n \tag{19}$$

Eq. (18) is expected to be more appropriate when dealing with ASPs consisting of single reactors in series, while Eq. (19) is likely to be more relevant when dealing with plug-flow ASPs, which are generally modeled as a series arrangement of continuous STRs, as explained in Section 4.1.

When using cost functions from literature, the variability of the money time value must be taken into consideration, especially when appraising such long-term projects: the actual money value is not the same than at the time of the correlation development. In fact, on one hand, inflation erodes the money purchasing power, while on the other hand, there is an opportunity to invest money and earn interests [76]. Since investment/capital cost functions express prices in the monetary units of the year they were developed, when relying on past records, these costs should be updated to reflect the changes in the economic conditions:

$$\frac{Cost(year\ 1)}{Cost(year\ 2)} = \frac{I(year\ 1)}{I(year\ 2)} \quad (20)$$

where I is the cost index. Various cost indices are commonly used by the chemical industry engineers to adjust for the effects of inflation [76]. Costs updating is not performed for the sole purpose of calculating the current ASP design charges, but also to make investment and exploitation costs commensurable, i.e. expressed in the same monetary units. Indeed, capital costs are generally invested as a lumped sum, while operating costs are paid on an annual or monthly basis. To take into consideration the variation of the money value (i.e. accounting for inflation and returns), all cash flow should be brought forwards or backwards to the same point in time so that they can be compared. Thereby, the ASP total costs should be formulated as a net present value or in equivalent annual worth. The economic life of WWTPs is commonly estimated to 20 years [28,29,35,36,38,39,41]. The effective interest rate is generally supposed in the range of 5–8% [28,29,34–36,38,39,41]. Rivas and coworkers [41] have also accounted for the expected inflation, considering a 6% inflation rate.

6.3. Maintenance costs

When addressing ASPs optimal design, maintenance costs should be integrated in the objective function since they depend on the capacity and size of equipment and constructions. Espírito-Santo and coworkers [29] estimated the reactors maintenance charges to 1% of the investment costs during the first 10 years, and 2% the following years. They also reported a maintenance cost function for the secondary settler. They neglected the maintenance expenses of equipment, but their life span was assessed at 10 years. El-Shorbagy and coworkers [28,34] used published cost functions to evaluate the maintenance costs of the settlers and the sludge pump. Ostace and coworkers [54] estimated the yearly maintenance costs of sensors and controllers to 20% of their purchasing cost.

In addition to the required material, maintenance requires human intervention. However, only El-Shorbagy and coworkers [28,34] have accounted for such expenses. They used cost functions from literature, expressing the settlers, aerators and sludge pumps charges in person-hour requirement per unit time. These costs were then multiplied by the wage rates in order to get monetary units.

7. Constraints

This section reviews the constraints used in the literature devoted to optimal control/design of ASPs. Constraints are generally employed to impose a minimum process performance, guarantee a physically relevant operation, or ensure that the mathematically feasible solutions remain within the limits of validity of the model, etc.

7.1. Discharge constraints

As noted previously, to ensure that the optimal solution provides an efficient treatment, constraints may be imposed on the pollutants concentrations in the effluent. Discharge constraints have been most commonly applied to COD, BOD_5, TSS, S_{NH}, S_{NO}, total nitrogen and total phosphorus, etc. Since environmental regulations vary from a country to another, and even within the same country, depending on the receiving waterbody, different discharge limits values have been reported in literature.

In the case of continuous ASPs, discharge limits are generally handled through inequality path constraints: instantaneous pollutants concentrations are constrained below the specified limits. This approach promotes environmental issues rather than economic concerns. However, in some studies (e.g. [40,79]), discharge limits were applied to the average rather than the instantaneous pollutants concentrations in the effluent. Thus, mathematically speaking, discharge limits were formulated as end-point rather than path constraints. This approach enables a more flexible and economical operation of the ASP, since pollutants concentrations are allowed to exceed the prescribed limits during peakflows. Additionally, it is more convenient in cases where compliance with discharge limits is assessed over long time frames, e.g. in Sweden and Spain [79].

In the case of batch and discontinuous ASPs such as the SBR process, since considerations are only about the system final state, discharge limits are always formulated as end-point constraints.

