I&EC research
Industrial & Engineering Chemistry Research

Article

pubs.acs.org/IECR

Model Adaptation for Real-Time Optimization in Energy Systems

Fernán J. Serralunga, Miguel C. Mussati, and Pio A. Aguirre*

INGAR Instituto de Desarrollo y Diseño (CONICET - UTN), Avellaneda 3657 (S3002GJC) Santa Fe, Argentina

Supporting Information

ABSTRACT: Real-time optimization (RTO) is widely used in industry to operate processes close to their maximum performance. The models used for RTO need to be adapted using real-time data to ensure feasibility of the model optimal inputs and convergence to the real plant optimal point. Heat and power systems are suitable for being optimized in real-time because of their fast dynamics and the benefits achievable by reacting to changes in power prices and steam demand. This work proposes a modifier adaptation strategy that exploits the structure of certain problems to make the adaptation faster and more reliable, which is proven to be particularly useful for heat and power systems. The adaptation is performed in the equations that predict efficiencies or performance of unit operations. By identifying the variables that modify each performance factor, the number of data sets needed for gradient correction is reduced. This makes the proposed strategy suitable for real-time optimization of processes with a large number of inputs. Two alternatives are proposed to implement the approach: gradient calculation by finite differences and quadratic regression using current and past data. The features and behavior of this approach are shown through two case studies: (i) a simple model with three processes, and (ii) a heat and power system of a sugar and ethanol plant. A comparison with other existent approaches shows a better performance in terms of operating cost and sensitivity to noise.

1. INTRODUCTION

A main concern in decision-making on industrial processes is to operate a plant at its maximum performance under changing scenarios. Nowadays, the availability of real-time process measurements and the improvements in processor capabilities encourage the use of computer-aided techniques to operate as close as possible to the desired optimal performance. Real-time optimization (RTO) exploits the plant's degrees of freedom to find the inputs that maximize a selected performance index, while satisfying operation, environmental, and safety constraints, and reacting to process, ambient, and prices disturbances. The process performance is optimized for a steady-state behavior, and the results can be sent as targets to a predictive control system or as set points to conventional control.

Figure 1 illustrates a traditional hierarchical structure for decision-making in industrial processes. In this scheme, RTO is placed as a link between scheduling (with a scale of days or weeks) and advanced control, normally model predictive control (MPC) (with scale of minutes). Recent publications address the integration of RTO and MPC in a single layer by adding a steady-state final cost in the MPC objective function[1,2] or by solving rigorously a dynamic real-time optimization problem (DRTO).[3,4]

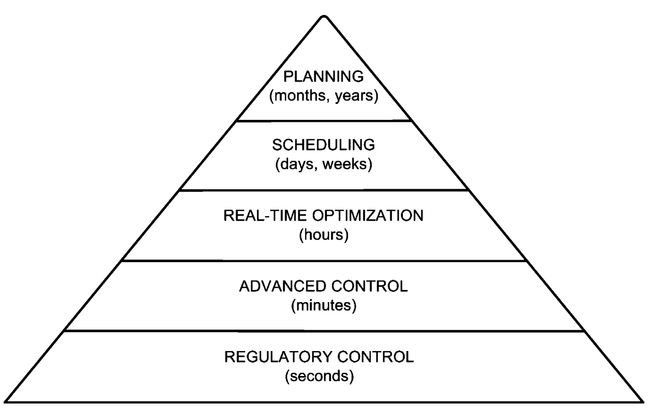

Figure 1. Decision-making hierarchy in an industrial plant.

The state of the art in RTO industrial applications has been investigated by Darby et al.[5] and Mansour and Ellis.[6] They identify five steps in the process: (i) steady-state detection, (ii) data validation, (iii) parameter adaptation, (iv) optimization, and (v) decision to update or not the settings of the advanced control system with the solution found in step iv. Data validation may include a data reconciliation and gross error detection step.

RTO optimization approaches are generally based on models. This means that the maximization of the process performance is calculated for a model of the process. Model-based RTO approaches require dealing with plant-model mismatch, which can lead to infeasibilities as well as to suboptimal solutions. The mismatch can be parametric or structural: in the former, the estimated parameter values are not the real ones; in the latter, the proposed model functions do not represent the real plant behavior even if the optimal parameters values are found. None of them can be avoided in industrial applications.

Different adaptation strategies using online process data are used to partially overcome this problem.[7] The traditional *two-step approach*[8] solves first a parameter estimation optimization problem, and then uses the updated parameters on the performance optimization step. Upon convergence and with a certain number of parameters, this approach guarantees feasibility, but it requires a larger number of parameters to reach the real plant optimum.

Received: December 28, 2012
Revised: October 24, 2013
Accepted: October 29, 2013
Published: October 29, 2013

dx.doi.org/10.1021/ie303621j | *Ind. Eng. Chem. Res.* 2013, 52, 16795−16810

Modifier adaptation approaches[9] replace the first optimization step of the two-step approach with the calculation of correction terms (*modifiers*). They can include biasing terms for constraints (*constraint-value modifiers*) and gradient correction terms for constraints and objective functions (*constraint-gradient* and *cost-gradient modifiers*, respectively). The modifiers are generally filtered using past and current estimations in order to reduce the effect of measurements noise. The quoted references show that, upon convergence, these methods reach a Karush-Kuhn-Tucker point of the real plant.

A third adaptation type, called *direct input*, does not solve optimization problems, but emulates a control system using as set points selected variables that guarantee the optimum. For example, NCO tracking[10,11] (where NCO stands for necessary conditions for optimality) defines as set points of the RTO system the Karush−Kuhn−Tucker conditions.

Uncertainty in measurements and results can also lead to infeasibilities. This problem is generally approached by backing-off the constraints (i.e., reducing the feasible region by defining more conservative constraints). Stochastic RTO approaches have also been developed.[12]

The results provided by the RTO scheme should be analyzed to decide if they will be implemented or not.[13] The decision is based on a statistical analysis of the calculated inputs in order to avoid unnecessary changes caused by the inherent variability of the measurements.

Model-free RTO approaches are described in the literature.[14] They need to make plant perturbations in several directions in order to determine experimentally the objective function gradient to calculate changes in a descent (or ascent) direction.

Heat and power systems often present some characteristics that make RTO a suitable alternative for operating them in an optimal way. Along the day, utilities systems face changes in steam demand, price of electric power (which can even change every few minutes according to grid demands in deregulated markets), and boiler and turbine efficiencies. The relatively fast dynamics of steam generation systems allow performing the optimization with a frequency on the order of minutes instead of the typical hours. Steady-state models can be used for optimization because of these fast dynamics. In addition, this behavior allows implementing a RTO scheme without the need of model predictive control to apply the solutions.

In this work, a modifier adaptation RTO scheme is proposed. It includes a strategy for correcting model gradients, which aims to reduce the dimension of the gradient calculation problem. It exploits a characteristic of many energy systems: all mass and energy balances can be modeled rigorously, without the need of empirical parameters; therefore, the equations to calculate them are free of structural and parametric mismatch. All the uncertainties are included in a subset of the model equations, which predicts the efficiency and performance of unit operations. Moreover, it is possible to identify which variables affect the performance of each process unit in order to reduce the number of data sets required to calculate real gradients (i.e., the number of past RTO cycles that are necessary to estimate the gradients of the current cycle). Multiple data sets are used for adapting modifiers. Yip and Marlin[15] proposed using multiple data sets as a way to increase the number of parameters to be adapted, while this work uses them to overcome structural mismatch and to reach the real plant optimal point.

Section 2 summarizes the current state of the art in modifier adaptation RTO and gradient correction methods. The proposed adaptation approach is presented in section 3. The application of the proposed RTO scheme is shown and discussed in section 4 through two case studies: (1) a simplified generic process and (2) a heat and power system of a sugar and ethanol plant.

2. MODEL ADAPTATION AND GRADIENT CORRECTION METHODS

2.1. Model Adaptation Strategies. The aim of real-time optimization is to find an input vector $\mathbf{u}^*$ that solves the following optimization problem:

$$\begin{aligned} &\min_{\mathbf{u}} Q_p(\mathbf{y_p}, \mathbf{u}) \\ &\text{s.t.} \\ &\mathbf{y_p} = \mathbf{f}_p(\mathbf{u}) \\ &\mathbf{g}(\mathbf{y_p}, \mathbf{u}) \le 0 \\ &\mathbf{u_l} \le \mathbf{u} \le \mathbf{u_u} \end{aligned} \tag{1}$$

where $\mathbf{u}$ and $\mathbf{y}_p$ are the vectors of process inputs and outputs, respectively; Q_p is the real process operation cost (a scalar function); $\mathbf{f_p}$ is the equations set that represents the *real* plant behavior; $\mathbf{g}$ is the process constraints set; and $\mathbf{u}_l$ and $\mathbf{u_u}$ are the lower and upper bounds sets on the inputs, respectively.

As the actual process behavior $\mathbf{f}_p(\mathbf{u})$ is unknown, the outputs are estimated from a process model:[16]

$$\begin{aligned} &\mathbf{f}(\mathbf{x}, \mathbf{u}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = 0 \\ &\mathbf{y} = \mathbf{h}(\mathbf{x}, \mathbf{u}, \boldsymbol{\alpha}, \boldsymbol{\beta}) \end{aligned} \tag{2}$$

where $\mathbf{y}$ is the vector of the estimated outputs; $\mathbf{x}$ are the state variables; $\mathbf{f}$ is the process model equations set; and $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are the vectors of fixed and adjustable parameters, respectively. Hereafter, model 2 will be denoted by the simplified expression $\mathbf{y} = \mathbf{y}(\mathbf{u},\boldsymbol{\alpha},\boldsymbol{\beta})$, and it will be implicitly included in cost and constraint functions as $Q(\mathbf{u},\boldsymbol{\alpha},\boldsymbol{\beta})$ and $\mathbf{g}(\mathbf{u},\boldsymbol{\alpha},\boldsymbol{\beta})$, respectively.

The two-step approach updates the adjustable parameters $\boldsymbol{\beta}$ in a first stage, and then uses these values in the cost optimization step.[17] In general, the first step requires solving the following minimization problem:

$$\begin{aligned} &\min_{\beta} \| \mathbf{y_m} - \mathbf{y} \| \\ &\text{s.t.} \\ &\mathbf{y} = \mathbf{y}(\mathbf{u}, \boldsymbol{\alpha}, \boldsymbol{\beta}) \end{aligned} \tag{3}$$

where $\mathbf{y_m}$ is the vector of the measured outputs. The resulting optimal parameter values $\boldsymbol{\beta}^*$ are then fixed, and used to find an input vector $\mathbf{u}^*$ that minimizes the cost Q:

$$\begin{aligned} &\mathbf{u}^* \in \arg\min_{\mathbf{u}} Q(\mathbf{u}, \boldsymbol{\alpha}, \boldsymbol{\beta}^*) \\ &\text{s.t.} \\ &\mathbf{g}(\mathbf{u}, \boldsymbol{\alpha}, \boldsymbol{\beta}^*) \le 0 \\ &\mathbf{u}_l \le \mathbf{u} \le \mathbf{u_u} \end{aligned} \tag{4}$$

Because of the unavoidable structural mismatch, it cannot be guaranteed that the previous steps lead to a point matching the plant KKT conditions. To do that, additional parameters for correcting the gradient should be added to the model. The ISOPE algorithm[18] (Integrated System Optimization and Parameter Estimation) corrects the objective function as follows:

$$Q_m = Q(\mathbf{u}, \boldsymbol{\alpha}, \boldsymbol{\beta}^*) - \boldsymbol{\lambda}^{\mathrm{T}} \cdot \mathbf{u} \tag{5}$$

where the vector of modifiers λ is calculated as[19]

$$\lambda = \left[\left[\frac{\partial \mathbf{y}}{\partial \mathbf{u}}\right]^{\mathrm{T}} - \left[\frac{\partial \mathbf{y}_m}{\partial \mathbf{u}}\right]^{\mathrm{T}}\right] \cdot \left[\frac{\partial Q}{\partial \mathbf{y}}\right] \tag{6}$$

Modifier adaptation does not adjust model parameters. Instead, it guarantees feasibility and optimality upon convergence by solving the following problem at each RTO cycle k:[16]

$$\begin{aligned} &\min_{\mathbf{u},\mathbf{y}} Q(\mathbf{u}, \boldsymbol{\alpha}, \boldsymbol{\beta}) + (\lambda^{\mathbf{Q}})^{\mathrm{T}} \cdot \mathbf{u} \\ &\text{s.t.} \\ &\mathbf{g}(\mathbf{u}, \boldsymbol{\alpha}, \boldsymbol{\beta}) + \boldsymbol{\varepsilon}^{G} + (\lambda^{G})^{\mathrm{T}}(\mathbf{u} - \mathbf{u}_k) \leq 0 \\ &\mathbf{u}_l \leq \mathbf{u} \leq \mathbf{u}_u \end{aligned} \tag{7}$$

Prior to the optimization step, the modifiers (i.e., correction factors for gradients and constraint values) are calculated:

$$\begin{aligned} &\varepsilon_i^{G} = [g_{p,i} - g_i(\mathbf{u}_k, \boldsymbol{\alpha}, \boldsymbol{\beta})], \quad i = 1, ..., n_g \\ &\lambda_i^{G} = \left[\frac{\partial g_{p,i}}{\partial \mathbf{u}} - \frac{\partial g_i(\mathbf{u}, \boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \mathbf{u}}\right]_{\mathbf{u}=\mathbf{u}_k}, \quad i = 1, ..., n_g \\ &\lambda^{Q} = \left[\frac{\partial Q_p}{\partial \mathbf{u}} - \frac{\partial Q(\mathbf{u}, \boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \mathbf{u}}\right]_{\mathbf{u}=\mathbf{u}_k} \end{aligned} \tag{8}$$

where n_g is the number of inequality constraints to be adapted.

