On-line implementation of nonlinear MPC: an experimental case study

Lino O. Santos[a], Paulo A.F.N.A. Afonso[a], José A.A.M. Castro[a], Nuno M.C. Oliveira[a], Lorenz T. Biegler[b,*]

[a] *Departamento de Engenharia Química, F.C.T., Universidade de Coimbra-Pólo II, Pinhal de Marrocos, 3030-290 Coimbra, Portugal*
[b] *Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh PA 15213, USA*

Abstract

A nonlinear model predictive control algorithm is implemented on-line to control the liquid level and temperature in a pilot plant CSTR, where an irreversible exothermic chemical reaction is simulated experimentally by steam injection. The dynamic behavior of the pilot plant reactor is represented using a mechanistic, first principle model and a comparison between off-line simulation and experimental data is presented. Several sources of model mismatch and unmeasured disturbances are identified that affect the quality of the model in representing the reactor dynamics. Despite these mismatches and disturbances, the closed loop system is able to track setpoint changes and reject disturbances quite well. In particular, the NMPC controller is demonstrated for different tuning parameters and under conditions of constraint saturation at unstable points. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Process control; Predictive control; Model-based control; Nonlinear models

1. Introduction

Model predictive control (MPC) with nonlinear process models (or NMPC) is being applied increasingly often in the chemical industry and many of these applications are viewed as essential for a number of process control problems. NMPC is needed especially for nonlinear, unsteady batch processes where a trajectory needs to be followed from the prediction of a nonlinear model. It is especially useful for processes operating at or near 'singular' points (e.g., input multiplicities and other sign changes in the process gains) that cannot be captured by linear controllers and where higher order information is needed. It is also necessary for processes with wide swings in operation, beyond the ranges of a local linearization. These characteristics are observed in many process problems including changeovers in continuous processes, tracking problems in startup and batch processes and the control of nonlinear reactors. For these systems, NMPC is a more natural and straightforward approach than gain scheduling or intermittent approximation by linear models. NMPC uses the nonlinear dynamic model to predict the effect of sequences of control steps on the controlled variables. The aim is to derive an optimal set of control steps, which will drive these output variables to the desired steady state setpoints, based on economic, safety, environmental or product considerations.

There are a number of NMPC studies that deal with experimental and industrial applications. For instance, in their survey Qin and Badgwell (2000) report 88 industrial NMPC applications. However, if one considers only those that incorporate the on-line solution of nonlinear dynamic optimization problems applied to first principle models, then their number dwindles to five. This clearly indicates the difficult nature of NMPC problems in this class. Moreover, Qin and Badgwell also mention the challenges and difficulties of dealing with the efficient solution of nonconvex optimization problems on-line, particularly with accurate, first principle process models. This is a key feature addressed in this experimental study.

*Corresponding author. Tel.: +1-412-268-2232.
E-mail addresses: lino@eq.uc.pt (L.O. Santos), paulo@estga.ua.pt (P.A.F.N.A. Afonso), eq1jmc@eq.uc.pt (J.A.A.M. Castro), nuno@eq.uc.pt (N.M.C. Oliveira), lb01@andrew.cmu.edu (L.T. Biegler).

This study considers the application of NMPC to an experimental pilot plant reactor apparatus. It involves a continuous stirred tank reactor (CSTR) with a capacity of 80 liters, fitted with a cooling jacket and a hydraulic stirring system. Here an exothermic zero order chemical reaction, A→B, is experimentally simulated in the vessel. This is done on-line by calculating the amount of live steam to inject into the vessel contents, in order to simulate the heat generated by the chemical reaction (Pinheiro & Kershenaum, 1999; Afonso, Oliveira, & Castro, 1996; Kershenbaum & Kittisupakorn, 1994). Since no reaction actually takes place, the system described in this study is more properly defined as a 'pseudo-reactor'. Nevertheless, it still features nonlinear dynamics and exhibits open loop instability under certain operating conditions.

There are a number of advantages to the particular experimental system on which our study was applied. First, the system is flexible, easily configured and safe. Second, it can handle exothermic reaction mechanisms of arbitrary kinetic complexity (in addition to the zero order reaction considered here), without the additional bother of handling hazardous chemicals. Moreover, as an experimental system (with steam valves, boilers, etc.) it incorporates unmeasured and unmodeled disturbances, model mismatches and many other real world limitations that often cannot be appreciated in a simulation study. Finally, the safety and relative simplicity of this system allows research to focus on the validation of advanced control and estimation algorithms under real world conditions, and to provide detailed case studies that are often not considered in more complex industrial systems.

This pilot plant has been the subject of several studies and was used as a benchmark to test control algorithms and other tools (Afonso et al., 1996; Afonso, Ferreira, & Castro, 1998; Afonso, 1998). In these studies the reactor dynamics were described by a reduced model structure: a global mass and energy balance to the reactor, representing the level and temperature evolution, and using experimental correlations to predict some of its parameters. This study involves a more detailed model of the reactor that includes the description of the jacket temperature dynamics and the evolution of the concentration of the reactant (Santos, Afonso, & Castro, 1998). The purpose of this work is to implement an NMPC algorithm on-line, using the formulation as described in Santos, Oliveira, and Biegler (1995) and in Santos (2001).

Section 2 presents a description of the operating conditions of the reactor, the strategy used to simulate the zero order chemical reaction and the reactor model development. Section 3 provides a brief description of the NMPC algorithm, and issues related to the on-line implementation. Section 4 provides details on NMPC performance as well as a comparison between experi-

mental and simulation results, while Section 5 concludes the paper.

2. Pilot plant reactor

As seen in Fig. 1, the reactor feed consists of two inlet water streams, with feed rates F_1 and F_2 and temperatures T_1 and T_2, while the outlet stream with flow rate F_3 flows by gravity through valve Vc5. The reactor features a jacket equipped with a spiral baffle and a hydraulic stirring system. The flow rate of the cooling fluid, F_j, with inlet temperature T_{j0} and outlet temperature T_{j2}, is determined by the aperture of the valve Vc3, λ_3. In order to provide a higher cooling capacity, the recirculating cooling water is purged continuously and fresh water is added. The speed of the agitator is controlled by manipulating valve Vc4. Both the reactor vessel and pipes are insulated to minimize heat losses. The reactor is equipped with standard equipment to measure and manipulate the process variables, along with a computer interface for both acquisition and control. A more detailed description of the pilot plant installation can be found in Afonso (1998).

