Model predictive control for the self-optimized operation in wastewater treatment plants: Analysis of dynamic issues

M. Francisco [a,*], S. Skogestad [b], P. Vega [a]

[a] Department of Computing and Automation, University of Salamanca, Salamanca, Spain
[b] Department of Chemical Engineering, Norwegian University of Science and Technology, Trondheim, Norway

ABSTRACT

This paper describes a procedure to find the best controlled variables in an economic sense for the activated sludge process in a wastewater treatment plant, despite the large load disturbances. A novel dynamic analysis of the closed loop control of these variables has been performed, considering a non-linear model predictive controller (NMPC) and a particular distributed NMPC-PI control structure where the PI is devoted to control the process active constraints and the NMPC the self-optimizing variables. The well-known self-optimizing control methodology has been applied, considering the most important measurements of the process. This methodology provides the optimum combination of measurements to keep constant with minimum economic loss. In order to avoid nonfeasible dynamic operation, a pre-selection of the measurements has been performed, based on the nonlinear model of the process and evaluating the possibility of keeping their values constant in the presence of typical disturbances.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

The efficiency of most wastewater treatment plants (WWTP) is an important issue still to be improved. In order to fulfill the imposed legal effluent requirements for large load variations the operating costs are usually higher than the actually needed. Optimization of WWTP would provide a significant cost reduction, but it has not been extensively considered yet. There are some works in the literature, but most of them only consider the problem from a heuristic viewpoint or stating a particular optimization problem. In Stare et al. (2007), different control strategies are proposed and compared in terms of the operating costs, which are evaluated but not optimized. Other works, such as Ingildsen et al. (2002), Machado et al. (2009) and Samuelsson et al. (2007), tackle the problem of reducing costs, but not in a systematic way. Some of them (Francisco et al., 2011; Rivas et al., 2008) also include plant design, and others are only focused on tanks aeration (Amand and Carlsson, 2012). Only Araujo et al. (2011, 2013) provides a comprehensive approach, performing a sensitivity analysis of optimal operation for the selection of the best control structure in term of costs and effluent quality. The work of Cadet et al. (2004) is similar but without considering the economics of the system.

The aim is to satisfy effluent quality regulations with reasonable economic expenses. The WWTP influent variations are rather large making the plant to work away from the optimal operation point, with the subsequent economic loss. One possible approach to overcome this is the re-optimization of the plant when some disturbances occur by applying Real Time Optimization techniques (Darby et al., 2011), which can be very demanding computationally, or perform the set-point optimization off-line (Machado et al., 2009; Guerrero et al., 2011). In this work, the approach considered is the selection of some controlled variables that when kept constant, the economic loss is small with respect to costs when the operation is re-optimized. The methodology used to find these variables is the self-optimizing procedure of Skogestad (2000). The WWTP model considered for its application is the widely used Benchmark Simulation Model No. 1 (BSM1), described in Alex et al. (2008).

The appropriate control structure selection is crucial for the optimal operation of plants. The decisions on which variables should be controlled, which should be measured, and which inputs should be manipulated are part of the control structure selection. Generally, these decisions are based on heuristic methods that cannot guarantee optimality, but in this work, self-optimizing control (SOC) is applied, which is a methodology for the selection of the best controlled variables that minimize operating costs, considering a steady state of the process. The initial quantitative ideas related to self-optimizing control are presented in Morari et al. (1980), and later, Skogestad (2000) defined the problem more

* Corresponding author. Tel.: +34 923294500x1926.
 E-mail addresses: mfs@usal.es (M. Francisco), skoge@ntnu.no (S. Skogestad), pvega@usal.es (P. Vega).

precisely, including also implementation error. The methodology is well-developed for linear models that generate quadratic optimization problems, and for that reason a nonlinear validation is needed. The "exact local method" (Halvorsen et al., 2003) provides the best controlled variables for both disturbances and implementation errors, and it was the first author that proposed linear combinations of measurements as controlled variables, determined by a matrix **H**. The problem of finding such combination may be reformulated as a quadratic optimization problem with linear constraints (Alstad et al., 2009), and the analytical solution to this problem provides a straightforward way to obtain the matrix **H** (Yelchuru and Skogestad, 2011). Specific worst-case and average loss minimization have been also proposed (Kariwala, 2007; Kariwala et al., 2008), and the use of branch and bound methods has been introduced to enable the application of the methodology to large-dimensional processes (Cao and Kariwala, 2008).

For processes with large load disturbances it is common that the steady state is difficult to reach. Then, the controlled variables can be determined and adapted on the basis of an algorithm that tracks the necessary conditions of optimality (NCO), making the SOC adaptive to operating conditions changes. Model-free NCO tracking procedure using finite perturbations to calculate the gradients has been developed in Srinivasan et al. (2008). The regression-based approach (Ye et al., 2013) and its extension to hierarchical control (Ye et al., 2014) provide a new methodology to determine CVs approximating the necessary conditions of optimality (NCO) in the whole operating region, achieving near-optimal operation globally, enlarging the operation region where the economic loss is acceptable. In Jäschke and Skogestad (2011), it is shown that NCO tracking in the optimization layer and SOC in the lower control layer are complementary methodologies because unexpected disturbances, which are not rejected by SOC, can be handled by the model free NCO tracking procedure. There are also other methodologies based on neighboring-extremal control (NEC) (Gros et al., 2009), where the gradients are evaluated by model based approaches, but eventually they are also local approximations. Another possibility to deal with large disturbances is the use of dynamic SOC, where the operational cost during transient response is taken into account. In Hu et al. (2012), a formulation of dynamic SOC which considers economic cost and set-point tracking cost at the same time has been developed, stating a multiobjective optimization problem equivalent to an optimal control problem.

The work presented in this article is one of the first approaches to the SOC of the Benchmark Simulation Model No. 1 (BSM1) of a WWTP, which is a complicated nonlinear process benchmark, whose optimization is a difficult task. Therefore, the local SOC approximation has been chosen as a starting point, in order to find possible difficulties in the methodology and to propose the basis for implementing a more complex structure, particularly the one described in Jäschke and Skogestad (2011). With this approach, the SOC variables control decreases the operating costs with respect to the single active constraints control, which is the first step to improve the plant economy and safety (Maarleveld and Rijnsdorp, 1970). Moreover, in this article the focus is more on checking the controllability of the SOC variables and active constraints by using an advanced controller, particularly a NMPC.

Although there are plenty of successful works of SOC (see e.g. Umar et al., 2012) the dynamic validation of results is usually performed by means of decentralized PI controllers, making a previous pairing with variables (Larsson et al., 2001; Araujo and Skogestad, 2008). In Alstad (2005), the dynamic performance has been improved adding compensators on the measurements to avoid right half plane zeros, and the effect of the basis vectors for the null space method on poles and zeros has been studied. In Baldea et al. (2008), a singular perturbation-based framework has been employed, which accounts for the time scale separation

present in the open loop dynamics of integrated plants, resulting in a controller design procedure that accounts for both economical optimality and dynamic performance. It is important to note that controllability can be improved by changing the matrix of combinations measurements **H** (Alstad, 2005), reducing the coupling and allowing for the implementation of decentralized PI controllers. Regarding to the active constraints control in SOC, only works dealing with dynamic validation of the control structures select a particular controller for that task. When there are active constraints only for manipulated variables, it is straightforward to keep constant the corresponding variable. For example, in Araujo and Skogestad (2008) the operation is optimal for maximum cooling in the heat exchangers for the ammonia synthesis process. In other situations, when constraints are active for some measurements, decoupled PI controllers based on the RGA matrix study are proposed (Alstad, 2005; Gera et al., 2013). In addition, if the set of active constraints changes depending on the disturbances affecting the process, the self-optimizing variables have to be recalculated following a systematic procedure, as in Manum and Skogestad (2012), that determines the different regions using a parametric program, based on a link with explicit model predictive control. Another approach is the implementation of a cascade control structure to satisfy both optimality and constraint requirements (Cao, 2003). In Jacobsen and Skogestad (2011), a methodology for finding active constraints regions is also proposed.

For the case of the WWTP, it has been proved that the set of active constraints does not change with the disturbances, but there is some coupling between active constraints control and the control of the self-optimizing variables. This coupling cannot be fully removed only by changing the matrix **H** of measurements combinations, and therefore in this work, a multivariable NMPC controller is considered as a novelty. The NMPC is a mature control strategy, and in this case an offset free formulation is considered to tackle the plant model mismatch and unknown disturbances based on Pannocchia and Rawlings (2003) and the extension to nonlinear MPC in Morari and Maeder (2012), using an augmented model with an additional integrating disturbance vector and adapting the MPC reference to the current disturbance estimate.

The first objective of this work is to find the self-optimized variables for the BSM1 following the simple procedure of Yelchuru and Skogestad (2011), which considers set-point and implementation errors and provides a set of optimal controlled variables as combination of the available measurements. The second objective is to evaluate the dynamic behavior of those variables by implementing a multivariable constrained nonlinear model predictive controller (NMPC). In particular, two approaches have been considered: one centralized MPC controlling the active constraints and self-optimized variables, and a distributed control structure with an NMPC controlling the self-optimized variables and local PI controllers for the active constraints control.

This article is structured as follows. First, the WWTP is described, in particular the activated sludge process. Then, the controlled variables selection methodology is explained, and the local methods for self-optimizing control. In the next section the methodology is applied to the BSM1, followed by the process controllability analysis with the description of the NMPC formulations and the distributed control structure. The article ends with a dynamic analysis and conclusions.

2. Description of the process

The purpose of a wastewater treatment plant (WWTP) is to process sewage and return clean water to the river. Activated sludge process (ASP) is a very important part of the cleaning procedure, and the Benchmark Simulation Model No. 1 (BSM1) (Alex et al.,
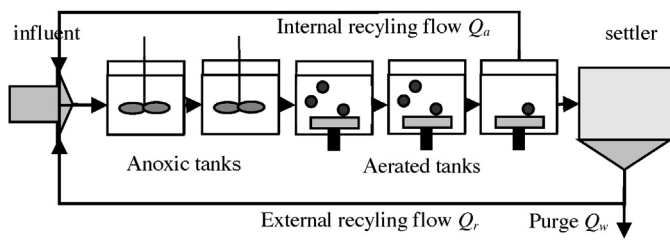
Fig. 1. Benchmark Simulation Model No. 1 (BSM1).

2008) has been used as a standard ASP model for performance assessment of control strategies and process optimization. The BSM1 consists of five biological reactors connected in series and one secondary settler. The reactors are modeled according to mass balances described in the Activated Sludge Model No. 1 (ASM1), developed by the IWAQ (International Association on Water Quality) (Henze et al., 1987).

As presented in Fig. 1, the nitrogen removal is achieved using a first denitrification step performed in the anoxic tanks, placed before the aerated tanks where the nitrification step is carried out. The first two anoxic tanks are assumed perfectly mixed and have a volume of $V_1 = V_2 = 1000\,\text{m}^3$. The rest of the three aerated tanks have a volume of $V_3 = V_4 = V_5 = 1333\,\text{m}^3$. Eight different processes are modeled, involving thirteen state variables at each bioreactor. An internal recycle (Q_a) from the last tank to the first one is used to supply the denitrification step with nitrate. In order to maintain the microbiological population, sludge from the settler is recirculated into the reactors by means of an external recycle (Q_r), and sludge excess is purged from the bottom of the settler (Q_w). For the secondary settler the one-dimensional ten-layer model implementing the double exponential settling velocity model has been used (Takacs et al., 1991). The total volume of the settler is $V_{dec} = 6000\,\text{m}^3$. The full dimensions of the BSM1, the value of the kinetic and stoichiometric parameters and the influent characteristics are defined in Alex et al. (2008).