The calculated modifiers can be affected by process measurements noise, which can increase results variability. Furthermore, if the modifiers are applied directly from eq 8 the correction could be excessive and affect the convergence of the RTO algorithm.[9] To diminish these undesired effects, an exponential filter can be applied to the calculated adaptation parameters.[7,20] Defining the vector of all constraints and gradients as $\mathbf{C} = (g_1...,g_{ng},\partial g_1/\partial \mathbf{u},...,\partial g_{ng}/\partial \mathbf{u},\partial Q/\partial \mathbf{u})$, and the vector of all modifiers $\boldsymbol{\Lambda} = (\varepsilon_1^{G},...,\varepsilon_{ng}^{G},(\lambda_1^{G})^{\mathrm{T}},...(\lambda_{ng}^{G})^{\mathrm{T}},(\lambda^{Q})^{\mathrm{T}})$, the expression for the filtered modifiers in RTO cycle k can be stated as

$$\Lambda_k := (I - K)\cdot\Lambda_{k-1} + K\cdot(\mathbf{C}_{p,k} - \mathbf{C}(\mathbf{u}, \boldsymbol{\alpha}, \boldsymbol{\beta})) \tag{9}$$

where $\mathbf{C}_{p,k}$ and $\mathbf{C}(\mathbf{u},\boldsymbol{\alpha},\boldsymbol{\beta})$ are, respectively, the real plant and model values for constraints and gradients, and K is a filtering matrix (in general, a diagonal matrix whose elements are between 0 and 1).

2.2. Gradient Correction Methods. ISOPE and modifier adaptation algorithms require an experimental determination of gradients, which is not a trivial task, and several techniques have been developed. Among them, the following ones can be mentioned:[21,22]

Finite Forward Differences (FFD). It requires small perturbations in the input variables around each steady state to calculate each derivative as follows:

$$\frac{\partial y}{\partial u} \approx \frac{y(u + \delta) - y(u)}{\delta} \tag{10}$$

In large (i.e., with a large number of inputs) or slow processes, this technique could be difficult to implement because of the time needed to perform the perturbations and to reach each steady state. Despite these limitations, it was proposed to use FFD to initialize gradient calculations until enough data is available for a dual control strategy.[23] Alternatively, a sinusoidal excitation is proposed.

Dual Control. The gradient is calculated based on previous steady-state data. A linear estimation of the derivatives for a property y can be done by solving the following equation system:

$$[(\mathbf{u}_k - \mathbf{u}_{k-1})(\mathbf{u}_k - \mathbf{u}_{k-2})...(\mathbf{u}_k - \mathbf{u}_{k-nu})]^{\mathrm{T}} \nabla \mathbf{y} \approx \begin{bmatrix} \mathbf{y}_k - \mathbf{y}_{k-1} \\ \mathbf{y}_k - \mathbf{y}_{k-2} \\ ... \\ \mathbf{y}_k - \mathbf{y}_{k-nu} \end{bmatrix} \tag{11}$$

where nu is the number of inputs.

Instead of updating the gradients as eq 11, it was also proposed to update cost and constraint gradient modifiers:[24]

$$[(\mathbf{u}_k - \mathbf{u}_{k-1})(\mathbf{u}_k - \mathbf{u}_{k-2})...(\mathbf{u}_k - \mathbf{u}_{k-nu})]^{\mathrm{T}} \boldsymbol{\lambda} \approx \begin{bmatrix} \boldsymbol{\varepsilon}_k - \boldsymbol{\varepsilon}_{k-1} \\ \boldsymbol{\varepsilon}_k - \boldsymbol{\varepsilon}_{k-2} \\ ... \\ \boldsymbol{\varepsilon}_k - \boldsymbol{\varepsilon}_{k-nu} \end{bmatrix} \tag{12}$$

where ε and λ are the modifier and the gradient modifier of a constraint or of the cost.

A difficulty may arise from the fact that the equations can be linearly dependent or the system ill-conditioned. To avoid this, a constraint has to be added to the RTO problem so that the new point will not cause ill-conditioning. This can be done by limiting the condition number of the equation system that the new input will produce. Analytical expressions for upper bounds on the errors due to truncation and noise were proposed by Marchetti et al.[9] Adding these new constraints proved to be more effective than limiting the condition number. Changes in measured or unmeasured disturbance variables can make inconsistent the gradient estimation until enough data are collected to make a new estimation with the actual conditions. This problem becomes critical when the number of inputs increases.

Alternatively, the output derivative matrix can be estimated by using a Broyden updating algorithm:

$$\mathbf{BR}_k = \mathbf{BR}_{k-1} + \frac{\mathbf{y}_k - \mathbf{y}_{k-1} - (\mathbf{BR}_{k-1})^{\mathrm{T}}\cdot(\mathbf{u}_k - \mathbf{u}_{k-1})}{\|\mathbf{u}_k - \mathbf{u}_{k-1}\|^2} \times (\mathbf{u}_k - \mathbf{u}_{k-1}) \tag{13}$$

This matrix has to be initialized at the beginning of the operation. This recursive method avoids the ill-conditioning problems that can appear in eqs 11 and 12, updating the gradient only in the direction of vector $(\mathbf{u}_k - \mathbf{u}_{k-1})$. However, it also needs additional constraints to limit truncation and noise errors, and to force the system to update the gradient in all directions of the inputs space.[25]

Bunin et al.[26,27] propose in recent works the use of a weighted least-squares regression for gradient estimation using past data. They also suggest the use of Lipchitz constants to bind the estimates; this includes fixing a partial derivative to 0 (very small values in practice) if an input is known not to influence a gradient. The use of a least-squares regression and the use of the plant knowledge to ignore the estimation of partial derivatives that are known to be zero were also approached by the authors of this paper.[28,29] These ideas will be detailed in section 3.

Model-Based Approaches. Among the model-based approaches, the following ones can be mentioned: parameter estimation-gradient calculation and neighboring extremals, self-optimizing control. They estimate the gradient on the basis of parameters variations using derivatives of the proposed model. Therefore, they cannot guarantee convergence under structural mismatch.

Dynamic Identification. The gradients of the steady-state model are identified during the transient between RTO runs using a dynamic (linear or nonlinear) model.

3. MODEL ADAPTATION USING PERFORMANCE EQUATIONS

3.1. Description of the Method. For certain process types, all structural and parametric mismatches can be grouped into a subset of the model equations, which are called hereafter *performance equations*. These equations, which can be used to predict reaction rates, boiler or turbine efficiencies, or heat exchange coefficients, among others, are formulated using experimental data, and can be corrected using online measurements. The rest of the model equations, which include mass, energy, and entropy balances, can be considered as structurally correct, although they can include, as a parameter, the result of a performance equation (for example, a boiler energy balance can include the efficiency calculated by the efficiency equation).

The real-time optimization problem to solve can be stated as

$$\min_{\mathbf{y},\mathbf{x},\mathbf{u},\boldsymbol{\eta}} Q(\mathbf{y}, \mathbf{u})$$

s.t.

$$\mathbf{f}(\mathbf{x}, \mathbf{u}, \boldsymbol{\eta}) = \begin{bmatrix} \mathbf{f}_1(\mathbf{x}, \mathbf{u}) \\ \mathbf{f}_2(\mathbf{x}, \mathbf{u}, \boldsymbol{\eta}) \end{bmatrix} = 0$$

$$\mathbf{y} = \mathbf{h}(\mathbf{x}, \mathbf{u})$$

$$\mathbf{p}_m(\mathbf{u}) = \boldsymbol{\eta}$$

$$\mathbf{g}(\mathbf{y}, \mathbf{x}, \mathbf{u}) \le 0$$

$$\mathbf{u}_\mathbf{l} \le \mathbf{u} \le \mathbf{u}_\mathbf{u} \qquad (14)$$

where $\mathbf{y}$ are the measured process outputs, $\mathbf{x}$ are the state variables, $\mathbf{u}$ are the process inputs (which can also be measured), $\mathbf{f}$ is the process model (mass, energy and entropy balances), $\mathbf{h}$ is the relation between outputs and inputs and states, and $\mathbf{g}$ are the inequality constraints. $\boldsymbol{\eta}$ are performance or efficiency factors, whose functionality with process inputs is estimated by an approximate model $\mathbf{p}_m(\mathbf{u})$. All the structural and parametric uncertainty in the model is concentrated in this model $\mathbf{p}_m$. It is assumed that the submodel $\mathbf{f}_2$ allows acquisition of the performance factors $\boldsymbol{\eta}$ if $\mathbf{y}$ and $\mathbf{u}$ are known.

Moreover, each performance factor η_j can be expressed as a function of new variables $\mathbf{v}_j$, which are in turn obtained as a function of $\mathbf{u}$:

$$\eta_j = p_{mj}(\mathbf{v}_j(\mathbf{u})) \quad j = 1, 2, ..., nj \qquad (15)$$

where nj is the number of performance equations in the system.

The following examples illustrate some of the possible forms or "structures" that variables $\mathbf{v}_j$ can take for an input vector $\mathbf{u} = (u_1, u_2, u_3)^T$:

•$\mathbf{v}$ is the same input vector $\mathbf{u}$:

$$\eta_1 = p_{m1}(u_1, u_2, u_3) \quad \mathbf{v}_1 = \mathbf{u}$$

•$\mathbf{v}$ is a subset of $\mathbf{u}$:

$$\eta_2 = p_{m2}(u_1, u_3) \quad \mathbf{v}_2 = (u_1, u_3)^T$$

$$\eta_3 = p_{m3}(u_2) \quad v_3 = u_2$$

•$\mathbf{v}$ is a linear combination of $\mathbf{u}$: $\mathbf{v}(\mathbf{u}) = P\cdot\mathbf{u}$, where P is an $nv \times nu$ matrix; nv is the dimension of $\mathbf{v}$. (This is a generalization of the two previous cases):

$$\eta_4 = p_{m4}(u_1 - u_2, u_3) \quad \mathbf{v}_4 = (u_1 - u_2, u_3)^T$$

•$\mathbf{v}$ is a continuous function of $\mathbf{u}$:

$$\eta_5 = p_{m5}(u_1 \cdot u_3) \quad v_5 = u_1 \cdot u_3$$

In the systems studied in this work, the dimension of the vectors $\mathbf{v}_j$ that affect each performance equation is lower than the dimension of the input vector $\mathbf{u}$. This is a typical situation in process systems formed by networks of unit operations (such as heat and power systems); the performance of each subprocess is not affected by other processes operating in parallel or downstream, but only by the inlet conditions to the subprocess itself.

If this happens, the gradient correction terms can be calculated in terms of $\mathbf{v}_j$ instead of $\mathbf{u}$. Therefore, the dimension of gradient estimation problems can be reduced with respect to eq 11 and 12 as fewer data sets (i.e., data from fewer RTO cycles) are necessary to estimate a gradient correction term for each equation. Furthermore, it avoids the experimental calculation of the gradient in directions of the space of inputs in which the directional derivative is known to be zero.

The model adaptation problem uses the available measurements to perform the necessary corrections. Having all outputs $\mathbf{y}$ measured and a proper model structure, this adaptation can be decomposed in three steps: (1) data reconciliation, (2) performance factors calculation, and (3) performance equations adaptation.

