Assessing the reliability of different real-time optimization methodologies

Article *in* The Canadian Journal of Chemical Engineering · December 2015

DOI: 10.1002/cjce.22402

CITATIONS
6

READS
152

4 authors:

Diego F. Mendoza
Universidad Autónoma del Caribe
15 PUBLICATIONS **61** CITATIONS
SEE PROFILE

Jose Eduardo A. Graciano
University of São Paulo
21 PUBLICATIONS **107** CITATIONS
SEE PROFILE

Fábio S. Liporace
Petróleo Brasileiro S.A.
32 PUBLICATIONS **171** CITATIONS
SEE PROFILE

Galo Carrillo Le Roux
University of São Paulo
88 PUBLICATIONS **703** CITATIONS
SEE PROFILE

Assessing the Reliability of Different Real-Time Optimization Methodologies

Diego Fernando Mendoza,[1] José Eduardo Alves Graciano,[2] Fabio dos Santos Liporace[3] and Galo Antonio Carrillo Le Roux[2]*

1. Universidad Autónoma del Caribe, Departamento de Ingeniería Mecánica, Barranquilla, Colombia
2. Universidade de São Paulo, Escola Politécnica, Departamento de Engenharia Química, São Paulo, SP-Brazil
3. CENPES/PETROBRAS S. A., Rio de Janeiro, Brazil

There is not a consensus about the benefits of implementing Real-Time Optimization (RTO) technologies to increase the profit of process plants. A lack of experimental and theoretical works which evaluate the scope and limitations of different RTO approaches makes it more difficult to have a sensible opinion about this topic. Most works available in the open literature that study different RTO approaches use few (often one) operation conditions to draw general conclusions about the virtues of a particular methodology. In the present work, we compare the performance of the classical two-step method with more recently proposed derivative-based methods (modifier adaptation, Integrated System Optimization Parameter Estimation (ISOPE), and an algorithm based on the Sufficient Conditions of Feasibility and Optimality (SCFO)) under different measurement noise, model mismatch, and disturbance using a Monte Carlo methodology. The results show that the classical RTO method can be reasonably reliable if provided with a model flexible enough to mimic the local process topology, a parameter estimation method suitable for handling measurement noise characteristics, and a method to improve the sample information quality. Implementing a derivative-based RTO method, in cases of evident model mismatch, should be considered only if the gap between the predicted and the real optimum is large enough and the level of measurement noise is low.

Keywords: real-time optimization (RTO), parameter estimation, modifier adaptation, model parameter adaptation

INTRODUCTION

The overall control system of a production process is a complex mechanism that should be segmented to obtain simpler tasks to implement. These tasks can be regrouped in a hierarchical structure, following a logical order, namely a functional or temporal order. The functional form deals with ensuring process safety, profitability, and product quality while the temporal form is applied in cases where differences between fast and slow variables are significant.[1]

The present contribution focuses on the functional hierarchical decomposition control (Figure 1, adapted from Darby et al.[2]), especially in the optimization layer (represented by the real-time optimization (RTO) layer in this chart). The RTO is inserted in the functional hierarchical control structure with the objective of providing ideal economic targets for the model predictive control (MPC) layer, which is expected to maintain the process under control at its maximum value of economic objective function.

The classical and more natural way to design the RTO layer is by using a first principles steady-state model to describe the process behaviour and to optimize an economical objective function subject to this phenomenological model. This strategy gained prominence in the late 1980s, when circumstances (namely equation-oriented modelling environments, computer process capability, and large-scale sparse matrix solvers) allowed the application of this kind of RTO.[2]

The basic idea of the "classical RTO method" (also called model parameter adaptation, MPA) is to update some key parameters of the steady-state model to reduce plant-model mismatch, using plant measurements,[3] and then to optimize the plant using the updated model. Despite using a high-fidelity plant model, this does not guarantee the absence of structural plant-model mismatch. Furthermore, measurement noise and incomplete plant information are important sources of uncertainty in the updated parameters, increasing the plant-model mismatch, and consequently leading to suboptimal operation points.

In 1985, Biegler and colleagues[4] discussed the use of simplified models to optimize complex models (which is the main idea behind the classical RTO method). They found that the plant-model mismatch derived from the simplified model may cause problems, since the mathematical optimum of the simplified model is likely to disagree with the plant optimum. Further, they showed that, to be adequate, a model must have the same Karush-Kuhn-Tucker (KKT) point as the real plant.

Forbes et al.[5] introduced the concept of model adequacy for the classical RTO method. They developed a procedure to determine whether a model is sufficiently flexible to represent a more complex model through a suitable choice of adjustable parameters. Nonetheless, the uncertainty in the parameters due to identifiability problems may mislead the classical RTO algorithm, since this method relies on parameter estimation to reduce plant-model mismatch.[6] Therefore, the classical RTO method cannot guarantee convergence to the true optimum under a large structural plant-model mismatch and small excitation in the operation conditions.[7]

* Author to whom correspondence may be addressed.
E-mail address: galoroux@usp.br
Can. J. Chem. Eng. 94:485–497, 2016

DOI 10.1002/cjce.22402
Published online 1 February 2016 in Wiley Online Library (wileyonlinelibrary.com).

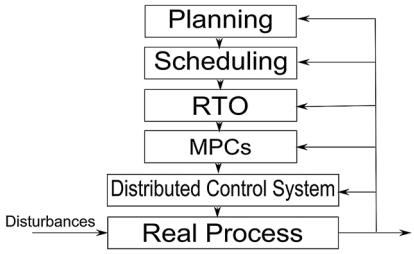

Figure 1. Process control hierarchy.

Other methods have been developed to supposedly make the RTO algorithm able to converge to the true plant optimum despite structural plant-model mismatch. The first, proposed by Roberts,[8] is a modification of the classical RTO method called 'integrated system optimization and parameter estimation' (ISOPE). In this methodology the parameter estimation and optimization steps are integrated, resulting in a modified economical objective function for the optimization step that is able to handle the structural mismatch problem, in cases when plant derivative can be calculated accurately.

The second method, the Modifier Adaptation method (MA),[7] differs from classical RTO in the way that plant information is used, since the measurements are employed to fulfill the necessary first-order optimality conditions (NOC) of the plant (using the so-called modifiers) without updating the parameters of the model. The MA scheme is able to calculate the plant optimum in the presence of plant-model mismatch provided that an accurate plant gradient is available, which is presently the main limitation for industrial applications.

Bunin et al.[9] proposed a method to tackle the plant-model mismatch problem called 'sufficient conditions for feasibility and optimality' (SCFO). This method combines the concepts of descent half-space and quadratic upper bound to derive sufficient conditions to guarantee the improvement of the plant objective function at each iteration, as well as the concepts of approximately active constraints and Lipschitz continuity to ensure constraint feasibility at each step. Although this method has a solid mathematical background to achieve its goals,[10] some of its assumptions are very difficult to meet in practice, such as knowledge of global Lipschitz constants, global quadratic upper bounds, and the exact value of the restrictions at the current iteration.

Bunin et al.[10] created an extension of the SCFO method for practical implementation. They proposed the use of a feasible region for the plant gradient to guarantee a descent region. The algorithm works within a region where the worst case ensures a decrease in the plant objective function without violating the constraints. However, Bunin and colleagues[10] state that it is unclear if applying SCFO is beneficial, since the SCFO algorithm may affect the convergence speed, especially when the RTO target is accurate (provided by the MPA for instance).

Due to the uncertainties of each RTO approach, there is no general consensus about the reliability of the different RTO methods for increasing the profit of a process plant.[2] Therefore, in this contribution, a Monte Carlo methodology is applied to evaluate the performance of each strategy under the same process uncertainties: parameter plant-model mismatch, measurement noise, and disturbances in the unmeasured variables; using the benchmark Williams-Otto reactor as a case study.

MATERIALS AND METHODS

This section presents the main characteristics and information flow for the different RTO methodologies compared in this paper. A detailed description of these algorithms can be found in the research articles cited in each section.

MPA Method

The structure of the classical RTO algorithm is presented in Figure 2. The RTO cycle starts with the steady-state detection module, responsible for analyzing the process measurements and deciding, based on statistical criteria, whether the plant is in steady state. Then, the stationary point goes through data reconciliation and gross error detection stage. Next, the screened information is used in the parameter estimation module to update the model parameters. Finally, the updated model is employed to find a new operation point that hopefully maximizes the plant's profit, which will be used as a set point by the process control layer.