2.1. Experimental reaction simulation

The zero order exothermic chemical reaction is experimentally simulated by injecting steam directly into the liquid inside the reactor. This strategy was also used by Kershenbaum and Kittisupakorn (1994). To simulate the heat released by the reaction, Afonso et al. (1996) obtained an experimental correlation that determines the power released by the injection of steam into
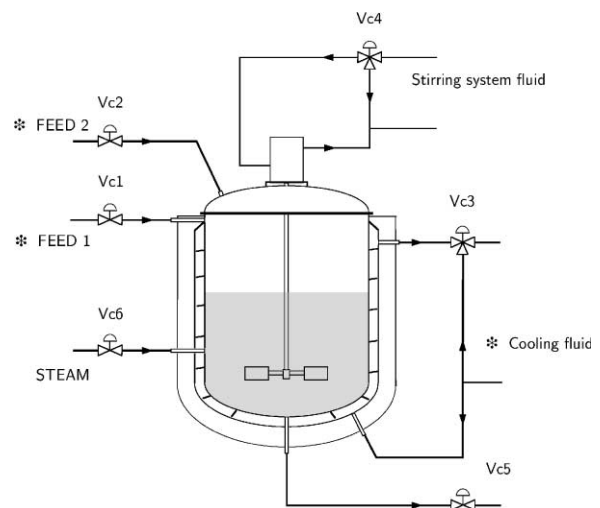

Fig. 1. Pilot plant reactor diagram. The fluid in the streams denoted by * is water.

the liquid inside the reactor, as a function of the aperture of valve Vc6

$$Q_{steam} = c e^{(d\lambda_6)},\tag{1}$$

where $c = 1203$ W and $d = 3.22$. Since the steam resources available are limited, this correlation is valid only for $0 \leqslant \lambda_6 \leqslant 0.9$. Setting Eq. (1) to the heat generation term from the reactor energy balance equation and solving for λ_6, leads to

$$\lambda_6 = \frac{1}{d}\left(\ln\frac{-r_A V(-\Delta H_r)}{c}\right),\tag{2}$$

where $(-\Delta H_r)$, V and r_A are the heat of reaction, the reactor liquid volume and the rate of consumption of reactant A ($-r_A = k_0 \exp\{-E_a/(RT_r)\}$), respectively. Although the kinetic model used in (2) is identical to the one used in the model for the NMPC calculations, there are several sources of uncertainty resulting from this experimental system (Afonso et al., 1996). For instance, the valve that controls the steam flow rate, Vc6, exhibits considerable hysteresis (see Graphic IV in Figs. 6 and 7) and the steam pressure is subject to continuous changes as the total load in the boiler oscillates. Here the steam pressure near the reactor is around 3 bar and the boiler pressure oscillates between 4.5 and 5 bar.

Thus, all of these operating difficulties can be viewed as uncertainty in the kinetic system. In principle, other reaction processes can be simulated using this type of experimental arrangement (Kershenbaum & Kittisupakorn, 1994). However, the zero order reaction simulation is simpler to implement because the rate of reaction does not depend on the reactant concentration. Moreover, the kinetics feature open loop unstable system behavior and this can severely affect the reactor temperature under certain operating conditions.

2.2. Pilot plant reactor model

The literature abounds with models that represent the dynamic behavior of continuous pilot plant reactors. Depending on the design, size and operating conditions, several modeling approaches can be adopted. Assuming perfect mixing with densities of the input and output streams identical to the liquid inside the reactor, the total reactor mass balance is given by

$$\frac{dV}{dt} = F_1 + F_2 - F_3,\tag{3}$$

where F_1 and F_2 are the inlet flows, F_3 is the outlet flow, and V is the reactor liquid volume. The volume is given by $V = V_0 + \pi r^2 h$, with $[V] = m^3$, $[h] = m$, $V_0 = 4.2 \times 10^{-3} m^3$, and the reactor radius is $r = 0.232$ m (Afonso et al., 1996). An experimental correlation determines the outlet flow F_3 as a function of the outlet valve aperture λ_5. This is used in the neighborhood of the nominal operating level $h = 0.30$ m,

which corresponds to a reactor liquid volume of 55 liters (Santos, 2000). The mass balance for reactant A is given by

$$\frac{dC_A}{dt} = \frac{F_1}{V}(C_{A1} - C_A) + \frac{F_2}{V}(C_{A2} - C_A) - k_0 e^{-E_n/(RT_r)}.\tag{4}$$

Assuming negligible heat generated by the agitator and identical heat capacities for the reactor liquid and feed streams, the dynamics of the reactor temperature can be described by

$$\frac{dT_r}{dt} = \frac{1}{\beta_r}(-Q_R + Q_G),\tag{5}$$

where Q_R is the rate of heat removal and Q_G is the rate of heat generated by the reaction, given by

$$Q_R = -\rho C_p F_1(T_1 - T_r) - \rho C_p F_2(T_2 - T_r) + UA(T_r - T_j),\tag{6}$$

$$Q_G = (-\Delta H_r)V k_0 e^{-E_n/(RT_r)}.\tag{7}$$

Here C_p is the specific heat capacity of the fluid, U is the overall heat transfer coefficient and A is the heat transfer area which is related to the liquid level and reactor radius r as: $A \simeq \pi r(r + 2h)$. Since the reaction is zero order, both the evolution of the reactant concentration and the reactor temperature are independent of C_A.

The jacket features a spiral baffle made of steel. In our model, we lump the mass of this metal strip, which is spirally wound around the vessel wall. Assuming a uniform temperature, the evolution of the jacket temperature is represented by

$$\frac{dT_i}{dt} = \frac{1}{\beta_j}(\rho_j C_{pj} F_j(T_{j0} - T_j) + UA(T_r - T_j)),\tag{8}$$

where C_{pj} is the specific heat capacity of the coolant, and F_j is the coolant flow rate. At each sampling time the temperature T_j in the right-hand side of Eq. (8) is approximated to the average of the inlet and outlet jacket stream temperature measurements: $T_j = (T_{j0} + T_{j2})/2$. Finally, the coefficients β_r and β_j in Eqs. (5) and (8) are given by

$$\beta_r = \rho C_p V + \alpha_r, \quad \beta_j = \rho_j C_{pj} V_j + \alpha_j,\tag{9}$$

where α_r and α_j represent the contribution of the wall and spiral baffle jacket thermal capacitances. This is an approximation to account for the influence of the metal wall and spiral baffle on the reactor and jacket temperature dynamics, *partitioned* here over both the liquid reactor and jacket energy balances. Note that Eqs. (5) and (8) were derived under the assumption of negligible heat losses. In fact, the reactor vessel is insulated and the heat losses are not significant for the range of temperatures of plant operation (Afonso, 1998).