Step 1: Data Reconciliation. The reconciled values $(\mathbf{y}_k, \mathbf{u}_k)$ for the RTO cycle k are obtained from the measured values $(\mathbf{y}_{mk}, \mathbf{u}_{mk})$ as follows:

$$(\mathbf{y}_k, \mathbf{u}_k) = \arg\min_{\mathbf{y},\mathbf{x},\mathbf{u}} (\mathbf{y}_{mk} - \mathbf{y})^T A(\mathbf{y}_{mk} - \mathbf{y}) + (\mathbf{u}_{mk} - \mathbf{u})^T B(\mathbf{u}_{mk} - \mathbf{u})$$

s.t.

$$\mathbf{f}_1(\mathbf{x}, \mathbf{u}) = 0$$

$$\mathbf{y} = \mathbf{h}(\mathbf{x}, \mathbf{u}) \qquad (16)$$

where A and B are the variance-covariance matrices for outputs and inputs, respectively.[20,30]

Step 2: Performance Factors Calculation. Using the optimal variables values computed from eq 16, the equation system

$$\mathbf{f}_2(\mathbf{x}, \mathbf{u}, \boldsymbol{\eta}) = 0 \qquad (17)$$

is solved to obtain $\boldsymbol{\eta}_k$. Alternatively, the determination of the performance factors can be combined with the data reconciliation step, solving them simultaneously.

Step 3: Performance Equations Adaptation. The performance equations can be adapted for each RTO cycle k by introducing a modifier adaptation structure (note that in this case it is applied to performance equations instead of cost and constraints):[28]

$$\eta_j = p_{mj}(\mathbf{v}_j) + \beta_{jk} + \boldsymbol{\lambda}_{jk}^{\mathrm{T}}\cdot(\mathbf{v}_j - \mathbf{v}_{j,k}) \quad j = 1, 2, ..., nj \tag{18}$$

where $\beta_{jk} = \eta_{jk} - p_{mj}(\mathbf{v}_{j,k})$ is the constraint modifier (following the biasing update structure proposed by Forbes and Marlin),[31] and $\boldsymbol{\lambda}_{jk} = \nabla_v \eta_j|_{\mathbf{v}_j = \mathbf{v}_{j,k}} - \nabla_v p_{m,j}|_{\mathbf{v}_j = \mathbf{v}_{j,k}}$ is the gradient modifier for performance indicator η_j.

In this case, the gradient can be estimated using data from the current and previous RTO cycles:

$$\nabla_v \eta_j \approx [\eta_{j,k} - \eta_{j,k-1} \quad \eta_{j,k} - \eta_{j,k-2} \quad \cdots \quad \eta_{j,k} - \eta_{j,k-nI}]\cdot \mathbf{V}_j^{-1}$$
$$j = 1, 2, ..., nj \tag{19}$$

where $\mathbf{V}_j[\mathbf{v}_{j,k} - \mathbf{v}_{j,k-1} \ \mathbf{v}_{j,k} - \mathbf{v}_{j,k-2} \ ... \ \mathbf{v}_{j,k} - \mathbf{v}_{j,k-n_{ij}}]$, and n_{ij} is the dimension of vectors $\mathbf{v}_j$ and $\nabla_v \eta_j$. This approach is useful if n_i is close to 1 and lower than n_u, thus allowing a reduction of the dimension of the gradient estimation problem.

All gradients with respect to the independent variables $\mathbf{u}$ can be calculated following the chain rule:

$$\frac{\mathrm{d}\eta_j}{\mathrm{d}\mathbf{u}} = \left(\frac{\mathrm{d}p_{mj}}{\mathrm{d}\mathbf{v}_j} + \lambda_{jk}\right)^{\mathrm{T}} \cdot \frac{\mathrm{d}\mathbf{v}_j}{\mathrm{d}\mathbf{u}}, \quad j = 1, ..., nj \tag{20}$$

If all the structural plant-model mismatch can be included into these performance equations and these equations are a function of the proposed variables $\mathbf{v}$, all cost and plant gradients can be known by calculating the parameters $\boldsymbol{\lambda}_j$, $j = 1, ..., nj$. Therefore, upon convergence, a KKT point of the adapted model will also be a plant KKT point.

3.2. Implementation of the Equations Correction.

Alternative 1: Correction + Exponential Filter. The model can be adapted using eqs 18 and 19. An exponential filter is applied to the calculated adaptation parameters:

$$\boldsymbol{\beta}_k := K_\beta\cdot(\boldsymbol{\beta}_k - \boldsymbol{\beta}_{k-1}) + \boldsymbol{\beta}_{k-1}$$
$$\boldsymbol{\lambda}_{jk} := K_{\lambda j}\cdot(\boldsymbol{\lambda}_{jk} - \boldsymbol{\lambda}_{j,k-1}) + \boldsymbol{\lambda}_{j,k-1}, \quad j = 1, ..., nj \tag{21}$$

where K are filtering matrices (diagonal matrices[28] whose elements are between 0 and 1). Alternatively, a moving average filtering can be used for parameters that are known to change slowly.

Maximum and minimum values can be defined for modifiers $\boldsymbol{\lambda}_{jk}$ to limit the effect of noise. However, if the model performance equations have a significant mismatch, setting such maximum and minimum values will not guarantee the convergence to the real plant optimum.

A dual control strategy similar to the one proposed by Marchetti et al.[9] can also be applied. It defines two feasible regions for the optimization problem on each RTO cycle k, which guarantee an upper bound in the error due to truncation and noise in the gradient estimation for cycle $k + 1$. Otherwise, constraints based on the condition number of matrices $\mathbf{V}_j$ can be added.[32]

The case studies presented in section 4 use a simplified strategy that adds new constraints that set a maximal value $\boldsymbol{\delta}$ to the changes in process inputs $\mathbf{u}$; that is,

$$\|\mathbf{u} - \mathbf{u}_k\| \leq \boldsymbol{\delta} \tag{22}$$

As this simplified strategy does not guarantee that the gradient calculation problem will be well-conditioned, an update criterion is defined for each parameter $\boldsymbol{\lambda}_j$: they will not be updated if the changes on the corresponding variables $\mathbf{v}_j$ are smaller than a predefined threshold.

Alternative 2: Quadratic Regression. Efficiency η_j can also be adapted using a generic function γ_j:

$$\eta_j = h_{mj}(\mathbf{u}_k) + \gamma_{jk}(\mathbf{v}_j(\mathbf{u})) \quad j = 1, 2, ...nj \tag{23}$$

The adaptation term γ_{jk} is obtained by regression using a moving horizon. For example, a quadratic function can be proposed; its coefficients are updated by least-squares weighted quadratic regression:

$$\min_{a_k, \mathbf{b}_k, C_k} \sum_{j=1}^{nj} \sum_{i=k-N}^{k} w_i(\beta_{ji} - \gamma_{ji})^2 = \sum_{j=1}^{nj} [\min_{a_{jk}, \mathbf{b}_{jk}, C_{jk}} \sum_{i=k-N}^{k} w_i(\beta_{ji} - \gamma_{ji})^2]$$

s.t.

$$\left.\begin{array}{l} \beta_{ji} = \eta_{j,i} - p_{m,j}(\mathbf{v}_i) \\ \gamma_{ji} = a_{jk} + \mathbf{b}_{jk}^{\mathrm{T}}\cdot\mathbf{v} + \mathbf{v}_{ji}^{\mathrm{T}} C_{ji} \mathbf{v}_{ji} \\ \mathbf{v}_{ji} = \mathbf{v}_j(\mathbf{u}_i) \\ a_j^{\mathrm{L}} \leq a_{jk} \leq a_j^{U} \\ \mathbf{b}_j^{\mathrm{L}} \leq \mathbf{b}_{jk} \leq \mathbf{b}_j^{\mathrm{U}} \\ C_j^{\mathrm{L}} \leq C_{jk} \leq C_j^{\mathrm{U}} \end{array}\right\} \begin{array}{l} j = 1, ..., nj \\ i = (k - N), ..., k \end{array} \tag{24}$$

The inclusion of the quadratic terms C_{jk} adds more degrees of freedom to the modifier adaptation scheme.[27] They are not strictly necessary to obtain the desired property of matching plant and model KKT conditions upon convergence. However, they may be useful to provide a better estimation of the gradients at the current cycle k if the correction term β_{ji} is not linear with $\mathbf{v}_{ji}$ (i.e., if the gradient is not constant). Furthermore, by including extra degrees of freedom, the approach can improve the fitting of the real functionality of the efficiency terms η_j, which will result in a better prediction of these terms in the economic optimization problem.

However, the use of these terms requires more information from the data than a linear estimation. They must not be included in the optimization problem (24) in the first RTO cycles when the system is initialized, as the problem would be underspecified and the parameters can take any value. In addition, the prediction of extrapolated values (i.e., values outside the range determined by the N data sets of problem (24)) can be inaccurate if the functionality of the correction term is not quadratic. To overcome this problem, the authors suggest using conservative values for the upper and lower limits C_j^{U} and C_j^{L} (see examples in section 4).

The weights w_i can be adopted following different criteria. In the case studies in section 4, the criterion is to adopt a weight of 1 to data of cycle k. This weight decays for older data (the older the data, the lower the weight). This strategy penalizes the probability of low frequency disturbances that can make data sets from previous cycles not useful for the current gradient estimation. At the same time, if the system is converging to the optimum, the data from the most recent cycles will probably be at a lower distance to the current point. Therefore, they are assigned a higher weight (in comparison with older data), as they provide more information for the local gradient calculation. Other authors suggest a weight based on the distance from the current point and on the noise level.[26]

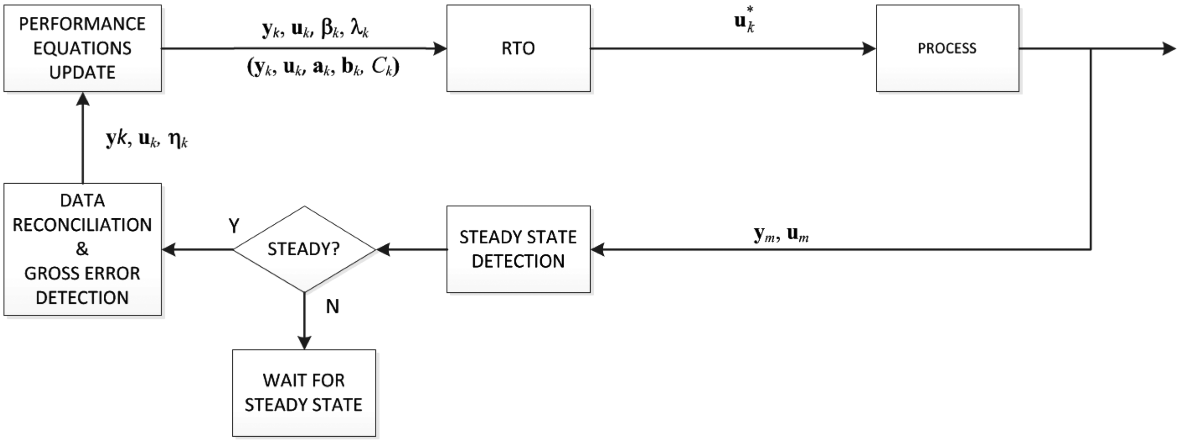

Figure 2. Scheme of the proposed RTO approach.

Although the polynomial regression approach does not require additional constraints to avoid ill-conditioning of the estimation problem,[26] the data set that is used for solving the problem (24) may not contain sufficient information for a valid estimation of all the proposed parameters. This can be particularly problematic when the RTO scheme converges to a point, because if no disturbances are present the system will remain in the convergence point (or within a small region around it) for several RTO cycles, and the quality of the calculated parameters will decrease. If necessary, a check on the quality of the estimates can be performed, and plant experiments can be forced to improve it.[33] In the case studies of section 4, this validation was not implemented for the sake of simplicity; nevertheless, the proposed approach showed a good performance.

Yip and Marlin[15] have suggested a structure similar to eq 24. The difference lies in that they propose a simultaneous data reconciliation and parameter estimation, which is solved for several data sets at the same time. In this paper, a decomposition of the problem is proposed, which leads to a faster computation and may be useful for certain types of processes, like the steam and power system analyzed in Case Study II. In addition, the approach here proposed seeks to obtain the real optimal plant cost by matching the KKT conditions, while the aforementioned authors have only suggested a procedure to adjust parameters that cannot be obtained using a single data set.

Figure 2 shows a diagram illustrating the proposed RTO structure. As pointed out in the Introduction section, it includes a previous steady-state detection step, which is only mentioned briefly in this work. There exists a vast literature addressing this issue.[34−36]

4. CASE STUDIES

In this section, two case studies are presented and discussed to show the behavior and performance of the proposed structures. First, the main methodological features of the proposed RTO scheme are explained and highlighted through a very simple case (Case Study I). Then, a more complex and realistic case study is considered, which consists of a heat and power system of a sugar and ethanol facility (Case Study II).