The basic statement of the optimization module can be written as:

$$\begin{aligned} u^* = \min_u \quad & \varphi(u,y) \\ \text{such that} \quad & y = F_p(u) \\ & g(u,y) \le 0 \end{aligned} \tag{1}$$

where φ is the process (economic) performance index, y is the plant output, $F_p(u)$ is the plant map, and $g(u,y)$ are the process constraints. In model-based RTO (MPA) the plant outputs are estimated from a mathematical model $F(u,\theta)$, locally fitted by the parameters θ in the parameter estimation module.

$$\begin{aligned} \min_u \quad & \varphi(u,\widehat{y}) \\ \text{such that} \quad & \widehat{y} = F(u,\theta) \\ & g(u,\widehat{y}) \le 0 \end{aligned} \tag{2}$$

The MPA method has common vulnerabilities, namely a lack of process information, plant-model mismatch, and numerical optimization issues.[11] However, it is the most used online optimization method by the industry.[2]

ISOPE Method

One of the difficulties with the optimization problem stated in Equation (2) is the mismatch between the model and the real plant. The ISOPE method was developed to handle the structural plant-model mismatch,[1] complementing the measurements used in the MPA method with plant derivative information to reduce the offset created by the structural mismatch. ISOPE still retains the parameter estimation and economic optimization steps used by

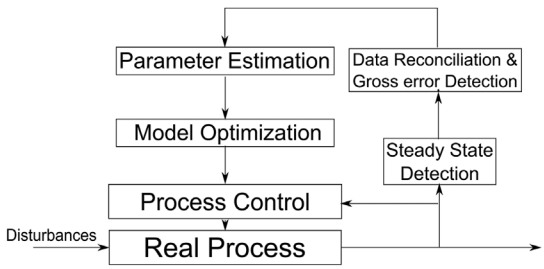

Figure 2. Classical RTO structure.

the MPA. However, ISOPE optimizes a modified economic function, adding a term coming from the parameter estimation step that allows a first-order correction.

ISOPE derivation starts by reformulating the RTO problem (Equation (2)), adding a penalty term (the so-called regularization term) to the economic performance index:

$$\begin{aligned} \min_{u,\theta} \quad & \phi(u, F(u,\theta)) + \rho\|u - v\|^2 \\ \text{such that} \quad & F(u,\theta) = F_p(u) \\ & g(v) \le 0 \\ & u = v \end{aligned} \tag{3}$$

where ρ is the regularization parameter and v are additional variables that allow Equation (3) to have essentially the same solution as the problem stated in Equation (1). The Lagrange function of the optimization problem given in Equation (3) is:

$$\begin{aligned} L(u,\theta,v,\xi,\mu,\lambda) = \phi(u,\theta) + \rho\|u - v\| + \xi^T(F(u,\theta) - F_p(u)) \\ +\mu^T g(v) + \lambda^T(u - v) \end{aligned} \tag{4}$$

where ξ, μ, and λ are Lagrange multipliers. The first-order optimality conditions applied to the Lagrange function are as follows:

$$2\rho(u - v) + \lambda + [\nabla_u F(u,\theta) - \nabla_u F_p(u)]^T \xi = 0 \tag{5a}$$

$$\nabla_v \phi(v,\theta) - 2\rho(u - v) - \lambda + \nabla_v g(v)^T \mu = 0 \tag{5b}$$

$$\nabla_\theta \phi(v,\theta) + \nabla_\theta F_p(u,\theta)^T \xi = 0 \tag{5c}$$

$$u - v = 0 \tag{5d}$$

$$F(u,\theta) - F_p(u) = 0 \tag{5e}$$

$$g(v) \le 0,\ \mu \ge 0,\ \mu^T g(v) = 0 \tag{5f}$$

The multipliers ξ and λ can be calculated from Equations (5a, 5c, 5d) as follows:

$$\xi = -[\nabla_\theta F(u,\theta) \nabla_\theta F(u,\theta)^T]^{-1} \nabla_\theta F(u,\theta) \nabla_\theta \phi(v,\theta) \tag{6}$$

$$\lambda = [\nabla_u F(u,\theta) - \nabla_u F_p(u)]^T \nabla_y \phi(u, F(u,\theta)) \tag{7}$$

Finally, the optimization problem solved in the ISOPE method is the modified model-based optimization problem:

$$\begin{aligned} \min_{v} \quad & \phi(u,\theta) - \lambda(u,\theta)^T v + \rho\|u - v\|^2 \\ \text{such that} \quad & y = F(u,\theta) \\ & g(v) \le 0 \end{aligned} \tag{8}$$

where $\lambda(u, \theta)$ is the multiplier given in Equation (7). This new optimization problem has the same optimality conditions as Equation (3). A comprehensive description of this formulation is given by Brdys and Tatjewski.[1] The basic ISOPE algorithm is shown in Figure 3.

ISOPE was derived assuming that the model is able to perfectly match plant outputs by updating model parameters

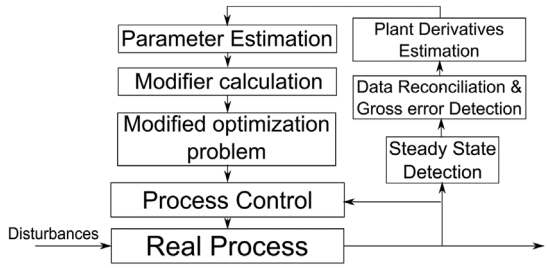

Figure 3. ISOPE structure.

(point parametric condition[8]) and that an accurate plant derivative is available. These crucial assumptions ensure that the solution obtained by the modified model-based optimization problem converges to the true plant optimum.[12] The main challenge in this method is the requirement of plant derivatives used to compute the modifiers' values, since the estimation of these quantities is considerably affected by measurement noise.[13]

MA Method

The idea behind the modifier adaptation (MA) method is to use measurements to correct the cost and constraint predictions between successive RTO iterations in such a way that the KKT point for the model coincides with the plant optimum.[6]

Given the real process model $F_p(u)$ and the RTO model $F(u)$, it is possible to construct a corrected model $F_c(u)$ similar to the real process model, Equation (9). The correction term, proposed in Equation (10), comes from a first-order Taylor series expansion of the discrepancy term around the current operation point (Equation (10)). The final corrected model is presented in Equation (11).

$$F_c(u) = F(u) + [F_p(u) - F(u)] \tag{9}$$

$$F_p(u) - F(u) = F_p(u^0) - F(u^0) + \underbrace{\left(\frac{\partial F_p}{\partial u}(u^0) - \frac{\partial F}{\partial u}(u^0)\right)}_{\lambda}^T (u - u^0) \tag{10}$$

$$F_c(u) = F(u) + \varepsilon + \lambda^T(u - u^0) \tag{11}$$

where ε and λ^T are the so-called modifiers, ε is the gap between the plant and predicted function values, and λ^T is the difference between the slopes, which is calculated as the difference between model and plant derivatives (see Equation (10)). A very useful graphical interpretation of these features is presented by Marchetti et al.[7]

The objective function and the constraints of the RTO problem are reformulated using this methodology. The problem is restated as follows:

$$\begin{aligned} \min \quad & \phi_c(u,\theta) = \phi_m(u,\theta) + \lambda_\phi^T(u - u_k) \\ \text{such that} \quad & G_c(u,\theta) = G_m(u,\theta) + \varepsilon + \lambda_G^T(u - u_k) \le 0 \end{aligned} \tag{12}$$

where the subscripts c and m are the corrected and original RTO models, respectively; ϕ is the economic objective function, and G is the set of inequality constraints.

The fundamental difference between the MA and ISOPE frameworks is how the modifiers are calculated and the parameter

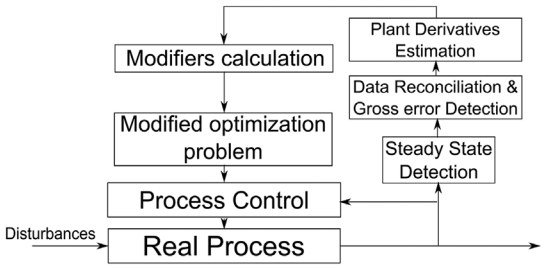

Figure 4. MA structure.

updating. In MA, the modifier is calculated from derivatives of the economic objective function with respect to inputs (u), while the ISOPE method uses the derivatives of output (y) with respect to inputs (u). In addition, parameters are updated during ISOPE iterations while MA uses a fixed parameter set during optimization; i.e. there is no parameter updating (see Figure 4).

SCFO Method

The SCFO method, initially proposed by Bunin et al.[9] and modified for practical implementation by Bunin et al.,[10] adapts nonlinear optimization theory to RTO problems. The method is designed to calculate the plant optimum without violating any "hard" constraint while improving the plant profit at each RTO iteration, executing a projection problem based on information of plant derivatives and topology. In other words, given a target (a possible future RTO point, predicted by any RTO algorithm; MPA for instance) the SCFO method implements a correction in this target, based on plant derivative information. The projection problem, given by Equation (13), minimizes the distance between the target (u^*_{k+1}) and the feasible point (u), subject to a bounded deviation ($\delta_{g,j}$) from the active constraints ($g_j(u_k) \geq -\varepsilon_j$) and an improvement in the objective function (δ_φ). These two restrictions work to maintain the solution found in the projection problem ($\bar{u}^*_{k+1}$) at the interior of the hard constraints region, given by $g_j(u_k) \geq -\varepsilon_j$, and to grant a profit improvement: $\nabla\varphi(u_k)^T(u-u_k) \leq -\delta_\varphi$. This behaviour is achieved within the region where the problem nonlinearities are closely approximated by the first-order local information (gradient information).

$$\begin{aligned} \bar{u}^*_{k+1} = \arg\min_{u} & (\|u - u^*_{k+1}\|(^2_2 \\ \text{such that} \quad & \nabla g_j(u_k)^T(u-u_k) \leq -\delta_{g,j} \forall j : g_j(u_k) \geq -\varepsilon_j \\ & \nabla\varphi(u_k)^T(u-u_k) \leq -\delta_\varphi \\ & u^L \leq u \leq u^U \end{aligned} \tag{13}$$

where the subscript k is the RTO iteration, the point u^*_{k+1} is the input target (calculated from the classical RTO approach in this work), $\bar{u}^*_{k+1}$ is the target projected into a feasible descent space, $\nabla\varphi$ and ∇g_j are, respectively, the plant derivative of the objective function and constraints with respect to the input variables, and δ are the minimal changes required in the projected direction.

