

## MULTIVARIABLE CLOSED-LOOP IDENTIFICATION WITH CONSTRAINED MPC CONTROL: CASE-STUDY ON A C3/C4 SPLITTER

Oscar A. Z. Sotomayor,<sup>1</sup> Darci Odloak<sup>1</sup> and Lincoln F. L. Moro<sup>2</sup>

<sup>1</sup> CESQ – Department of Chemical Engineering Polytechnic School of the University of São Paulo, Brazil

<sup>2</sup> Advanced Control Group PETROBRAS, Brazilian Oil Company, Rio de Janeiro, Brazil

### Abstract

This paper deals with the model re-identification in closed-loop systems with already existing MPC controllers. It is assumed that the controller has a two-layer structure where the upper layer performs a simplified economic optimization in which the economic objective is represented as a linear combination of the process inputs. This is the case of several commercial MPC packages applied in industry. This work focuses on the case where the existing process model shows signs of deterioration and there is benefit in obtaining a new model. It is proposed a methodology where the test signal is introduced in the coefficients of the objective function of the economic layer. The approach allows the continuous operation of the system as the process constraints and product specification can be satisfied during the test. The application of the method is illustrated by simulation on a C3/C4 splitter of the oil industry. The method is simple to be implemented in existing commercial packages, and the results show that the method has a good potential to be applied in practice.

**Keywords:** Closed-loop identification, model predictive control, model maintenance, control performance assessment, distillation column.

### I. INTRODUCTION

Model predictive control (MPC) strategies have become the dominant approach in advanced process control applications (Qin & Badgwell, 2003). Return on investment and substantial benefits are generated directly from the ability of MPC to ensure that the plant operates at its most profitable constraints. One of the key elements of MPC is the process model, which is used to predict the future behavior of the system output along a prediction horizon that is long enough to encompass all the process dynamics that need to be taken into account in the computation of the control law. Thus, it becomes clear that a poor process model may result in poor output predictions, and a decrease in the control performance, which may jeopardize the economic benefits associated to the control strategy. Consequently, an accurate model is a key factor in the performance of the MPC controller.

In general, from 2 to 3 years after commissioning, these controllers rarely keep performing as

they were initially designed (Treiber *et al.*, 2003). The main cause of the performance deterioration of MPC is related to the model deterioration resulting from changes in the dynamic behavior of the plant or persistent unmeasured disturbances that force the plant to a different operating point. Changes in the dynamics of the plant may result from wearing, fouling, catalyst decay, etc, or changes in the operating conditions or product specifications. Usual disturbances include the environment temperature, variations on the feedstock impurities or operating problems in an upstream process. Minor disturbances include process and instrumentation noises, which are always present up to a limited level. In general, the above listed problems intensify with time and tend to accentuate the plant/model error, leading to a significant degradation in the performance of the control system. In order to keep the performance of the MPC controller at an acceptable level, it is essential to carry out the MPC re-commissioning in a periodic basis,



which means to re-identify the process model and, if necessary, to retune the MPC controller considering the new model. However, due to the plant commitment to production goals, model re-identification means, in most cases, to develop a new model based on plant data obtained in closed-loop conditions.

Closed-loop identification is a research subject with growing interest in the last decade (Hjalmarsson *et al.*, 1996; Forssell & Ljung, 1999a; Landau, 2001). Important aspects on model identification have been studied and several identification strategies have been proposed. Production goals and safety aspects are the usual incentives to research in closed-loop identification in order to build a model that is relevant to process control (Landau, 2001). As the controller diminishes the effects of disturbances and keeps the process within the designed operating ranges, the resulting process data can generate more reliable models without significant uncertainties. In general, closed loop identification reduces the problem of gain directionality as the controller introduces the necessary correlation between the inputs in order to excite the outputs in the adequate magnitudes and directions (Andersen & Kümmel, 1992; Li & Lee, 1996). However, the correlation between the inputs, and the significant correlation between inputs and outputs become the fundamental problem in the closed-loop identification procedure, as these correlations may result in systematic error or bias in the estimated parameters. Moreover, the inherent reduction of the excitation signal resulting from the presence of the controller may result in a poor signal/noise relation, which may produce random error and a significant variance in the model parameters. Despite these disadvantages, the closed-loop identification is carrying a lot of attention in the last years and it has been considered a serious alternative to the controller retrofitting when the error plant/model is significant.

All closed-loop identification can be classified as direct, indirect or joint input-output methods (Ljung, 1999a). Both indirect and joint input-output methods require prior knowledge of the controller or assume that it has a certain LTI structure. Obviously, these methods are not suitable for MPC applications, because MPC presents nonlinear and time-variant features. For this sort of control strategy, the direct method is the recommended choice for closed-

loop identification (Rivera & Flores, 1999; Zhu & Butoyi, 2002). Some works have used the indirect and the two-stage identification methods, a variant of joint input-output method, in unconstrained MPC systems, e.g. see Saffer & Doyle (2002). Unconstrained MPC is basically a LTI controller, but it is very little used in industry because it does not offer the economic benefits associated to its constrained counterpart.