The methodology followed to analyze the RTO system performance is the same as used in other publications in this area.[37] Two models are used: one model, called *real plant,* is assumed to represent exactly the system to be optimized. It is used to simulate the plant's steady state and to evaluate the real value of the objective function. A second model, called *RTO model*, has a structural and parametric mismatch with respect to the real plant. A set of the real plant results, called *measurements*, is used to adapt the RTO model. The optimal calculated inputs are used in the real plant, which allows evaluation of the real cost and acquisition of new measurements for the next RTO cycle.

To evaluate and compare the performance of the proposed and other RTO adaptation approaches, the Extended Design Cost criterion was used.[38] It evaluates the total loss of profit of a given RTO approach, the real optimal plant cost being the reference used to calculate the maximum profit that can be achieved by optimization.

4.1. Case Study I. A Simple System.

A scheme of the system is shown in Figure 3. The process consists of three unit operations, with two degrees of freedom (F_1 and F_2). Each unit operation is characterized by efficiency η_i, which is known to be a function of the corresponding flow F_i.

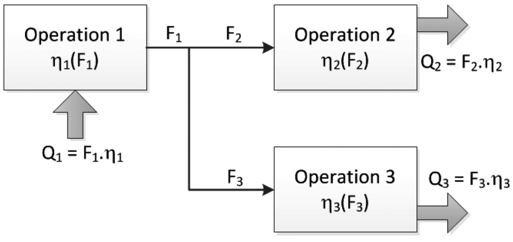

Figure 3. System modeled in Case Study I

Table 1 shows the efficiency equations for each operation.

The measurements from the real plant are F_1, F_2, Q_1, Q_2, and Q_3. As there is no redundancy in the measurements, there is no need of a data reconciliation step. In a RTO cycle k, the steps for adaptation/optimization are as follows.

1. Balances.

$$F_{3,k} = F_{1,k} - F_{2,k} \tag{25}$$

2. Performance Factors Calculations.

$$\eta_{j,k} = \frac{Q_{j,k}}{F_{j,k}}, \quad j = 1, 2, 3 \tag{26}$$

Alternative 1:

Table 1. Performance Equations for Case Study I

efficiency	real plant	RTO model
η_1	$10000/(85 - 20\cdot\exp(-(F_1 - 80)/190)$	$10000/(58 - 0.08\cdot F_1)$
η_2	$65 - 0.006\cdot(F_2 - 80)^2 - 24/F_2$	$65 - 0.06\cdot F_2$
η_3	$80 - 0.006\cdot(F_3 - 35)^2 - 12/F_3$	$50 + 0.04\cdot F_3$

3.a. Model Adaptation.

$$\begin{aligned} &\beta_{jk} = \eta_{j,k} - p_{mj}(F_{j,k}) \\ &\textit{if } |F_{j,k} - F_{j,k-1}| > \delta, \\ &\lambda_{jk} = \frac{\eta_{j,k} - \eta_{j,k-1}}{F_{j,k} - F_{j,k-1}} - p'_{mj}(F_{j,k}), \quad j = 1, 2, 3 \\ &\lambda_{jk} = f_j\cdot\lambda_{jk} + (1 - f_j)\cdot\lambda_{j,k-1} \\ &\textit{end if} \\ &\lambda_{jk} = \max(\lambda_j^{\mathrm{L}}, \min(\lambda_j^{\mathrm{U}}, \lambda_{jk})) \end{aligned} \tag{27}$$

4.a. Cost Minimization.

$$\begin{aligned} &\min_{\mathbf{F},\mathbf{Q},\eta} Q_1 - Q_2 - Q_3 + \rho\sum_{j=1}^{3}(F_j - F_{j,k})^2 \\ &\text{s.t.} \\ &F_1 - F_2 - F_3 = 0 \\ &Q_j - F_j\cdot\eta_j = 0, \quad j = 1, 2, 3 \\ &\eta_j = p_{mj}(F_j) + \beta_{jk} + \lambda_{jk}\cdot(F_j - F_{j,k}), \quad j = 1, 2, 3 \\ &\mathbf{F}_L \le \mathbf{F} \le \mathbf{F}_U \\ &|F_j - F_{j,k}| \le M \end{aligned} \tag{28}$$

Alternative 2:

3.b. Model Adaptation.

$$\begin{aligned} &\beta_{jk} = \eta_{j,k} - p_{mj}(F_{j,k}) \\ &(a_{jk}, b_{jk}, c_{jk}) \\ &\quad = \arg\min\left(\sum_{i=k-N}^{k} w_i(\beta_{ji} - a_j - b_j\cdot F_{ji} - c_jF_{j,i}^2)^2\right) \\ &\text{s.t.} \\ &a_j^{\mathrm{L}} \le a_{jk} \le a_j^{\mathrm{U}} \quad j = 1...nj \\ &b_j^{\mathrm{L}} \le b_{jk} \le b_j^{\mathrm{U}} \\ &c_j^{\mathrm{L}} \le c_{jk} \le c_j^{\mathrm{U}} \end{aligned} \tag{29}$$

4.b. Cost Minimization:

$$\begin{aligned} &\min_{\mathbf{F},\mathbf{Q},\eta} Q_1 - Q_2 - Q_3 + \rho\sum_{j=1}^{3}(F_j - F_{j,k})^2 \\ &\text{s.t.} \\ &F_1 - F_2 - F_3 = 0 \\ &Q_j - F_j\cdot\eta_j = 0 \\ &\eta_j = p_{mj}(F_j) + a_{jk} + b_{jk}\cdot F_j, \quad j = 1, 2, 3 \\ &\mathbf{F}^{\mathrm{L}} \le \mathbf{F} \le \mathbf{F}^{\mathrm{U}} \\ &|F_j - F_{j,k}| \le M \end{aligned} \tag{30}$$

5. Apply Optimization Results and Wait for Next Steady State: $k := k + 1$. The regression step (3b) is implemented without the quadratic coefficients c_j from $k = 0$ to $k = 7$. For $k \ge 7$, these coefficients c_j are included.

Table 2 shows the values of the parameters used in the RTO scheme.

Table 2. Parameters Used in Case Study 1

parameter	value	parameter	value
F^{L}	$[80, 55, 20]^{\mathrm{T}}$	b^{U}	$[7, 7, 7]^{\mathrm{T}}$
F^{U}	$[120, 90, 50]^{\mathrm{T}}$	c^{L}	$[-0.1, -0.1, -0.1]^{\mathrm{T}}$
f	$[0.7, 0.7, 0.7]^{\mathrm{T}}$	c^{U}	$[0.1, 0.1, 0.1]^{\mathrm{T}}$
N	$\min(k, 10)$	ρ	5
w_i	0.9^{k-i}	δ	0.3
a^{L}	$[-500, -500, -500]^{\mathrm{T}}$	M	3
a^{U}	$[500, 500, 500]^{\mathrm{T}}$	λ^{L}	$[-0.9, -0.9, -0.9]^{\mathrm{T}}$
b^{L}	$[-7, -7, -7]^{\mathrm{T}}$	λ^{U}	$[0.9, 0.9, 0.9]^{\mathrm{T}}$

Alternatives 1 and 2 were evaluated for 40 RTO cycles. Each alternative is run under two scenarios: in the first one, no noise is present in the measurements; in the second one, both flows F_1 and F_2 and outputs Q_1, Q_2, and Q_3 are perturbed with Gaussian noise of mean 0 and standard deviation 0.1.

The results are compared with an RTO approach without gradient adaptation (i.e., Alternative 1 with $\boldsymbol{\lambda} = 0$), and with a modifier adaptation algorithm where the gradients with respect to inputs F_1 and F_2 are calculated for the three outputs Q_1, Q_2, and Q_3. A Broyden update is used to calculate experimental gradients in this implementation (called hereafter Broyden-MA). To control gradient variance and offset, additional constraints must be included in the Broyden-MA approach.[25] The constraints used in this implementation are

$$(F_1 - F_{1,k})^2 + (F_2 - F_{2,k})^2 \le 1 \tag{31}$$

$$\text{abs}\left(\alpha_k^{\mathrm{T}}\begin{pmatrix} F_1 - F_{1,k} \\ F_2 - F_{2,k} \end{pmatrix}\right) \ge \sqrt{0.5\cdot\alpha_k^{\mathrm{T}}\cdot\alpha_k} \tag{32}$$

where $\boldsymbol{\alpha}_k^{\mathrm{T}} = ((-F_{2,k} + F_{2,k-1}),(F_{1,k} - F_{1,k-1}))$, that is, a vector normal to the two last inputs.

Constraint 32 generates two feasible regions as usual in all dual control modifier adaptation approaches. The optimization must be solved for each feasible region, and the best of the two solutions is selected and sent to the control system to be applied in the plant. The exponential filter for updating all gradient modifiers in Broyden-MA was selected as 0.6.

The model and all the mentioned approaches were implemented in General Algebraic Modeling System (GAMS).[39] The optimization problems were solved using CONOPT 3.[40]

The initial point chosen for the study was $[F_{1,0} = 90, F_{2,0} = 60]$. The objective function value computed at this point is 5907.75. The real optimal point is obtained by solving the optimization problem of the real plant; the computed optimal solution is $[F_1 = 105.59, F_2 = 71.09]$, which corresponds to a cost of 5520.809.

The system is initialized with two experimentation steps in order to allow gradient estimations. They correspond to cycles 1 and 2: $[F_{1,1} = 90.5, F_{2,1} = 60]$, $[F_{1,2} = 90.5, F_{2,2} = 60.5]$. In the last cycles ($k > 15$) of Alternative 1, as the model approaches the optimum, the changes can be too small to update gradients (see eq 27). To overcome this situation, if the gradient of a performance equation was not updated in cycle k, then an additional constraint is added in cycle k + 1:

$$F_i \geq F_{i,k} - 1\cdot\mathrm{sign}(\mathrm{d}Q/\mathrm{d}F_i) \tag{33}$$

Equation 33 indicates that the direction chosen for experimentation is a descent direction for the objective function, according to the model and the corrections made.

Figures 4, 5, and 6 show the evolution for Alternative 1, Alternative 2, and Broyden-MA approaches, respectively. All

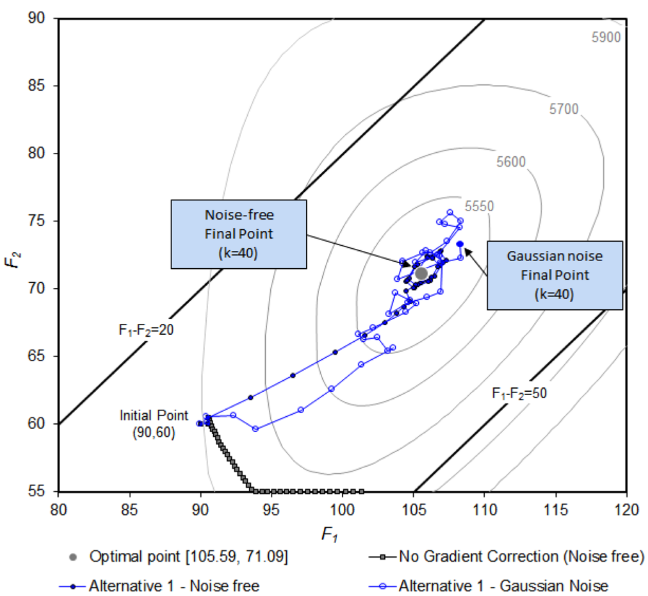

Figure 4. Alternative 1 evolution (performance equation gradient correction).

cases are compared with the traditional approach (i.e., without gradient correction). It can be observed that the traditional approach leads the plant to an operating point different to the real optimum. For noise-free cases, Alternatives 1 and 2 converge to the vicinity of the optimum, while Broyden-MA cannot reach the real minimum cost in the number of RTO cycles analyzed in

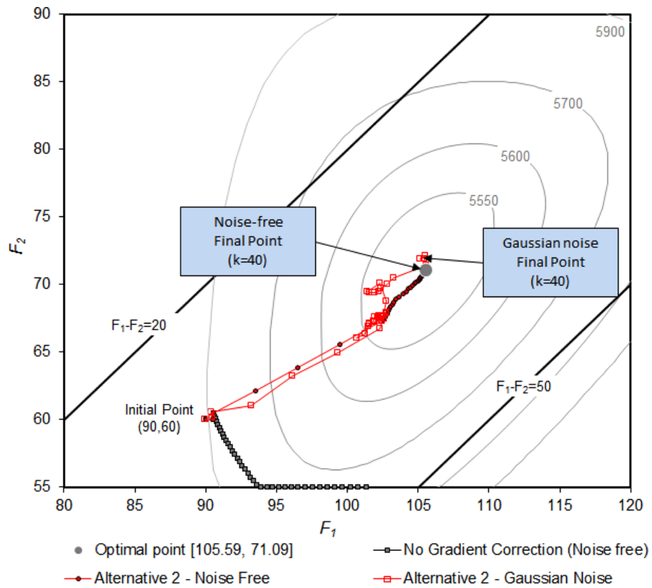

Figure 5. Alternative 2 evolution (performance equation corrected with quadratic regression).