The need for accurate real process derivatives limits the practical implementation of this algorithm. For this reason the authors modified the projection problem to work within a feasible region given by the derivative of the real process. This region can be obtained assuming a certain local structure for the function,[10] or in a less rigorous approach, it may be calculated by adding an uncertainty region around the estimated gradient. The modified projection problem, using the estimated gradient and the uncertainty region, is given by Equation (14).

$$\begin{aligned} \bar{u}^*_{k+1} = \arg\min_{u,S,s_\phi} & (\|u - u^*_{k+1}\|(^2_2 \\ \text{such that} \quad & \sum_{i=1}^{nu} s_{ji} \leq -\delta_{g,j} \\ & \frac{\partial \underline{g}_j}{\partial u_i}\Big|_{uk}(u_i - u_{k,i}) \leq s_{ji} \ \forall j : g_j(u_k) \geq -\varepsilon_j \\ & \frac{\partial \overline{g}_j}{\partial u_i}\Big|_{uk}(u_i - u_{k,i}) \leq s_{ji} \ \forall j : g_j(u_k) \geq -\varepsilon_j \\ & \sum_{i=1}^{nu} s_{\varphi,i} \leq -\delta_\varphi \\ & \frac{\partial \underline{\varphi}}{\partial u_i}\Big|_{uk} \ (u_i - u_{k,i}) \leq s_{\varphi,i} \\ & \frac{\partial \bar{\varphi}}{\partial u_i}\Big|_{uk} \ (u_i - u_{k,i}) \leq s_{\varphi,i} \end{aligned} \tag{14}$$

where s represents slack variables responsible for ensuring the choice of direction for the worst case; $\frac{\partial \underline{g}}{\partial u_i}$ and $\frac{\partial \overline{g}}{\partial u_i}$ are the lower and upper bounds of the constraint derivatives; and $\frac{\partial \underline{\varphi}}{\partial u_i}$ and $\frac{\partial \overline{\varphi}}{\partial u_i}$ are the lower and upper bounds for the objective function derivatives. The main structure of the algorithm is presented in Figure 5, where the target calculation corresponds to the MPA solution and the projection problem is performed by solving Equation (14).

Plant Derivative Estimation

The plant derivative is estimated from process measurements using Boryden's approximation formula:

$$B_k = B_{k-1} + \left[(y_k - y_{k-1}) - (B_{k-1}(u_k - u_{k-1})^T)\right] \cdot \frac{(u_k - u_{k-1})M}{((u_k - u_{k-1})M(u_k - u_{k-1})^T)} \tag{15}$$

where B is the matrix of estimated derivatives, u is the vector of input variables, y is the vector of outputs, and M is a scaling (diagonal) matrix.[14] The indices k and $k-1$ indicate the current and previous steady state points, respectively. In this work the Broyden method is preferred to methods such as finite differences (FD) or dynamic model identification (DMI) on the basis of practical applicability, since FD and DMI require large numbers of upsets or depend on the availability of dynamic plant information, which is difficult and costly to achieve in a real process plant.[12]

The dual approach, proposed by Rodger and Chachuat,[15] is implemented in MA and ISOPE algorithms to improve the plant derivatives estimated by the Broyden method, enforcing minimal perturbation in different directions (to get better information at

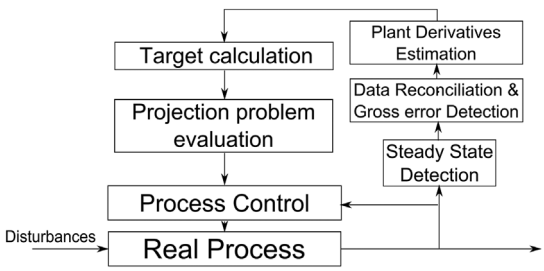

Figure 5. SCFO structure.

each step), and maximum step length (to avoid the "peak phenomenon," as discussed by Rodger[14]). This approach is implemented by a set of constraints (Equations (16, 17)), which determine two possible regions for the solution search.

$$-w_k(u-u_k)+\sqrt{w_k{}^T B^{-1} w_k}\le 0 \qquad (u-u_k)^T\Gamma(u-u_k)\le 1 \tag{16}$$

$$+w_k(u-u_k)-\sqrt{w_k{}^T B^{-1} w_k}\le 0 \qquad (u-u_k)^T\Gamma(u-u_k)\le 1 \tag{17}$$

where u_k and u are the actual and future RTO points, w_k is the unitary vector orthogonal to the last two RTO points, B is the parameter matrix for the minimum upset, and Γ is the parameter matrix of maximum step length. In this work the values of B and Γ are $diag([500.4])$ and $diag([400.15])$ respectively.

In the dual approach, the economic optimization problem is divided into two problems, one using Equation (16) and other with Equation (17). Then, these problems are solved in parallel and the best result is implemented. A graphical interpretation can be found in Rodger.[16]

CASE STUDY: THE WILLIAMS AND OTTO REACTOR

The Williams and Otto CSTR (continuous stirred tank reactor) is a well-known case study used for the development and comparison of RTO strategies by several authors.[7,18,19] This process is illustrated in Figure 6. The reactor is fed with Fa and Fb (pure streams of components A and B, respectively). These components react, producing an intermediate component C, which reacts with another B molecule to produce the desired products P and E. There is a side reaction between components C and P, producing byproduct G which has zero commercial value and is waste. The reactions and their kinetics are given in Equation (18).

$$\begin{aligned} A+B &\rightarrow C & k1 &= \eta p1\ \exp(-Ea1/(T_R+273.15)) \\ B+C &\rightarrow P+E & k2 &= \eta p2\ \exp(-Ea2/(T_R+273.15)) \\ P+C &\rightarrow G & k3 &= \eta p3\ \exp(-Ea3/(T_R+273.15)) \end{aligned} \tag{18}$$

where Ea is the activation energy and ηp is the pre-exponential factor. These values are given in Table 1.

The process is modelled at steady state by the mass balances, using the reactor temperature (T_R) and flow rate of B (Fb) as controlled variables, and keeping the flow rate of reactant A (Fa) and the mass holdup (W) at 1.8275 kg/s and 2105 kg, respectively. The economic objective is to maximize the profit given by

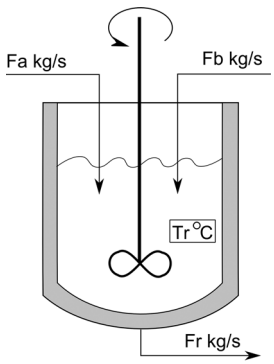

Figure 6. Williams and Otto reactor.

Table 1. Plant parameters for experimental design

Parameters	Region 1	Region 2	Region 3	Region 4
Ea 1	6666.7	6666.7	6666.7	6666.7
Ea 2	8333.3	8444.3	8333.3	8333.3
Ea 3	11 111	11 101	11 111	11 111
Fa (kg/s)	1.8275	1.8275	1.8275	2.2000
$\eta\ p\ 1$		1.6599×10^{6}		
$\eta\ p\ 2$		7.2117×10^{8}		
$\eta\ p\ 3$		2.6745×10^{12}		

Equation (19).

$$\phi = 1143.38X_PF_R + 25.92X_EF_R - 76.23F_A - 114.34F_B \tag{19}$$

where X_P and X_E are the mass fractions of P and E in the reactor outlet stream (F_R).

To analyze the performance of each RTO methodology under structural plant-model mismatch, a simpler kinetic (approximated model) is proposed to describe the original system.

$$\begin{aligned} A+2B &\rightarrow P+E & k1 &= \eta 1\ \exp(-\nu 1/(T_R+273.15)) \\ A+B+P &\rightarrow G & k2 &= \eta 2\ \exp(-\nu 2/(T_R+273.15)) \end{aligned} \tag{20}$$

where ν is the activation energy and η is the pre-exponential factor. Both of these values are estimated by the parameter estimation module.