In closed-loop identification, theoretical studies show that, in order to achieve necessary and sufficient conditions for process identifiability, an external excitation signal is required (Gustavsson *et al.*, 1977). The external excitation signal should be persistently exciting across the frequency range of interest (Godfrey, 1993), and must be independent of any process disturbance. In MPC systems, external excitation is a dither signal that can be injected in the manipulated variables (e.g. Kline *et al.*, 2006) or in the controlled variable set-points (e.g. de Klerk & Craig, 2003) or both. The main disadvantage of this approach is that there is no guarantee that the process constraints and product specifications will be attended during the execution of the identification test. On the other hand, an insufficient excitation may compromise the identification requirements. To face these problems, some authors have proposed a new class of excitation methods for MPC that can be called as internal excitation methods. They are either based on an adequate modification of the objective function of the controller or through the inclusion of an excitation constraint. Sagias & Nikolaou (2001) propose the extension of the cost function of the conventional MPC to address both control and identification objectives. The approach considers injection of a dither signal into the new cost function trying to force a predefined sequence of moves on the manipulated inputs. With this strategy, the control problem remains a QP, i.e. no new optimization code is needed. Genceli & Nikolaou (1996) propose the MPC-I (model predictive control and identification) where the persistent excitation of the inputs is imposed as a constraint in the control optimization problem. The resulting problem is no longer a QP and it is also non-convex, which limits the practical application of the method. Another crucial shortcoming of the MPC-I approach is that they cannot easily applied to MPC packages already installed,



because they need some modifications on the code of the control program. This is certainly not available for most of the practitioners.

The main goal of this work is to establish a general closed-loop identification methodology to be applied to commercial MPC packages, considering that the

identification step must be performed in an industrial environment where production cannot be interrupted and the process operation must be kept inside an operating window defined by product specification and equipment constraints.

## II. THE PROPOSED METHODOLOGY

Most of the commercial MPC packages are implemented in a two-layer structure as represented in Figure N° 1, with both layers executed with the same sampling period. The upper layer usually corresponds to a simplified steady state economic optimization, and the lower layer stands for a MPC algorithm in which the outputs are controlled in specified zones or ranges (Maciejowski, 2002) instead of fixed references. In the economic optimization we search for the optimum steady state values of the system inputs (input targets). This is done by optimizing a linear or quadratic economic objective subject to bound constraints in the system inputs and outputs. The outputs at the optimal steady state are computed through a steady state linear model and using the available output prediction in the MPC algorithm. The optimal input targets are sent to the MPC algorithm, where the control cost is extended to include a term that weights the distance between the present value of the input and the optimal target.

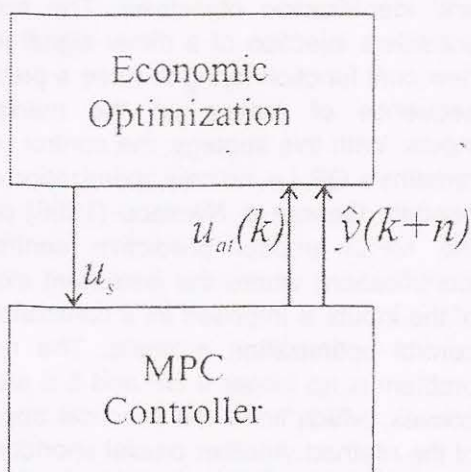


Figure N° 1. Typical MPC layered structure

A reasonable consideration is that when the process model is poor, the economic optimization layer will change the input target to the MPC layer quite often, and so, the input targets could be viewed as a possible test signal in the closed-loop identification

procedure. However, to assume that these natural moves on the input targets will be sufficiently exciting is a questionable matter.

Taking advantage of the MPC layered structure (Figure N° 1), and in order to guarantee the necessary excitation of the process, which we are trying to re-identify, we propose a method where the test signal is introduced as a multiplier factor of the economic coefficients in the objective function of the economic optimization layer. For this purpose, the economic optimization problem is written as follows:

$$\min_{\Delta u_s, \delta_y} - (W_1 \otimes W_{exc}) \Delta u_s + \|W_2 \Delta u_s\|_2^2 + \|W_3 \delta_y\|_2^2 \quad (1)$$

subject to:

$$\begin{aligned} \Delta u_s &= u_s - u_{at} \\ y_s &= G_0 \Delta u_s + \hat{y}(k+n/k) \\ u_{\min} &\leq u_s \leq u_{\max} \\ y'_{\min} &\leq y_s + \delta_y \leq y'_{\max} \end{aligned} \quad (2)$$

where:

$u_{at}$  vector of manipulated inputs at present time  $k$

$u_s$  vector of targets of the manipulated inputs

$y_s$  vector of predicted outputs at steady state  
 $\hat{y}(k+n/k)$  prediction of the controlled output at instant  $k+n$  ( $n$  is the model horizon or settling time of the dynamic process model) computed at time  $k$ .

$\delta_y$  vector of slack variables for the controlled outputs

$G_0$  steady state gain matrix of the process

$W_1$  vector of economic coefficients of the manipulated inputs, which may be considered as the partial derivatives of the true economic function in relation to the inputs

$W_2$  matrix of weights of the manipulated input moves



$W_3$  matrix of weights of the output slack variable

$W_{exc}$  time-variant weighting vector of the economic coefficients (persistently exciting test signals)

$u_{min}$   $u_{max}$  bounds of the manipulated inputs (the same as used in the control layer)

$y'_{min}$   $y'_{max}$  bounds of the controlled outputs (which zones can be smaller than or equal to the ones used in the MPC layer)

In Equation (1), operator  $\otimes$  denotes the Hadamard (or Schur) product. On the above equations, the slack variable is included to guarantee existence of a solution to the optimization problem defined in (1) and (2). If  $W_{exc}$  is set equal to a time-invariant vector of ones, the excitation procedure is ended and we have the original economic optimization problem.