A summary of the main process variables and model parameters is listed in Table 1. Since the kinetic

Table 1
List of the model variables and parameters

Variables	
$F_1, F_2, T_1, T_2, h, T_{j0}, T_{j2}, F_j, \lambda_5, T_r$	Measured
$T_j, C_A, F_3, V, C_{A1}, C_{A2}$	Calculated from the model selected by the user
Parameters	
U, α_j, α_r	To be estimated
$C_p, C_{pj}, (E_a/R), V_j, k_0, (-\Delta H_r), \rho, \rho_j$	Known *a priori*

parameters are *imposed* for the zero order kinetics, the most relevant model parameters to be estimated are U, α_r and α_j. Using dynamic tests, previous studies (Santos et al., 1998; Santos, 2001) found that α_r can be set to zero, while *typical* estimates of α_j and U at the nominal level operating conditions, $h = 0.30$ m, are approximately 7×10^5 J/K and 900 W/(m^2 K), respectively. The plot of the heat generation and removal terms (7) and (6), shows two steady states (Fig. 2 and Table 3): one stable at $T_r \simeq 31.5°C$ and an unstable steady state at $T_r \simeq 40°C$. Here the slope of Q_R is maintained constant by adjusting the coolant flow rate F_j. Table 2 gives the data for the remaining parameters and variables.

From the linearized form of the model equations at conditions given in Tables 2 and 3, the Jacobian of the model has an eigenvalue with a positive real part over an operating region of reactor temperatures. Moreover, the real part of this eigenvalue increases as the reactor temperature increases. Therefore, the operation of the reactor at higher reactant conversions can be more problematic because it is easier for small disturbances to destabilize the system.

3. Nonlinear MPC

In this section, we describe the basic framework for the NMPC algorithm and the implementation of the algorithm with the pilot plant reactor.

3.1. Control problem formulation

The nonlinear MPC framework considered here requires a mechanistic model of the process of the form

$$\dot{x} = f(x, u, d; \theta), \tag{10}$$

$$y = g(x; \theta) \tag{11}$$

with f and g twice continuously differentiable, where $x \in R^{n_s}$ is the state vector, $u \in R^{n_m}$ is the control vector, $d \in R^{n_d}$ is the disturbance vector, $\theta \in R^{n_\theta}$ is the parameter vector and $y \in R^{n_o}$ is the vector of output variables. An optimization formulation with different output and input predictive horizon lengths (denoted by p and m respectively, with $p \geqslant m$) is used to solve the model
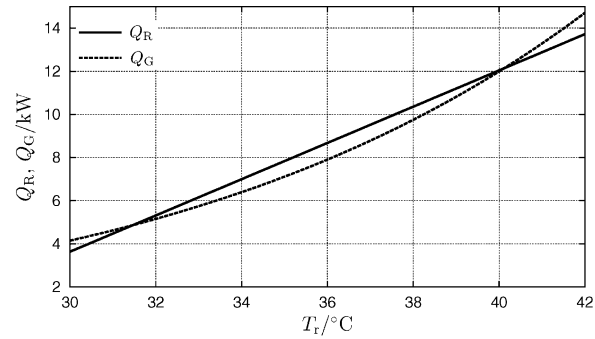

Fig. 2. Heat generation and heat removal curves (Q_G and Q_R) using the data from Table 2.

Table 2
Model data

C_{A1}	2.0 (mol/l)
C_{A2}	10.0 (mol/l)
C_p, C_{pj}	4184.0 (J/(kg K))
F_1	0.0 (l/min)
F_2, F_3	4.0 (l/min)
E_a/R	10080 (K)
k_0	6.20×10^{14} (mol/(m^3 s))
r	0.232 (m)
T_1	21.0 (°C)
T_2	21.0 (°C)
T_{j0}	26.0 (°C)
U	900.0 (W/(m^2 K))
V_j	0.014 (m^3)
α_j	7.0×10^5 (J/K)
α_r	0.0 (J/K)
$(-\Delta H_r)$	33488.0 (J/mol)
ρ, ρ_j	1000.0 (kg/m^3)

Table 3
Multiple steady states and the respective cooling flow rate

Steady states	Lower	Upper
h	0.30	0.30 (m)
C_A	7.82	4.60 (mol/l)
T_r	31.5	40.1 (°C)
T_j	28.0	28.0 (°C)
F_j	14.0	48.8 (l/min)

(10–11) over the predictive horizon p, where the state equations are integrated inside each sampling interval (Santos et al., 1995).

Our NMPC approach includes a number of desirable features. In particular, we apply a Newton-type algorithm which exploits the structure of NMPC problems. Compared to general purpose nonlinear programing solvers, this leads to solution of NMPC problem in far fewer iterations (see Oliveira & Biegler, 1995). Moreover, as described in Santos et al.

(1995), multiple shooting is applied to the mechanistic model (10–11) in order to obtain NMPC solutions for unstable systems. This is essential because direct integration methods often fail on open loop unstable systems.

In addition, this formulation is extended to include integral action as described in Oliveira and Biegler (1995). This is necessary to eliminate the steady state offset in the process outputs resulting from step disturbances and to compensate for the model mismatch. Moreover, in addition to input, state, and output constraints, this model predictive control formulation can handle output terminal constraints, control move rate constraints, as well as relaxations that convert these to soft constraints.