Parameter Estimation Module

In our analysis we consider a perfect and an approximate model: Equations (18, 20), respectively. In both cases, all kinetic parameters (pre-exponential factors and activation energies) are estimated using the product compositions Xp, Xe, and the Fa flow rate as measurements. This is because it is very unlikely that a real plant has online measurements of all product compositions (online composition measurements are very expensive). The objective function is a simple least squares with a diagonal matrix equal to 1. Furthermore, the last 3 historical points in the RTO path are used in the parameter estimation module, as suggested and implemented by Pfaff[19] to increase the amount of information.

EXPERIMENTAL DESIGN

The present section aims to design a comprehensive experiment to evaluate the performance of the algorithms over a wide range of different situations. For this reason, we consider 5 process characteristics that can modify the evaluation of an RTO algorithm. The first two problems, measurement noise and initial parameter values, are related to the parameter estimation module. Both may deteriorate the parameter estimation and change the RTO path, resulting in different performances for the same RTO algorithm. The influence of these random variables is assessed through Monte Carlo simulations, where 500 RTO trials are carried out using different initial parameter values and measurement noise, sampled following uniform and normal distributions (see Appendix A).

The third and fourth problems are disturbances presented in measured and unmeasured variables. These process characteristics are simulated in the plant by changing the parameter listed in Table 1, which results in the 4 regions depicted by Figure 7. The first and second disturbance steps correspond to changes in the values of the kinetic constants (unmeasured disturbances), which

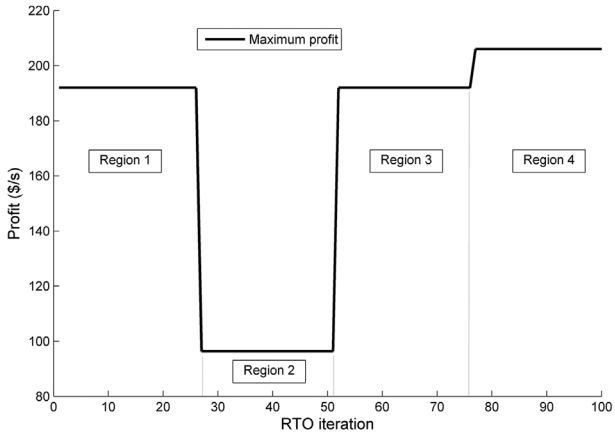

Figure 7. Optimum profile with respect to disturbances.

may be associated, for example, with a decrease in catalyst performance, while the last one is due to a sudden increase in the feed flow rate Fa (measured disturbance).

The fifth problem affecting RTO performance is the structural mismatch between plant and model. Two cases are considered in the experimental design: the perfect model when both plant and model are represented by Equation (18), and the approximate model when the plant is given by Equation (18) and the model by Equation (20). In both cases the plant is simulated according to the parameters described in Table 1, while the model parameters are estimated by the parameter estimation module.

In short, the Monte Carlo simulations were performed for each RTO algorithm using measurement noise levels of 0 % and 0.5 %, with the perfect and approximated models (total = 16 experiments). In each MC experiment, 500 RTO trials were conducted (with 100 iterations in each RTO) starting from the same nominal point. Three disturbance scenarios are assumed at iterations 25, 50, and 75, creating 4 different regions (see Figure 7).

The performances of the RTO methodologies are compared through three statistics computed from the profit error: root mean square error, average profit loss (absolute value), and frequency of obtaining profit loss < 1 % in the last 5 RTO iterations of each region (%). In this work the profit error is defined as the difference between the instantaneous profit using the set points calculated by the RTO and the true optimum in each region defined in Figure 7.

Appendix B shows the performance of each RTO method under perfect conditions. These experiments are important to illustrate that the algorithms work well under ideal conditions and their implementation is correct.

RESULTS

Results for the Perfect Model

Figure 8 presents the results of the four RTO methods using noise-free measurements and the perfect model. In this figure the frequency distribution of the economic objective function is denoted by the colour scale.

The behaviour shown in Figure 8 and the dispersion metric presented in Table 2 indicate that the MPA method presents the lowest scattering profile, since this method is not influenced by the uncertainty in the derivative caused by the Broyden approximation that affects all the derivative-based

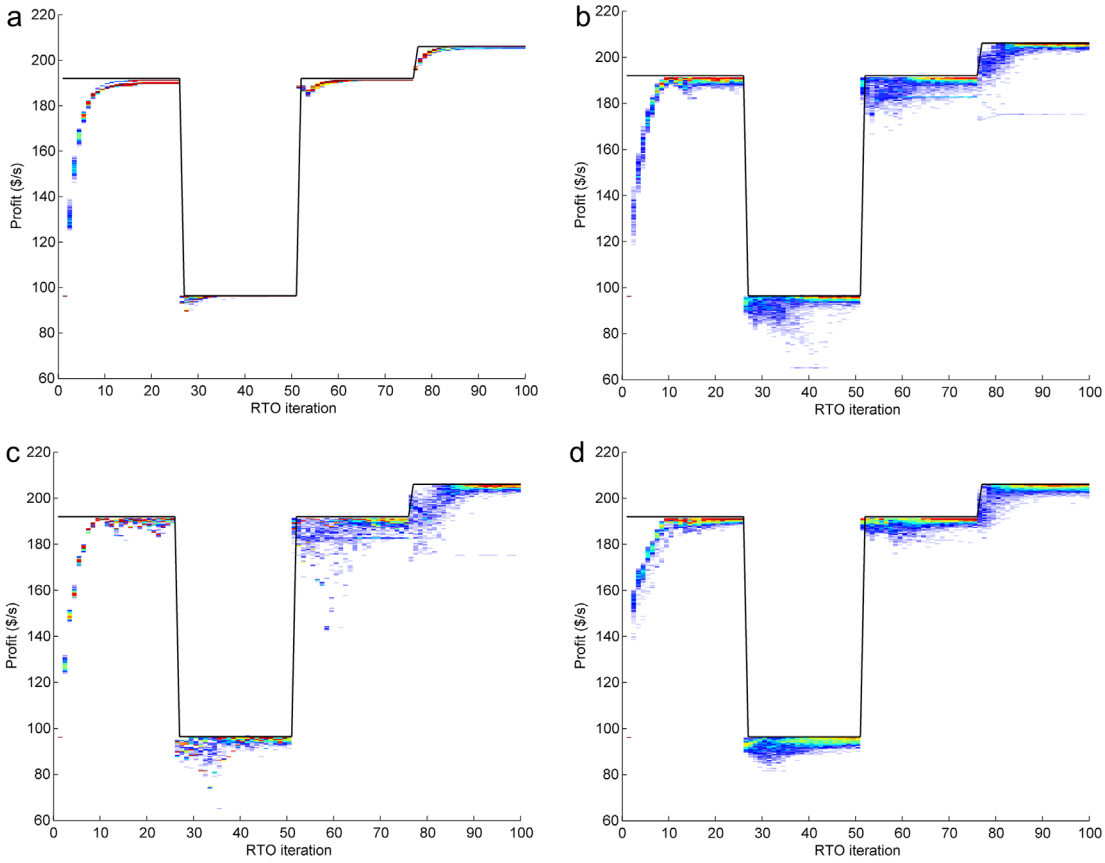

Figure 8. MC experiments using noise-free measurements and the perfect model: (a) MPA, (b) MA, (c) ISOPE, and (d) SCFO.

Table 2. Root mean square error for MC experiments using noise-free measurements and the perfect model

Method	Region 1	Region 2	Region 3	Region 4
MPA	8.15	4.59	8.36	9.13
MA	8.44	7.35	10.61	11.58
ISOPE	8.15	6.31	10.89	12.24
SCFO	8.68	5.07	8.81	9.83

methods tested. In this category the SCFO exhibits the lowest dispersion.

The frequency of obtaining the optimum profit (within 1 %) in the last 5 RTO iterations is shown in Table 3. The MPA methodology follows the optimum plant operation path along the different plant upsets. In this case, the information quality as well as the model structure allow the parameter estimation routine to identify a topology converging to the "true" optimum in a few RTO cycles ($\sim$15 cycles on average), even after plant disturbances.

Regarding profit loss during the RTO, the path followed by MPA is the most cost-effective (on average 3.04 U.S. dollars (USD)/s), since it presents lower profit loss than any derivative-based method tested. SCFO shows the best result for the first region (Table 4) because it has the largest first step among the methods; however, its average profit loss is 4.64 USD/s.

The results for MC simulations with the perfect model and measurement noise are shown in Figure 9. A comparison of the statistics of the RTO performance using noisy measurements (Tables 5–7) with previous noise-free measurements (Tables 2–4) indicates a lower performance of the RTO methodologies due to measurement noise.

Comparing the RTO methods with and without measurement noise shows that noise always increases profit loss (Tables 4, 7). As in the noise-free case, the MPA has the lowest profit loss on average. This loss is even lower than for the derivative-based methods using perfect measurements.

Results for the Approximated Model

These experiments assess the behaviour of the RTO methodologies under structural plant-model mismatch. The results obtained in the Monte Carlo simulations using the approximate model and noise-free measurements are depicted in Figure 10. The structural mismatch does not allow the convergence of the MPA method to the true optimum in all regions, which is confirmed by the low frequency of obtaining profit losses < 1 % (Table 9). In contrast, the derivative-based methods are able to handle the structural mismatch in all tested regions, as can be observed in Figure 10 and Table 9.