As a result of the solution to the problem defined in (1) and (2), we obtain the input target  $u_s$  that is passed to the MPC layer, which solves the following QP:

$$\begin{aligned} \min_{\Delta u} & \sum_{i=1}^p \|Q(\hat{y}(k+i/k) - y_{sp})\|_2^2 \\ & + \sum_{j=1}^m \|R\Delta u(k+j-1/k)\|_2^2 \\ & + \sum_{j=1}^m \left\| R_u \left( u_{oi} + \sum_{i=1}^j \Delta u(k+i-1/k) - u_s \right) \right\|_2^2 \end{aligned} \quad (3)$$

subject to:

$$\begin{aligned} -\Delta u_{max} & \leq \Delta u(k+j-1/k) \leq \Delta u_{max}, \quad j=1, \dots, m \\ u_{min} & \leq u_{oi} + \sum_{i=1}^j \Delta u(k+i-1/k) \leq u_{max}, \quad j=1, \dots, m \end{aligned} \quad (4)$$

where:

- $\hat{y}$  output predictions
- $y_{sp}$  set-point of the system output
- $\Delta u$  vector of control moves
- $m$  control horizon
- $p$  prediction horizon
- $Q$  diagonal weight matrix of the controlled outputs
- $R$  diagonal weight of matrix of the manipulated inputs

$R_u$  diagonal matrix that weights the distance between the computed input and the optimum target.

In this controller, the zone control strategy is implemented as follows. For each output  $y_j$ , we observe the prediction at time instant  $k+i$ . We may have the following cases:

(a)  $y_{j,min} \leq \hat{y}_j(k+i/k) \leq y_{j,max}$

In this case, the error on this output at time instant  $k+i$  can be ignored by the controller. This means that weight  $Q$  should be made equal to zero and the set-point can assume any value.

(b)  $\hat{y}_j(k+i) > y_{j,max}$

In this case, the controller should bring the output back to the upper boundary of the control zone. Then, we should select the output set point such that  $y_{j,sp} = y_{j,max}$ , and weight  $Q$  should be made equal to the value obtained in the controller tuning procedure.

(c) Analogously, if  $\hat{y}_j(k+i) < y_{j,min}$

then, this output should be brought back to the lower boundary of the control zone. In this case we make  $y_{j,sp} = y_{j,min}$  and weight  $Q$  is selected as in the previous case.

The proposed method to excite the system in closed-loop can be considered as an internal excitation method, and it can be easily implemented in existing MPC packages. Each test signal is designed with a different shape, such that the cross-correlation between the inputs is minimized. We will show, in the example section, that, with this approach, the inputs can be adequately excited, the feedback effect on the test data is minimized and the requirements about the test signal for the system to be identifiable are reached. Also the approach attends the process safety requirements and the specifications of the products are attended satisfactorily.

The other steps of the proposed methodology follow same steps of the identification methodology that is usually applied to industrial processes (Ljung, 1999a): design of the test signal, selection of the model structure, computation of the model parameters and model validation. The development of each of these steps is illustrated in the case study of closed-loop identification that is presented in the next section.