7.2. Concentration of suspended solids in the bioreactors

Many authors have imposed constraints to the TSS concentration in the bioreactors. While the lower bound constraint aims at ensuring an acceptable treatment efficiency, the upper bound constraint was used for different operational purposes.

Hakanen and coworkers [18,19] who considered an ASP consisting of reactors in series, have constrained the TSS concentration in the last reactor to be below 6000 g m^{-3}. Hakanen and coworkers [20] has restricted its variation range between 2000 and 6000 g m^{-3}. In fact, these authors have considered steady influent conditions. Thereby, such upper bound constraints allow obtaining a robust optimal solution: maintaining a low TSS concentration in the bioreactors allows avoiding the settler overload in case of peakflows.

Rivas and coworkers [40,41] limited the TSS concentration in the final reactor to 3500 g m^{-3}. In fact, the sludge settling rate in the settler decreases for high TSS concentrations. However, the perfect point-settler approach used by these authors does not account for this effect (contrary to the Takács and coworkers' model [31]). Therefore, in this case, the use of this constraint is very relevant since it ensures that the mathematically feasible solutions do not result in an overload of the settler, a situation that is not taken into account in their settler model.

Guerrero and coworkers [10,55] and Ostace and coworkers [54] controlled the TSS concentration using a feedback PI-controller manipulating the sludge wastage flow rate. They set its setpoint value to 4500, 2500 and 3850 g m^{-3} respectively. These three studies aimed at determining the optimal ASP operating strategy that could be achieved using different control systems (closed-loop optimization). According to these authors, fixing the TSS concentration allows a better comparison of the performance achieved by each control strategy.

Many other studies have also imposed constraints on the TSS concentration in the bioreactors, however without justifying the

reasons. Espírito-Santo and coworkers [29] constrained it between 800 and 6000 g m^{-3}, Iqbal and Guria [42] between 2000 and 9000 g m^{-3}, El-Shorbagy and coworkers [34] between 3000 and 5000 g m^{-3}, and Descoins and coworkers [56] between 1000 and 9000 g m^{-3}.

7.3. Duration of the aeration-on/aeration-off periods in sequentially aerated bioreactors

Kim and coworkers [2], dealing with the optimal control of an AAS, have constrained the duration of the aeration-on and aeration-off periods between 0.5 and 4 h. They argued that a minimum air-on period of 0.5 h is necessary to avoid the washout of nitrifying bacteria. Their strategy proved convenient since their bench-scale AAS achieved good COD and S_{NH} removal efficiencies, even under shock loads.

Chachuat and coworkers [6,33,51], focusing on the optimal control of an AAS, have used similar constraints. They imposed a maximum aeration-on duration of 1 or 2 h in order to prevent the propellers from early wear or damage. The maximum aeration-off duration was set at 2 h. In fact, a too large non-stirring period leads to the sedimentation of the bacterial flocs, and hence, invalidates the STR assumption by reducing the reactor capacity. Moreover, the sludge blanket may develop anaerobic conditions that are not accounted for in the ASM1 model that they have used. Minimum aeration-on and aeration-off periods of 15 min were used to avoid too frequent cycling of the turbines, which increases the operational life of the aeration equipment. Moreover, the minimum aeration-on duration ensures a convenient re-mixing of the activated sludge after a large air-off period. Further studies focusing on AASs have used similar constraints as Chachuat and coworkers [9,49,50,80].

7.4. Dissolved oxygen concentration in the bioreactors

When optimizing ASPs operation/design, unless if suitable constraints are used, the optimal solution could correspond to conditions of low DO concentration in the aerobic reactor(s), about 0.2–0.4 g m^{-3} [11,47]. In fact, in the framework of ASM models, such conditions allow simultaneous nitrification and denitrification at low aeration costs, i.e. an excellent treatment at low operating charges. Although the results are mathematically correct, they are physically irrelevant since in practice, ASPs cannot achieve an efficient treatment in this range of DO concentrations. In fact, as discussed in [11], the accuracy of ASM1 in this range of DO conditions is questionable. Thereby, many authors have imposed constraints on the DO concentration in the reactors so as to restrict the solutions space to physically relevant operating conditions.