Table 3. Frequency of achieving < 1 % profit loss in the last 5 RTO iterations of each region. MC experiments using noise-free measurements and the perfect model

Method	Region 1	Region 2	Region 3	Region 4
MPA	100	100	100	100
MA	72.16	43.36	60.48	84.80
ISOPE	55.00	39.24	56.16	79.24
SCFO	86.48	28.72	76.84	67.08

Table 4. Average profit loss for MC experiments using noise-free measurements and the perfect model

Method	Region 1 (USD/s)	Region 2 (USD/s)	Region 3 (USD/s)	Region 4 (USD/s)
MPA	8.50	0.78	1.33	1.56
MA	8.90	5.44	6.30	5.57
ISOPE	9.33	4.23	6.88	7.88
SCFO	7.59	3.78	3.06	4.13

The scattering presented by MPA and SCFO tends to be alike in every region. This dispersion is lower than for the MA and ISOPE under similar conditions (Table 8). However, the lower average profit loss corresponds to the path followed by MA in the first region, MPA in the second region, and SCFO in the third and fourth regions (Table 10). On average SCFO presents the best economic results.

Figure 11 shows the outcome of the MC simulations for the case using the approximate model and measurement noise of 0.5 %. The results indicate an increasing scattering of the RTO path compared to the case with the same structural model mismatch and noise-free measurements (Tables 8, 11). The derivative-based methods are also more sensitive to noise than the MPA method. Indeed, the profit loss increases by $\sim$45 % for the MA method and 36 % for the ISOPE and SCFO methods, in comparison with a decrease of $\sim$18 % for MPA under the same conditions (Tables 10, 13).

The MPA, as observed in the MC simulations using the approximate model and noise-free measurements, presents offsets between the predicted and actual optimum and in the first three regions. This behaviour reduces the frequency of obtaining profit losses < 1 % in these regions using MPA when compared to the derivative-based methods (Table 12). However, this method outperforms the derivative-based approaches in region 4, where the offset is not present.

A comparison of the approximate model with its noise-free counterpart reveals a constant increase in the scattering in each region, similar to that observed in the perfect model simulation with and without noise.

DISCUSSION

The Monte Carlo simulations using a structural perfect model show that MPA performs better than derivative-based methods in the presence of disturbances and measurement noise. This result is due in part to the experimental conditions fulfilling the assumptions made for the parameter estimation method, since the Least Squares estimator is able to handle noisy data with characteristics such as measurements being independently and normally distributed with zero mean.[20] On the other hand, poor performance of the Least Squares estimator, and poor overall performance of the MPA method, should be expected in cases where measurements are corrupted with gross errors or are non-independently distributed. In this case redescending or appropriate likelihood estimators should be used.[21]

The approximate model used in the second analysis fulfills the adequacy criterion of Forbes et al.[5] since there is at least one set of parameters that predicts the same optimal point as the plant (at least for regions 1 and 3). However, Marchetti[6] points out that for this set of parameters, the model outputs differ from the plant output, becoming unlikely to converge on the "ideal" set of parameters through a parameter estimation and optimization approach (MPA method).

The results obtained in the MC simulations using the approximate model and noise-free measurements suggest that MPA is unable to converge to the true optimum. As a consequence, the

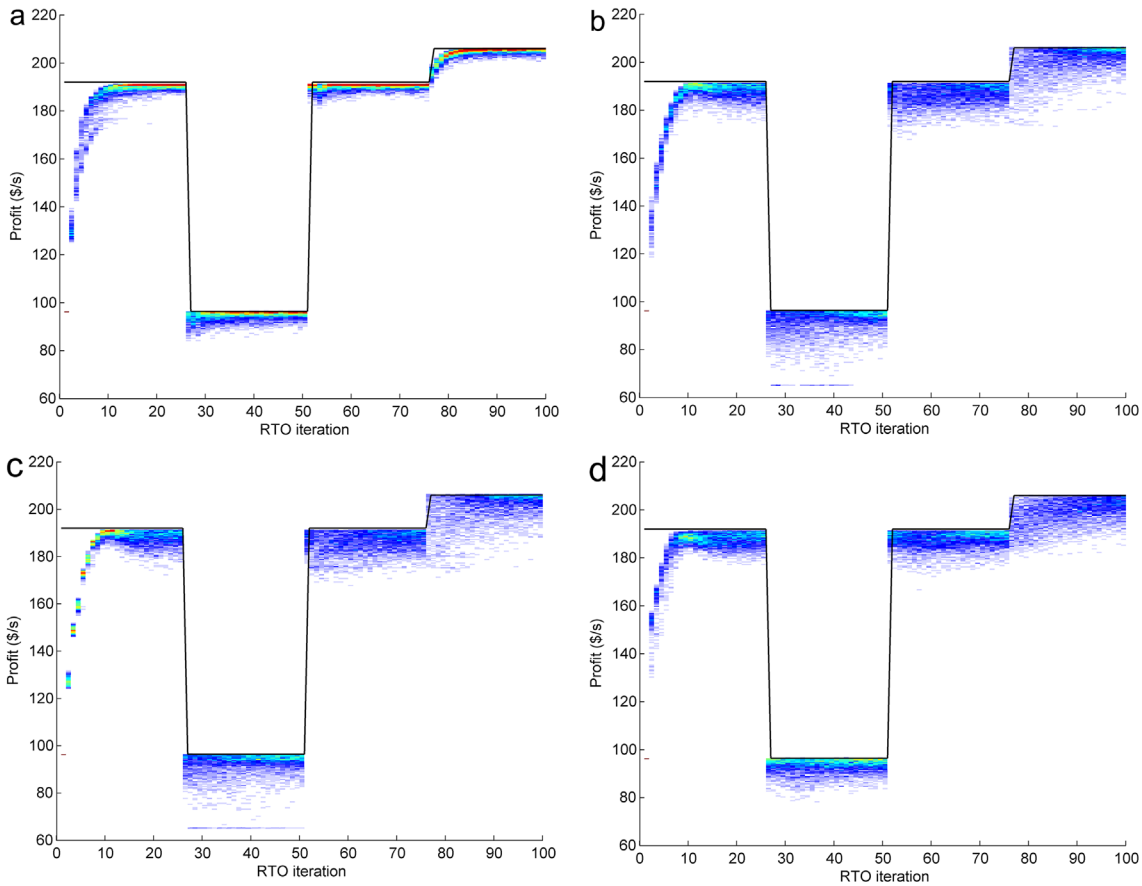

Figure 9. MC experiments using noisy measurements (0.5 %) and the perfect model: (a) MPA, (b) MA, (c) ISOPE, and (d) SCFO.

derivative-based RTO methods (SCFO and MA) have better economic performance. On the other hand, MPA shows better economic performance than derivative-based methods in cases where optimization runs under measurement noise and plant-model mismatch. Also in this case, in the fourth region, MPA does not present an offset because the model is able to simulate the process topology.

The improvement observed in the MPA method is related to the parameter estimation module, since the upsets introduced by the measurement noise are sufficient to increase the sample distribution quality, obtaining better information. Similar results could be achieved by introducing the dual methodology to the MPA approach. Figure 12 shows a comparison between MPA with and without the dual methodology. It is important to note that the dual approach decreases the bias observed in the RTO path followed by the MPA method, consequently decreasing the profit loss ~28 % compared to MPA without the use of the dual approach.

The derivative-based methods present better results than the MPA method only in the case of model mismatch and noise-free measurements. In particular, the SCFO method presents the best economic performance among the derivative-based methods, followed by the MA and then the ISOPE method. This fact indicates that SCFO is better designed to handle the uncertainty introduced by the Broyden estimation. The comparison between MA and ISOPE shows slightly better performance for MA, indicating that the parameter estimation module is not necessary for this approach type.