The combination of these features leads to the following nonlinear programing formulation which we solve at every time index i (Santos, 2001):

$$\min_{\tilde{X}, U, \varepsilon} \Upsilon_i(\tilde{Y}, U, \varepsilon) = \Psi_i(\tilde{Y}, U) + P_i(\varepsilon) \tag{12}$$

s.t. $\dot{x} = f(x, u, d; \theta), \quad y = g(x; \theta)$
$u(t) = u(t_{i-1+m}), \ t \in [t_{i+m}, t_{i+p}]$
$\tilde{x}_{i+k} = \phi(\tilde{x}_{i+k-1}, u_{i+k-1}),$
$k = 1, \ldots, p \tag{13}$

$z_{i+k} = z_{i+k-1} + K_1(y_{i+k} - y_{\text{sp},i+k}),$

$k = 1, \ldots, p \tag{14}$

$y_{i+p} - y_{\text{sp},i+p} \leqslant \textbf{Tol} \tag{15}$

$\tilde{X}_{\text{L}} \leqslant \tilde{X} \leqslant \tilde{X}_{\text{U}} \tag{16}$

$\tilde{Y}_{\text{L}} - \varepsilon_{\text{y}} \leqslant \tilde{Y} \leqslant \tilde{Y}_{\text{U}} + \varepsilon_{\text{y}} \tag{17}$

$U_{\text{L}} - \varepsilon_{\text{u}} \leqslant U \leqslant U_{\text{U}} + \varepsilon_{\text{u}} \tag{18}$

$\Delta U_{\text{min}} - \varepsilon_{\Delta\text{u}} \leqslant \Delta U \leqslant \Delta U_{\text{max}} + \varepsilon_{\Delta\text{u}} \tag{19}$

$\varepsilon \geqslant 0, \tag{20}$

where the subscripts $_{\text{sp}}$, $_{\text{L}}$ and $_{\text{U}}$ stand for *setpoint*, *lower* and *upper bounds*, respectively. The objective function (12) is defined such that

$$\Psi_i(\tilde{Y}, U) = \sum_{k=1}^{p} \tilde{e}_{i+k}^{\text{T}} Q_{\text{y}k} \tilde{e}_{i+k}$$

$$+ \sum_{k=1}^{m} \left(u_{i+k-1} - u_{\text{r},i+k-1} \right)^{\text{T}} \times Q_{\text{u}k}(u_{i+k-1} - u_{\text{r},i+k-1}),$$

where the subscript $_{\text{r}}$ stands for *reference*, $Q_{\text{u}k}$ and $Q_{\text{y}k}$ are diagonal weighting matrices, and $\tilde{e}_{i+k} = \tilde{y}_{\text{sp},i+k} - \tilde{y}_{i+k}$. In (12) $P_i(\varepsilon)$ is the penalty term when constraint relaxation is requested, and ε is a measure of the original constraint violations on the outputs, inputs

and control move rates, defined by

$$\varepsilon = \begin{bmatrix} \varepsilon_{\text{y}}^{\text{T}} & \varepsilon_{\text{u}}^{\text{T}} & \varepsilon_{\Delta\text{u}}^{\text{T}} \end{bmatrix}^{\text{T}}.$$

For instance, if the penalty term is defined according to the exact penalty formulation, it follows that $P_i(\varepsilon) = r^{\text{T}}\varepsilon$, where r is the vector of penalty parameters of appropriate size defined by: $r = \rho \cdot [1 \cdots 1]^{\text{T}}$, $\rho \in R^+$. The *augmented* vectors $\tilde{X}, \tilde{Y}, U$ and ΔU are defined by

$$\tilde{X} = \begin{bmatrix} \tilde{x}_{i+1} \\ \vdots \\ \tilde{x}_{i+p} \end{bmatrix}, \quad \tilde{Y} = \begin{bmatrix} \tilde{y}_{i+1} \\ \vdots \\ \tilde{y}_{i+p} \end{bmatrix}, \quad U = \begin{bmatrix} u_i \\ \vdots \\ u_{i+m-1} \end{bmatrix}$$

and

$$\Delta U = \begin{bmatrix} \Delta u_i^{\text{T}} & \Delta u_{i+1}^{\text{T}} & \cdots & \Delta u_{i+m-1}^{\text{T}} \end{bmatrix}^{\text{T}},$$

where $\Delta u_{i+k} = u_{i+k} - u_{i+k-1}$, $k = 2, \ldots, m-1$. The notation $\tilde{x}_{i+k}$ and $\tilde{y}_{i+k}$ denote the *augmented* state and output vectors using the additional state equations (14) for integral action:

$$\tilde{x}_{i+k} = \begin{bmatrix} x_{i+k} \\ z_{i+k} \end{bmatrix}, \quad \tilde{y}_{i+k} = \begin{bmatrix} y_{i+k} \\ z_{i+k} \end{bmatrix}.$$

In (14), K_1 is a constant diagonal matrix that determines the speed of the response of the integrator element. Vectors ΔU_{min} and ΔU_{max} in (19) are defined as follows:

$$\Delta U_{\text{min}} = \begin{bmatrix} \Delta u_{\text{min}}^{\text{T}} & \cdots & \Delta u_{\text{min}}^{\text{T}} \end{bmatrix}^{\text{T}},$$

$$\Delta U_{\text{max}} = \begin{bmatrix} \Delta u_{\text{max}}^{\text{T}} & \cdots & \Delta u_{\text{max}}^{\text{T}} \end{bmatrix}^{\text{T}},$$

where $\Delta u_{\text{min}} < 0$ and $\Delta u_{\text{max}} > 0$, $\Delta u_{\text{min}}, \Delta u_{\text{max}} \in R^{n_{\text{m}}}$. Although we assume ΔU_{min} and ΔU_{max} to be constant over the entire input predictive horizon, the implementation of variable profile bounds is also straightforward. Moreover, to obtain stable, bounded solutions, we employ a multiple shooting formulation where $\phi(\tilde{x}_{i+k-1}, u_{i+k-1})$ is obtained through the integration of (10) inside the sampling interval $t \in (t_{i+k-1}, t_{i+k})$ only. The constraints (13) are then satisfied directly within the optimization problem, and the states and outputs converge to continuous profiles over the predictive horizon. Finally, the vector **Tol** in (15) is the vector of tolerances for the output terminal constraints, which again promotes the stability of the solution. The control problem formulation (12–20) is implemented in a computational framework coded in Fortran 77. A detailed description of this NMPC control formulation can be found in Santos (2001).

3.2. Reactor control settings

In this study, output terminal constraints, control move rate constraints and constraint relaxations were turned off. The NMPC control configuration uses as output variables $y^{\text{T}} = [h \ T_{\text{r}}]$ and controls $u^{\text{T}} = [F_{\text{j}} \ F_3]$. Table 4 gives the operating constraints that are

Table 4
Operating constraints on the state and control variables

States	Lower	Upper
h	0.08	0.41 (m)
C_A	0	None (mol/l)
T_r	0	None (°C)
T_j	0	None (°C)
Controls	Lower	Upper
F_3	0	12 (l/min)
F_j	0	76 (l/min)

incorporated into the control algorithm. The lower bound on the level avoids a zero value, which can cause a failure of the numerical integrator. The level upper bound prevents the reactor from overflowing; the upper bounds on the controls are specified based on the pilot plant operating limits. The results presented in Section 4 were obtained assuming that the model is perfect and all the state variables are measured. The control tuning parameters selected in this study are given in Section 4.