Some studies [29,35,36,50] held the DO concentration in the classic ASP above 2 g m^{-3}. Balku and Berber [9] and Balku [50] who investigated the optimal control of AASs, have constrained the average DO concentration during the aeration-on periods to be above 2 g m^{-3} and 3 g m^{-3} respectively. They argued that such high DO concentrations are necessary to limit the growth of filamentous bacteria, which may cause serious operating problems such as bulking. Rivas and coworkers [40,41], dealing with ASPs configurations consisting of reactors in series, have fixed the DO concentrations in the aerobic reactors to 2 g m^{-3}. Hreiz and coworkers [47] considered time-varying aeration rates in their superstructure consisting of a single reactor. They constrained the average DO concentration to be higher than 2 g m^{-3} (aerobic conditions) or lower than 0.1 g m^{-3} (anoxic conditions) during every cycle. Hreiz and coworkers [48] dealt with a superstructure encompassing four reactors in series. Similar constraints than previously were applied to DO concentrations, apart the last reactor upstream of the clarifier, where only aerobic conditions were allowed. In fact, if anoxic conditions occur in the last reactor, it is possible that denitrification continues to occur within the settler, and hinders sludge settling due to the generation of nitrogen bubbles. To the authors' knowledge, no current settler model allows accounting for this phenomenon, i.e. for the effects of the operating conditions on sludge settleability.

7.5. Alkalinity

The biomass involved in ASPs requires a pH in the range of 6-8 to survive and proliferate. Moreover, low pH may lead to corrosiveness problems and bulking, while high pH favors scale occurrence in pipes. Some of the biochemical processes occurring in ASPs have an impact on pH, especially nitrification which leads to an important decrease in the medium alkalinity. According to ASM1 [22], if alkalinity drops below 1 mol m^{-3}, the pH may become unstable and decreases to low values. Thereby, Hakanen and coworkers [18,19] considered adding sodium bicarbonate to maintain the medium alkalinity within a suitable range, between 1.5 and 2 mol m^{-3} according to these authors. Espírito-Santo and coworkers [29] constrained alkalinity in the reactor to be in the range of 6–8 mol m^{-3}, although no chemicals addition was considered for its regulation.

7.6. Sludge retention time

For an efficient gravity settling in the clarifier, bacteria must be flocculated, i.e. aggregated into enough large units. The bio-flocculation is partly due to the production of extra-cellular polymers which fix the bacteria together. In ASPs, biomass requires a minimum solids retention time (SRT) of about 3–4 days to develop a sticky slime layer. On the other hand, too large SRTs lead to poorly settleable flocs for different reasons among which the overgrowth of filamentous bacteria. Therefore, El-Shorbagy and coworkers [34] constrained SRT between 4 and 27 days, and Descoins and coworkers [56] between 5 and 35 days. Hreiz and coworkers [47,48] imposed a sludge age between 4 and 30 days. In addition to guaranteeing good floc settleability, the most important purpose of the upper bound was to limit the sludge mineralization since it alters the sludge calorific value. Indeed, a constant calorific value of the sludge (i.e. independent of its content in organic matter) of 15 MJ per kg of dry sludge was assumed.

7.7. Reactors size

When addressing the optimal design of ASPs, the reactor(s) volume(s), taken as a DV(s), is generally constrained by simple upper and lower bounds. In order to restrict the feasible solutions space, large values of lower bounds can be used, since small reactors do not allow attaining good treatment efficiencies. However, in some studies, more physical constraints were applied to the reactors size. El-Shorbagy and coworkers [28,34] (following [66]) imposed three restrictions to the bioreactor size:

$$V_L \geqslant \frac{\text{Oxygen requirement}}{0.1} \tag{21}$$

$$V_L \geqslant \frac{1000\, Q_{Air}}{Lim_U} \tag{22}$$

$$V_L \leqslant \frac{1000\, Q_{Air}}{Lim_L} \tag{23}$$

where V_L is expressed in m^3, Q_{air} in $m^3\,min^{-1}$ and the oxygen requirement in kg-O_2 h^{-1}. Eq. (21) expresses the fact that the

maximum volumetric oxygen transfer rate that can be economically achieved by actual aerators is around 0.10 kg-O_2 m^{-3} h^{-1}. Eq. (22) limits the aeration rate so as to prevent bioflocs shearing (a Lim_U value of 90 m^3/(min.1000 m^3) was used). Eq. (23) indicates the minimum air flow rate to be injected in order to ensure mixing (a value of Lim_L of 20 m^3/(min.1000 m^3) was used). For similar reasons, Dey and coworkers [65] have restricted the power input of the surface aerator between 10 W m^{-3} [72] and 90 W m^{-3} [66]. Hreiz and coworkers [47,48] used Eq. (24) to set the upper bound for the air flow rate. However, they assumed a Lim_U value of 50 m^3-/(min.1000 m^3), so as the maximum air flow rate represents the capacity of aerators typically used nowadays.