Table 5. Root mean square error for MC experiments using noisy measurements (0.5 %) and the perfect model

Method	Region 1	Region 2	Region 3	Region 4
MPA	9.95	5.19	8.59	9.34
MA	9.40	8.61	10.81	12.54
ISOPE	8.95	9.11	11.14	13.33
SCFO	9.97	5.09	6.36	9.91

Table 6. Frequency of achieving < 1 % profit loss in the last 5 RTO iterations of each region. MC experiments using noisy measurements (0.5 %) and the perfect model

Method	Region 1	Region 2	Region 3	Region 4
MPA	75.44	44.96	79.08	80.84
MA	28.08	11.20	23.36	32.12
ISOPE	25.72	12.20	22.08	28.56
SCFO	25.32	18.60	28.96	17.96

Table 7. Average profit loss for MC experiments using noisy measurements (0.5 %) and the perfect model

Method	Region 1 (USD/s)	Region 2 (USD/s)	Region 3 (USD/s)	Region 4 (USD/s)
MPA	9.72	2.96	1.98	2.87
MA	11.72	8.42	8.14	10.14
ISOPE	11.60	8.08	8.63	11.41
SCFO	9.97	5.09	6.36	9.91

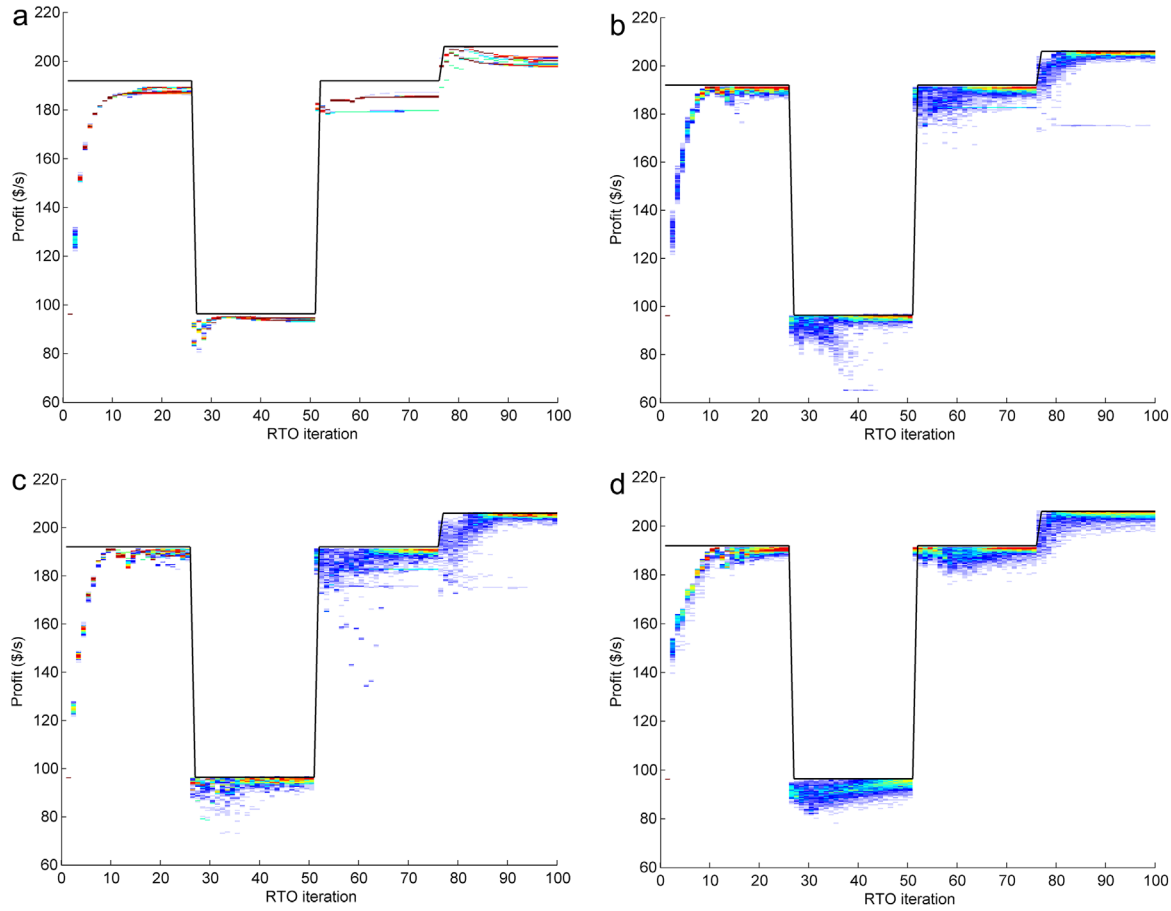

Figure 10. MC experiments using noise-free measurements and approximate model: (a) MPA, (b) MA, (c) ISOPE, and (d) SCFO.

In general, the results show that derivative-based methods are more sensitive to measurement noise than the classical MPA method. In fact, the conditions used in the numerical experiments are especially difficult for the derivative-based methods, since a random sample of parameters (in the first iteration (Appendix A)) is likely to produce significant plant-model mismatch from the first RTO iteration. Therefore, the approximation given by Broyden is prone to produce a misleading search direction from the beginning of the RTO iterations. Another factor affecting the performance of Broyden's method is the set of "drastic" changes in process topology induced by the sudden disturbances added to the experiments during the RTO iterations. For instance, note that the dispersions obtained for MA and ISOPE in regions 1 and 3 are very different, although the plant parameters are the same in both regions (Figures 8, 10). In the first region the algorithm starts from a unique point where the derivatives are estimated by the model, at a corner point where the derivative module is large and points approximately to the optimum solution; region 3 starts from several different points (end points of the second region). Compared to region 1, the starting points of region 3 are placed in a flatter area, which decreases the Broyden method's derivative prediction.

The influence of measurement noise on the gradient prediction is analyzed by simply calculating a sequence of gradient approximations under different levels of noise using Broyden's method. The quality of the gradient estimate is evaluated using the

Table 8. Root mean square error for MC experiments using noise-free measurements and the approximate model

Method	Region 1	Region 2	Region 3	Region 4
MPA	8.00	4.58	8.22	9.22
MA	8.32	7.61	10.74	11.47
ISOPE	8.09	5.90	11.25	12.22
SCFO	8.45	5.51	8.99	9.66

Table 9. Frequency of achieving <1 % profit loss in the last 5 RTO iterations of each region. MC experiments using noise-free measurements and the approximate model

Method	Region 1	Region 2	Region 3	Region 4
MPA	1.12	0.00	0.00	0.00
MA	73.40	42.28	62.72	81.44
ISOPE	57.80	37.76	54.28	74.56
SCFO	73.52	13.04	64.28	65.24

Table 10. Average profit loss for MC experiments using noise-free measurements and the approximate model

Method	Region 1 (USD/s)	Region 2 (USD/s)	Region 3 (USD/s)	Region 4 (USD/s)
MPA	10.90	3.15	7.64	5.97
MA	8.92	5.61	6.44	5.31
ISOPE	9.84	3.77	7.55	7.72
SCFO	9.28	5.77	3.74	3.94

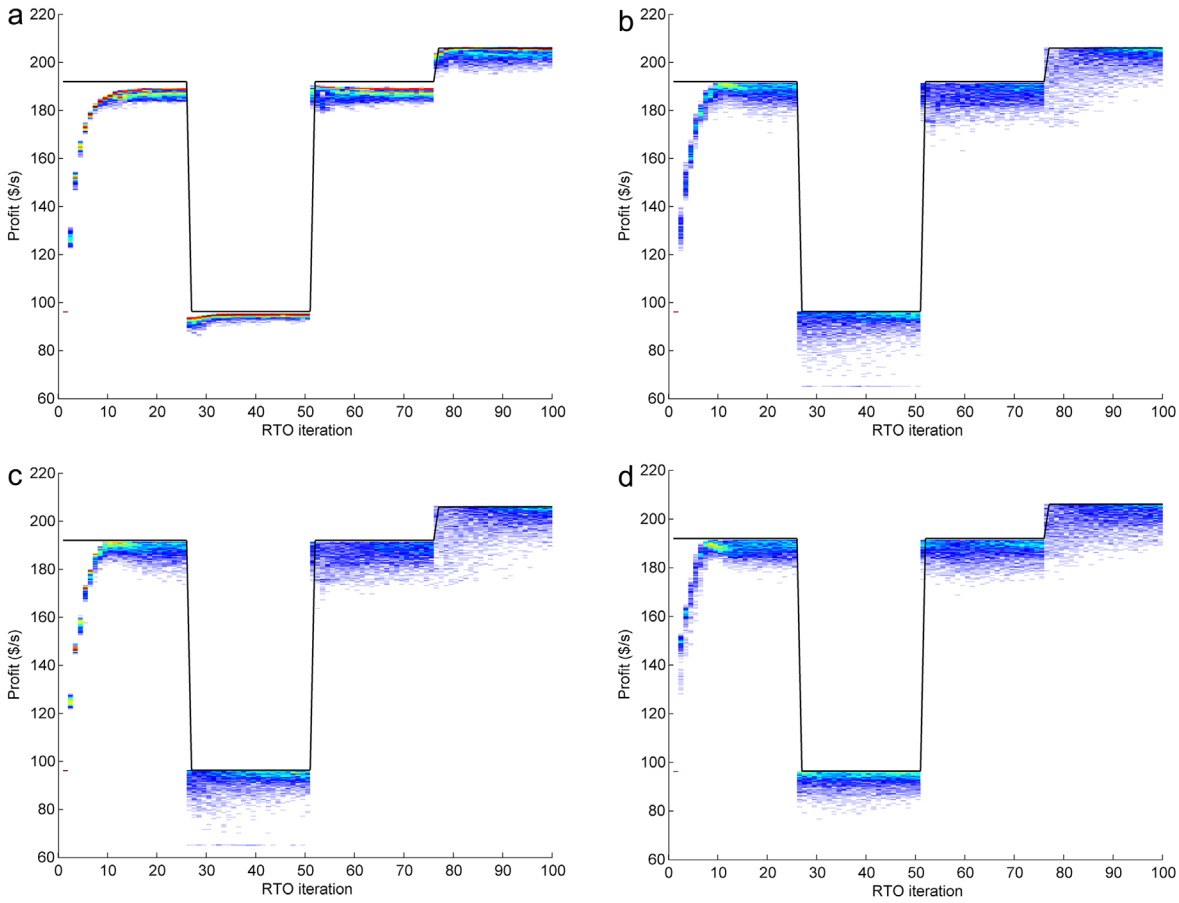

Figure 11. MC experiments using noisy measurements (0.5 %) and approximate model: (a) MPA, (b) MA, (c) ISOPE, and (d) SCFO.

angle and the norm ratio between the predicted and true gradients. Figure 13 shows the influence of measurement noise on these two characteristics for a sequence of four RTO iterations, starting from the same point and converging toward the optimum.