3. Controlled variables selection

In this section a summary of the plant-wide design procedure is given (top-down part), where the controlled variables selection is a fundamental issue. The main steps of the methodology are described below.

• Define operational objectives and constraints:

The first step is the selection of a scalar cost function that comprises the whole operational plant costs. A typical cost function for a nonproductive plant includes only energy balances. The operational constraints imposed by process safety, environmental regulations, control limitations, and product specifications, must be also identified.

• Determine the steady state optimal operation:

In this work we assume that the economics of the plant are primarily determined by steady state behavior, so the steady state degrees of freedom are the same as the economic degrees of freedom. The number of steady state degrees of freedom can be obtained counting the number of dynamic manipulated variables and subtracting the number of degrees of freedom with no steady state effect. These variables can be identified from the flow sheet of the process. Particularly, the most relevant degrees of freedom are the ones that affect the cost function. In the SOC methodology, the

number of steady state degrees of freedom determines the number of steady-state controlled variables that need to be selected.

In order to state the optimization problem to determine the process optimal operation, the most relevant disturbances are selected previously evaluating the sensitivity of the cost function to each disturbance and choosing those with the highest sensitivity.

Then, the following optimization problem is stated, considering nominal disturbances $\mathbf{d}_0$:

$$\min_{\mathbf{u}_0} \quad J_0(\mathbf{x}, \mathbf{u}_0, \mathbf{d}_0) \tag{1}$$

subject to:

$$g_1(\mathbf{x}, \mathbf{u}_0, \mathbf{d}_0) = 0 \,(\text{Steady state model equations})$$

$$g_2(\mathbf{x}, \mathbf{u}_0, \mathbf{d}_0) \leq 0 \,(\text{process constraints and others})$$

where $\mathbf{x}$ is the state vector, $\mathbf{u}_0$ the steady state degrees of freedom and $\mathbf{d_0}$ the disturbances vector.

• Selection of the active constraints for control:

The active constraints found when solving (1) must be controlled tightly for optimal operation (active constraints control), consuming part of the degrees of freedom. For further application of the SOC methodology, as it is based on local analysis, it is assumed that the set of active constraints does not change for all typical disturbances. Although this assumption is important, it is satisfied for many processes allowing for a direct application of SOC. If this is not the case, the next steps of the methodology must be performed for each set of active constraints (Jacobsen and Skogestad, 2011).

• Select primary controlled variables (CV):

In this point, the identification of as many economic controlled variables as the number of remaining degrees of freedom is performed. Then, close-to-optimal operation is achieved with constant nominal set-points, even when disturbances appear.

Firstly, all candidate measurements must be identified, together with the expected static measurement error. The measurements can also include manipulated variables (e.g. flow rate measurements) or measured disturbances. Then, the economic controlled variables (also called self-optimized variables) can be selected as single measurements or combination of measurements, defined by a coefficient matrix $\mathbf{H}$ with real constant terms, where $\mathbf{c}$ is the vector of controlled variables and $\mathbf{y}$ is the vector of available independent measurements:

$$\mathbf{c} = \mathbf{H} \cdot \mathbf{y} \tag{2}$$

The selection can be based on a qualitative approach, following heuristic rules, or based on a quantitative approach, such as the local approach based on the Taylor expansion of the loss function around the equilibrium nominal point $\mathbf{u}_{opt}(\mathbf{d})$. In order to achieve near-optimal operation without the need to re-optimize the process when disturbances occur, the loss must be minimized:

$$L(\mathbf{u}, \mathbf{d}) = J_c(\mathbf{u}, \mathbf{d}) - J_{opt}(\mathbf{u}_{opt}(\mathbf{d}), \mathbf{d}) = \frac{1}{2}[\mathbf{u} - \mathbf{u}_{opt}(\mathbf{d})]^T \mathbf{J}_{uu}[\mathbf{u} - \mathbf{u}_{opt}(\mathbf{d})] \tag{3}$$

where J_c is the cost value when the set-point is kept constant, and J_{opt} is the optimum cost re-optimizing for the corresponding $\mathbf{d}$, $\mathbf{u}_{opt}$ is the optimum value for $\mathbf{u}$ and $\mathbf{J}_{uu}$ is the Hessian of the cost function. Some particular techniques based on that are the maximum gain rule (Halvorsen et al., 2003), the nullspace method (Alstad and Skogestad, 2007) and the exact local method (Alstad et al., 2009).

• Nonlinear validation of losses:

The procedure described is of local nature, hence the proposed controlled variables are not globally optimal and the actual non-linear losses must be calculated in order to validate the results. If nonfeasibilities appear for some disturbances, the set of selected controlled variables should be dismissed.

• Dynamic analysis of the selected CV:

The last step of the methodology is the dynamic analysis of the control of selected CV, by means of classical PIDs or more advance multivariable controllers, when typical disturbances are applied to the process. In this case, the steady state degrees of freedom $\mathbf{u}_0$ are the manipulated variables of the controllers, whose objective is to maintain the self-optimized variables $\mathbf{c}$ at the optimal set-points.

4. Local methods for self-optimizing control

Exact local methods give loss expressions for the worst and average case of disturbances. Although CV can be selected as a subset of the available measurements, lower loss is achieved by selecting CV as linear combinations of measurements. For that reason, a combination matrix $\mathbf{H}$ is defined as in (2), and it can be found through minimization of the following expression (Halvorsen et al., 2003; Alstad et al., 2009) that considers combined disturbances and measurement errors and it is valid for average and worst case of disturbances:

$$\min_{\mathbf{H}} \left\| \mathbf{J}_{uu}^{1/2}(\mathbf{HG}^y)^{-1}\mathbf{HY} \right\|_F \qquad (4)$$

where

$$\mathbf{F} = \mathbf{G}_d^y - \mathbf{G}^y\mathbf{J}_{uu}^{-1}\mathbf{J}_{ud}; \quad \mathbf{Y} = [\ \mathbf{FW}_d \quad \mathbf{W}_n\ ]; \quad \mathbf{y} = \mathbf{G}^y\mathbf{u} + \mathbf{G}_d^y\mathbf{d}$$

$\mathbf{W}_d$ and $\mathbf{W}_n$ are scaling matrices for disturbances and implementation errors, $\mathbf{G}^y$ and $\mathbf{G}_d^y$ are the process transfer matrices (linearized model), and $\mathbf{J}_{uu}, \mathbf{J}_{ud}$ are the Hessians:

$$\mathbf{J}_{uu} = \frac{\partial^2 J}{\partial \mathbf{u}^2}, \quad \mathbf{J}_{ud} = \frac{\partial^2 J}{\partial \mathbf{u}\partial \mathbf{d}}$$

Although this optimization problem seems to be nonconvex, it can be reformulated as a convex problem for the case when $\mathbf{H}$ has no particular structure:

$$\min_{\mathbf{H}} \left\| \mathbf{HY} \right\|_F$$
$$\text{s.t.} \quad \mathbf{HG}^y = \mathbf{J}_{uu}^{1/2} \qquad (5)$$

An analytical solution to this problem is given by Alstad et al. (2009) and it has been simplified by Yelchuru and Skogestad (2011):

$$\mathbf{H}^T = (\mathbf{YY}^T)^{-1}\mathbf{G}^y\mathbf{Q} \qquad (6)$$

where $\mathbf{Q}$ is any nonsingular matrix of $n_c \times n_c$ (n_c is the number of controlled variables).

5. Methodology applied to the BSM1

In this section, the described methodology is applied to the ASP in the BSM1 benchmark.

A. Operational objectives and constraints

The operational objectives of the WWTP include operational costs and other process and regulations constraints. The cost function defined in Alex et al. (2008) has been considered:

$$J = k_E(AE + PE + ME) + k_D SP \qquad (7)$$

where PE is the pumping energy cost, AE is the aeration energy, ME is the mixing energy and SP is the sludge production cost; k_E

Table 1
Constraints.

Effluent constraints	Constraints on the manipulated variables
$COD_e \leq 100\ (\mathrm{gCOD/m^3})$	$0 \leq K_L a_{1-5} \leq 360\ (\mathrm{m^3/d})$
$BOD_{5,e} \leq 10\ (\mathrm{gBOD/m^3})$	$Q_w \leq 1844.6\ (\mathrm{m^3/d})$
$TN_e \leq 18\ (\mathrm{gN/m^3})$	$Q_a \leq 92230\ (\mathrm{m^3/d})$
$TSS_e \leq 30\ (\mathrm{gSS/m^3})$	$Q_r \leq 36892\ (\mathrm{m^3/d})$
$S_{NH_e} \leq 4\ (\mathrm{gN/m^3})$	

Table 2
Disturbances.

	$Q^{(in)}$ (m³/d)	$COD^{(in)}$ (g/m³)	$TSS^{(in)}$ (g/m³)	T (°C)
$\mathbf{d}_0$	18,446	381.19	211.26	15
$\mathbf{d}_1$	21,320	332.81	183.38	15
$\mathbf{d}_2$	40,817	204.85	115.61	15
$\mathbf{d}_3$	19,746	353.87	195.89	15
$\mathbf{d}_{42}$	44,453	140.43	77.52	15
$\mathbf{d}_5$	20,851	347.56	198.73	15

and k_D are the weights representing the energy price and sludge disposal cost respectively. The expressions for all terms in (7) are the following:

$$PE = \frac{1}{\Theta}\int_{t_0}^{t_0+\Theta}(0.004 \cdot Q_a(t) + 0.008 \cdot Q_r(t) + 0.05 \cdot Q_w(t))dt \qquad (8)$$

$$AE = \frac{S_{O,sat}}{\Theta \cdot 1.8 \cdot 1000}\int_{t_0}^{t_0+\Theta}\sum_{k=1}^{k=5}V_k \cdot K_L a_k(t)dt \qquad (9)$$

$$ME = \frac{24}{\Theta}\left(\int_{t_0}^{t_0+\Theta}\sum_{k=1}^{5}\begin{bmatrix} 0.005 \cdot V_k & if & K_L a_k(t) < 20d^{-1} \\ 0 & & \text{otherwise} \end{bmatrix}\right) \qquad (10)$$

$$SP = \frac{1}{\Theta}\left(0.75 \cdot \int_{t_0}^{t_0+\Theta}TSS_w \cdot Q_w(t) \cdot dt\right) \qquad (11)$$

where Q_a is the internal recycling flow, Q_r is the external recycling flow, Q_w is the purge flow, $K_L a$ is the mass transfer coefficient for oxygen, $S_{O,sat}$ is the oxygen saturation concentration, and Θ is the range of time where the indexes are evaluated. The previous equation for sludge production represents the total solid flow from wastage and assuming that the amount of solids accumulated in the system over the period of time considered is null due to the steady state operation.

The constraints needed for process operability are listed in Table 1, where COD is the chemical oxygen demand, BOD_5 is the 5 day biological oxygen demand, TSS is the total suspended solids concentration, and TN is the total nitrogen concentration, all measured in the effluent.

B. Degrees of freedom and most relevant disturbances

Using the information given in the process flowsheet in Fig. 1, it can be found that there are eight manipulated variables that correspond to eight degrees of freedom ($\mathbf{u}$): $Q_a, Q_r, Q_w, K_L a^{(1-5)}$. The liquid levels in the reactor tanks are assumed to be constant at maximum capacity.