3.3. On-line implementation

In the experimental apparatus, the outlet reactor and coolant flow rates are manipulated through PI controllers, which manipulate the valves Vc5 and Vc3, respectively (Fig. 1). Here, the setpoints of these two PI controllers are determined by the NMPC algorithm. However, as the outlet flow rate is not measured, an experimental correlation is used to predict F_3 as a function of λ_5 (see $f(\lambda_5)$ in Fig. 3) (Santos, 2000). This value is then compared with the corresponding NMPC setpoint to calculate the error for the outlet flow rate PI controller. Also, since there are no concentration measurements, the prediction of C_A made at the previous instant is used to initialize the model at each sampling time. The operating constraints in Table 4 are also incorporated in the control problem, except that a lower bound of 10 l/min has to be set on the coolant flow rate. This lower bound on F_j prevents the flow meter from getting stuck at low flow rates.

The NMPC framework runs simultaneously with another program that implements all of the PI controllers as well as the simulation algorithm for the chemical reaction. The sampling time is 2 s for the PI controllers and the reaction simulation, and at least 15 s for the NMPC controller. Both programs are implemented in Fortran 77, and run on a Linux Pentium 133 MHz laptop. For the experiments in this study, the CPU time to solve the NMPC control problem at every time index averages between 2 and 3 s. An overview of the experimental control arrangement with a description of the data flow between the plant, the NMPC frame-
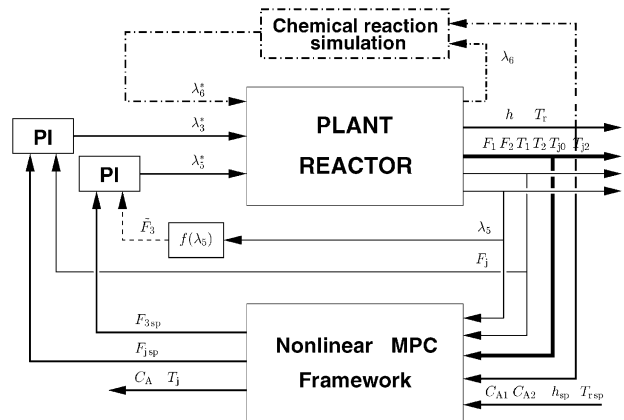

Fig. 3. Diagram of the data flow between the plant, the NMPC framework, the PI controllers and the chemical reaction simulation algorithm.

work, the PI controllers and the chemical reaction simulation algorithm is given in Fig. 3.

4. Experimental results

We now consider a number of experimental control scenarios for this system. In particular, we consider setpoint changes so that we can *operate at an unstable point* with a higher reactant conversion. Next we consider the *influence of saturation constraints* at this point and the ability of the NMPC control to deal with this feature. Finally, we consider the performance of the NMPC controller in dealing with the *effects of unmeasured disturbances and model mismatch*. All of the simulation results presented in this subsection were obtained with only a single feed to the reactor (i.e., $F_1 = 0$) and the agitator speed set constant at 145 rpm.

4.1. Operating at an unstable point

We first consider a sequence of setpoint changes in the reactor temperature. The aim of these changes is to drive the reactor to operating conditions around the unstable steady state and to get a higher rate of conversion of reactant A. Based on simulation tests, output and input diagonal weighting matrices $Q_{yk} = \mathrm{diag}(5 \times 10^2, 10^5)$ and $Q_{uk} = \mathrm{diag}(10^{-1}, 10^{-3})$, $k = 1, \ldots, p$, are used to improve the conditioning of the control problem. Also, the integral action in our NMPC controller is turned off.

The effect of the tuning parameters is significant at this operating point. For instance, under steady state operating conditions, without any setpoint change or large disturbances, the control variables exhibit a ringing behavior around the nominal operating value. On the other hand, if we penalize the control moves (by increasing Q_u) we obtain sluggish servo and regulatory

responses. It turns out that by simultaneously increasing the value of the penalty weights on the errors, Q_{yk}, the closed loop response speed is recovered.

Figs. 4 and 5 show the reactor closed loop response to a sequence of temperature setpoint step changes, as well as to a step disturbance in the inlet jacket temperature. The profiles were obtained with predictive horizons $(p, m) = (20, 5)$ and a sampling time $\Delta t = 30$ s. This corresponds to a predictive time horizon of 10 min. Note that the effect of the controller action to comply with the setpoint changes starts up 10 min before the setpoint change is formally considered, because setpoint profiles in the predictive control framework are updated in response to *scheduled* setpoint changes.

4.2. Influence of saturation constraints

From Fig. 4, we see that the coolant flow rate reaches its upper operating constraint at time $t \simeq 130$ min. At this point, no more cooling capacity is available to control the temperature rise. On the other hand, we see that as F_j increases from $t \simeq 100$ min up to its maximum operating value, the controller stabilizes the reactor by reducing the residence time, and this keeps the level at a value below its setpoint. Moreover, if we decrease the weighting parameter in Q_{yk} corresponding to the level setpoint, the controller drives the level to an even lower value. As a result, this behavior reduces the demand for additional cooling capacity for the reactor.

On the other hand, if additional cooling capacity is not needed, more conventional closed loop behavior is observed. For instance, in Fig. 5 we introduce a disturbance in the reactor inlet jacket temperature T_{j0}. This temperature is decreased from 26 to 10°C at time $t = 165$ min and the coolant flow rate decreases significantly in order to maintain a zero offset on the reactor temperature. Moreover, because of these more favorable operating conditions the controller actuates on F_3 to eliminate the level offset ($165 \leqslant t \leqslant 173$ min). Here, maintaining the temperature at its setpoint value requires a much lower coolant flow rate. Also, as seen in Fig. 5, the temperature setpoint change at $t = 217$ min is easily handled by the NMPC controller. This permits the reactor to operate at a higher reactant conversion, with the system remaining closed loop stable.