For the noise-free case, the maximum average angle between the predicted plant gradient using Broyden's method and the true gradient is $< 2^{\circ}$, meaning that the Broyden method approximation is close to the true local direction of maximum function increase.

Table 11. Root mean square error for MC experiments using noisy measurements (0.5 %) and the approximate model

Method	Region 1	Region 2	Region 3	Region 4
MPA	8.10	4.63	8.53	9.45
MA	9.25	8.79	10.77	12.46
ISOPE	8.88	8.43	11.27	13.40
SCFO	9.36	6.60	9.88	11.18

Table 12. Frequency of achieving < 1 % profit loss in the last 5 RTO iterations of each region. MC experiments using noisy measurements (0.5 %) and the approximate model

Method	Region 1	Region 2	Region 3	Region 4
MPA	1.44	0.00	1.92	49.44
MA	27.48	13.72	23.60	31.28
ISOPE	27.32	11.84	20.44	28.96
SCFO	25.20	15.96	26.56	25.16

Moreover, for this case the predicted derivative norms are similar to the true one, indicated by a norm ratio close to 1 in Figure 13b. Under these conditions (noise-free and good initial guess), Broyden's approximation shows a reasonable estimation of the plant gradient.

The increment in the angle between the true and estimated gradients confirms the high sensitivity of Broyden's method to measurement noise (Figures 13a, c, e, g). Additionally, the increase of scattering at each RTO step indicates high sensitivity to information degradation (i.e. measurement noise). This behaviour can be better appreciated in cases with 0.5 % and 1 % noise, where the norm ratios of the derivatives are highly scattered, affecting the length of steps taken toward the optimum by the derivative-based RTO routine.

The high sensitivity to measurement noise of Broyden's gradient estimation (even for measurement noises as low as 0.5 %) is a serious issue for its implementation in real life problems, and the reason why several alternative approaches have

Table 13. Average profit loss for MC experiments using noisy measurements (0.5 %) and the approximate model

Method	Region 1 (USD/s)	Region 2 (USD/s)	Region 3 (USD/s)	Region 4 (USD/s)
MPA	11.47	2.67	4.84	3.75
MA	11.66	8.11	8.15	10.25
ISOPE	11.98	7.49	8.74	11.10
SCFO	9.88	5.51	6.64	8.95

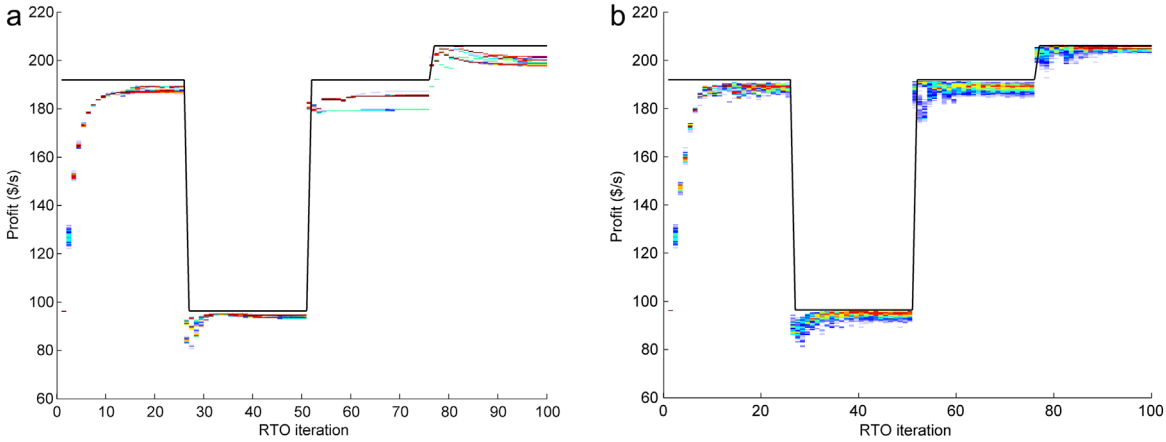

Figure 12. Comparison between MPA with approximate model and free measurement noise: (a) MPA without dual approach; (b) MPA with dual approach.

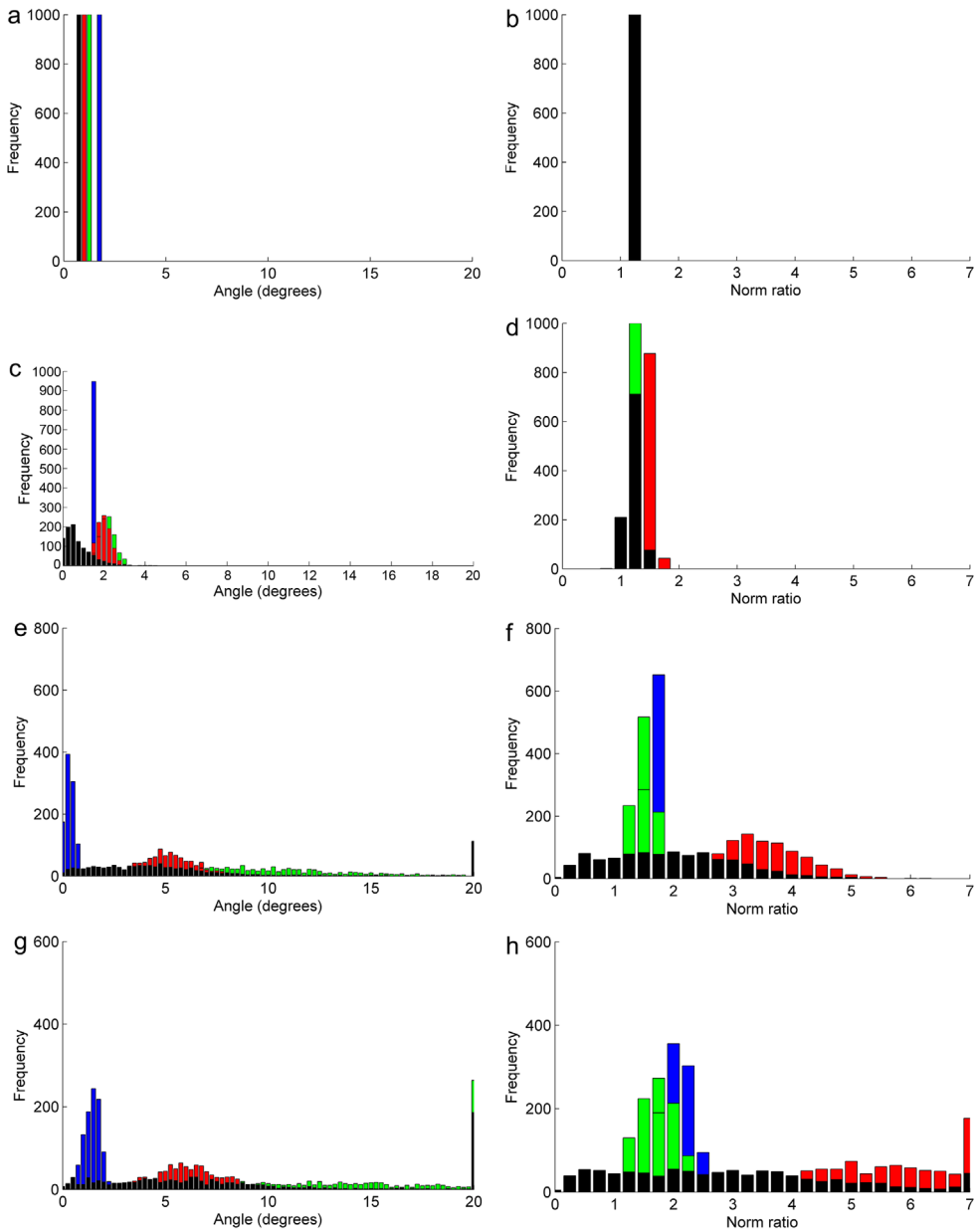

Figure 13. Derivative analysis: (a, c, e, g) angle distribution between true and predicted gradient; (b, d, f, h) norm ratio distribution between true and predicted gradient. Point 1 = blue, point 2 = green, point 3 = red, point 4 = black. Noise level: 0 % (a, b), 0.05 % (c, d), 0.5 % (e, f), 1 % (g, h).

been devised to improve plant gradient estimation.[12] An interesting alternative is to take advantage of transient periods to get more information from the plant.[22] Some techniques are known to use this information in the identification of linear or nonlinear dynamic models used to predict plant gradients.[12] These techniques can be implemented without affecting the basic (derivative-based) steady-state RTO scheme, which will probably improve the performance of these methods.