The disturbances selected are some of the most important inputs to the plant, which are the influent flow $Q^{(in)}$, the chemical oxygen demand in the influent $COD^{(in)}$, and the total suspended solids in the influent $TSS^{(in)}$. The total nitrogen in the influent $TN^{(in)}$ and the process temperature T have not been considered in the SOC methodology in order to simplify the results, but its inclusion is straightforward. The weather profile events and derived disturbances are presented in Table 2: $\mathbf{d}_0$ corresponds to the nominal load conditions (average during dry weather), $\mathbf{d}_1$ are the average load

Table 3
Steady state operating point.

	Optimal point results	Measurement range	Measurement noise
Cost	406.54		
Q_r (m³/d)	20,320	0–100,000	2500
Q_a (m³/d)	0		
Q_w (m³/d)	172.05		
$K_La^{(1)}$ (1/d)	133.98		
$K_La^{(2)}$ (1/d)	122.12		
$K_La^{(3)}$ (1/d)	97.95		
$K_La^{(4)}$ (1/d)	90.11		
$K_La^{(5)}$ (1/d)	91.08	0–360	9.000
$S_O^{(1)}$ (g/m³)	0.1323	0–10	0.250
$S_O^{(2)}$ (g/m³)	0.1744	0–10	0.250
$S_O^{(3)}$ (g/m³)	0.1442	0–10	0.250
$S_O^{(4)}$ (g/m³)	0.1426	0–10	0.250
$S_O^{(5)}$ (g/m³)	0.1671	0–10	0.250
$S_{NO}^{(1)}$ (g/m³)	1.4725	0–20	0.500
$S_{NO}^{(2)}$ (g/m³)	2.4598	0–20	0.500
$S_{NO}^{(3)}$ (g/m³)	3.3600	0–20	0.500
$S_{NO}^{(4)}$ (g/m³)	4.4497	0–20	0.500
$S_{NO}^{(5)}$ (g/m³)	6.0781	0–20	0.500
$MLSS$ (g/m³)	5385.5	0–10,000	250.00
TN_e (g/m³)	13.45		
COD_e (g/m³)	70.75		
BOD_{5e} (g/m³)	5.028		
S_{NHe} (g/m³)	4		
TSS_e (g/m³)	30		
$Q^{(in)}$ (m³/d)	18,446	0–100,000	2500.0
$COD^{(in)}$ (gCOD/m³)	381.19	0–1000	25.000
$TSS^{(in)}$ (gSS/m³)	211.26	0–10,000	250.00

Table 4
Nonlinear losses for different measurements as candidates to primary controlled variables.

Candidate	Nominal optimal value	d_1	d_2	d_3	d_{42}	d_5
$S_O^{(1)}$	0.132	0.0039	0.0762	0.0054	0.1362	1.1058
$S_{NO}^{(1)}$	1.472	0.0049	0.2599	0.0006	0.5340	0.3918
$S_{NH}^{(1)}$	15.39	0.4284	Inf	0.1870	Inf	4.8878
$S_O^{(2)}$	0.174	0.0045	0.6107	0.0008	0.1764	0.1109
$S_{NH}^{(2)}$	12.86	0.1732	35.642	0.0859	Inf	4.3033

E. Selection of the primary controlled variables

The expression (6) with matrix **Q** selected as the identity has been considered to obtain the optimal set of five unconstrained self-optimizing control variables associated with the unconstrained degrees of freedom as a combination of measurements determined by matrix **H**. The expected range of disturbances for the activated sludge process has been taken from Araujo et al. (2013) as representative values for scaling:

$$\mathbf{W}_d = \begin{pmatrix} 22371 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 110 \end{pmatrix}$$

In this work, a prescreening of measurements has been performed, very useful to avoid infeasibilities when controlling the CV variables obtained as a result of the SOC methodology (Larsson et al., 2001). The economic losses have been calculated with Eq. (3) for different weather conditions using the nonlinear mathematical model of the process, considering only individual measurements taken out of the relevant variables presented in Table 3, and the results are shown in Table 4. More precisely, the results of this table have been obtained solving different optimization problems (1) with cost function (7), keeping constant at its steady state value one measurement at a time and applying the different load disturbances (weather events). The loss is obtained calculating the difference between these results and the costs obtained solving (1) without the constraint on the corresponding measurement (i.e. re-optimizing the process).

The prescreening consists of removing the primary CV candidate variables that make the process infeasible for some load disturbances, which in this case study are $S_{NH}^{(1)}$ and $S_{NH}^{(2)}$. Infeasibility arises because the influent to the plant does not provide enough concentration of the corresponding compound to satisfy the nominal reference and keep the optimal operation.

Then, based on this study, several sets of measurements have been considered providing different combination matrices **H**, showing three of them in this article (Table 5). Note that additionally, some inputs and disturbances have also been considered as

values during the rainy weather, d_2 are the average load values only for a rain event (extracted from the rain BSM1 disturbances), d_3 are the average load during the whole period for storms, d_{42} are the average values during the second storm (of BSM1 storm influent), d_5 are the average values for one year with average temperature.

C. Steady state optimal operation

The nominal operating point for the process can be seen in Table 3, for the most relevant variables. The measurement range and noise only is specified for the variables susceptible to be part of the measurements sets determined later. The optimization of the process has also been performed for different disturbances, always showing the same three active constraints. The results have been obtained with the *interior-point* method using *fmincon* function of MATLAB®. Multiple starting points for the optimization have been considered due to the local nature of the algorithm, in order to get an optimum close to the global solution.

D. Active constraints control

The optimization shows that three constraints are active: total effluent suspended solids, TSS_e (upper limit), effluent ammonium S_{NHe} (upper limit), and internal recycling flow Q_a (lower limit). Two of them are output active constraints, so they will be associated to two degrees of freedom, remaining 5 available degrees of freedom ($n_u = 5$). The other is an input constraint, $Q_a = 0$. The fact that $Q_a = 0$ can be explained by the fact that the recycling sludge from the secondary settler returns quite sufficient organic matter and nitrate for denitrification. For this process, the set of active constraints does not change with typical disturbances, so the control structure obtained with the self-optimizing procedure is fixed. The three active constraints must be controlled to ensure optimal operation.

Table 5
Measurements sets.

Set 1 (**H₁**)	$S_O^{(1)}, S_{NO}^{(1)}, S_O^{(5)}, S_{NO}^{(5)}, Q^{(in)}, COD^{(in)}, TSS^{(in)}, K_La^{(5)}, Q_r$
Set 2 (**H₂**)	$S_O^{(1)}, ..., S_O^{(5)}, S_{NO}^{(1)}, ..., S_{NO}^{(5)}, Q^{(in)}, COD^{(in)}, TSS^{(in)}, MLSS, K_La^{(5)}, Q_r$
Set 3 (**H₃**)	$S_O^{(1)}, ..., S_O^{(5)}, S_{NO}^{(1)}, ..., S_{NO}^{(5)}, Q^{(in)}, COD^{(in)}, TSS^{(in)}, K_La^{(5)}, Q_r$

Table 6
Nonlinear losses for different combination matrices and disturbances.

	d_1	d_2	d_3	d_{42}	d_{43}	d_5
H₁	Inf	Inf	Inf	Inf	Inf	Inf
H₂	0.223	Inf	0.127	Inf	Inf	1.229
H₃	0.038	0.627	0.069	1.182	0.821	0.300

Table 7
Sensitivity of measurements in H.

	$S_{-O}(1)$	$S_{-O}(2)$	$S_{-O}(3)$	$S_{-O}(4)$	$S_{-O}(5)$
$\|\|\mathbf{H}_3(:, i)\|\|$	2.0437	1.4526	1.5448	1.1585	0.8972
	$S_{-NO}(1)$	$S_{-NO}(2)$	$S_{-NO}(3)$	$S_{-NO}(4)$	$S_{-NO}(5)$
$\|\|\mathbf{H}_3(:, i)\|\|$	10.8114	8.4629	10.4801	6.5360	3.9557
	$Q^{(in)}$	$COD^{(in)}$	$TSS^{(in)}$	$K_La^{(5)}$	Q_r
$\|\|\mathbf{H}_3(:, i)\|\|$	1.2374	0.8393	1.7221	2.0495	4.7614

measurements. The range and noise level (included in matrix $\mathbf{W}_n$) for each measurement are provided in Table 3. In order to select the most suitable, the general principle stating that increasing the number of measurements improves the control has been followed. The nonlinear losses have been obtained for each set (see Table 6). The combination of measurements provided by $\mathbf{H}_1$ is rejected because it gives no feasible solutions when trying to maintain the nominal references with the least demanding disturbance $\mathbf{d}_1$. In the case of $\mathbf{H}_2$, where more measurements have been added, there are still some cases providing infeasibility for the largest disturbances. Finally, for matrix $\mathbf{H}_3$, where MLSS has been removed as a measurement, there are no infeasibilities, so this is the selected combination and the corresponding matrix $\mathbf{H}_3$ is:

$$\mathbf{H}_3^T = \begin{pmatrix} -2.05e-4 & 0.0932 & 0.0678 & 0.0877 & 0.0807 \\ -1.94e-4 & 0.0466 & 0.0807 & 0.0548 & 0.0468 \\ -1.97e-4 & 0.0479 & 0.0487 & 0.0927 & 0.0495 \\ -1.70e-4 & 0.0345 & 0.0351 & 0.0413 & 0.0685 \\ -1.68e-4 & 0.0348 & 0.0354 & 0.0411 & 0.0337 \\ -5.92e-5 & 0.2882 & 0.1530 & 0.2078 & 0.2064 \\ 7.90e-5 & 0.2092 & 0.2304 & 0.0975 & 0.1071 \\ -5.70e-5 & 0.2002 & 0.2087 & 0.3017 & 0.0777 \\ -7.57e-5 & 0.1012 & 0.0993 & 0.1481 & 0.1680 \\ 4.76e-5 & -0.0719 & -0.0824 & -0.0971 & -0.0659 \\ -3.23e-8 & 5.448e-6 & 5.627e-6 & 5.451e-6 & 3.111e-6 \\ -1.02e-6 & 4.280e-4 & 2.604e-4 & 2.777e-4 & 3.679e-4 \\ -1.97e-7 & 6.133e-5 & 6.402e-5 & 8.102e-5 & 7.133e-5 \\ -8.54e-6 & -1.045e-3 & -9.714e-4 & -2.487e-3 & -3.620e-3 \\ 7.62e-8 & 1.888e-5 & 1.943e-5 & 2.188e-5 & 1.671e-5 \end{pmatrix}$$

For numerical calculation of the sensitivity matrices $\mathbf{F}$, disturbance variations of $\pm 1\%$ have been considered, and the linearized model of the process has been obtained using SIMULINK®.

In order to analyze the influence of each measurement in the self-optimized variables, the 2-norm of the corresponding column of $\mathbf{H}_3$ is presented in Table 7, where the magnitude of the elements of $\mathbf{H}_3$ has been normalized such that $\|\|\mathbf{H}_3\|\|_F = 1$ and the matrix has been scaled with respect to the measurements such that the new scaled matrix $\bar{\mathbf{H}}_3 = \mathbf{H} \cdot \mathbf{D}_y$ where $\mathbf{D}_y = diag(span(y))$ and $span(y)$ is the measurement range in Table 3. From the economic perspective, the nitrate and nitrite concentrations in all reactors are the most important measurements in the linear combinations for self-optimizing, and the less relevant are the oxygen concentrations in the reactors.

6. Process controllability analysis

Once the economically optimal controlled variables have been selected, the closed loop control dynamic behavior has been evaluated. The methodology followed for obtaining the linear combinations of measurements is based on steady state models, and they can have a complex dynamic behavior, making this study an essential step for a successful implementation. In this work, some advanced control structures have been considered to keep the selected the optimal CV at optimal set-points despite influent disturbances, based on multivariable nonlinear MPC (NMPC) and PI control.