4.3. Effect of unmeasured disturbances and model mismatch

There are a number of unmeasured and unmodeled disturbances associated with this experimental system. Firstly, note the oscillatory behavior on the controls F_j and F_3, for instance in Graphic III of Figs. 4 and 6. This behavior can be attributed to oscillations on the flow
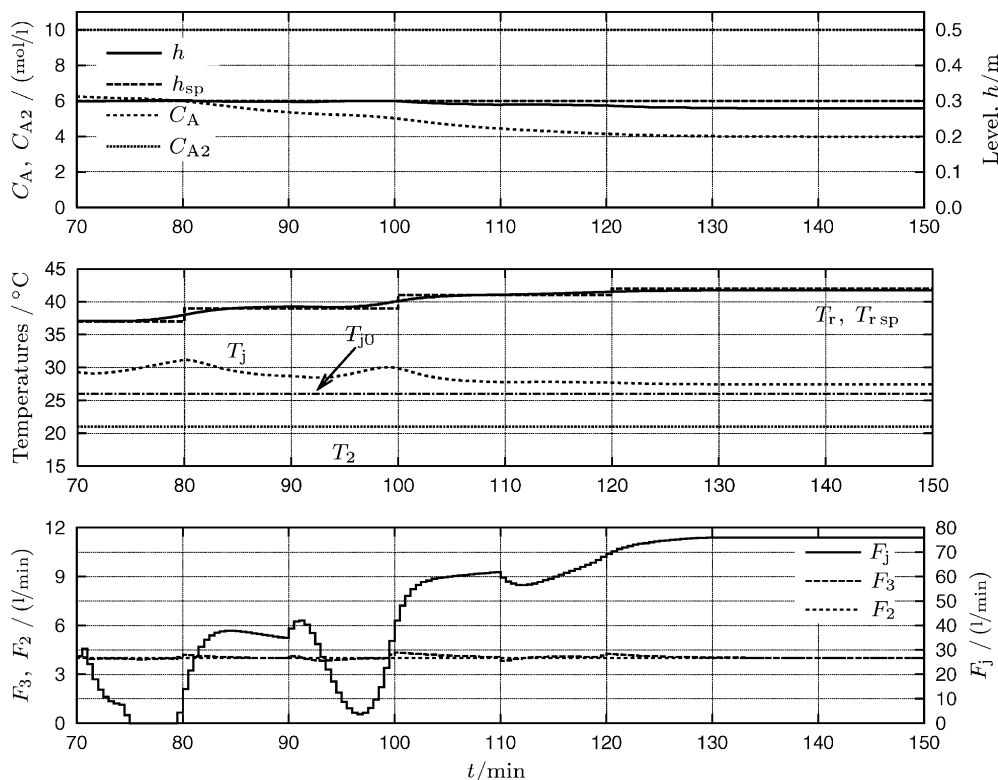

Fig. 4. Reactor closed loop response to a sequence of step changes in the reactor temperature: $\Delta T_{rsp} = +2$°C at $t = 80$ and 100 min, and $\Delta T_{rsp} = +1$°C at $t = 120$ min (continued on Fig. 5).
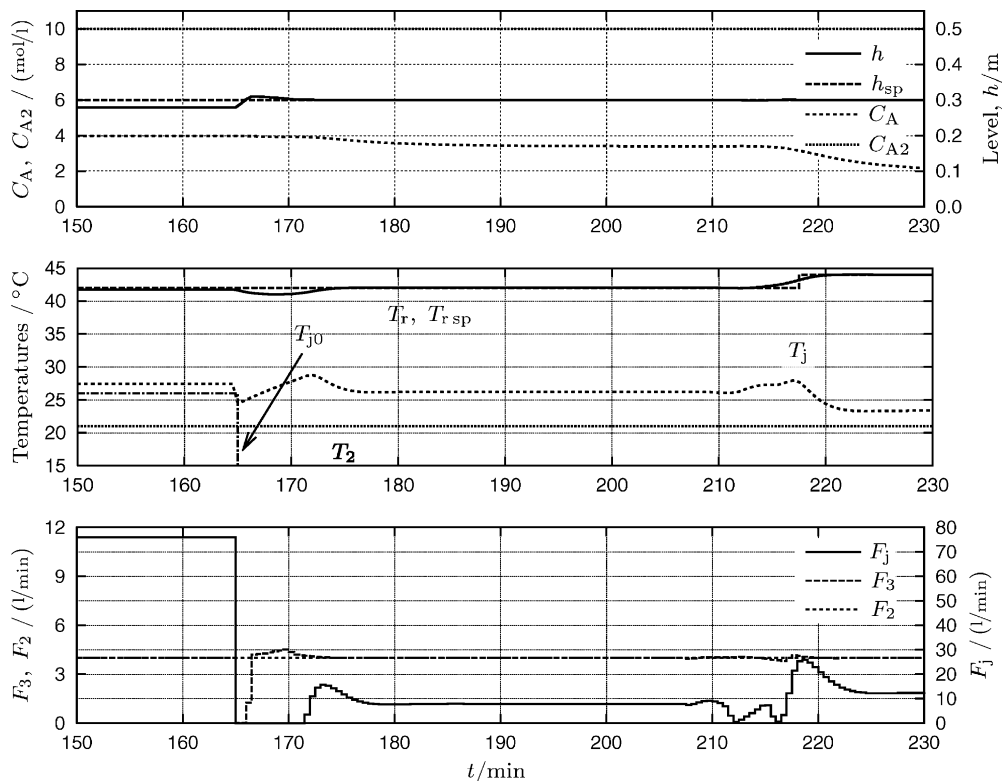
Fig. 5. (continued from Fig. 4). Reactor closed loop response to a step change in the inlet jacket temperature with $\Delta T_j = -16°C$ at $t = 165$ min, and to a step change in the reactor temperature setpoint, $\Delta T_{rsp} = +2°C$, at $t = 217$ min.

measurements and to the fact that the PI controllers are not fast enough to reach their setpoints within the NMPC sampling time. Since the model is reinitialized at each sampling time, this leads to frequent variations of the PI controller setpoints. Secondly, as explained in Section 3.3, there is no measurement available for F_3; this variable is inferred from the valve position for Vc5. The arrangement made to control F_3, along with the hysteresis associated with valve Vc5 is a source of permanent oscillations. Lastly, there are also periodic disturbances on the feed F_2 (see Graphic I in Figs. 6 and 7) because of pressure variations in the water storage tank. Nevertheless, despite these unmeasured disturbances, Figs. 4–7 show that the NMPC controller performs very well in tracking the setpoints and rejecting disturbances.