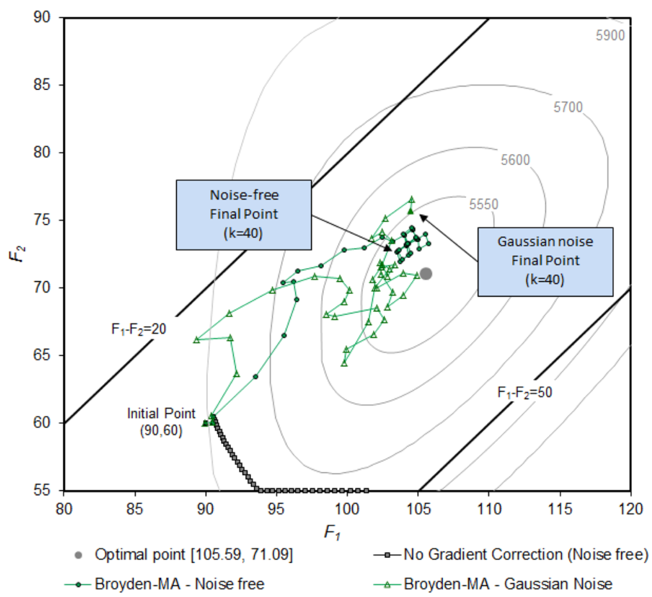

Figure 6. Broyden-MA evolution.

this case. The proposed alternatives also show a better performance than Broyden-MA when Gaussian noise is considered. As seen in the figures, Alternative 2 has a slower convergence to the optimum than Alternative 1, but shows a smaller variability of the results when it approaches the optimum. This is a desirable behavior of the system since the changes in the vicinity of the optimum generate small or negligible benefits (or even an increase in the operating cost due to the error in gradient estimation).

Figure 7 plots the real plant cost evolution for all analyzed adaptation strategies. For this case study, it shows that the proposed alternatives perform better than the compared approaches in this case study.

The performance of each RTO approach is evaluated using the Extended Design Cost (EDC) criterion.[38] The lower the EDC is, the higher is the benefit obtained by the RTO system. The EDC criterion is formulated as follows:

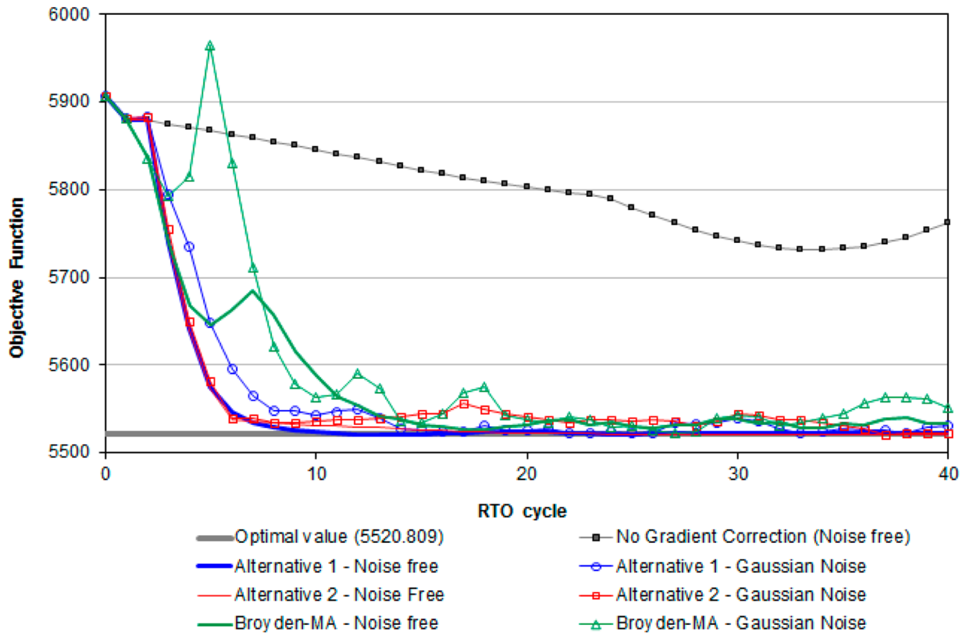

Figure 7. Cost evolution for Alternative 1, Alternative 2, Broyden-MA, and no gradient correction.

$$\mathrm{EDC} = \int_{T_0}^{T_f} (Q_{r,k} - Q_r^*)\,dt \approx \frac{1}{2}(Q_{r,k_0} - Q_r^*) + \sum_{k=k_0+1}^{k_f-1} (Q_{r,k} - Q_r^*) + \frac{1}{2}(Q_{r,k_f} - Q_r^*) \quad (34)$$

where Q_r^* is the real plant optimal cost and $Q_{r,k}$ is real plant cost at cycle k. T_0 and T_f are initial and final times, which correspond to RTO cycles k_0 and k_f, respectively. The results are summarized in Table 3. The EDC is analyzed for the whole set of cycles (i.e., from $k = 0$ to $k = 40$) and also for the period from $k = 5$ to $k = 40$. The first comparison allows an evaluation of the total loss of profit of the RTO approach, while the second one analyzes how the system behaves when the plant is in the vicinity of the optimum (this is true in the gradient correction approaches; without gradient correction, the approach converges to other operating point). The following criterion is considered to set $k = 5$ in the latter case: at $k = 5$ for noise-free cases, using both Alternative 1 and 2, the plant reached a cost differing less than 1% from the real plant optimal cost. According to this criterion, the evolution of the system can be divided in two stages: in the first, the plant cost at cycles k $Q_{r,k}$ is significantly different (i.e., > 1%) from the optimal cost Q_r^* ; in the second, the cost is close to the desired objective function value, and therefore, the cost variation after each RTO cycle is small. In this second stage, the difference between the plant cost and the optimal plant cost are due to gradient estimation errors and data variability.

Table 3. Extended Design Cost for All Studied Scenarios. Case Study I

scenarios	EDC from $k = 0$ to 40 absolute	EDC from $k = 0$ to 40 relative	EDC from $k = 5$ to 40 absolute	EDC from $k = 5$ to 40 relative
no action	15477.6	100	13542.9	100
Noise-Free Scenarios				
no gradient correction	11244.4	72.65	9453.1	69.80
Broyden-MA	2384.7	15.41	1090.1	8.05
Alternative 1	1402.4	9.06	123.7	0.91
Alternative 2	1426.3	9.22	150	1.11
Gaussian Noise in Data				
Broyden-MA	3320.6	21.45	1661.5	12.27
Alternative 1	1969.1	12.72	499.9	3.69
Alternative 2	1898.8	12.27	587.3	4.34

The relative EDC values presented in Table 3 are calculated by eq 35:

$$\text{Relative EDC} = \mathrm{EDC}/\mathrm{EDC}_{NA} \cdot 100 \quad (35)$$

where EDC_{NA} is the Extended Design Cost if no optimization is applied to the process.

The Extended Design Cost evaluation shows that Alternatives 1 and 2 outperform the Broyden-MA and traditional strategies. The difference is more noticeable for $k > 5$, when the proposed alternatives have converged to the vicinity of the optimum. As expected, the traditional no-gradient correction approach (NGC) shows the worst performance since significant structural mismatch between the real plant and the RTO model is present in this case.

Gradient correction makes Broyden-MA perform much better than NGC. However, it fails to converge to the real plant optimum in the number of RTO cycles specified in this case. The reason is that the gradients are calculated with respect to all process inputs, including those whose partial derivative is 0. In addition, Alternatives 1 and 2 make use of the process knowledge to reduce the dimension of the gradient (for example in efficiency factor η_3, whose gradient is calculated as a function of $F_3 = F_1 - F_2$, instead of obtaining independently the partial derivatives with respect to F_1 and F_2). Therefore, the Broyden-MA implementation has more sources of potential errors, which cause a slower convergence.

The choice of the filtering parameter value can be critical to the performance of the adaptation approach. Figure 8a shows the variation of the EDC with the gradient modifier exponential filter for Alternative 1 and Broyden-MA, as well as the variation of EDC with the weighting factor w_i for Alternative 2 (see Table 2).

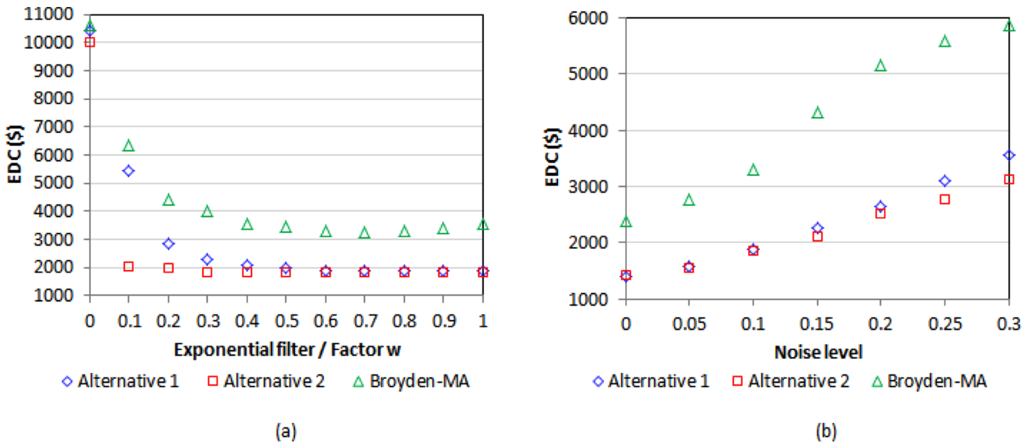

Figure 8. (a) Extended Design Cost (EDC) vs exponential filter (Alternative 1 and Broyden-MA), and vs weighting factor (Alternative 2). (b) EDC vs noise level.

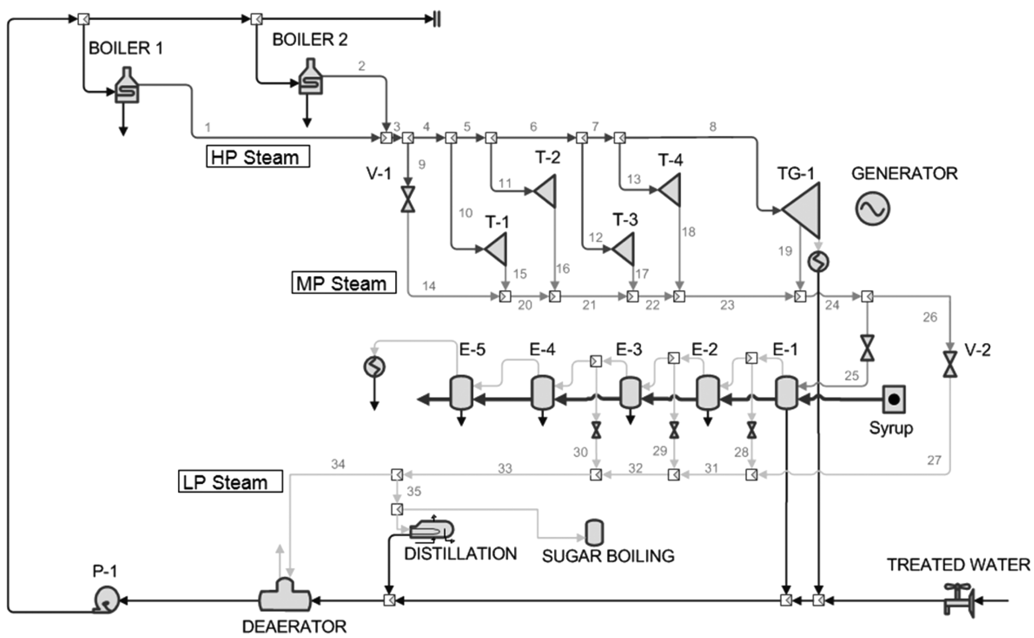

Figure 9. Sugar and ethanol heat and power system diagram.