A. Nonlinear model predictive control formulation

The first control structure considered is a centralized multivariable NMPC for controlling the active constraints and the self-optimized variables, with the following objective function:

$$\min_{\Delta\mathbf{u}(k+i|k) \ 0<i<H_c} V(k) = \sum_{i=H_w}^{H_p} \left\|\mathbf{y}(k+i|k) - \mathbf{r}(k+i)\right\|_{\mathbf{Q}}^2$$
$$+ \sum_{i=0}^{H_c-1} \left\|\Delta\mathbf{u}(k+i|k)\right\|_{\mathbf{R}}^2 + \left\|\mathbf{y}(k+H_p|k) - \mathbf{r}(k+H_p)\right\|_{\mathbf{P}}^2 \quad (12)$$

s.t.

$$\mathbf{u}_{min} < \mathbf{u}(k+i|k) < \mathbf{u}_{max} \quad i = 0, \ldots, H_c - 1$$
$$\mathbf{y}_{min} < \mathbf{y}(k+i|k) < \mathbf{y}_{max} \quad i = H_w, \ldots, H_p$$
$$\Delta\mathbf{u}_{min} < \Delta\mathbf{u}(k+i|k) < \Delta\mathbf{u}_{max} \quad i = 0, \ldots, H_c - 1$$

where y is the vector of controlled outputs, $\mathbf{u}$ the vector of manipulated variables, $\mathbf{r}$ is the set-points vector, $\mathbf{P}$ is the terminal penalty weight matrix, $\mathbf{Q}$ is the output weighting matrix, $\mathbf{R}$ is the move suppression weighting matrix, all of them positive definite, H_c is the control horizon, H_w and H_p are the initial and final prediction horizons respectively, $\mathbf{u}_{min}$ and $\mathbf{u}_{max}$ are the lower and upper bounds for the manipulated variables, $\mathbf{y}_{min}$ and $\mathbf{y}_{max}$ are the lower and upper bounds for the output variables, $\Delta\mathbf{u}_{min}$ and $\Delta\mathbf{u}_{max}$ are the lower and upper bounds for the control moves. The index k denotes the current sampling point, $\mathbf{y}(k+i|k)$ is the predicted output at time $k+i$, depending of measurements up to time k, and $\Delta\mathbf{u}(k+i|k)$ are the future control moves at time $k+i$, depending of measurements up to time k. The use of a terminal penalty weight matrix $\mathbf{P}$ ensures closed loop stability (Mayne et al., 2000).

The predictions are obtained using the following nonlinear discrete time prediction model of the process along the prediction horizon:

$$\mathbf{x}(k+1|k) = \mathbf{f}(\mathbf{x}(k|k), \mathbf{u}(k|k), \mathbf{d}(k|k))$$
$$\mathbf{y}(k|k) = \mathbf{g}(\mathbf{x}(k|k), \mathbf{u}(k|k), \mathbf{d}(k|k)) \quad (13)$$

where $\mathbf{x}$ is the vector of measured or estimated states, $\mathbf{u}$ is the vector of manipulated variables and $\mathbf{d}$ is the vector of measured disturbances (including feedforward action), $\mathbf{f}$ and $\mathbf{g}$ are vector functions that represent the mathematical model of the process (nonlinear differential equations).

B. Offset free nonlinear model predictive control formulation

In order to cope with model plant mismatch and unmeasured disturbances, another formulation for the NMPC has been considered, stating the following optimization problem:

$$\min_{\mathbf{x}_s, \mathbf{u}_s, \mathbf{u}(k+i|k) \ 0<i<H_{p-1}} V(k)$$
$$= \sum_{i=0}^{H_p-1} \left(\left\|\mathbf{x}(k+i|k) - \mathbf{x}_s\right\|_{\mathbf{Q}}^2 + \left\|\mathbf{u}(k+i|k) - \mathbf{u}_s\right\|_{\mathbf{S}}^2 \right)$$
$$+ \left\|\mathbf{x}(k+H_p|k) - \mathbf{x}_s\right\|_{\mathbf{P}}^2 \quad (14)$$

s.t.

$$\mathbf{x}_s = \mathbf{f}(\mathbf{x}_s, \mathbf{u}_s, \mathbf{v}_0)$$
$$\mathbf{r}(k) = \mathbf{g}(\mathbf{x}_s, \mathbf{u}_s, \mathbf{v}_0)$$
$$\mathbf{x}(t+1|k) = \mathbf{f}(x(t|k), \mathbf{u}(t|k), \mathbf{d}(k), \mathbf{v}_0) \quad t = 1, \ldots, H_p - 1$$
$$\mathbf{v}_0 = \hat{\mathbf{v}}(k)$$
$$\mathbf{x}_0 = \hat{\mathbf{x}}(k)$$

$$\mathbf{u}_{min} < \mathbf{u}(k+i|k) < \mathbf{u}_{max} \quad i = 0, \ldots, H_p - 1$$
$$\mathbf{x}_{min} < \mathbf{y}(k+i|k) < \mathbf{x}_{max} \quad i = 0, \ldots, H_p - 1$$

where $(\mathbf{x_s}, \mathbf{u_s})$ are the state and manipulated variables vectors that determine the target equilibrium, $\mathbf{v}$ is the vector of an additional unknown disturbance that captures the model plant mismatch and other unmeasured disturbances, $\mathbf{x}_{min}$ and $\mathbf{x}_{max}$ are the lower and upper bounds for the state variables, $\mathbf{x}(t+1|k)$ are the predicted states along the prediction horizon, $\mathbf{Q}$ is in this case the states weighting matrix, $\mathbf{S}$ is the manipulated variables weighting matrix.

In this formulation, the NMPC penalizes deviations of the state and inputs from a new reference, called target equilibrium, obtained within the iterative receding horizon NMPC optimization problem. As can be seen from the constraints above, this target is a steady state obtained including the effect of an unmeasured disturbance $\mathbf{v}$. In turn, the prediction model is augmented with the mentioned unmeasured disturbance model. Note that the unmeasured disturbance $\mathbf{v}$ is considered to affect the process through the plant dynamics, so $\mathbf{v}$ is a vector with n_u elements, where n_u is the number of manipulated variables. Moreover, as in the SOC the controlled variables may include a direct feed-through part, the unmeasured disturbances are also considered in the output equations:

$$\mathbf{x}(k+1|k) = \mathbf{f}(x(k|k-1), \mathbf{u}(k), \mathbf{d}(k), \mathbf{v}(k|k-1))$$
$$\mathbf{v}(k+1|k) = \mathbf{v}(k|k-1) \tag{15}$$
$$\mathbf{y}(k+1|k) = \mathbf{g}(\mathbf{x}(k+1|k-1), \mathbf{u}(k+1), \mathbf{d}(k+1), \mathbf{v}(k+1|k))$$

This disturbance is estimated from the measured process variables using an estimator. This estimator provides the initial point for the predictions of the next receding horizon step of the NMPC.

$$\hat{\mathbf{x}}(k+1|k) = \mathbf{f}(\hat{\mathbf{x}}(k|k-1), \mathbf{u}(k), \mathbf{d}(k)\hat{\mathbf{v}}(k|k-1)) + \mathbf{L}_x(\mathbf{y}_p(k) - \mathbf{g}\left(\hat{\mathbf{x}}(k|k-1), \mathbf{u}(k), \mathbf{d}(k), \mathbf{v}(k|k-1)\right))$$

$$\hat{\mathbf{v}}(k+1|k) = \hat{\mathbf{v}}(k|k-1) + \mathbf{L}_d(\mathbf{y}_p(k) - \mathbf{g}\left(\hat{\mathbf{x}}(k|k-1), \mathbf{u}(k), \mathbf{d}(k), \mathbf{v}(k|k-1)\right))$$

$$\tag{16}$$

where $\mathbf{y}_p$ is the real plant measurement vector and $\mathbf{L}_x$, $\mathbf{L}_d$ are the estimator gain matrices.

The implementation of this observer in this article is a steady state Extended Kalman Filter, where the gain matrices $\mathbf{L}_x$ and $\mathbf{L}_d$ are obtained using a linearization of the process (13) at the working steady state. For more details see Pannocchia and Rawlings (2003) and Morari and Maeder (2012).

C. Model predictive control applied to the BSM1

For the application of the described NMPC formulations to the SOC of the BSM1, the measured disturbances are $\mathbf{d} = (Q^{(in)}, COD^{(in)}, TSS^{(in)})$, the manipulated variables are $\mathbf{u} = (Q_r, K_La^{(1-5)}, Q_w)$ and $\mathbf{x}$ is a vector comprising 145 states corresponding to the full BSM1 description of the plant (5 reactors and secondary settler) (Alex et al., 2008). The internal prediction model of the NMPC controller is the full BSM1 mathematical model, which is simulated over a prediction horizon integrating the nonlinear differential equations of the model, and choosing the coincidence points with a sampling period of $T_s = 0.052$ days. Model plant mismatch is also possible by modifying some parameters of the BSM1 model. The NMPC constraints are:

$$0 \leq K_La^{(1-5)} \leq 360\,\text{d}^{-1}$$
$$Q_w \leq 1844.6\,\text{m}^3/\text{d}$$
$$Q_a \leq 92230\,\text{m}^3/\text{d} \tag{17}$$
$$Q_r \leq 36892\,\text{m}^3/\text{d}$$

D. Distributed NMPC-PI control structure

The use of the centralized NMPC is a straightforward solution to control the process because the interactions are automatically tackled by the controller when it performs predictions. Nevertheless,

the number of tuning parameters is rather high, making the tuning a difficult task. This motivates the search for more simple control structures such as distributed NMPC-PI controllers. The control of active constraints is crucial for the optimal operation of the plant, and this control structure has the advantage that if the NMPC fails, the PI controllers still keep the active constraints in the desired values. On the other hand, the use of decoupled PI controllers can be a complex task due to the interactions between variables and it has been not considered in this work.

Regarding the practical choice of manipulated variables, note that typically the mass transfer coefficients of oxygen in each reactor $K_La^{(1-5)}$ are considered as manipulated variables, to avoid including the details of the aeration systems and reactors geometry. However, in this work the volumetric air flow rates are also considered as more realistic variables for dynamic assessment, considering the following correlation equations (Flores-Tlacuahuac et al., 2009).

$$K_La^{(j)} = 240(1 - e^{-1.2 \cdot Q_{air}^{(j)}})\quad j = 1, \ldots, 4 \tag{18}$$

$$K_La^{(5)} = 240(1 - e^{-0.55 \cdot Q_{air}^{(5)}}) \tag{19}$$

Then, the corresponding NMPC constraints are replaced with $0 \leq Q_{air_{1-5}} \leq 10\,\text{m}^3/\text{d}$ if Q_{air} flows are considered as manipulated variables.

In order to select the most appropriate loops for PI control in the decentralized NMPC-PI control structure, the Relative Gain Array (RGA) matrix has been calculated with an open-loop linearized model of the plant around the nominal optimum operating point, with $u = \begin{bmatrix} Q_r & Q_{air}^{(1)} & Q_{air}^{(2)} & Q_{air}^{(3)} & Q_{air}^{(4)} & Q_{air}^{(5)} & Q_w \end{bmatrix}$ and $y = \begin{bmatrix} TSS_e & S_{NHe} & c_1 & c_2 & c_3 & c_4 & c_5 \end{bmatrix}$, where the variables c_i are the components of the vector $\mathbf{c}$, as combinations of measurements.