8. Conclusion

Nowadays, the activated sludge process (ASP) is the most commonly used biological secondary treatment for nutrient removal from sewage water. Given the strengthening of the environmental regulations, it is necessary to improve ASPs operation and design, but also from an economic point of view, to reduce their operating costs. However, even for expert engineers, determining the optimal operating strategy for ASPs remains quite difficult and laborious given the complexity of the underlying biochemical phenomena and their interaction, the large number of operating parameters, and the variety of objectives to deal with (e.g. enhancing the treatment efficiency, minimizing the exploitation charges). In this context, dynamic optimization reveals to be a powerful tool for assisting and supporting designers in determining the optimal operating conditions for existing WWTPs, or simultaneously predicting the optimal design and operation for future plants.

However, the real optimality of the computed solution strongly depends on the optimization problem formulation. For example, the preference between alternative operating strategies is closely related to the objective functions used like cost functions, for which there is unfortunately no commonly approved expressions. Moreover, there is no model that can exactly predict the complex biological phenomena occurring in ASPs. The temporal variations of the wastewater characteristics, which constitute the main disturbance for the ASP operation, cannot be precisely known or estimated, since it depends on weather-related phenomena such as rainfall. For these reasons, engineers may lack confidence regarding the reliability of the computed solution, and even the true potential of optimization. In an attempt to settle these misgivings and to help engineers choosing appropriate and reliable formulations for their optimization problem, this paper has reviewed the literature devoted to optimal control and design of ASPs. The most important issues that should be addressed in the order to get reliable and realistic solutions have been discussed: (1) Managing the inevitable mismatch between the model predictions and the real ASP plant. (2) Dealing with the unpredictable variations in the wastewater characteristics. (3) Accounting for the slowest dynamic processes occurring in ASPs, of which the characteristic time is of the order of several days, so as to get a solution guaranteeing a sustainable plant functioning. (4) Appropriately selecting the decision variables and the flowsheet structure in order to simplify the problem formulation from a mathematical perspective: hence, the chances that the optimizer determines a satisfying solution are increased. (5) Conveniently choosing the cost functions correlations. The limits and potentials of the expressions used in literature are discussed, so as to help designers selecting appropriate and reliable correlations. (6) Successfully selecting the mathematical constraints in order to guarantee physically relevant operations, and ensure that the feasible solutions remain within the limits of validity of the model.

Acknowledgments

The authors gratefully thank CARNOT Institutes of Nancy (ICEEL – Lorraine University) and Marseille (STAR – Aix-Marseille University) for the funding of the project oWWTP, and Mr. Wassim Hraiz for the picture provided for the journal cover.