The noise level is the same as the base case (0.1 for efficiencies and flows). Each point plotted in Figure 8a is the average value from 50 runs of the case study (40 RTO cycles each, with the same starting point for all runs) with random noise generation. It can be observed that Alternatives 1 and 2 do not show a significant sensibility to parameter values greater than 0.6. The results for Broyden-MA suggest an optimal value for the filtering parameter between 0.6 and 0.8.

The noise level can also impact on the real-time optimization performance. Figure 8b shows the variation of EDC for each implementation when the noise level changes. Again, each point plotted is the average value resulting from 50 runs of the case study; in this case the filtering parameters are kept constant for each alternative at the value used in the study. As expected, when the noise level is increased, all the alternatives show a higher EDC. For higher noise levels, Alternative 2 performs better than Alternative 1, while for low noise levels their performances are similar. Broyden-MA shows a poorer performance for all noise levels.

4.2. Case Study II. A Heat and Power System for a Sugar and Ethanol Facility.

A model of an energy system for a sugar and ethanol plant was developed in GAMS and solved using CONOPT 3. It includes two bagasse boilers, four backpressure turbines for cane milling, an extraction-condensing steam turbogenerator, a five-effect evaporator system, and other steam demands. A diagram of the modeled system is shown in Figure 9. The model equations are included in the supplied Supporting Information. A model description can also be found in Serralunga et al.[29] The nomenclature associated with Case Study II can be found in Table 4.

If the steam header pressures and the syrup final concentration are fixed, the plant has 6 degrees of freedom. The six inputs selected to control the plant can be, for instance, TG-1 extraction flow, TG-1 condensing flow, vapor bleeding flows from effects 1, 2, and 3, and the difference in steam flows between Boiler 1 and Boiler 2. If modifier adaptation or ISOPE approaches were implemented in a dual control strategy following the methods mentioned in section 2, at least six data sets would be necessary to set the additional constraints to avoid ill-conditioning or to

Table 4. Nomenclature for Case Study II

symbol	item
F	flow (kg/h)
P	pressure (Pa)
T	temperature (K)
X^{med}	solids mass fraction
C_p	heat capacity (kJ/kg·K)
V	evaporator steam used in next effect (kg/h)
Q	boiler fuel use (kJ/h)
W	turbine power (kW)
c	evaporation coefficient (kg/m²·°C)
Subscripts	
cb	boiler
tb	backpressure turbine
tg	extraction-condensing turbine stage
e	evaporator effect
Superscripts	
o	outlet stream
sat	saturated (water or steam)

limit the gradient estimation error. This number can be high enough to discourage the implementation of the mentioned dual control strategies, as these constraints can limit in excess the feasible operating points at each RTO cycle, and plant disturbances can occur with a frequency higher than 6 RTO cycles, which can make invalid the past gradient estimations. Then, the proposed approach based on performance equations is suggested in this case.

Enthalpy and entropy values for water and steam streams, as well as saturation temperatures and pressures, are estimated as Irving and Liley.[41] Boiling point elevation (BPE) and heat capacity (C_p) for syrup are calculated according to Camargo et al.[42] All these equations as well as mass, energy, and entropy balances are free of error in the analyzed case. The errors in the model are included in the prediction of boiler efficiencies, turbogenerator isentropic efficiencies, and evaporation coefficients in each evaporator effect.

Evaporation coefficients are equivalent to heat transfer coefficients; they indicate the amount of steam produced per unit of heat transfer area and per degree of temperature difference between the heating steam and the boiling cane syrup. The equations used for the RTO model prior to correction are of the type of the Dessin equation.[43] The expression for calculating the real plant evaporation coefficients is intentionally different to the Dessin equation (see Table 5). While each coefficient is explicitly a function of temperatures and syrup Brix degrees, the gradient correction is performed with respect to the live steam flow used to heat the first evaporator stage. It is assumed that vapor bleeding is maximized and, therefore, temperatures and Brix degrees for the optimal configuration can be calculated as a function of the inlet steam flow.

Turbogenerator second stage efficiency for the real plant is a function of two variables: stage steam flow and inlet temperature (which depends on the efficiency of the first stage and the temperature and pressure of the first stage admission steam). However, the RTO model and the gradient correction strategy consider negligible the functionality with temperature, correcting the gradient only with respect to the steam flow.

Model and real plant equations are shown in Table 5.

The objective function accounts for a fuel cost assigned to bagasse burnt, and the income due to electric power export.

Four scenarios are studied. The first one, which is used to check the convergence of the proposed approach to the real plant optimum, considers a constant steam demand for distillation and other uses. In the second, steam demand can vary between two RTO cycles. In the period between two cycles, the change in demand is managed by adjusting steam production in boilers 1 and 2, keeping constant the calculated optimal difference between these two steam flows. In the third scenario, electric power price can change every four cycles. The fourth scenario combines the changes in demand and power price of scenarios 2 and 3, respectively.

The "measured" data and the variance of each measured variable are listed in Table 6. The subindexes in flows and temperatures indicate the stream number, according to Figure 9. As bagasse mass flow to boilers is difficult to be measured with precision,[44] boiler efficiencies should be calculated through an indirect method, based on air and flue gas temperatures, CO_2 and O_2 content in flue gas, and an estimation of losses due to radiation and unburnt bagasse. The variances for boiler efficiencies shown in Table 6 correspond to the accuracy of that whole indirect calculation.

Table 6. Measured Values and Variances. Case Study II

variable	measurement points	σ^2
flow	$F_1,F_2,F_8,F_9,F_{19},F_{25},F_{26},F_{28},F_{29},F_{30},F_{35}$	0.09 (t/h)²
efficiency	η_{cb1},η_{cb2}	0.36 (%)²
temperature	T_1,T_2,T_8	0.25 (°C)²
TG electric power	W_{TG}	625 (kW)²
evaporator effect steam temperature	$T_1^{sat},T_2^{sat}T_3^{sat},T_4^{sat},T_5^{sat}$	0.25 (°C)²

As there exists redundancy in measurements, data reconciliation is performed. The performance factors indicated in Table 5 are calculated simultaneously with reconciliation. The optimization problem solved in this step is given by eq 16, being $(\mathbf{y}_{mk},\mathbf{u}_{mk})$ and $(\mathbf{y}_k,\mathbf{u}_k)$ the sets of measured and reconciled variables from Table 6, respectively; and $\mathbf{h}(\mathbf{y},\mathbf{u},\boldsymbol{\eta})$ the whole energy system model. For this case, matrices A and B were assumed to be diagonal matrices, whose nonzero elements are the variances of the corresponding measured variables.

Table 5. Model and Real Plant Equations. Case Study II

performance factors	real plant ($\eta_j = p(x_j)$)	RTO model ($\eta_j = p_m(x_j)$)
Boiler 1 efficiency	$89.975 - 0.001(140 - F^0_{cb1})^2 - 10^{-6}(F^0_{cb1})^3$	$92 - 0.005(150 - F^0_{cb1})^2$
Boiler 2 efficiency	$91 - 0.001(155 - F^0_{cb2})^2 - 1.2 \times 10^{-6}(F^0_{cb2})^3$	$92 - 0.005(150 - F^0_{cb2})^2$
TG Stage 1 efficiency	$53 - 0.0023(120 - F_{tg1})^2 + 1 \times 10^{-6}(F^0_{tg1})^3$	$70 - 0.005(130 - F^0_{tg1})^2$
TG Stage 2 efficiency	$66 - 0.0018(50 - F^0_{tg2}) + 2.2 \times 10^{-6}(F^0_{tg2})^3 + 10^{-2}T^0_{tg1}$	$70 - 0.004(41 - F^0_{tg2})^2$
evaporation coefficients ($e = 1, ..., 5$)	$(0.0012 - 0.0001e)(96 - X^{med}_e)(T^{sat}_{e-1} - 54) - 0.2(X^{med}_e/100)^2 - 0.5\exp(-(T^{sat}_{e-1} - 50)/100)$	$0.001(100 - X^{med}_e)(T^{sat}_{e-1} - 54)$

Table 7 shows the process values that are fixed in this case study. Table 8 lists the initial operating point for all examined scenarios.

Table 7. Fixed values. Case Study II

variable	value
high pressure steam (bar)	80
medium pressure steam (bar)	10
low pressure steam (bar)	0.8
TG condenser pressure (bar)	0.35
boilers steam temperature (°K)	650
juice flow (t/h)	500
initial Brix (°Bx)	15
final Brix (°Bx)	65
evaporator area, $e = 1, ..., 5$ (m^2)	2000
turbines efficiency, tb = 1, ..., 4 (%)	50
turbines power, tb = 1, ..., 4 (kW)	500

Table 8. Initial Operating Point. Case Study II

variable	value
F^o_{cb1}(t/h)	70
F^o_{cb2} (t/h)	158.9
F^i_{tg1} (t/h)	100
F^o_{tg1} (t/h)	70
VB_{e1} (t/h)	139.0
VB_{e2} (t/h)	12.4
VB_{e3} (t/h)	15.8
steam demand d_0 (t/h)	200
power cost $c_{W,k}$ ($/kWh)	0.018

The following step is a model adaptation. Alternative 2 was selected since more stable results than Alternative 1 were observed in the vicinity of the optimum in Case Study I, as well as because the quadratic regression of the data provides a correction that is valid in a wider region than the gradient correction of Alternative 1.

The adaptation procedure is formulated as follows:

$$\beta_{j,k} = \eta_{j,k} - p_{mj}(F_{j,k})$$

$$(a_{jk}, b_{jk}, c_{jk}) = \arg\min\left(\sum_{i=k-15}^{k} 0.9^i\cdot(\beta_{ji} - \gamma_{ji})^2\right)$$

s.t.

$$\begin{aligned}
&\gamma_{1i} = a_{1k} + b_{1k}\cdot F^o_{cb1,k} + c_{1k}\cdot(F^o_{cb1,k})^2\\
&\gamma_{2i} = a_{2k} + b_{2k}\cdot F^o_{cb2,k} + c_{2k}\cdot(F^o_{cb2,k})^2\\
&\gamma_{3i} = a_{3k} + b_{3k}\cdot F^o_{tg1,k} + c_{3k}\cdot(F^o_{tg1,k})^2\\
&\gamma_{4i} = a_{4k} + b_{4k}\cdot F^o_{tg2,k} + c_{4k}\cdot(F^o_{tg2,k})^2\\
&\gamma_{ji} = a_{jk} + b_{jk}\cdot V_{0,k} \quad j = 5, ..., 9\\
&a_j^L \le a_{jk} \le a_j^U \quad j = 1, ..., nj\\
&b_j^L \le b_{jk} \le b_j^U \quad j = 1, ..., nj\\
&c_j^L \le c_{jk} \le c_j^U \quad j = 1, ..., nj
\end{aligned} \tag{36}$$

It can be noticed that boiler and turbine efficiencies are corrected with a quadratic function, while evaporation coefficients are modified with a linear function of the live steam to effect 1. Parameters c_{jk} are only allowed to be different to 0 for RTO cycles with $k > 4$.

The real plant model allows a lower temperature difference in each evaporator effect. The reason for using this constraint back-off in the RTO model is to ensure feasibility: uncertainty in measurements together with plant-model mismatch can lead to RTO solutions that are infeasible for the real system, as it was mentioned in the Introduction section.