$RGA(\omega = 0)$

$$= \begin{pmatrix} -0.0001 & 0.0001 & -0.000 & 0.000 & -0.000 & -0.000 & 1.000 \\ -0.012 & -1.11 & 0.249 & 0.245 & 0.719 & 0.907 & 0.000 \\ 1.041 & -0.012 & 0.027 & -0.063 & -0.002 & 0.009 & -0.0001 \\ -0.051 & 10.798 & -8.888 & -0.403 & -2.928 & 2.472 & 0.0004 \\ 0.036 & -7.678 & 12.691 & -4.038 & 0.796 & -0.809 & -0.0004 \\ -0.055 & -0.634 & -3.799 & 7.634 & -1.888 & -0.259 & 0.0003 \\ 0.040 & -0.367 & 0.719 & -2.376 & 4.303 & -1.320 & -0.0002 \end{pmatrix}$$

The frequency dependent RGA is also important for the consideration of the process dynamics in the variables pairing. Then, the RGA for the typical frequency of the disturbances ($\omega_c \approx 2\pi$ rad/d) is:

$\left| RGA(\omega = 2\pi) \right|$

$$= \begin{pmatrix} 0.0060 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.9943 \\ 0.1504 & 0.3825 & 0.1135 & 0.1268 & 0.3681 & 0.4736 & 0.0048 \\ 1.1166 & 0.1373 & 0.0148 & 0.0929 & 0.0382 & 0.1528 & 0.0097 \\ 0.7405 & 3.5592 & 2.9638 & 0.0410 & 0.7868 & 1.5123 & 0.0441 \\ 0.7447 & 2.3291 & 4.8387 & 1.3538 & 0.1015 & 0.7338 & 0.0474 \\ 0.3854 & 0.7773 & 1.3380 & 2.8094 & 0.4815 & 0.1372 & 0.0292 \\ 0.3908 & 0.4006 & 0.1552 & 0.8248 & 1.1759 & 1.2454 & 0.0264 \end{pmatrix}$$

At the view of both RGA matrices, the following pairings have been selected: TSS_e (controlled variable)–Q_w (manipulated variable), and $S_{NH,e}$ (controlled variable)–Q_{air_5} (manipulated variable), which is coherent with the process functioning by looking at the process layout. Although this is the best possible pairing, note that

Table 8
Control structures.

Control structure		Controlled variables	Manipulated variables
Centralized NMPC		$\mathbf{c} = \mathbf{H}_3 y$, S_{NHe}, TSS_e	$Q_{air_1} - Q_{air_5}$, Q_r, Q_w
Distributed NMPC-PI	NMPC	$\mathbf{c} = \mathbf{H}_3 y$	$Q_{air_1} - Q_{air_4}$, Q_r
control	PI No. 1	TSS_e,	Q_w
	PI No. 2	S_{NHe}	Q_{air_5}

some coupling is still detected when choosing Q_{air_5} as manipulated variable.

Other possible pairings could arise from the RGA study, but they have not been considered because of the poor dynamic results. Then, for the rest of variables, a multivariable NMPC has been implemented, where the internal model of this NMPC includes the PI controllers in the predictions. This fact is important in order to produce reliable predictions for the NMPC, assuming that the PI controllers are properly tuned. Finally, it can be observed from the magnitude of some of the elements of the RGA that the plant is ill-conditioned, and therefore difficult to control.

As for PI controllers tuning, the guidelines of Skogestad (2003) have been followed, with some practical considerations due to the complexity of the process. The tuning parameters are:

$$K_p = \frac{1}{K} \frac{\tau_1}{\tau_c + \theta} \tag{20}$$

$$T_i = \min(\tau_1, 4(\tau_c + \theta)) \tag{21}$$

where K_p is the proportional gain of the controller, T_i is the integral time of the controller, τ_1 is the time constant, K is the gain, and θ is the time delay of the first order system identified from a Q_w step response of $TSSe$, and τ_c is the desired closed loop time constant, which is in fact a tuning parameter.

7. Dynamic analysis

In this section, the closed loop performance of the process is evaluated in the presence of step disturbances corresponding to the different weather events of Table 2 and some of the dynamic sets of

disturbances provided in Alex et al. (2008) for the BSM1. The control structures considered are the centralized NMPC and the distributed NMPC-PI described above (see Table 8 for a summary). In order to simplify the control structures, particularly to implement real independent PI controllers in the distributed NMPC-PI control structure, the measurements Q_{air_5} and Q_r have been removed from the linear combinations of measurements that conform the self-optimized variables, thus obtaining a new combination matrix $\mathbf{H}_4$ analogous to $\mathbf{H}_3$. The inclusion of Q_{air_5} as measurement would imply a permanent communication of the PI controller with the NMPC that in this way has been removed. On the other hand, although the volumetric air flows Q_{air} have been considered here as more realistic manipulated variables, dynamic responses considering $K_L a$ factors as manipulated variables are similar if the controllers are properly tuned, because the additional nonlinearity of (18) and (19) is embedded in the NMPC predictions. In order to perform also more realistic simulations, disturbance $TN^{(in)}$ has been added as an additional unmeasured disturbance, although it has not been considered when obtaining the SOC combination matrices.

The references for the five self-optimized variables obtained from the economical optimization of the plant are the following: $c_1 = -0.001$; $c_2 = 2.099$; $c_3 = 1.864$; $c_4 = 1.982$; $c_5 = 1.356$. As for the self-optimized variable c_1, it has not been considered for control because its significance in costs is negligible, at the view of the corresponding row of the measurements combination matrix. In Figs. 2–4 the dynamic responses for centralized NMPC control of the self-optimized variables and active constraints are presented, together with control actions. The figures show that for all
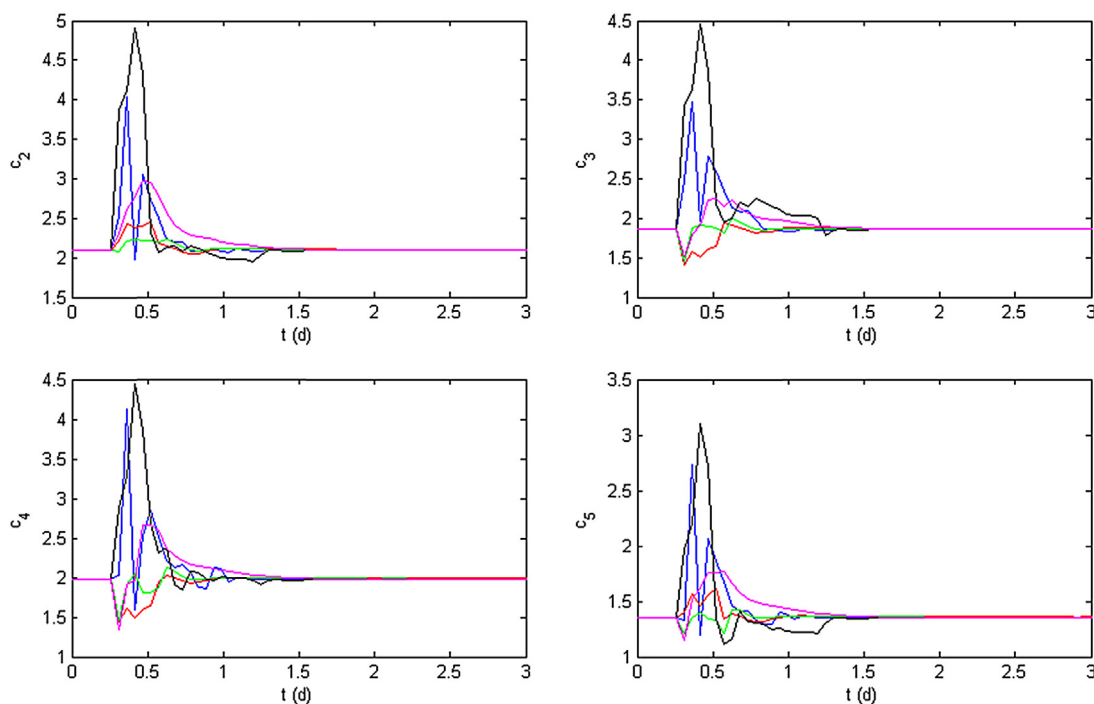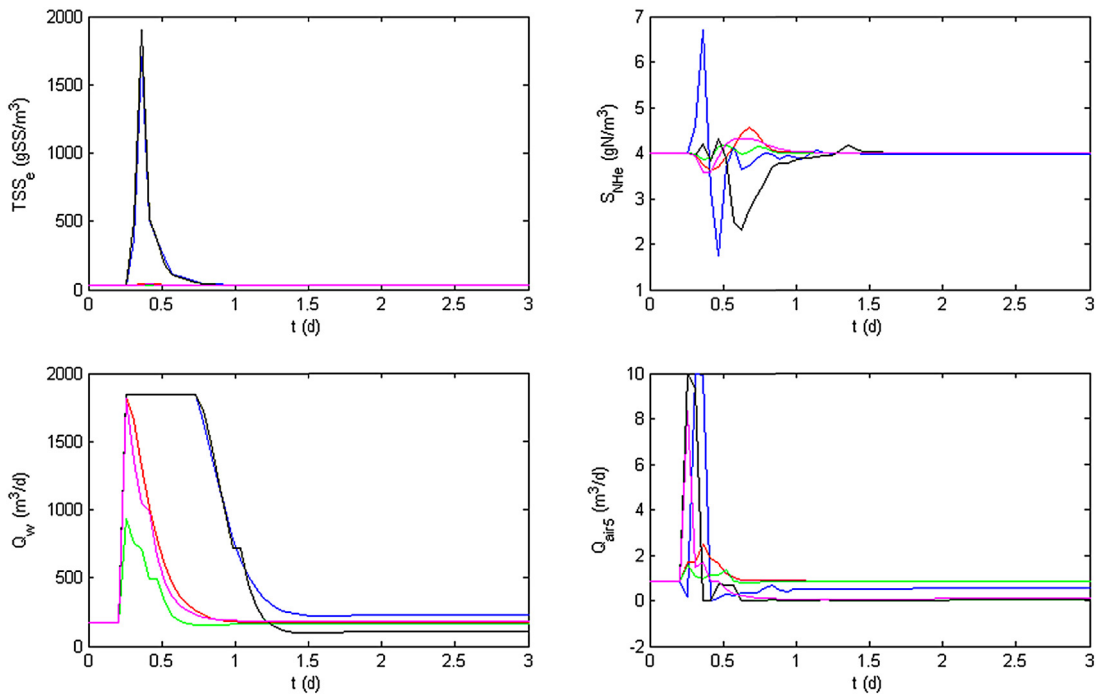