References

[1] M.A.Z. Coelho, C. Russo, O.Q.F. Araújo, Optimization of a sequencing batch reactor for biological nitrogen removal, Water Res. 34 (2000) 2809–2817.
[2] H. Kim, T.J. McAvoy, J.S. Anderson, O.J. Hao, Control of an alternating aerobic–anoxic activated sludge system–part 2: optimization using a linearized model, Control Eng. Pract. 8 (2000) 279–289.
[3] S.M. Souza, O.Q.F. Araújo, M.A.Z. Coelho, Model-based optimization of a sequencing batch reactor for biological nitrogen removal, Bioresour. Technol. 99 (2008) 3213–3223.
[4] F.J. Fernández, M.C. Castro, M.A. Rodrigo, P. Cañizares, Reduction of aeration costs by tuning a multi-set point on/off controller: a case study, Control Eng. Pract. 19 (2011) 1231–1237.
[5] R.J. Seviour, P.H. Nielsen, Microbial ecology of activated sludge, IWA Publishing, 2010.
[6] B. Chachuat, N. Roche, M.A. Latifi, Dynamic optimisation of small size wastewater treatment plants including nitrification and denitrification processes, Comput. Chem. Eng. 25 (2001) 585–593.
[7] J. Luo, L.T. Biegler, Dynamic optimization of aeration operations for a benchmark wastewater treatment plant, Preprints of the 18th IFAC World Congress, Milano, Italy, 2011.
[8] C.G. Moles, G. Gutierrez, A.A. Alonso, J.R. Banga, Integrated process design and control via global optimization: a wastewater treatment plant case study, Chem. Eng. Res. Des. 81 (2003) 507–517.
[9] S. Balku, R. Berber, Dynamics of an activated sludge process with nitrification and denitrification: start-up simulation and optimization using evolutionary algorithm, Comput. Chem. Eng. 30 (2006) 490–499.
[10] J. Guerrero, A. Guisasola, R. Vilanova, J.A. Baeza, Improving the performance of a WWTP control system by model-based setpoint optimization, Environ. Model. Softw. 26 (2011) 492–497.
[11] S. Rigopoulos, P. Linke, Systematic development of optimal activated sludge process designs, Comput. Chem. Eng. 26 (2002) 585–597.
[12] M. Schlüter, J.A. Egea, L.T. Antelo, A.A. Alonso, J.R. Banga, An extended ant colony optimization algorithm for integrated process and control system design, Ind. Eng. Chem. Res. 48 (2009) 6723–6738.
[13] J.A. Egea, M. Rodríguez-Fernández, J.R. Banga, R. Martí, Scatter search for chemical and bio-process optimization, J. Global Optim. 37 (2007) 481–503.
[14] O. Exler, L.T. Antelo, J.A. Egea, A.A. Alonso, J.R. Banga, A Tabu search-based algorithm for mixed-integer nonlinear problems and its application to integrated process and control system design, Comput. Chem. Eng. 32 (2008) 1877–1891.
[15] V. Bhaskar, S.K. Gupta, A.K. Ray, Applications of multi-objective optimization in chemical engineering, Rev. Chem. Eng. 16 (2000) 1–54.
[16] A. Ghosh, S. Dehuri, Evolutionary algorithms for multi-criterion optimization: a survey, Int. J. Comput. Inf. Sci. 2 (2004) 38–57.
[17] A. Konak, D.W. Coit, A.E. Smith, Multi-objective optimization using genetic algorithms: a tutorial, Reliab, Eng. Syst. Saf. 91 (2006) 992–1007.
[18] J. Hakanen, K. Sahlstedt, K. Miettinen, Simulation-based interactive multiobjective optimization in wastewater treatment, in: Int. Conf. Eng. Optim., Rio de Janeiro, Brazil, 2008.
[19] J. Hakanen, K. Sahlstedt, K. Miettinen, Wastewater treatment: new insight provided by interactive multiobjective optimization, Decis. Support. Syst. 51 (2011) 328–337.
[20] J. Hakanen, K. Sahlstedt, K. Miettinen, Wastewater treatment plant design and operation under multiple conflicting objective functions, Environ. Modell. Software 46 (2013) 240–249.
[21] W.-M. Xie, R. Zhang, W.-W. Li, B.-J. Ni, F. Fang, G.-P. Sheng, H.-Q. Yu, J. Song, D.-Z. Le, X.-J. Bi, C.-Q. Liu, M. Yang, Simulation and optimization of a full-scale carrousel oxidation ditch plant for municipal wastewater treatment, Biochem. Eng. J. 56 (2011) 9–16.
[22] M. Henze, C.P. Grady, W. Gujer, G. Marais, T. Matsuo, Activated Sludge Model No. 1. Technical Report 1, IAWQ, London, 1987.
[23] W. Gujer, M. Henze, T. Mino, M. Van Loosdrecht, Activated sludge model No. 3, Water Sci. Technol. 39 (1999) 183–193.
[24] W. Gujer, M. Henze, T. Mino, T. Matsuo, M.C. Wentzel, G.V.R. Marais, The activated sludge model No. 2: biological phosphorus removal, Water Sci. Technol. 31 (1995) 1–11.
[25] M. Henze, W. Gujer, T. Mino, T. Matsuo, M.C. Wentzel, G.V.R. Marais, M.C.M. Van Loosdrecht, Activated sludge model no 2d, ASM2d, Water Sci. Technol. 39 (1999) 165–182.
[26] J.A. Egea, D. Vries, A.A. Alonso, J.R. Banga, Global optimization for integrated design and control of computationally expensive process models, Ind. Eng. Chem. Res. 46 (2007) 9148–9157.
[27] F. Fang, B. Ni, W. Li, G. Sheng, H. Yu, A simulation-based integrated approach to optimize the biological nutrient removal process in a full-scale wastewater treatment plant, Chem. Eng. J. 174 (2011) 635–643.