The real-time optimization problem is formulated as follows:

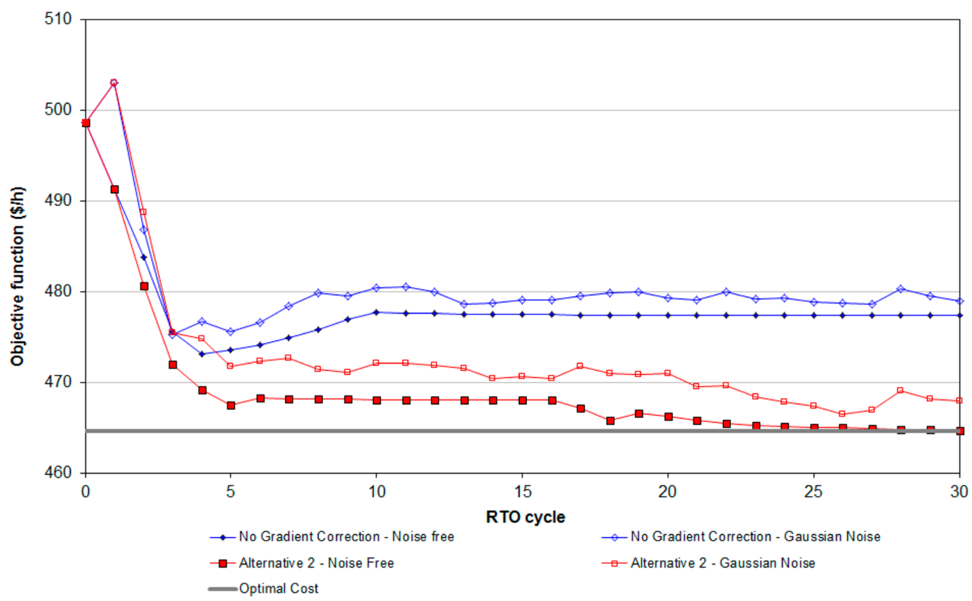

Figure 10. Case Study II, scenario 1. Objective function for constant steam demand and constant power price.

dx.doi.org/10.1021/ie303621j | Ind. Eng. Chem. Res. 2013, 52, 16795−16810

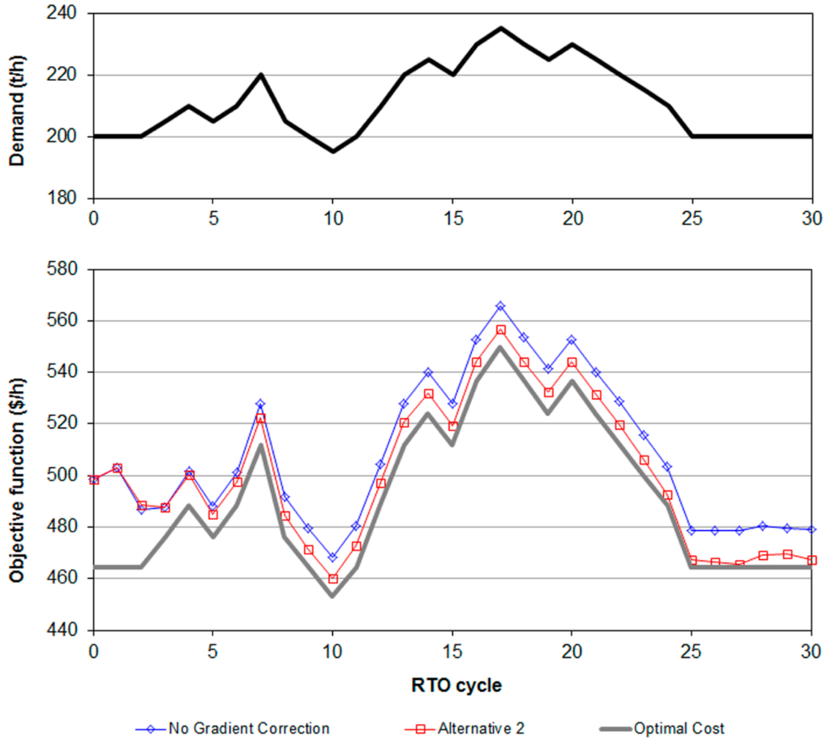

Figure 11. Case Study II, scenario 2. Objective function for variable steam demand and constant power price.

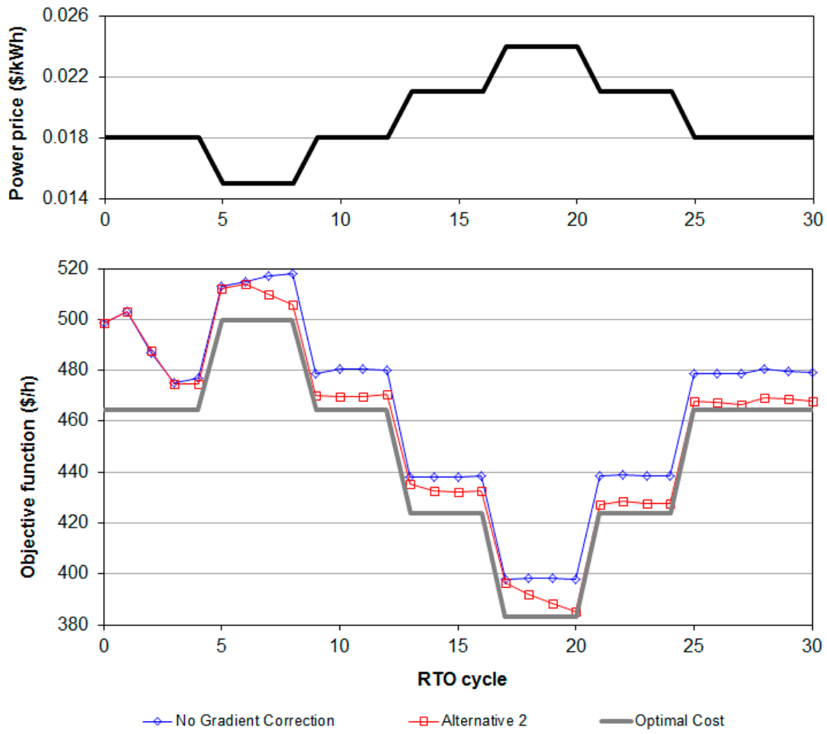

Figure 12. Case Study II, scenario 3. Objective function for constant steam demand and variable power price.

dx.doi.org/10.1021/ie303621j | *Ind. Eng. Chem. Res.* 2013, 52, 16795−16810

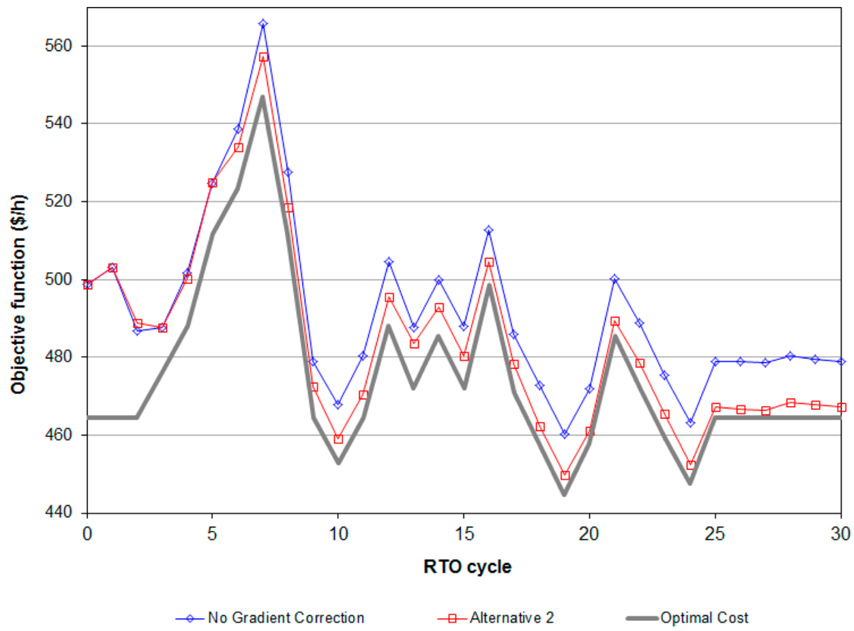

Figure 13. Case Study II, scenario 4. Objective function for variable steam demand and variable power price.

$$
\begin{aligned}
&\min\ Q_1 + Q_2 - c_{W,k}W_{TG} \\
&\text{s.t.} \\
&\mathbf{h}(\mathbf{y}, \mathbf{u}, \eta) = 0 \\
&\eta_j = p_m(\mathbf{x}_j) + \gamma_{jk}(\mathbf{x}_j) \\
&F_{35} = d_k \\
&\| F_{cb1} - F_{cb1,k} \| \le 5 \\
&\| F_{cb2} - F_{cb2,k} \| \le 5 \\
&\mathbf{u}^L \le \mathbf{u} \le \mathbf{u}^U \\
&\mathbf{y}^L \le \mathbf{y} \le \mathbf{y}^U
\end{aligned} \tag{37}
$$

where **h** is the whole model, p_m is the original model's performance equation, and γ_{jk} is the performance equation adaptation function in cycle k. $\mathbf{y}^L$ and $\mathbf{y}^U$ are lower and upper bounds for **y**. $c_{W,k}$ and d_k are the power price and steam demand in cycle k, respectively.

Figure 10 shows the resulting cost evolution for 30 RTO cycles for scenario 1 together with the constant power price and steam demand profiles. Noise-free cases are also shown for this scenario in order to illustrate the convergence of the RTO approaches. It can be observed that the proposed strategy converges to the real plant optimal cost in the noise-free case, while a strategy without gradient correction (called hereafter NGC) converges to a different operating point, as expected. When measurements noise is present, the performance decreases in both approaches, but the proposed strategy shows a significant comparative improvement with respect to the other one.

Although the gradient correction neglects the dependence of TG stage 2 efficiency with extraction steam temperature (which is a function of the TG inlet flow and steam conditions), convergence to the vicinity of the real-plant optimum is achieved. This is because the effect of this variable is small. In fact, the impact of stage 1 extraction temperature on this efficiency is considered implicitly in the biasing term a_{4k}, and not in the gradient correction terms.

Figure 11 shows the results obtained for scenario 2, together with the variable steam demand profile along the RTO cycles. It is assumed that Gaussian noise is present in the measurements in all examined scenarios. The proposed strategy outperforms NGC in each of the RTO cycles (except in $k = 2$, i.e. when there is few data for adapting the performance equations) for the simulated demand profile and the selected tuning parameters (N, w^i).

Figure 12 shows the results obtained for scenario 3, together with the variable power price profile. The suggested strategy performs in a better way, except for cycle $k = 2$, that is, when there is few data for adapting the performance equations. This behavior can also be observed in scenario 1 (Figure 10), as prices and demand are the same for all scenarios for $0 \le k \le 4$. It can also be seen that after a change in the power price in cycle k, the cost in k + 1 is slightly farther from the optimum than the previous cost. This is due to the fact that the optimal inputs were calculated for the price corresponding to the previous cycle.

Figure 13 shows the results obtained for scenario 4, which combines variable demand and variable price profiles. Again, the suggested strategy performs better than the traditional one for all optimization cycles except for $k = 2$.

Table 9 shows the Extended Design Cost evaluation for all scenarios. The total deviation from the real plant optimum is evaluated for all RTO cycles and for the period elapsed from $k = 6$ to the end, when the system is considered to be correctly initialized (i.e., with enough data to obtain accurate values in the correction terms) and close to the convergence to the real plant optimum in the proposed adaptation strategy. The results show significant improvements by applying the suggested gradient correction approach. Comparison of noise-free and Gaussian noise cases for scenario 1 show that the performance of Alternative 2 is highly affected by noise. Indeed, when convergence has been reached ($k \ge 6$), the increase in EDC computed for the Gaussian noise case is almost 3 times larger than the noise-free case (125 vs 48.5). Relative EDC values are very similar for scenarios 2, 3, and 4, with improvements of around 50% if all cycles are considered, and 60−70% for $k \ge 6$.

Table 9. Extended Design Cost for all Studied Scenarios in Case Study II

scenarios	EDC from $k = 0$ to 30 absolute	EDC from $k = 0$ to 30 relative to NGC	EDC from $k = 6$ to 30 absolute	EDC from $k = 6$ to 30 relative to NGC
Scenario 1, Noise Free				
no gradient correction	397.1	100.0	301.1	100.0
Alternative 2	124.7	31.4	48.5	16.1
scenario 1				
no gradient correction	470.3	100.0	353.0	100.0
Alternative 2	243.0	51.7	131.4	37.2
scenario 2				
no gradient correction	494.6	100.0	362.1	100.0
Alternative 2	275.4	55.7	147.5	40.7
scenario 3				
no gradient correction	479.9	100.0	376.4	100.0
Alternative 2	227.4	47.4	127.9	34.0
scenario 4				
no gradient correction	491.2	100.0	338.7	100.0
Alternative 2	262.5	53.4	111.3	32.9

The results obtained in this case study indicate that the proposed approach can be implemented in processes having a high number of inputs due to the simplicity of the implementation and the better performance with respect to the traditional (no gradient correction) approach.

5. CONCLUSIONS AND FURTHER WORK

An alternative modifier adaptation approach was proposed for real-time optimization. It exploits the fact that, for certain types of processes, the plant-model mismatch can be concentrated in a set of performance or efficiency equations, while rigorous mass, energy and entropy balances are considered as exact. These performance equations are often a function of a set of variables that has fewer elements than the set of all process inputs; therefore, the number of RTO cycles that must provide data for the gradient estimation problem and for the additional dual control constraints is reduced with respect to those suggested in the literature.

The application of the proposed approach to systems presenting such structure, like heat and power systems, results in faster and more reliable adaptation, and consequently, in an improved real-time optimization performance in terms of operating cost and sensitivity to noise.