Fig. 2. Self-optimized variables using the centralized NMPC control structure, for disturbances $\mathbf{d}_1$ (red), $\mathbf{d}_2$ (blue), $\mathbf{d}_3$ (green), $\mathbf{d}_{42}$ (black), $\mathbf{d}_5$ (magenta). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
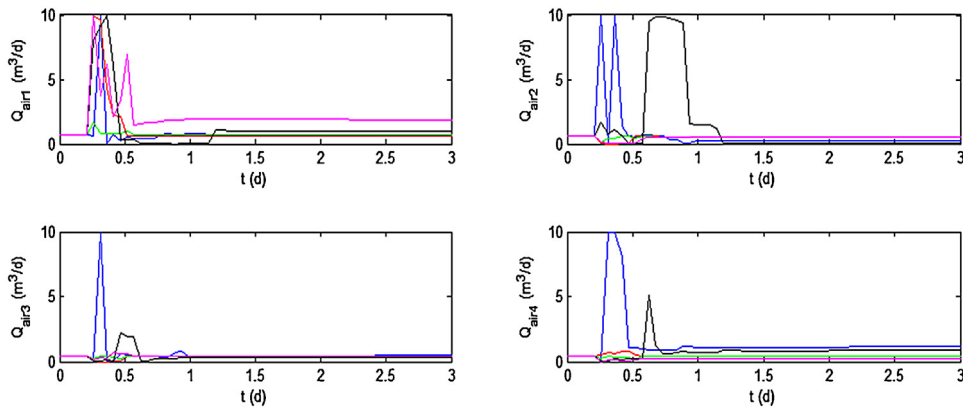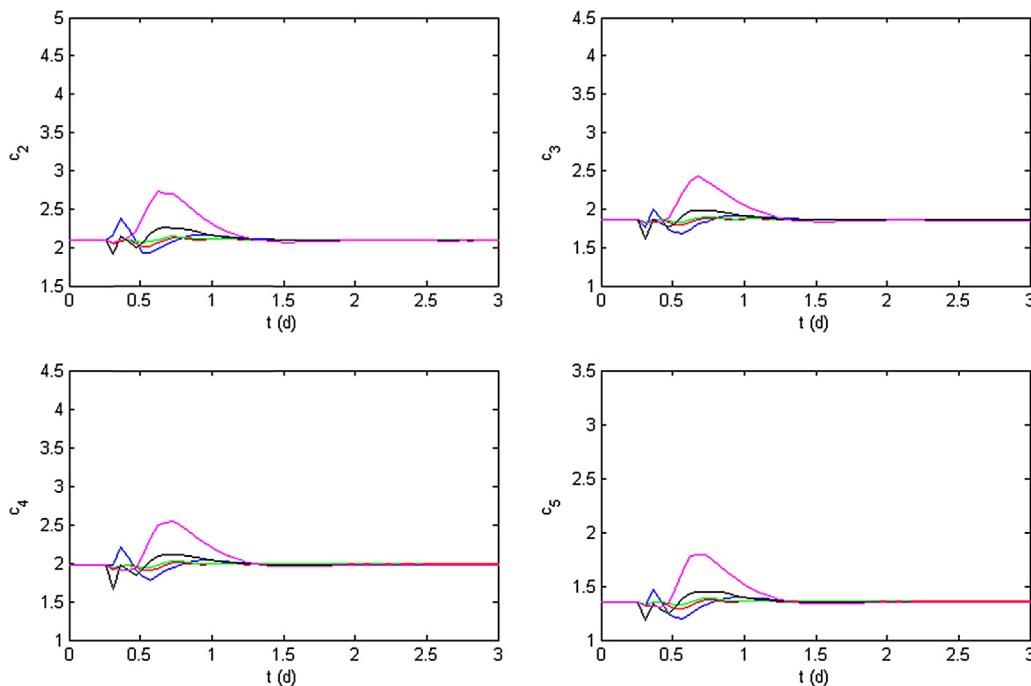
Fig. 3. Manipulated variables for the centralized NMPC control structure, for disturbances $\mathbf{d}_1$ (red), $\mathbf{d}_2$ (blue), $\mathbf{d}_3$ (green), $\mathbf{d}_{42}$ (black), $\mathbf{d}_5$ (magenta). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 4. Active constraints control for the centralized NMPC control structure for disturbances $\mathbf{d}_1$ (red), $\mathbf{d}_2$ (blue), $\mathbf{d}_3$ (green), $\mathbf{d}_{42}$ (black), $\mathbf{d}_5$ (magenta). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

step disturbances applied, the set-point tracking is good, with a settling time of approximately 1.5 days. The overshoot is larger for the rain and storm events, because they are the most demanding disturbances. For these weather events, the $TSS^{(in)}$ concentration decreases and the influent flow $Q^{(in)}$ increases considerably, moving the process operating point far from nominal conditions. The complex nonlinearities involved in the process also deteriorates the performance, and the presence of right-half plane zeros may limit the bandwidth of the controller.

The controller tuning has been performed considering the following guidelines. The weights on the outputs have the ability of direct control efforts toward the tracking of a particular output. In this case, the interest is in the self-optimized variables in order to minimize economic losses, so $\mathbf{Q} = \text{diag}(1\ 1\ 2\ 2\ 2\ 2\ 2)$. For the weights on the rate of change of inputs, the tuning has been performed by trial and error, starting from a small value until closed loop stability is achieved. A small further detuning has been performed, in order to give some robustness to the controller and then $\mathbf{R} = \text{diag}(0.1\ 0.014\ 0.014\ 0.014\ 0.014\ 0.014\ 0.005)$. The NMPC horizons are $H_w = 1$, $H_p = 5$ and $H_c = 1$. No analytical

tuning has been performed because most of the techniques available are only for linear prediction models and mainly first order prediction models.

In Figs. 5–7, the dynamic responses for the distributed NMPC-PI control of the self-optimized variables and active constraints are presented, together with some of the control actions. The figures also show good set-point tracking for all disturbances, with approximately the same settling time than in the centralized NMPC responses. The main difference is in the overshoot, which is considerably smaller than in the case of the centralized NMPC, for all controlled variables excepting TSS_e, and all weather events, particularly for the most demanding disturbances. The PI control of the active constraints in a fast time scale and its consideration in the NMPC predictions, allows for a better transient of the self-optimized variables, which are in a slower time scale. Another reason that explains this behavior is that the tuning of a centralized NMPC is complex due to the large number of parameters. For the distributed control structure, the NMPC is easier to tune, and the PI controllers can be tuned using the SIMC rules or other standard techniques.

Fig. 5. Self-optimized variables using the distributed NMPC-PI control structure, for disturbances d_1 (red), d_2 (blue), d_3 (green), d_{42} (black), d_5 (magenta). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
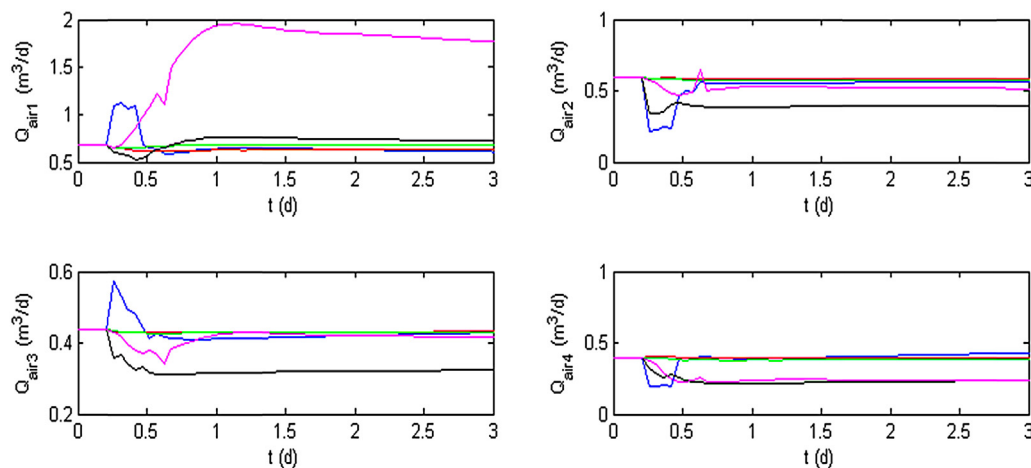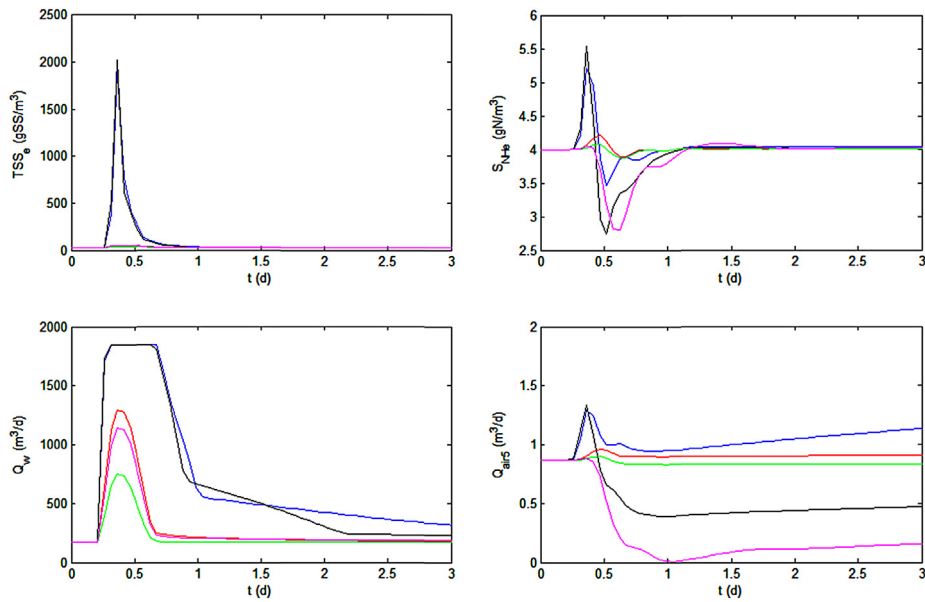

Fig. 6. Manipulated variables for the distributed NMPC-PI control structure, for disturbances d_1 (red), d_2 (blue), d_3 (green), d_{42} (black), d_5 (magenta). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The tuning parameters for the distributed NMPC-PI control are $Q = \text{diag}(2\ \ 2\ \ 2\ \ 2\ \ 2)$, $R = \text{diag}(0.5\ \ 1\ \ 1\ \ 1\ \ 1)$ and the same MPC horizons than for the centralized NMPC. The tuning parameters for the PI control No. 1 are $K_p = -54.8$, $T_i = 27.4$ and $K_p = -0.2$, $T_i = 0.08$ for No. 2, and they have been selected by SIMC rules with the following considerations. The τ_c parameter for PI No. 1 has been selected equal to the system effective time delay (Skogestad, 2003), but the integral time has been increased considerably due to the natural integrating properties of this loop. For PI No. 2, the τ_c has been selected to approximately four times the system effective time delay, with a final fine tuning, in order to have a less oscillating response.

In order to complete the dynamic analysis, simulations for different sets of time varying disturbances (5 days) taken from BSM1 have been performed. In Figs. 8–10 a performance comparison of the centralized NMPC and the distributed NMPC-PI structure for dry weather conditions is presented, together with the open loop

behavior. At the view of the figures, the controller improves the set-point tracking with respect to the open loop, but no further adjustment is possible because of the large varying influent. In Table 9, the performance indices for rain weather dynamic disturbances are presented. By looking at this table, the cost for open loop is much smaller than for any of the control structures, but this is not a desirable behavior because of the large violations of the active constraint for TSS_e. The cost for the case of either distributed NMPC-PI or centralized NMPC control are smaller than the active constraints control, because these control structures also regulate the self-optimized variables. Theoretically, the centralized control should provide also smaller cost than the distributed NMPC-PI control because it takes into account all the interactions in a comprehensive fashion, but for these disturbances, due to the tuning difficulties and the effect of transients the cost is a bit larger.