[28] W. El-Shorbaghy, A. Arwani, R.L. Droste, Optimal sizing of activated sludge process with ASM3, Int. J. Civ. Environ. Eng. 11 (2011) 19–55.
[29] I.A.C.P. Espírito-Santo, E.M.G.P. Fernandes, M.M. Araújo, E.C. Ferreira, How wastewater processes can be optimized using LOQO, Lect. Notes Econ. Math. Syst. 563 (2006) 435–455.
[30] Z. Li, R. Qi, B. Wang, Z. Zou, G. Wei, M. Yang, Cost-performance analysis of nutrient removal in a full-scale oxidation ditch process based on kinetic modeling, J. Environ. Sci. 25 (2013) 26–32.
[31] I. Takács, G.G. Patry, D. Nolasco, A dynamic model of the clarification-thickening process, Water Res. 25 (1991) 1263–1271.
[32] B. Chachuat, Méthodologie d'optimisation dynamique et de commande optimale des petites stations d'épuration à boues activées (Ph.D. dissertation), Institut National Polytechnique de Lorraine, 2001 [in French].
[33] B. Chachuat, N. Roche, M.A. Latifi, Long-term optimal aeration strategies for small-size alternating activated sludge treatment plants, Chem. Eng. Process. 44 (2005) 593–606.
[34] W. El-Shorbaghy, N. Nawras, R.L. Droste, Optimization of A2O BNR processes using ASM and EAWAG Bio-P models: model formulation, Water Qual. Res. J. Can. 46 (2011) 13–27.
[35] I.A.C.P. Espírito-Santo, E.M.G.P. Fernandes, M.M. Araújo, E.C. Ferreira, On the secondary settler models robustness by simulation, WSEAS Trans. Inf. Sci. Appl. 3 (2006) 2323–2330.
[36] I.A.C.P. Espírito-Santo, E.M.G.P. Fernandes, M.M. Araújo, E.C. Ferreira, Wasted sludge treatment contributions in the WWTP total cost, WSEAS Trans. Inf. Sci. Appl. 4 (2007) 655–662.
[37] G.A. Ekama, J.L. Barnard, F.W. Günthert, P. Krebs, J.A. McCorquodale, D.S. Parker, E.J. Wahlberg, Secondary settling tanks: theory, modeling, design and operation, Technical report 6, International Association on Water Quality (IAWQ), 1978.
[38] N. Alasino, M.C. Mussati, N. Scenna, Wastewater treatment plant synthesis and design, Ind. Eng. Chem. Res. 45 (2007) 7497–7512.
[39] N. Alasino, M.C. Mussati, N. Scenna, P. Aguirre, Wastewater treatment plant synthesis and design: combined biological nitrogen and phosphorus removal, Ind. Eng. Chem. Res. 49 (2010) 8601–8612.
[40] A. Rivas, I. Irizar, E. Ayesa, Model-based optimisation of wastewater treatment plants design, Environ. Modell. Software 23 (2008) 435–450.
[41] A. Rivas, E. Ayesa, A. Galarza, A. Salterain, Application of mathematical tools to improve the design and operation of activated sludge plants. Case study: the new WWTP of Galindo-Bilbao. Part I: optimum design, Water Sci. Technol. 43 (2001) 157–165.
[42] I. Iqbal, C. Guria, Optimization of an operating domestic wastewater treatment plant using elitist non-dominated sorting genetic algorithm, Chem. Eng. Res. Des. 87 (2009) 1481–1496.
[43] L. Costa, I.A.C.P. Espírito-Santo, E.M.G.P. Fernandes, R. Denysiuk, Using a genetic algorithm to solve a bi-objective WWTP process optimization, Oper. Res. Proc. (2011) 359–364.
[44] B. Beraud, J.-P. Steyer, C. Lemoine, E. Latrille, Optimization of WWTP control by means of multi-objective genetic algorithms and sensitivity analysis, Comput. Aided Chem. Eng. 25 (2008) 539–544.
[45] J.B. Copp, The COST simulation benchmark. Description and simulator manual, Office for Official Publications of the European Communities, Luxembourg, 2002.