The approach was developed by considering the characteristics of heat and power systems. Nevertheless, it can be applied to other types of processes, in particular to those that involve networks of several pieces of equipment or unit operations, for example in a plant-wide real-time optimization scheme.

Two alternatives have been developed based on performance equations adaptation: (i) a numerical gradient estimation combined with an exponential filter, and (ii) a quadratic regression of past data. This regression allows adapting the equations with more terms (in addition to the bias and gradient correction terms of other approaches), expanding the region where the adaptation is valid.

The two case studies analyzed in this work show that the proposed approach converges to the real plant optimum, like other gradient correction strategies. Moreover, in the examined scenarios it outperforms other approaches. Case Study II, which deals with the heat and power system of a sugar and ethanol plant, shows that performance equations adaptation can be applied to real-time optimization of heat and power systems, and that it is a suitable alternative for systems with a large number of inputs.

The method does not imply necessarily the involvement of simplified models. On the contrary, detailed models based on first principles and physical properties are adequate to implement this approach, as the sources of structural mismatch can be detected and related to an empirical performance equation.

The application of the proposed strategy requires a previous modeling and identification effort, in order to build a model which separates the performance equations from the rest of the rigorous balances, and to detect which variables impact on each performance index in a significant way.

Performance equations adaptation does not exclude the use of strategies like ISOPE or the original modifier adaptation method. Moreover, the proposed approach can be combined with them or included as the gradient estimation strategy, while following the procedures and dual control criteria of those methods. The advantage in this case is that the model structure and the proper use of the data would reduce the gradient estimation problem, in comparison with the calculation of each gradient with respect to all process inputs, which does not make use of the knowledge of the process to be optimized.

Another alternative that can be studied is to update the performance equations with an independent parameter estimator (as done in certain control applications),[45] which can work at a higher frequency than the real-time optimizer.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The GAMS input files with the implemented RTO approach are available to the reader upon request to the authors by e-mail. The Supporting Information includes the equations used in the sugar and alcohol energy system model of subsection 4.2 as well as the bounds in the operating variables of this system. This material is available free of charge via the Internet at http://pubs.acs.org..

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: paguir@santafe-conicet.gov.ar. Tel.: +54-342-4535568.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support from CONICET and Soteica Latinoamérica S.A.

■ REFERENCES

(1) Souza, G. D.; Odloak, D.; Zanin, A. C. Real Time Optimization (RTO) with Model Predictive Control (MPC). In *10th International Symposium on Process Systems Engineering: Part A*; Rita Maria de Brito Alves, C. A. O. do N.; Evaristo Chalbaud Biscaia, J. Eds.; Computer Aided Chemical Engineering; Elsevier: Amsterdam, The Netherlands, 2009; Vol. 27, pp 1365−1370.

(2) Adetola, V.; Guay, M. Integration of Real-time Optimization and Model Predictive Control. *J. Process Control* **2010**, *20*, 125−133.

(3) Biegler, L. T. Technology Advances for Dynamic Real-Time Optimization. In *10th International Symposium on Process Systems Engineering: Part A*; Rita Maria de Brito Alves, C. A. O. do N.; Biscaia, E. C., Eds.; Computer Aided Chemical Engineering; Elsevier: Amsterdam, The Netherlands, 2009; Vol. 27, pp 1−6.

(4) Biegler, L. T.; Zavala, V. M. Large-scale Nonlinear Programming Using IPOPT: An Integrating Framework for Enterprise-wide Dynamic Optimization. *Comput. Chem. Eng.* **2009**, *33*, 575−582.

(5) Darby, M. L.; Nikolaou, M.; Jones, J.; Nicholson, D. RTO: An Overview and Assessment of Current Practice. *J. Process Control* **2011**, *21*, 874−884.
(6) Mansour, M.; Ellis, J. E. Methodology of On-line Optimisation Applied to a Chemical Reactor. *Appl. Math. Modell.* **2008**, *32*, 170−184.
(7) Chachuat, B.; Srinivasan, B.; Bonvin, D. Adaptation Strategies for Real-time Optimization. *Comput. Chem. Eng.* **2009**, *33*, 1557−1567.
(8) Chen, C. Y.; Joseph, B. On-line Optimization Using a Two-Phase Approach: An Application Study. *Ind. Eng. Chem. Res.* **1987**, *26*, 1924−1930.
(9) Marchetti, A.; Chachuat, B.; Bonvin, D. A Dual Modifier-Adaptation Approach for Real-Time Optimization. *J. Process Control* **2010**, *20*, 1027−1037.
(10) Srinivasan, B.; Biegler, L. T.; Bonvin, D. Tracking the Necessary Conditions of Optimality with Changing Set of Active Constraints Using a Barrier-penalty Function. *Comput. Chem. Eng.* **2008**, *32*, 572−579.
(11) Jäschke, J.; Skogestad, S. NCO Tracking and Self-Optimizing Control in the Context of Real-time Optimization. *J. Process Control* **2011**, *21*, 1407−1416.
(12) Zhang, Y.; Monder, D.; Forbes, J. F. Real-Time Optimization Under Parametric Uncertainty: A Probability Constrained Approach. *J. Process Control* **2002**, *12*, 373−389.
(13) Zhang, Y.; Nadler, D.; Forbes, J. F. Results Analysis for Trust Constrained Real-Time Optimization. *J. Process Control* **2001**, *11*, 329−341.
(14) Srinivasan, B.; Bonvin, D. Convergence Analysis of Iterative Identification and Optimization Schemes. *Am. Control Conf.* **2003**, 1956−1961.
(15) Yip, W. S.; Marlin, T. E. Multiple Data Sets for Model Updating in Real-time Operations Optimization. *Comput. Chem. Eng.* **2002**, *26*, 1345−1362.
(16) Marchetti, A.; Chachuat, B.; Bonvin, D. Modifier-Adaptation Methodology for Real-Time Optimization. *Ind. Eng. Chem. Res.* **2009**, *48*, 6022−6033.
(17) Forbes, J. F.; Marlin, T. E. Design Cost: A Systematic Approach to Technology Selection for Model-Based Real-Time Optimization Systems. *Comput. Chem. Eng.* **1996**, *20*, 717−734.
(18) Roberts, P. D. Coping with Model-Reality Differences in Industrial Process Optimisation—A Review of Integrated System Optimisation and Parameter Estimation (ISOPE). *Comput. Ind.* **1995**, *26*, 281−290.
(19) Tatjewski, P. Iterative Optimizing Set-Point Control—The Basic Principle Redesigned. *Proceedings of the 15th IFAC World Congress, Barcelona, Spain*, 2002.
(20) Narasimhan, S.; Jordache, C. *Data Reconciliation & Gross Error Detection: An Intelligent Use of Process Data*; Gulf Professional Publishing: Houston, TX, 1999.
(21) Mansour, M.; Ellis, J. E. Comparison of Methods for Estimating Real Process Derivatives in On-line Optimization. *Appl. Math. Model.* **2003**, *27*, 275−291.
(22) Srinivasan, B.; François, G.; Bonvin, D. Comparison of Gradient Estimation Methods for Real-time Optimization. In *21st European Symposium on Computer Aided Process Engineering*; Pistikopoulos, E.N., M. C. G., Kokossis, A. C., Eds.; Computer Aided Chemical Engineering; Elsevier: The Netherlands, 2011; Vol. *29*, pp 607−611.
(23) Marchetti, A. G.; Basualdo, M. A New Dual Modifier-Adaptation Approach for Real-Time Optimization with Inaccurate Models. In *22nd European Symposium on Computer Aided Process Engineering*; Bogle, I. D. L., Fairweather, M., Eds.; Computer Aided Chemical Engineering; Elsevier: The Netherlands, 2012; Vol. *30*, pp. 942−946.
(24) Marchetti, A. G. A New Dual Modifier-adaptation Approach for Iterative Process Optimization with Inaccurate Models. *Comput. Chem. Eng.* **2013**, -.
(25) Rodger, E.; Chachuat, B. Design Methodology of Modifier Adaptation for On-Line Optimization of Uncertain Processes. In *Proceedings of 18th IFAC World Congress*; Milano, Italy, **2011**.
(26) Bunin, G. A.; François, G. Exploiting Local Quasiconvexity for Gradient Estimation in Modifier-Adaptation Schemes. In *2012 American Control Conference*; Montreal, Canada, June 27−29, **2012**.
(27) Bunin, G. A.; François, G.; Bonvin, D. From Discrete Measurements to Bounded Gradient Estimates: A Look at Some Regularizing Structures. *Ind. Eng. Chem. Res.* **2013**, *52*, 12500−12513.
(28) Serralunga, F.; Mussati, M. C.; Aguirre, P. A. An Alternative Real-time Optimization Algorithm with Modifier Adaptation: Application to Heat and Power Systems. In *22nd European Symposium on Computer Aided Process Engineering*; Bogle, I. D. L., Fairweather, M., Eds.; Computer Aided Chemical Engineering; Elsevier: The Netherlands, 2012; Vol. *30*, pp 367− 371.
(29) Serralunga, F.; Mussati, M. C.; Aguirre, P. A. Real-time Optimization of Energy Systems in Sugar and Ethanol Facilities: A Modifier Adaptation Approach. In *11th International Symposium on Process Systems Engineering*; Karimi, I. A., Srinivasan, R., Eds.; Computer Aided Chemical Engineering; Elsevier: The Netherlands, 2012; Vol. *31*, pp. 375− 379.
(30) Romagnoli, J.; Sanchez, M. C. *Data Processing and Reconciliation for Chemical Process Operations*; Process Systems Engineering; Academic Press: San Diego, CA, 1999; Vol. *2*.
(31) Forbes, J. F.; Marlin, T. E. Model Accuracy for Economic Optimizing Controllers: The Bias Update Case. *Ind. Eng. Chem. Res.* **1994**, *33*, 1919−1929.
(32) Brdys, M. A.; Tatjewski, P. *Iterative Algorithms for Multilayer Optimization Control*; 1st ed.; Imperial College Press: London, 2005.
(33) Yip, W. S.; Marlin, T. E. Designing Plant Experiments for Real-time Optimization Systems. *Control Eng. Pract.* **2003**, *11*, 837−845.
(34) Cao, S.; Rhinehart, R. R. An Efficient Method for On-line Identification of Steady State. *J. Process Control* **1995**, *5*, 363−374.
(35) Bhat, S. A.; Saraf, D. N. Steady-State Identification, Gross Error Detection, and Data Reconciliation for Industrial Process Units. *Ind. Eng. Chem. Res.* **2004**, *43*, 4323−4336.
(36) Roux, G. A. C. L.; Santoro, B. F.; Sotelo, F. F.; Teissier, M.; Joulia, X. Improving Steady-State Identification. In *18th European Symposium on Computer Aided Process Engineering*; Braunschweig, B., Joulia, X., Eds.; Computer Aided Chemical Engineering; Elsevier: Amsterdam, The Netherlands, 2008; Vol. *25*, pp 459−464.
(37) Yip, W. S.; Marlin, T. E. The Effect of Model Fidelity on Real-time Optimization Performance. *Comput. Chem. Eng.* **2004**, *28*, 267−280.
(38) Zhang, Y.; Forbes, J. F. Extended Design Cost: a Performance Criterion for Real-time Optimization Systems. *Comput. Chem. Eng.* **2000**, *24*, 1829−1841.
(39) *GAMS Language Guide*, release 2.25, version 92; GAMS Development Corporation: Washington D.C., 1997.
(40) Drud, A. *CONOPT 3 Solver Manual*; ARKI Consulting and Development A/S: Bagsvaerd, Denmark.
(41) Irvine Jr., T. F.; Liley, P. E. *Steam and Gas Tables with Computer Equations*; Academic Press: San Diego, CA, 1984.
(42) Camargo, C. A.; Ushima, A. H.; Ribeiro, A. M. M.; Souza, M. E. P.; dos Santos, N. F. *Manual de Recomendações-Conservação de Energia Na Indústria Do Açúcar e Álcool*; IPT: São Paulo, 1990.
(43) Hugot, E. *Handbook of Cane Sugar Engineering*; 3rd ed.; Elsevier Science Publishers B.V: Amsterdam, The Netherlands, 1986.
(44) Golato, M. A.; Franck Colombres, F. J.; Aso, G.; Correa, C. A.; Paz, D. Metodología de Cálculo de La Eficiencia Térmica de Generadores de Vapor (Methodology to Calculate Thermal Efficiency of Steam Boilers). *Rev. Ind. Agríc. Tucumán* **2008**, *85*, 17−31.
(45) Adetola, V.; Guay, M. Parameter Convergence in Adaptive Extremum-seeking Control. *Automatica* **2007**, *43*, 105−110.