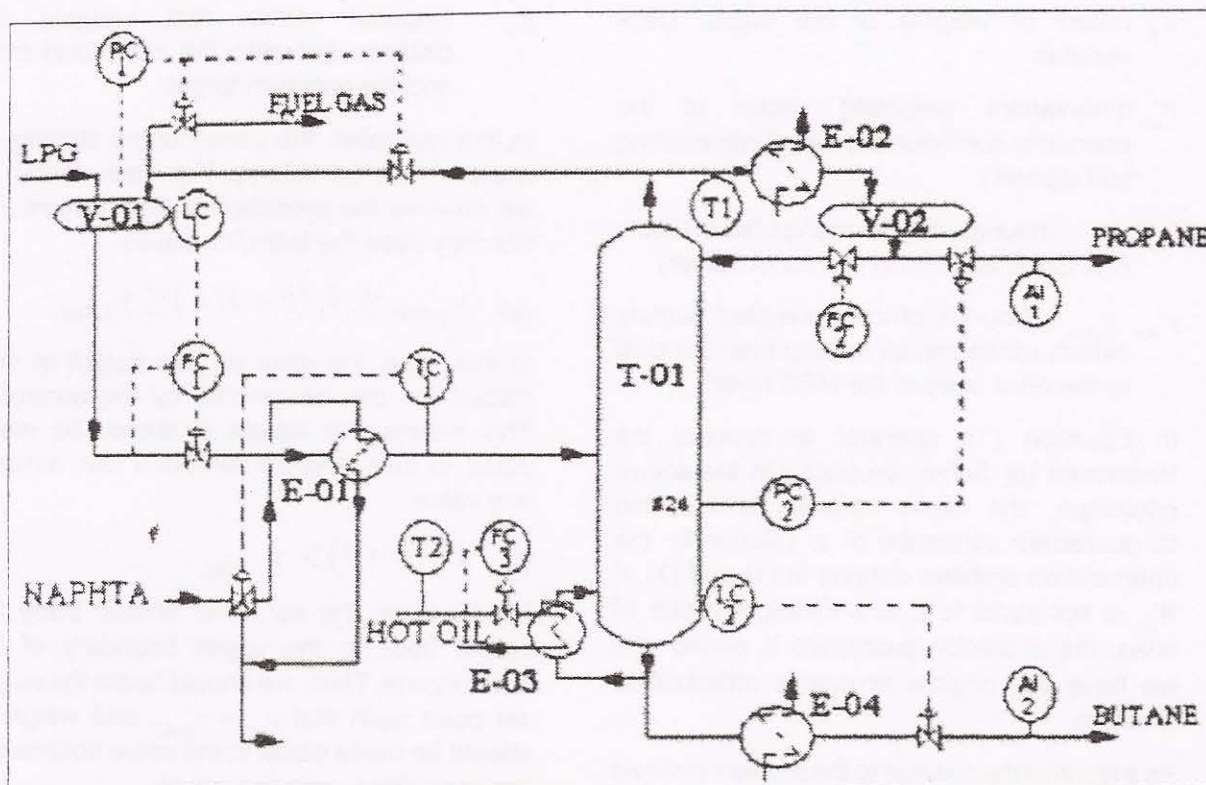


Figure N° 2. Schematic diagram of the C3/C4 splitter

### III. CASE-STUDY: INDUSTRIAL C3/C4 SPLITTER

This work considers the process studied by Porfírio et al. (2003). It is a distillation column of the fluid catalytic cracking unit at the Petrobras's Presidente Bernardes Refinery of Cubatão (RPBC), Brazil. In this process, the C3 stream (propane and propene) is separated from a C4 stream (butane, butene and other hydrocarbons with four atoms of carbon). A schematic diagram of the distillation column with some of the regulatory control loop (PID) is shown in Figure N° 2.

In Figure N° 2, T-01 is the distillation column, E stands for heat exchanger and V designates a process vessel. The feed stream is liquefied petroleum gas (LPG), which comes from the top of a debutanizer column that separates the LPG from gasoline. The C3 stream is produced as the top stream of the distillation column and the C4 stream is produced as the bottom stream of the column. AI1 and AI2 are analyzers that infer the contents of C4 in the propane stream and the contents of C3 in the butane stream, respectively.

This process is controlled by a commercial MPC package as detailed in section 2. The output variables  $y_1$  and  $y_2$  are the temperature (T1) at the first stage of the top section of the distillation column and the percentage of C3 (AI2) in the bottom stream, respectively. The input variables  $u_1$  and  $u_2$  are the flowrate (F3) of hot oil to the reboiler and the reflux flowrate (F2) to the top of the column, respectively.

In this study, based on Porfírio et al (2003), it is assumed that the "real" process and the "old" process model (as used in the MPC system) are given by:

Real process:

$$Gp = \begin{bmatrix} \frac{(-2.218e-4)s + (5.656e-4)}{s^2 + 3.4948s + 0.5902} e^{-1s} & \frac{(0.7413e-4)s - (1.452e-3)}{s^2 + 2.6987s + 0.4023} e^{-1s} \\ \frac{(-1.135e-3)s - (1.235e-3)}{s^2 + 1.6380s + 0.09852} e^{-1s} & \frac{(-0.3e-3)s + (0.2e-2)}{s^2 + 2.4298s + 0.0651} e^{-1s} \end{bmatrix} \quad (5)$$



Old process model:

$$Gm = \begin{bmatrix} \frac{(0.4227e-4)s + (0.1094e-4)}{s^2 + 0.1090s + 0.0243} e^{-1s} & \frac{(-1.2055e-4)s - (0.03824e-3)}{s^2 + 0.1342s + 0.0111} e^{-1s} \\ \frac{(-0.0873e-3)s - (0.1116e-3)}{s^2 + 0.1317s + 0.0073} e^{-1s} & \frac{(1.3e-3)s + (0.7e-2)}{s^2 + 2.2605s + 0.1366} e^{-1s} \end{bmatrix} \quad (6)$$

The main parameters of the controller are taken from Porfirio *et al.* (2003). Other tuning parameters are:  $R_u = \text{diag}[10,10]$ ,  $W_1 = [500, -500]$ ,  $W_2 = \text{diag}[200,200]$ ,  $W_3 = \text{diag}[1000,1000]$ ,  $u_{min} = \text{diag}[2800,1500]$ ,  $u_{max} = \text{diag}[3250,2000]$ ,  $y'_{min} = \text{diag}[0.80,48.25]$ ,  $y'_{max} = \text{diag}[0.95,49.75]$ . The sampling time is  $\Delta t = 1$ . As it can be noted from the input and

output bounds, the operating window of the C3/C4 splitter is quite narrow, which is a characteristic of high-purity distillation systems.

### 3.1 Test signal design and generation of the dataset

The design of the test signal plays a major role in the excitation and identification procedures of the process system. In the closed loop identification strategy proposed here, the purpose is to design a persistently exciting test signal that produces a persistent excitation of the process input.

One of the persistently exciting test signals most used in industrial practice is the pseudo-random binary sequence (PRBS) (Godfrey, 1993). The energy content of this signal is distributed uniformly over the entire frequency range. However, in many situations, as the present one, we are interested in a particular frequency (or frequency range) other than the broad spectrum. The generalized binary noise (GBN) proposed by Tulleken (1990) has a similar approach as the PRBS except an additional parameter related to the switching probability. This allows manipulation of the power spectrum of the test signal, such that most energy can be concentrated in the low-frequency range. Another advantage of GBN is that the signal length is flexible and it has a minimum crest factor. Due to these facilities, GBN is considered as the most suitable signal for control-relevant identification of industrial processes (Zhu, 2001).