To show this performance in more detail, we modified the weighting matrices of the NMPC controller to cope with a different performance scenario: $Q_{yk} = \text{diag}(10^6, 10^6)$, $Q_{uk} = \text{diag}(1.25, 1)$, $k = 1, ..., p$. Now the penalties on both outputs are the same and the penalties on the control moves are substantially increased, in an attempt to reduce the ringing behavior of the inputs. Also, the integral action was turned on with integral constants $K_1 = \text{diag}(0.005, 0.3)$ to compensate for unknown disturbances and model mismatch and to eliminate the

offset in the level and temperature. Finally, the length of the predictive horizon was reduced in order to reduce the computation time. This was done to avoid longer computation delays in the closed loop system, which would deteriorate the overall closed loop performance. These delays increase the possibility of unstable closed loop behavior and more oscillation on the manipulated variables. Due to limitations on the available computational power, predictive horizon lengths were set to $(p, m) = (10, 2)$ with a sampling time $\Delta t = 15$ s. This corresponds to a predictive time horizon of only 2.5 min, which is much less than the time horizons presented in Section 4.1. Again, about 2.5 s are required to solve each NMPC problem.

Figs. 6 and 7 show the response of the process variables under NMPC control to a sequence of changes on the reactor temperature setpoint T_{rsp}, followed by level setpoint changes after time $t = 165$ min. Note that the NMPC controller tracks these setpoint changes quickly and accurately. On the other hand, this performance also allows us to identify a number of sources of model mismatch. For example, Graphic IV in Figs. 6 and 7 shows the evolution of the steam valve aperture calculated from Eq. (2) and the corresponding measurement. This valve (Vc6) is slow to respond and exhibits significant hysteresis, requiring the PI controller
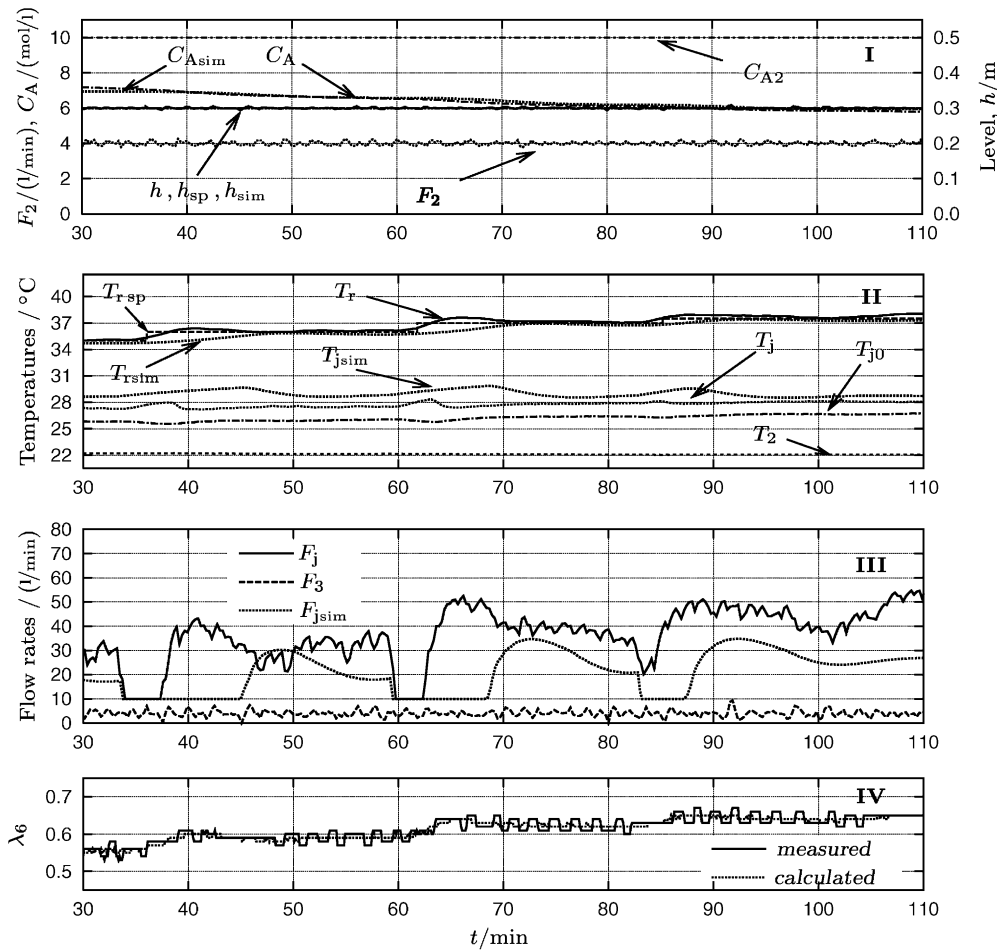
Fig. 6. Response to a sequence of changes of the reactor temperature setpoint ($35 \rightarrow 36 \rightarrow 37 \rightarrow 37.5°$C—graphic II). Labels with the subscript $_{sim}$ identify off-line simulation profiles. Graphic IV shown the evolution of the steam value aperture to simulate the chemical reaction (continued on Fig. 7).

to track the valve aperture calculated from the kinetic simulation algorithm.

Also, Graphics I, II and III in Figs. 6 and 7 show the evolution of the process variables. The comparison of the experimental profiles with the ones obtained by off-line simulation (labeled with the subscript $_{sim}$) and with the same control tuning parameters, shows plant/model mismatch, especially for the reactor temperature profiles (Graphic II). This is also evident from the difference between the profiles of the coolant flow rates (Graphic III), F_j and F_{jsim}. Although this difference can be somewhat reduced by refining the parameters α_j and U, it is also necessary to take into account the effect of the recirculation of the coolant (see Fig. 1). Due to current operating limitations it is not possible to keep the inlet jacket temperature constant, even with the help of the purge scheme in the circulation loop. Thus, as T_r increases T_{j0} also increases (as seen in Graphic II of Fig. 6). The effect of this increase in reactor temperature therefore requires a higher coolant flow rate, and this is indeed enforced by the NMPC controller.

5. Conclusions

A Newton-type NMPC controller is implemented on a pilot plant reactor. This reactor deals with a simulated zero order reaction and features highly nonlinear behavior. Moreover, under certain operating conditions it can become open loop unstable. Here a model is developed and validated with experimental data. Although it appears there is a good comparison between simulated and experimental results, there are several sources of unmeasured disturbances as well as a significant degree of plant/model mismatch within this system. Despite these challenges the NMPC controller performs very well in tracking the setpoints and rejecting disturbances. This is done through the rapid on-line solution of nonlinear programs, which incorporate first principle models, stable problem formulations, integral action and superior methods for constraint handling.