CONCLUSIONS

The performance of the classical RTO method (MPA) and three derivative-based methods (MA, ISOPE, SCFO) were compared under the influences of measurement noise, plant-model mismatch, and process disturbances. The main findings in this analysis are as follows.

The MPA method provides the best performance for the perfect model case among all the methods compared. This method shows the lowest profit loss in the scenarios studied. The key point in this method is to provide a model flexible enough to generate the local process topology, and a parameter estimation method capable of minimizing the overfitting caused by a lack of practical identifiability.[21] On the other hand, for the approximate model experiments, MPA presents better results than the derivative-based methods when there is measurement noise or a specific method (e.g. dual methodology) improving the quality of sample information.

The plant derivative predicted by Broyden's method is highly sensitive to measurement noise and to initial estimates of the derivatives. The SCFO method is more suitable for handling this kind of uncertainty, presenting the best economic results. Comparing MA and ISOPE shows that the parameter estimation module is less important than the derivative quality for this kind of approach.

This paper started with an inquiry on the reliability of the RTO methods. The results show that a classical RTO method can be reasonably reliable if provided with a model flexible enough to mimic process topology, a parameter estimation method suitable for handling process noise characteristics, gross errors, and lack of model identifiability, and a method to generate process upsets to improve the sample information quality (dual methodology).

The implementation of a derivative-based RTO method in cases of evident model mismatch should be considered only if the gap between the predicted and the real optimum is large enough (which is impossible to know a priori) and the level of measurement noise is low. Furthermore, another aspect to be considered is the need to implement better techniques for estimating the plant gradient using transient information.

There are aspects such as optimization method and steady-state detection that are not addressed in this contribution, which are important for the overall performance of an RTO implementation and should be the topic of new contributions. However, the aspects addressed can be considered the distinctive marks of the RTO methodologies, since-steady state detection and optimization can be essentially the same for all these methodologies.

APPENDIX A

Upper and lower bounds for the initial parameter values used in the first RTO iteration of each Monte Carlo simulation. The set of parameters is uniformly raffled.

Table A1. Parameter bounds used in the parameter estimation

Parameter bounds used in perfect model simulations

	A 1	Ea 1	A 2	Ea 2	A 3	Ea 3
Upper	2.7554×10^{12}	13 333	5.2000×10^{17}	16 667	3.6099×10^{18}	22 216
Lower	1.2884×10^{3}	3333	2.6853×10^{4}	4167	4.3589×10^{4}	5554

Parameter bounds used in approximate model simulations

	η 1	ν 1	η 2	ν 2
Upper	1.7183×10^{9}	9289	1.3291×10^{15}	14 304
Lower	6.6979×10^{6}	6866	1.5076×10^{11}	10 573

The measurement noise is simulated by function *randn* MATLAB with standard deviation = 0 and 0.5 % error:

$$z = m + error \ .m \ .\text{randn}() \tag{A1}$$

where z is the measurement contaminated with noise, m is the measurement without noise, and error = 0 or 0.005.

APPENDIX B

This section presents the behaviour of the optimization routines implemented using the approximate model and the perfect model. All RTOs start with the same parameter values, using noise-free measurements and ideal derivatives.

According to Figure B1 all RTO methods achieve the true optimum when accurate measurements are available, even in the presence of model mismatch. The only exception is the MPA,

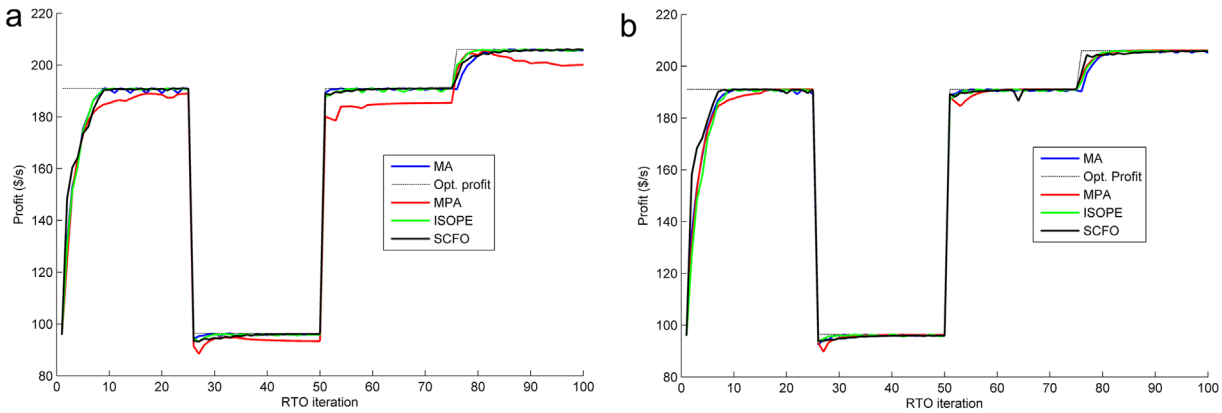

Figure B1. Algorithm results for ideal conditions: (a) RTO path using approximated model, and (b) RTO path using perfect model.

which presents offset in the case using the approximate model. These results show the basic behaviour of the algorithms assessed in this paper.

REFERENCES

[1] M. A. Brdys, P. Tatjewski, *Iterative Algorithms for Multilayer Optimization Control*, Imperial College Press, London **2005**.

[2] M. L. Darby, M. Nikolaou, J. Jones, D. Nicholson, *J. Process. Contr.* **2011**, *21*, 874.

[3] I. P. Miletic, T. E. Marlin, *Ind. Eng. Chem. Res.* **1998**, *37*, 3670.

[4] L. T. Biegler, I. E. Grossmann, A. W. Westerberg, *Comput. Chem. Eng.* **1985**, *9*, 201.

[5] J. F. Forbes, T. E. Marlin, J. F. MacGregor, *Comput. Chem. Eng.* **1994**, *18*, 497.

[6] A. Marchetti, *Modifier-Adaptation Methodology for Real-Time Optimization*, PhD thesis, École Polytechnique Fédérale de Lausanne, Lausanne **2009**.

[7] A. Marchetti, B. Chachuat, D. Bonvin, *Ind. Eng. Chem. Res.* **2009**, *48*, 6022.

[8] P. D. Roberts, *Int. J. Syst. Sci.* **1979**, *10*, 719.

[9] G. A. Bunin, G. François, D. Bonvin, "Sufficient Conditions for Feasibility and Optimality of Real-Time Optimization Schemes - I. Theoretical Foundations," *arXiv*, **2013**, accessed on 29 January 2015, http://arxiv.org/abs/1308.2620.

[10] G. A. Bunin, G. François, D. Bonvin, "Sufficient Conditions for Feasibility and Optimality of Real-Time Optimization Schemes - II. Implementation Issues," *arXiv*, **2013**, accessed on 29 January 2015, http://arxiv.org/abs/1308.2625.

[11] A. D. Quelhas, N. J. C. de Jesus, J. C. Pinto, *Can. J. Chem. Eng.* **2013**, *91*, 652.

[12] M. Mansour, J. E. Ellis, *Appl. Math. Model.* **2003**, *27*, 275.

[13] A. S. Lubansky, Y. L. Yeow, Y.-K. Leong, S. R. Wickramasinghe, B. Han, *AIChE J.* **2006**, *52*, 323.

[14] G. A. Bunin, G. François, D. Bonvin, *Ind. Eng. Chem. Res.* **2013**, *52*, 12500.

[15] E. A. Rodger, B. Chachuat, "Design Methodology of Modifier Adaptation for On-Line Optimization of Uncertain Processes," *18th IFAC World Congress*, International Federation of Automatic Control, Milano, 28 August – 2 September, **2011**.

[16] E. Rodger, *Dual Modifier Adaptation Methodology For the Online Optimization of Uncertain Processes*, MSc thesis, McMaster University, Hamilton, ON **2010**.

[17] J. F. Forbes, T. E. Marlin, *Comput. Chem. Eng.* **1996**, *20*, 717.

[18] Y. Zhang, D. Nadler, J. F. Forbes, *J. Process Contr.* **2001**, *11*, 329.

[19] G. C. Pfaff, *Generating Information for Real-Time Optimization*, MSc thesis, University of Alberta, Edmonton, AB **2001**.

[20] Y. Bard, *Nonlinear Parameter Estimation*, Academic Press, New York **1974**.

[21] J. E. Graciano, D. F. Mendoza, G. A. C. Le Roux, *Comput. Chem. Eng.* **2014**, *64*, 24.

[22] W. Bamberger, R. Isermann, *Automatica* **1978**, *14*, 223.

Manuscript received December 16, 2014; revised manuscript received June 18, 2015; accepted for publication June 24, 2015.