Based on guidelines provided by Zhu (2001) and a priori knowledge of the process ( $Gm$ ), two independent GBN signals of magnitude

$\pm 1$  with mean switching times set to 29.4 and 26.13min, respectively, are applied to the system simultaneously as indicated in section II. The duration of the excitation test is 4500min, resulting 4500 samples of input-output data, which are shown in Figure N° 3.

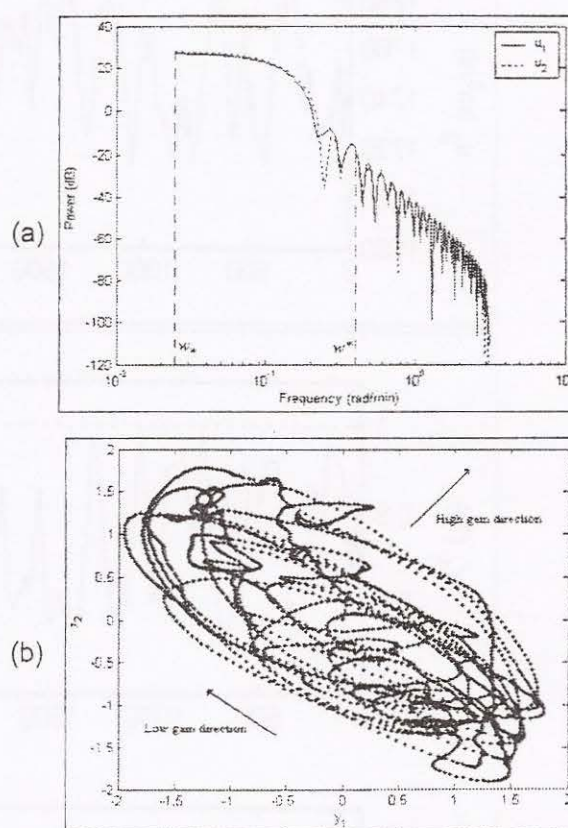


Figure N° 4. (a) power spectrum of the inputs, (b) excitation of output gain directions.

As can be seen in Figure N° 3, the excitation signal that was introduced in the distillation column did not compromise the operating objectives, as both the inputs and outputs were inside the respective boundaries. The strategy of the zone control minimizes the feedback effect on the identification data, because during most of the time of the test, the control action aiming at controlling the outputs was not activated. With this strategy the closed-loop identification procedure approaches the open-loop direct method procedure, which ignores the presence of the controller.