[46] J. Alex, L. Benedetti, J.B. Copp, K.V. Gernaey, U. Jeppsson, I. Nopens, M.N. Pons, L. Rieger, C. Rosen, J.P. Steyer, P. Vanrolleghem, S. Winkler, Benchmark Simulation Model No. 1 (BSM1), IWA taskgroup on Benchmarking of Control Strategies for WWTPs, Technical Report, 2008.
[47] Rainier Hreiz, Nicolas Roche, Brahim Benyahia, M.A. Latifi, Multi-objective optimal control of small-size wastewater treatment plants, Chem. Eng. Res. Des. http://dx.doi.org/10.1016/j.cherd.2015.06.039.
[48] R. Hreiz, N. Roche, B. Benyahia, M.A. Latifi, Multi-objective optimal control and design of large-size wastewater treatment plants, ECCE congress 2015, Nice, France, Extended version in preparation.
[49] M. Fikar, B. Chachuat, M.A. Latifi, Optimal operation of alternating activated sludge processes, Control Eng. Pract. 13 (2005) 853–861.
[50] S. Balku, Comparison between alternating aerobic-anoxic and conventional activated sludge systems, Water Res. 41 (2007) 2220–2228.
[51] B. Chachuat, N. Roche, M.A. Latifi, Optimal aeration control of industrial alternating activated sludge plants, Biochem. Eng. J. 23 (2005) 277–289.
[52] T.A. Doby, D.H. Loughlin, F.L. De Los Reyes, J.J. Ducoste, Optimization of activated sludge designs using genetic algorithms, Water Sci. Technol. 45 (2002) 187–198.
[53] J.A. Egea, I. Gracia, Dynamic multiobjective global optimization of a waste water treatment plant for nitrogen removal, in: 7th Vienna Int. Conf. on Math. Model., 2012, pp. 374–379.
[54] G.S. Ostace, J.A. Baeza, J. Guerrero, A. Guisasola, V.M. Cristea, P.S. Agachi, J. Lafuente, Development and economic assessment of different WWTP control strategies for optimal simultaneous removal of carbon, nitrogen and phosphorus, Comput. Chem. Eng. 53 (2013) 164–177.
[55] J. Guerrero, A. Guisasola, J. Comas, I. Rodrígez-Roda, J.A. Baeza, Multi-criteria selection of optimum WWTP control setpoints based on microbiology-related failures, effluent quality and operating costs, Chem. Eng. J. 188 (2012) 23–29.
[56] N. Descoins, S. Deleris, R. Lestienne, E. Trouvé, F. Maréchal, Energy efficiency in waste water treatments plants: optimization of activated sludge process coupled with anaerobic digestion, Energy 41 (2012) 153–164.
[57] N. Brand, A. Ostfeld, Optimal design of regional wastewater pipelines and treatment plant systems, Water Environ. Res. 83 (2011) 53–64.
[58] G. Fu, D. Butler, S.-T. Khu, Multiple objective optimal control of integrated urban wastewater systems, Environ. Modell. Software 23 (2008) 225–234.
[59] D.H. Nguyen, M.A. Latifi, F. Lesage, F. Mhlanga, M. Mulholland, C. Buckely, Récent Progrès en Génie des Procédés 104, 2013, SFGP, Paris, France.
[60] M.C. Mussati, S.F. Mussati, N. Alasino, P. Aguirre, N. Scenna, Optimal synthesis of activated sludge wastewater treatment plants for nitrogen removal, in: 2nd Mercosur Congr. on Chem. Eng., 4th Mercosur Congr. on Process Syst. Eng., Rio De Janeiro, Brazil, 2005.
[61] S.E. Scuras, A. Jobbagy, C.P. Leslie Grady Jr., Optimization of activated sludge reactor configuration: kinetic considerations, Water Res. 35 (2001) 4277–4284.
[62] J.-L. Vasel, Contribution à l'étude des transferts d'oxygène en gestion des eaux (Ph.D. dissertation), Fondation Universitaire Luxembourgeoise, 1988 [in French].
[63] P. Painmanakul, G. Hébrard, Effect of different contaminants on the α-factor: local experimental method and modeling, Chem. Chem. Eng. Res. Des. 86 (2008) 1207–1215.
[64] I. Seyssiecq, J.-H. Ferrasse, N. Roche, State-of-the-art: rheological characterisation of wastewater treatment sludge, Biochem. Eng. J. 16 (2003) 41–56.