A number of issues need to be considered for future work. To achieve better performance it may be necessary to consider on-line parameter estimation. However, before moving to this level of sophistication,
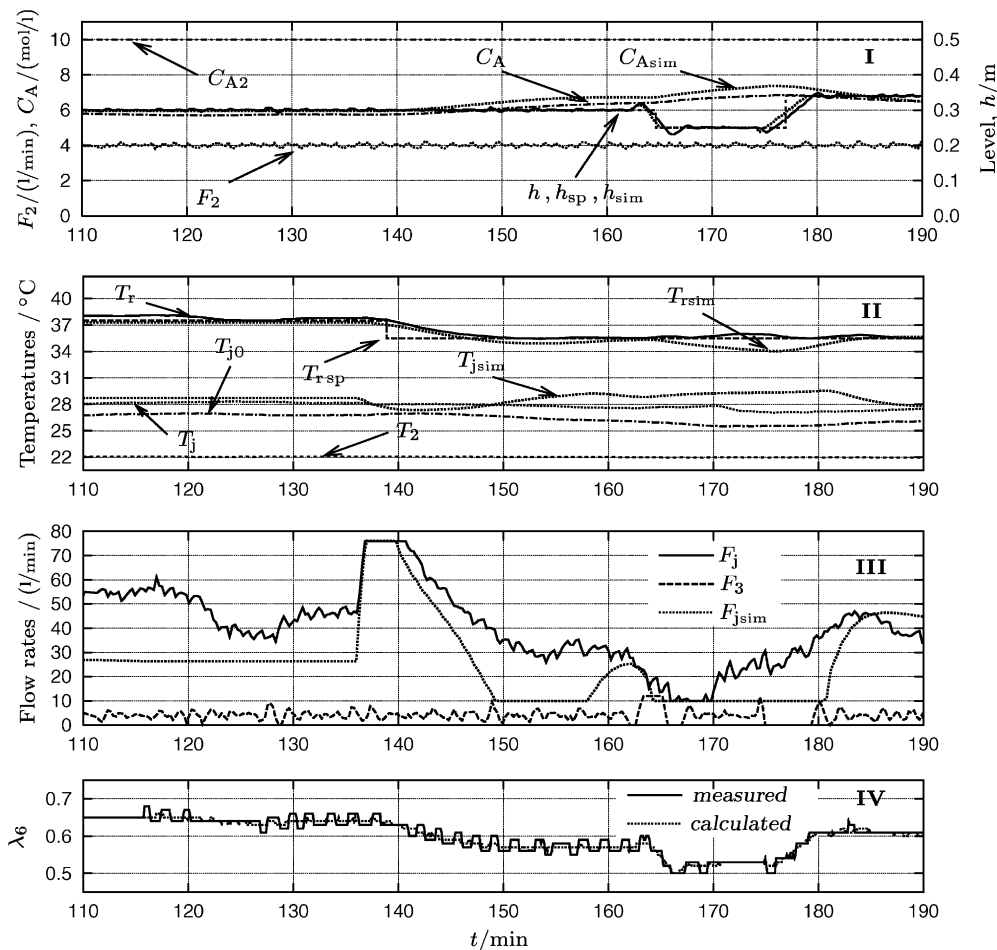
Fig. 7. (continued from Fig. 6). Response to a reactor temperature setpoint change ($37.5 \rightarrow 35.5^{\circ}$C—graphic II) followed by a sequence of changes of the level setpoint after $t = 165$ min ($0.30 \rightarrow 0.25 \rightarrow 0.34$ m—graphic I). Labels with the subscript $_{sim}$ identify off-line simulation profiles. Graphic IV shows the evolution of the stream valve aperture to simulate the chemical reaction.

a number of plant improvements can be made: using an intermediate jacket temperature sensor to obtain better jacket temperature information; using a flow meter to measure the outlet flow rate F_3; modifying the cooling system to guarantee a constant and lower inlet coolant temperature, T_{j0}; and using a faster computer for simulation and control. Nevertheless, despite these operating difficulties and plant/model mismatch, the NMPC closed loop system follows the setpoints and rejects disturbances quite well.

References

Afonso, P. A. F. N. A. (1998). *Produção assistida por computador na indústria dos processos químicos*. Ph.D. thesis, Universidade de Coimbra, Coimbra, Portugal.

Afonso, P. A. F. N. A., Ferreira, J. M. L., & Castro, J. A. A. M. (1998). Sensor Fault Detection and Indentification in a pilot plant under process control. *Trans IChemE, 76*(Part A), 490–497.

Afonso, P. A. F. N. A., Oliveira, N. M. C., & Castro, J. A. A. M. (1996). Model predictive control of a pilot plant reactor with a simulated exothermic reaction. *Computers & Chemical Engineering, 20*(Part B), S769–S774.

Kershenbaum, L. S., & Kittisupakorn, P. (1994). The use of a partially simulated exothermic (PARSEX) reactor for experimental testing of control algorithms. *Trans IChemE, 72*(Part A), 55–63.

Oliveira, N. M. C., & Biegler, L. T. (1995). Newton-type algorithms for nonlinear process control. Algorithms and Stability Results. *Automatica, 31*(2), 281–286.

Pinheiro, C. I. C., & Kershenbaum, L. S. (1999). Model predictive control of a pilot plant reactor temperature in a CSRT pilot plant operating at an unstable steady-state. *Computers & Chemical Engineering Supplement*, pp. S859–S862.

Qin, S. J., & Badgwell, T. A. (2000). An overview of nonlinear model predictive control applications. In: Allgower, F., Zheng, A. (Eds.), *Nonlinear model predictive control*, Switzerland: Birkhauser.

Santos, L. O. (2001). *Multivariable predictive control of nonlinear chemical processes.* Ph.D. thesis, Universidade de Coimbra, Coimbra, Portugal.

Santos, L. O., Oliveira, N. M. C., Biegler, L. T. (1995). Reliable and efficient optimization strategies for nonlinear model predictive control. In: Rawlings, J. B. (Ed.), *Proceedings of the Fourth IFAC Symposium on Dynamics and Control of Chemical Reactors,* *Distillation Columns and Batch Processes, DYCORD'95, Helsingør,* Denmark (pp. 33–38). Oxford: Elsevier Science.

Santos, L. O., Afonso, P. A. F. N., & Castro, J. A. A. M. (1998). Modeling an industrial scale pilot plant reactor. In: Ribeiro, F. R. Alves, S. S. (Eds.), *Proceedings of the Seventh International Chemical Engineering Conference CHEMPOR'98* (pp. 687–694). Lisboa, Portugal.