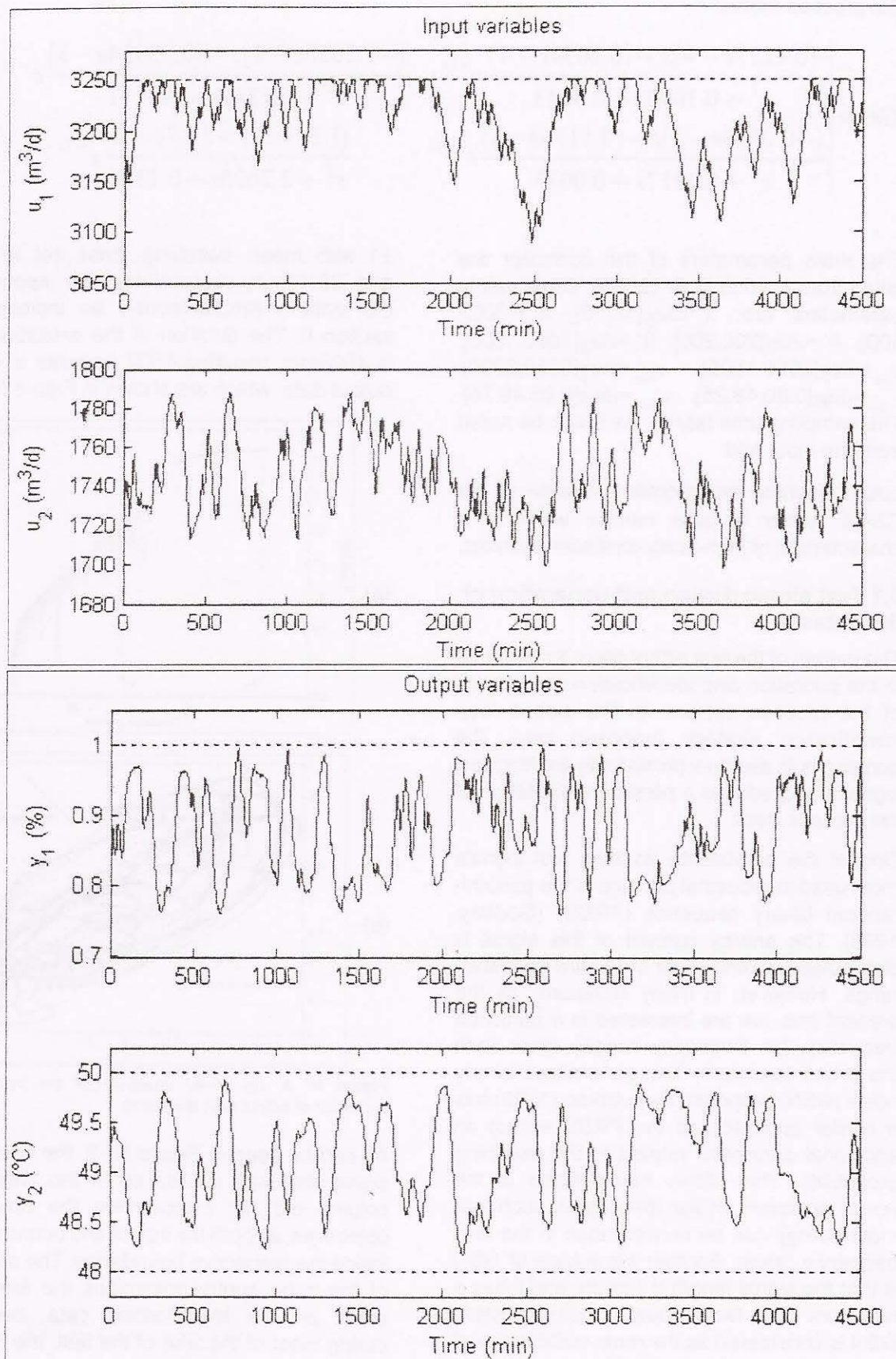


Figure N° 3. Input and output dataset of the C3/C4 splitter

As usual, in any model identification procedure, the simulation data represented in Figure N° 3 should be normalized in order to make the inputs and outputs of the system equally important. This is done by subtracting from each input and output its computed mean and dividing the result by the standard



deviation of that variable. Linear trends were also removed. Finally, the dataset is filtered with a low pass filter aiming to concentrate the identification results in the frequency range of interest. A portion of this dataset (first 3000 samples) is used for identification purposes and another part (remaining 1500 samples) to subsequently validate the estimated models.

To verify the quality of the identification dataset, Figure N° 4 presents the power spectrum of the inputs. The selected part of the spectrum ( $w_1$ ,  $w_2$ ) indicates the desired frequency range where the model will be identified. In fact, a good model in the low and mid frequency ranges is essential for MPC control purposes (Gopaluni *et al.*, 2003). Here, dips at frequencies outside of the indicated bandwidth and other high-frequency contributions were not included in the identification procedure. Figure N° 3 shows also that low and high output gain directions are well excited, allowing a good estimate in both directions.

### 3.2 Model structure selection

All identified models have certain bias (or deterministic error) and variance (or stochastic error). The best model is obtained by minimizing the total error. But, there is a tradeoff

between both parts of the error. By using FIR models, bias error can be minimized (because of the reduced model structure and order limitations) but it can result in higher variance due to the larger number of parameters. On the other hand, by using low-order parametric models, variance error is reduced but bias error is inevitably larger.

In many industrial MPC controllers, the stochastic model is fixed a priori and the 'best approximate process model' within some chosen model structure needs to be estimated. In the present paper, it is considered that the process can be represented by MISO transfer function models, with the stochastic model parameterized as unitary, of the following form:

$$\hat{y}_j(t, \rho_j) = \sum_{i=1}^{n_u} \hat{G}_{j,i}(s, \rho_{j,i}) u_i(t - \theta_{j,i}) + \varepsilon_j(t), \quad j = 1, \dots, n_y \quad (7)$$

where  $\hat{G}_{j,i}(s, \rho_{j,i})$  is the  $(j,i)$  transfer function defined as:

$$\hat{G}_{j,i}(s, \rho_{j,i}) = \frac{\hat{B}_{j,i}(s)}{\hat{A}_{j,i}(s)} = \frac{\sum_{k=0}^{m_{j,i}} \hat{b}_{j,i,k} s^k}{\sum_{k=0}^{n_{j,i}} \hat{a}_{j,i,k} s^k}, \quad \hat{a}_{j,i,n_{j,i}} = 1, \quad n_{j,i} \geq m_{j,i} \quad (8)$$

$u(t)$  is the input vector,  $\hat{y}(t, \rho)$  is the model output,  $\theta_{j,i}$  is the time-delay between the  $i$  input and the  $j$  output,  $\varepsilon(t)$  is the residual or total model error (bias plus variance),  $n_u$  and  $n_y$  are the number of inputs and outputs, respectively, and

$$\rho_{j,i} = \left[ \hat{b}_{j,i,m_{j,i}} \dots \hat{b}_{j,i,0} \hat{a}_{j,i,(n_{j,i}-1)} \dots \hat{a}_{j,i,0} \right]^T \in \mathbb{R}^{p_{j,i}}$$

with  $p_{j,i} = n_{j,i} + m_{j,i} + 1$ , where  $n_{j,i}$  and  $m_{j,i}$  denotes the denominator and numerator orders of  $\hat{G}_{j,i}(s, \rho_{j,i})$ , respectively. Therefore, the parameter vector is

$$\rho_j = \left[ \rho_{j,1}^T \dots \rho_{j,n_u}^T \right]^T \in \mathbb{R}^{p_j \times 1}, \quad p_j = \sum_{i=1}^{n_u} p_{j,i} \quad (9)$$

### 3.3 Identification procedure

The goal is to build a model as defined in Equation (7) based on closed-loop sampled data, focusing on the parameters of each transfer function  $\hat{G}_{j,i}(s, \rho_{j,i})$  rather than on the model error appearing in (7). The pure time-delay  $\theta_{j,i}$  is supposed to be known a priori and to be a multiple of the sampling time  $\Delta t$ . Several procedures have been proposed, which are developed as extensions of the existing procedures that deal with SISO systems and that allow the identification of transfer functions with a common denominator. This approach may not be realistic in several practical applications. In this work, it is used the CONTSID toolbox (Garnier *et al.*, 2006) that presents new methods in the time-domain to deal with processes described by MISO systems with several transfer functions with different denominators.