The dynamic performance evaluation using the full time varying disturbance profiles for different weather conditions take the

Fig. 7. Active constraints control for the distributed NMPC-PI control structure for disturbances $\mathbf{d}_1$ (red), $\mathbf{d}_2$ (blue), $\mathbf{d}_3$ (green), $\mathbf{d}_{42}$ (black), $\mathbf{d}_5$ (magenta). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
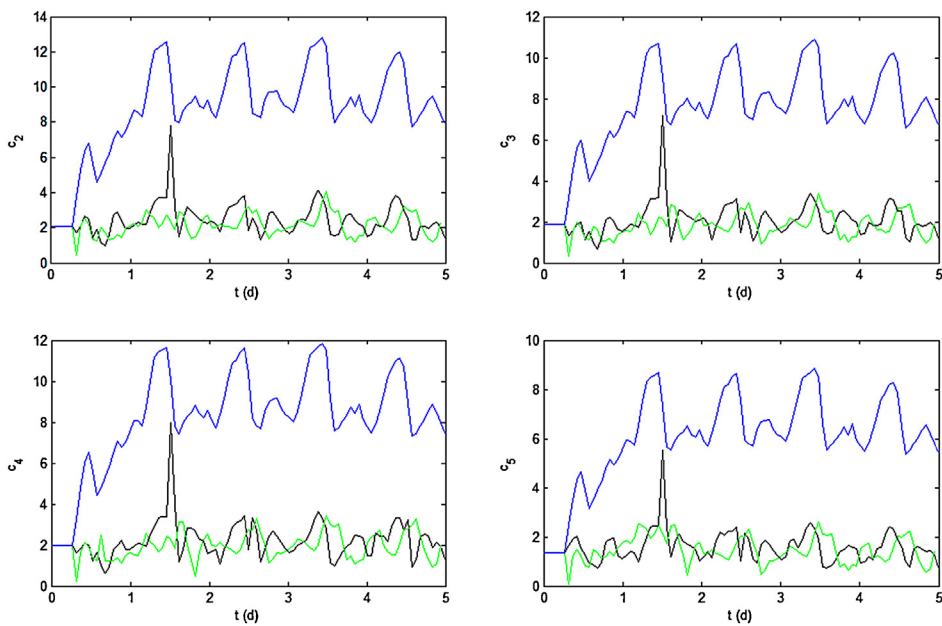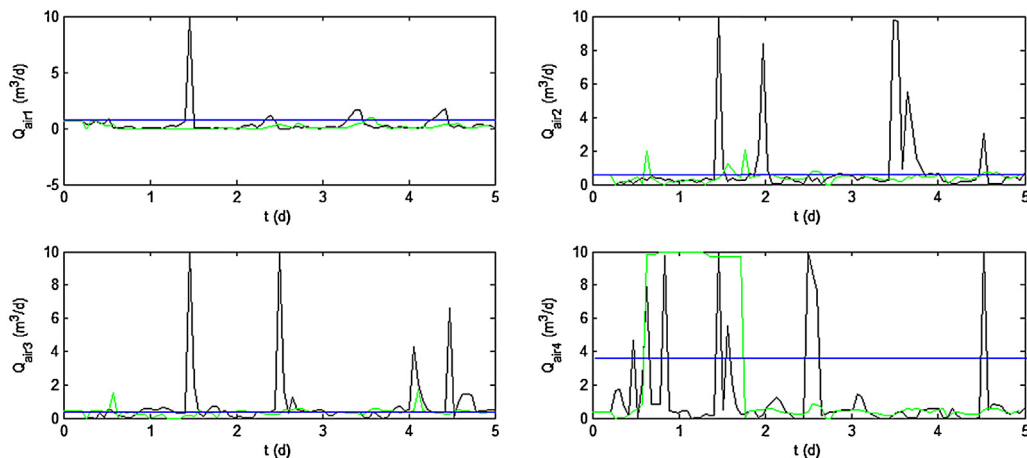

Fig. 8. Self-optimized variables control for the distributed NMPC-PI control structure (green), the centralized NMPC (black) and open loop behavior (blue) for dry weather disturbances. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 9
Performance indices for rain weather dynamic disturbances.

Control structure (rain weather dynamic profile)	Cost	PE	AE	ME	SP
Open loop	400.5293	169.3954	2763.1	0	1.7075e+06
Only active constraint control	575.4935	181.9706	2359.4	123.9801	4.1952e+06
Centralized NMPC	574.0025	293.3383	2368.3+03	147.1967	4.0151e+06
Distributed NMPC-PI	545.4043	180.3966	2320.0+03	80.6949	3.9138e+06

Table 10
Performance indices for rain event step disturbances.

Control structure (step disturbance $\mathbf{d}_3$)	Total cost	PE	AE	ME	SP
Centralized NMPC	616.5386	250.6723	2.8666e+03	4.5353	4.1947e+06
Distributed NMPC-PI	684.5272	186.9036	2.8303e+03	0	5.1622e+06
Only active constraint control	684.7230	186.9023	2.8327e+03	0	5.1620e+06

Fig. 9. Manipulated variables for the distributed NMPC-PI control structure (green), the centralized NMPC (black) and open loop behavior (blue) for dry weather disturbances. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
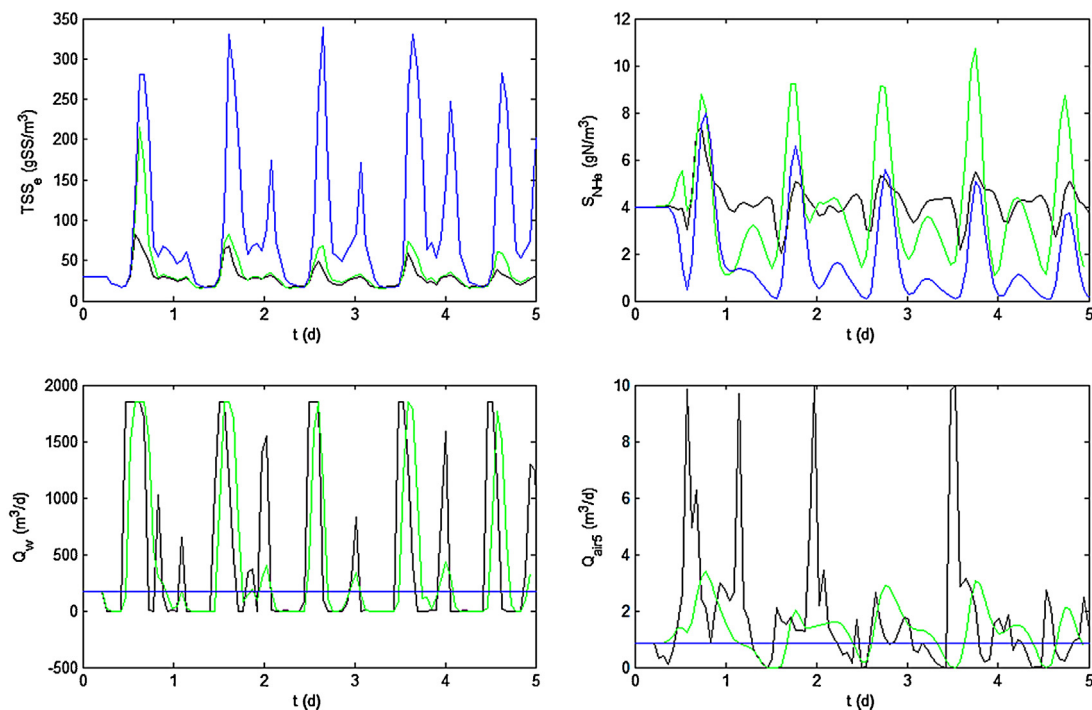
Fig. 10. Active constraints control for the distributed NMPC-PI control structure (green), the centralized NMPC (black) and open loop behavior (blue) for dry weather disturbances. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 11
Performance indices for storm weather step disturbances.

Control structure (step disturbance d_4)	Total cost	PE	AE	ME	SP
Centralized NMPC	417.4922	170.4853	2.7226e+03	0	1.9639e+06
Distributed NMPC-PI	419.2097	170.6695	2.7222e+03	0	1.9856e+06
Only active constraint control	419.2689	170.6694	2.7229e+03	0	1.9856e+06

plant to a nonsteady state operating point, and then the requirements for SOC are not fully satisfied. The slow dynamics of the process and the varying conditions in the influent prevent from reaching a complete steady state, and consequently the SOC results are only extrapolated to this situation as an approximation. Therefore, an analogous comparison has been performed with the rain event step disturbance d_3 (Table 10), and the storm weather step disturbance d_4 (Table 11). These step disturbances represent a load step change from the mean dry weather conditions

to the mean values for the mentioned weather conditions. As expected, both the centralized NMPC and distributed NMPC-PI provide smaller cost, showing the advantage of controlling the self-optimized variables. In these cases, the centralized NMPC cost is smaller than the distributed NMPC-PI cost, because it is an easier operating scenario for the process. Then, it is easier to tune the controller in this latter case than in the case of varying influent, and therefore in Tables 10 and 11 the cost is smaller with the centralized NMPC. The time horizon for evaluating all the cost indices
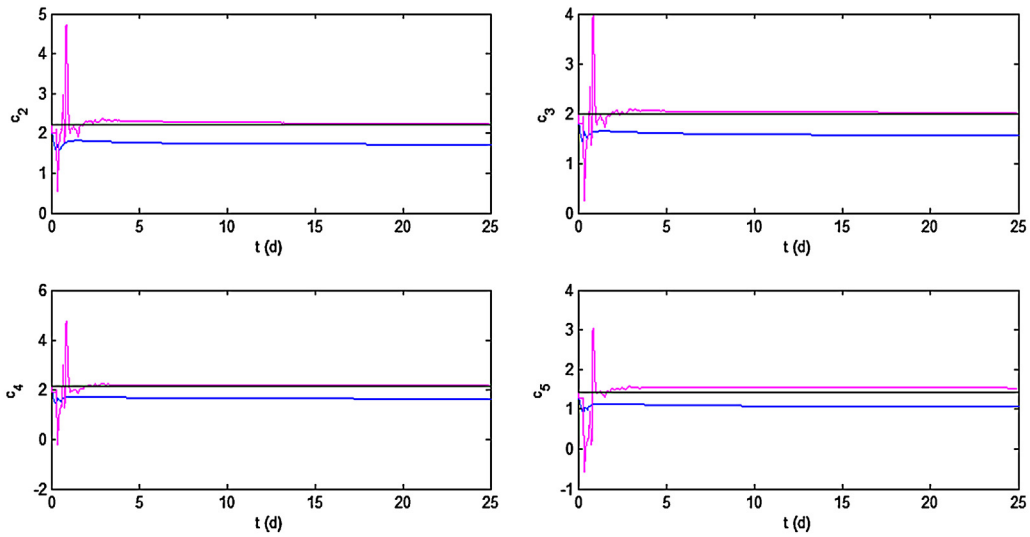
Fig. 11. Self-optimized variables when disturbance $\mathbf{d}_3$ is applied, using centralized NMPC (blue: standard NMPC formulation – magenta: offset free NMPC formulation – black: references). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
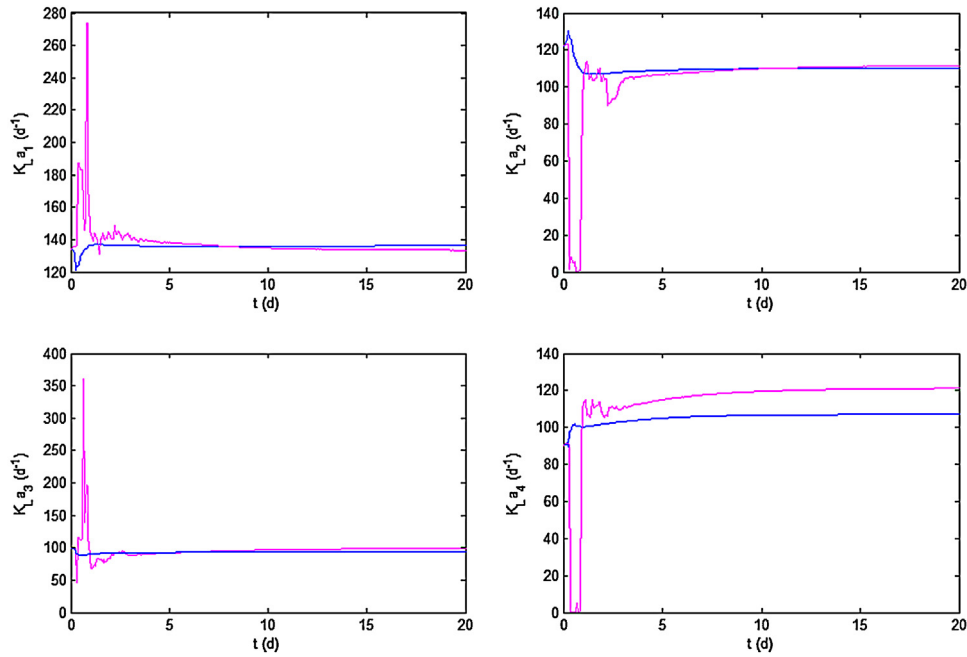
Fig. 12. Manipulated variables when disturbance $\mathbf{d}_3$ is applied, using centralized NMPC (blue: standard NMPC formulation – magenta: offset free NMPC formulation). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
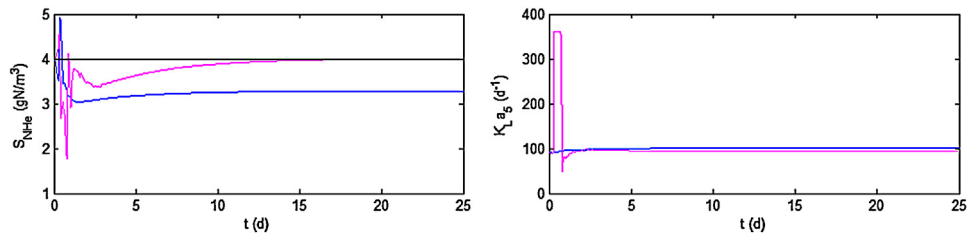
Fig. 13. Control of active constraint S_{NHe} when disturbance $\mathbf{d}_3$ is applied, using centralized NMPC (blue: standard NMPC formulation – magenta: offset free NMPC formulation – black: reference). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