[65] A. Dey, D. Truax, B. Magbuna, Optimization of operating parameters of intermittent aeration-type activated sludge process for nitrogen removal: a simulation-based approach, Water Environ. Res. 83 (2011) 636–642.
[66] C.P.L. Grady Jr., G.T. Daigger, H.C. Lim, Biological Wastewater Treatment, Marcel Dekker, Inc., New York, USA, 1999.
[67] A. Abusam, K.J. Keesman, G. Van Straten, H. Spanjers, K. Meinema, Sensitivity analysis in oxidation ditch modelling: the effect of variations in stoichiometric, kinetic and operating parameters on the performance indices, J. Chem. Technol. Biotechnol. 76 (2001) 430–438.
[68] A. Stare, D. Vrečko, N. Hvala, Z. Strmčnik, Comparison of control strategies for nitrogen removal in an activated sludge process in terms of operating costs: a simulation study, Water Res. 41 (2007) 2004–2014.
[69] S. Gillot, B. De Clercq, D. Defour, F. Simoens, K. Gernaey, P.A. Vanrolleghem, Optimization of wastewater treatment plant design and operation using simulation and cost analysis, in: Proceedings 72nd Annual WEF Conference and Exposition. New Orleans, USA, 1999.
[70] S. Gillot, A. Héduit, Prédiction des capacités d'oxygénation en eau claire des systèmes d'insufflation d'air, Technical report FNDAE 31, 2004 [in French].
[71] B. Marrot, A. Barrios-Martinez, P. Moulin, N. Roche, Experimental study of mass transfer phenomena in a cross-flow membrane bioreactor: aeration and membrane separation, Eng. Life Sci. 5 (2005) 409–414.
[72] B.E. Rittman, P.L. McCarty, Environmental Biotechnology: Principles and Application, Mc-Graw Hill, New York, 2001.
[73] P.D. Zakkour, M.R. Gaterell, P. Griffin, R.J. Gochin, J.N. Lester, Anaerobic treatment of domestic wastewater in temperate climates: treatment plant modelling with economic considerations, Water Res. 35 (2001) 4137–4149.
[74] P. Samuelsson, B. Halvarsson, B. Carlsson, Cost-efficient operation of a denitrifying activated sludge process, Water Res. 41 (2007) 2325–2332.
[75] P.A. Vanrolleghem, U. Jeppsson, J. Carstensen, B. Carlsson, G. Olsson, Integration of wastewater treatment plant design and operation – a systematic approach using cost functions, Water Sci. Technol. 34 (1996) 159–171.
[76] R. Turton, R.C. Bailie, W.B. Whiting, J.A. Shaeiwitz, Analysis, Synthesis, and Design of Chemical Processes, Prentice Hall, New Jersey, 1998.
[77] P. Vermeire, Economische optimalisatie van waterzuiveringsstations, Ontwikkeling van investeringskostenfuncties voor Vlaanderen, Engineers Thesis, Faculty of Agricultural and Applied Biological Sciences, University Gent, Belgium, 1999 [in Dutch].
[78] S. Gillot, P. Vermeire, P. Jacquet, H. Grootaerd, D. Derycke, F. Simoeons, P.A. Vanrolleghem, Integration of wastewater treatment plant investment and operating costs for scenario analysis using simulation, in: Proc. 13th Forum Appl. Biotechnol., Med. Fac. Landbouww, Univ. Gent., Belgium, 1999.
[79] L. Åmand, B. Carlsson, Optimal aeration control in a nitrifying activated sludge process, Water Res. 46 (2012) 2101–2110.
[80] B. Holenda, E. Domokos, Á. Rédey, J. Fazakas, Aeration optimization of a wastewater treatment plant using genetic algorithm, Optim. Control Appl. Meth. 28 (2007) 191–208.
[81] H. Liu, K.-H. Chang, C. Yoo, Multi-objective optimization of cascade controller in combined biological nitrogen and phosphorus removal wastewater treatment plant, Desalin. Water Treat. 43 (2012) 138–148.