The output-error (OE) method from the CONTSID toolbox is used to find the parameter vector  $\rho_j$  (9) for the C3/C4 splitter, with  $n_u = n_y = 2$ ,  $n_{j,i} = 2$  and  $m_{j,i} = 1$ . The identification is carried out off-line using the identification dataset and considering the values of the parameters of the old process model (6) as the initial estimation for the identification algorithm. The models are evaluated based on the following performance criteria:

$$FIT = 100 \times \left[ 1 - \frac{\text{norm}(y_j - \hat{y}_j)}{\text{norm}(y_j - \text{mean}(y_j))} \right] \quad (10)$$

$$R_T^2 = 1 - \frac{\text{var}(y_j - \hat{y}_j)}{\text{var}(y_j)} \quad (11)$$



where is the true system output and the model output. Coefficient  $FIT$  indicates the percentage of the output variation that can be associated to the model, while coefficient  $R_r^2$  measures how well the model output explains the behavior of the system output, and this parameter will be close to 1 in low noise conditions.

Simulation results of the new and old models are presented in Figure N° 5. Performance indicators and visual inspection show that the new model matches much better the system outputs than the available model in the MPC package. Loss function and Akaike's final prediction error (FPE) associated to each new model are also shown.

### 3.4 Model validation

Here, the goal is to confirm whether the obtained model is an accurate representation of the process and to provide a basis for possible re-identification. The models can be validated in a variety of ways. Residual correlation test, Bode frequency response, pole-zero plots and cross-validation are often used. For the case of the C3/C4 splitter, we will discuss some aspects as model error modeling, step response, cross-validation and interaction analysis.

#### 3.4.1 Model error modeling

It consists in building models that describe the dynamics relating the inputs to the residual, in order to verify, in the frequency domain, if any essential unmodeled dynamics are left. This idea has attracted much interest in the past decade, motivated by the linear robust control theory, and because it provides more freedom in investigating the residuals than the classical residual correlation test (Ljung, 1999b).

From Equation (7), we can interpret the residual as:

$$\varepsilon_j(t, \rho_j^e) = y_j(t) - \hat{y}_j(t, \rho_j) = \sum_{i=1}^{n_j} \hat{G}_{j,i}^e(s, \rho_{j,i}^e) u_i(t - \theta_{i,j}) \quad (12)$$

where  $\hat{G}_{j,i}^e(s, \rho_{j,i}^e)$  is the model of the error corresponding to model  $\hat{G}_{j,i}(s, \rho_{j,i})$ , with denominator and numerator orders denoted by  $n_{j,i}^e$  and  $m_{j,i}^e$  ( $n_{j,i}^e \geq m_{j,i}^e \gg n_{j,i}$ ) respectively, and parameter vector  $\rho_{j,i}^e$  defined in a similar way as in Equation (9).

In this study, the equation-error (EE) method from the CONTSID toolbox is used to compute the parameter vector  $\rho_{j,i}^e$ , with  $n_{j,i}^e = m_{j,i}^e = 10$ . This is performed using the residual as the output and the system inputs taken from the

identification dataset as the inputs. Figure N° 6 shows the Bode plot of the models of the error and their confidence regions (3 standard deviation of the normal distribution). As it can be observed, the new models  $\hat{G}_{j,i}^e(s, \rho_{j,i}^e)$  are satisfactory within 99.6% probability at the frequency range of interest. The new model is compared with the old model, which contains large uncertainties.

#### 3.4.2 Step response

In the plant-friendly identification procedure for MPC, step response is a logical approach for model validation. Figure N° 7 shows the step response comparison between the old model and the new model. As it can be noted, the new model captures the steady-state gains and the time constants of the process correctly. This means that using the new model the "dynamic matrix", which contains the step response coefficients corresponding to the inputs to the C3/C4 process, will be more successful than using the old model.

#### 3.4.3 Cross-validation

It is one of the most important and revealing tests for model validation. In this case, the validation dataset, which was not used in the identification procedure, is used to compare the model predictions. Cross-validation results from the old and new models and its respective performance indicators are presented in Figure N° 8. These results show that the new model reproduces almost perfectly the dominant dynamics of the C3/C4 system.

#### 3.4.4 Interaction analysis

As a final test, it is studied the interaction behavior of the C3/C4 models as a function of the frequency range in which we are interested. Here, we use the relative dynamic gain array number (RDGA-number) as an interaction measure, which is based on the Bristol's RGA, defined as (Skogestad & Postlethwaite, 1996):

$$RDGA - number(j\omega) = \left\| RDGA(j\omega) - I \right\|_{sum} \quad (13)$$

where:

$$RDGA(j\omega) = G(j\omega) \otimes \left( G(j\omega)^\dagger \right)^T \quad (14)$$

with  $(\dagger)$  denoting the Moore-Penrose pseudo-inverse. If the RDGA-number is close to zero, there is no interaction in the system (or it is negligible).



As noticed in Figure N° 9, the new model has almost the same RDGA-number profile as the C3/C4 true model, i.e. there is interaction in the whole frequency range of interest, which plays an important role in MPC control. The RDGA-number of the old model reveals less dynamic interaction than in the real plant. From medium to high frequencies, the interaction level corresponding to the old model decreases rapidly which may be inadequate for the purpose of multivariable control.