is $\Theta = 5$ days, based on the BSM1 disturbances data sets for one week.

In the previous results, although no plant model mismatch has been considered in order to stress other aspects of the methodology, the formulation (14) can be applied in all NMPC controllers,

giving an offset free response. In other to show the effectiveness of that formulation, a plant model mismatch has been considered varying the "real" process temperature while keeping constant the temperature of the model at $T = 15\,^{\circ}$C. In particular, the real plant has been represented with the BSM1 model with an increase of

3 °C in temperature, affecting to the value of all kinetic parameters because of their temperature dependences. The prediction model is the original BSM1 with $T = 15$ °C. The results of Figs. 11–13 show a comparison of performance of both formulations, for a rain event step disturbance $\mathbf{d}_3$. The use of the new formulation reduces the existing plant model mismatch offset considerably with respect to the NMPC formulation (12). The results are obtained in this case with $\mathbf{H}_3$ matrix and $K_L a$ as manipulated variables.

8. Conclusions

In this work, the SOC methodology has been applied to find the optimum controlled variables as a combination of measurements in a wastewater treatment plant. A prescreening of the most suitable measurements to avoid unfeasibilites when large load disturbances appear has been performed. The dynamic controllability of these variables has also been studied, by implementing a multivariable centralized nonlinear MPC, and a distributed control structure with PI control for the active constraints and a nonlinear MPC for the self-optimizing variables. The results show that both control structures give good set-point tracking, despite of a long transient particularly for the most severe disturbances. The distributed NMPC-PI control shows better performance because of the separate treatment of the different time scales of the process and the easier tuning of the controllers. In particular, the overshoots for the self-optimized variables are considerably smaller than those for the centralized NMPC. Finally, a dynamic evaluation with the full profile of disturbances for different weather conditions has been performed. For this latter case, the process never reaches a steady state operating point, as it is supposed in SOC methodology, and because of that costs only decrease a small amount when controlling the self-optimized variables. In order to overcome this behavior, the application of NCO tracking in an upper optimization layer is proposed as future work. Another possibility is the consideration of transients costs in the SOC cost function, arising a tradeoff between good dynamic performance and economics.

Acknowledgment

References

Alex JL, Benedetti L, Copp J, Gernaey KV, Jeppsson U, Nopens I, Pons MN, Rosen C, Steyer JP, Vanrolleghem P, Winkler S, 2008. Benchmark Simulation Model no.1 (BSM1). Tech. Report no. LUTEDX/(TEIE- 7229)/1-62/(2008).

Alstad V, Skogestad S. Null space method for selecting optimal measurement combinations as controlled variables. Ind Eng Chem Res 2007;46:846–53.

Alstad V, Skogestad S, Hori ES. Optimal measurement combinations as controlled variables. J Process Control 2009;19:138–48.

Alstad V. PhD Thesis Studies on selection of controlled variables PhD Thesis. Norwegian University of Science and Technology; 2005.

Amand L, Carlsson B. Optimal aeration control in a nitrifying activated sludge process. Water Res 2012;46(7):2101–10.

Araujo A, Skogestad S. Control structure design for the ammonia synthesis process. Comput Chem Eng 2008;32:2920–32.

Araujo A, Gallani S, Mulas M, Olsson G. Systematic approach to the design of operation and control policies in activated sludge systems. Ind Eng Chem Res 2011;50:8542–57.

Araujo A, Gallani S, Mulas M, Skogestad S. Sensitivity analysis of optimal operation of an activated sludge process model for economic controlled variable selection. Ind Eng Chem Res 2013;52(29):9908–21.

Baldea M, Araujo M, Skogestad S, Daoutidis D. Dynamic considerations in the synthesis of self-optimizing control structures. AIChE J 2008;54(7):1830–41.

Cadet C, Beteau JF, Carlos Hernandez S. Multicriteria control strategy for cost/quality compromise in wastewater treatment plants. Control Eng Pract 2004;12:335–47.

Cao Y. Self-optimizing control structure selection via differentiation. In: European Control Conference (ECC 2003); 2003. p. 445–53.

Cao Y, Kariwala V. Bidirectional branch and bound for controlled variables selection: Part I. Principles and minimum singular value criterion. Comput Chem Eng 2008;32(10):2306–19.

Darby ML, Nikolau M, Jones J, Nicholson D. RTO: An overview and assessment of current practice. J Process Control 2011;21:874–84.

Flores-Tlacuahuac A, Hernandez M, López-Negrete R. Bifurcation behavior of a large scale waste water treatment plant. Ind Eng Chem Res 2009;48:2605–15.

Francisco M, Vega P, Álvarez H. Robust integrated design of processes with terminal penalty model predictive controllers. Chem Eng Res Des 2011;89:1011–24.

Gera V, Panahi M, Skogestad S, Kaistha N. Economic plantwide control of the cumene process. Ind Eng Chem Res 2013;52:830–46.

Gros S, Srinivasan B, Bonvin D. Optimizing control based on output feedback. Comput Chem Eng 2009;33:191–8.

Guerrero J, Guisasola A, Vilanova R, Baeza J. Improving the performance of a WWTP control system by model-based set-point optimisation. Environ Model Softw 2011;26:492–7.

Halvorsen IJ, Skogestad S, Morud JC, Alstad V. Optimal selection of controlled variables. Ind Eng Chem Res 2003;42:3273–84.

Henze M, Grady CPL Jr, Gujer W, Marais GVR, Matsuo T. Activated Sludge Model No. 1, IAWQ Scientific and Technical Report No. 1, IWAQ, UK; 1987.

Hu W, Mao J, Xiao G, Kariwala V. Selection of Self-optimizing controlled variables for dynamic processes. In: 8th IFAC Symposium on Advanced Control of Chemical Processes; 2012.

Ingildsen P, Olsson G, Yuan Z. A hedging point strategy – balancing effluent quality, economy and robustness in the control of wastewater treatment plants. Water Sci Technol 2002;45(4–5):317–24.

Jacobsen MG, Skogestad S. Active constraint regions for optimal operation of chemical processes. Ind Eng Chem Res 2011;50:11226–36.

Jäschke J, Skogestad S. NCO tracking and self-optimizing control in the context of real-time optimization. J Process Control 2011;21:1407–16.

Kariwala V, Cao Y, Janardhanan S. Local self-optimizing control with average loss minimization. Ind Eng Chem Res 2008;47:1150–8.

Kariwala V. Optimal measurement combination for local self-optimizing control. Ind Eng Chem Res 2007;46(11):3629–34.

Larsson T, Hestetun K, Hovland E, Skogestad S. Self-optimizing control of a large-scale plant: the Tennessee Eastman process. Ind Eng Chem Res 2001;40:4889–901.

Maarleveld A, Rijnsdorp J. Constraint control on distillation columns. Automatica 1970;6:51–8.

Machado VC, Gabriel D, Lafuente J, Baeza JA. Cost and effluent quality controllers design based on the relative gain array for a nutrient removal WWTP. Water Res 2009;43(20):5129–41.

Manum H, Skogestad S. Self-optimizing control with active set changes. J Process Control 2012;22:873–83.

Mayne DQ, Rawlings JB, Rao CV, Sckokaert POM. Constrained model predictive control: Stability and optimality. Automatica 2000;36:789–814.

Morari M, Maeder U. Nonlinear offset-free model predictive control. Automatica 2012;48:2059–67.

Morari M, Stephanopoulos G, Arkun Y. Studies in the synthesis of control structures for chemical processes. Part I: Formulation of the problem. Process decomposition and the classification of the control task. Analysis of the optimizing control structures. Am Inst Chem Eng J 1980;26(2):220–32.

Pannocchia G, Rawlings JB. Disturbance models for offset-free model predictive control. AIChE J 2003;49(2):426–37.

Rivas A, Irizar I, Ayesa E. Model-based optimization of wastewater treatment plants design. Environ Model Softw 2008;23:435–50.

Samuelsson P, Halvarsson B, Carlsson B. Cost-efficient operation of a denitrifying activated sludge process. Water Res 2007;41:2325–32.

Skogestad S. Plantwide control: the search for the self-optimizing control structure. J Process Control 2000;10:487–507.

Skogestad S. Simple analytic rules for model reduction and PID controller tuning. J Process Control 2003;13:291–309.

Srinivasan B, Biegler LT, Bonvin D. Tracking the necessary conditions of optimality with changing set of active constraints using a barrier-penalty function. Comput Chem Eng 2008;32:572–9.

Stare A, Vrecko D, Hvala S, Strmcnik S. Comparison of control strategies for nitrogen removal in an activated sludge process in terms of operating costs: a simulation study. Water Res 2007;41:2004–14.

Takacs I, Patry G, Nolasco D. A dynamic model of the clarification-thickening process. Water Res 1991;25(10):1263–71.

Umar LM, Hu W, Cao Y, Kariwala V. Rangaiah GP, Kariwala V, editors. Selection of controlled variables using self-optimizing control method: recent developments and applications. Chichester, UK: John Wiley & Sons, Ltd; 2012.

Ye L, Cao Y, Li Y, Song Z. Approximating necessary conditions of optimality as controlled variables. Ind Eng Chem Res 2013;52:798–808.

Ye L, Cao Y, Ma X, Song Z. A novel hierarchical control structure with controlled variable adaptation. Ind Eng Chem Res 2014;53:14695–711.

Yelchuru R, Skogestad S. Optimal controlled variable selection with structural constraints using MIQP formulations. In: 18th IFAC World Congress; 2011.