#### IV. ASSESSMENT OF MPC PERFORMANCE BASED ON PREDICTABILITY OF MODEL ERROR

Several methods have been proposed for MPC performance assessment (Sotomayor & Odloak, 2006). In this section, a methodology is proposed, which uses results from the cross-validation step. The method is based on Thornhill et al. (1999) that propose a method for evaluation of the regulatory control performance. However, instead of evaluating the controller behavior, the methodology is used to evaluate the model by analyzing the time series of the observed error and to verify the existence of a predictable pattern beyond the control horizon for each of the outputs of the models. The comparison is performed in term of variances of the residues.

The approach considers that if the MPC is to perform well, then its internal model should be able to predict the output of the process along a prediction horizon  $p$ , i.e. the model error should have no predictable components or it should not be distinguishable from a random walk stochastic process after time instant  $t+p$ . Thus, the proposed methodology

demands a method to make predictions of the error model  $p$ -steps ahead of the present time. Here, it is used an autoregressive time series model of the following form: (15)

$$\hat{\varepsilon}_j(t+p) = a_0 + a_1 \varepsilon_j(t) + a_2 \varepsilon_j(t-1) + \dots + a_{n_a} \varepsilon_j(t-n_a+1)$$

where  $\varepsilon_j(t)$  is the model error described in Equation (12)  $n_a$  and is the order of the autoregressive model. Parameters  $a_i$  in (15) can be fitted to a given dataset of  $n$  samples of the model error using a least squares fit procedure:

$$\begin{bmatrix} a_0 & a_1 & \dots & a_{n_a} \end{bmatrix}_j^T = X_j^\dagger Y_j \quad (16)$$

where: (17)

$$X_j = \begin{bmatrix} 1 & \varepsilon_j(1) & \varepsilon_j(2) & \dots & \varepsilon_j(n_a) \\ 1 & \varepsilon_j(2) & \varepsilon_j(3) & \dots & \varepsilon_j(n_a+1) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & \varepsilon_j(n-p-n_a+1) & \varepsilon_j(n-p-n_a+2) & \dots & \varepsilon_j(n-p) \end{bmatrix}$$

$$Y_j = [\varepsilon_j(n_a+p) \quad \varepsilon_j(n_a+p+1) \quad \dots \quad \varepsilon_j(n)]^T \quad (18)$$

The difference between the actual and predicted model errors is the residue  $r_\varepsilon(t)$ , which mean and variance provide relevant information regarding the predictability of the model, or the model behavior.

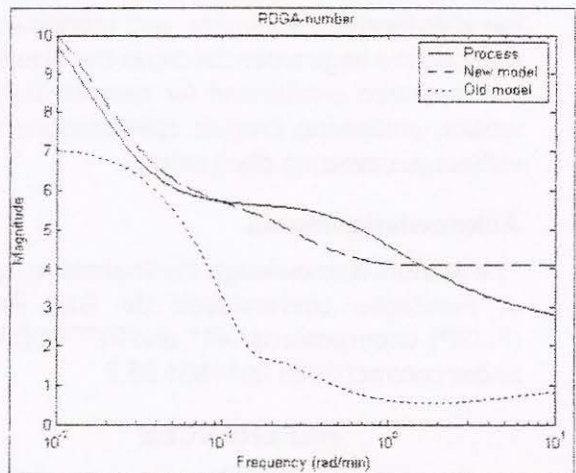


Figure N° 9. RDGA-number plot for the C3/C4 splitter

In the present application, the model error, the residue and the prediction of the model error of the old and new models are shown in Figure N° 10, which are generated using the validation dataset and parameter values  $n_a=30$  and  $n=1500$  (although only the first 500 points are plotted in Figure N° 10). Parameter  $p$  is the same as used in the MPC controller. In general, the new model presents much smaller variability of the residues and, therefore, a superior performance than the current model used in the controller. Particularly, the major variability of the residues of the old model corresponding to

output  $y_2$  indicates that it is a very poor model, as it was demonstrated by the results of the previous section. The methodology shown above can be used as a tool to decide the need or not of re-identifying the model of the process.

#### V. CONCLUSIONS

In this work, it has been proposed a methodology to re-identify the process model in a closed loop strategy for multivariable industrial processes,



which are controlled by MPC packages based on the two layer structure. The test signal is introduced in the economic coefficients of objective function of the economic layer of the controller. The test signal is then translated into optimal targets to the inputs of the system in the MPC algorithm, which uses the concept of zone control for the outputs.

The proposed approach was tested by simulation in an industrial C3/C4 separation column. The results show that the proposed method is very promising in terms of future practical application. The main feature of the method is that considerable time is saved in the identification procedure, and manpower is reduced in a large extent because the algorithm can operated unattended for several days or weeks, preserving product specifications and without jeopardizing plant safety.

### Acknowledgements

The authors acknowledge the financial support of Fundação Universidade de São Paulo (FUSP), under process 1411, and PETROBRAS under contract 0050.0014884.05.2.

### REFERENCES

- Andersen, H.W., M. Kümmel (1992). Evaluating estimation of gain directionality – Part 2: A case study of binary distillation. *Journal of Process Control*, **2**(2): 67-86.
- de Klerk, E., I.K. Craig (2003). Multivariable closed-loop system identification of plants under model predictive control. In: *13th IFAC Symposium on System Identification (SYSID 2003)*, CD Rom, Rotterdam, The Netherlands.
- Forssell, U., L. Ljung (1999). Closed-loop identification revisited. *Automatica*, **35**(7): 1215-1241.
- Garnier, H., M. Gilson, O. Cervellin (2006). Latest developments for the Matlab CONTSID toolbox. In: *14th IFAC Symposium on System Identification (SYSID 2006)*, CD Rom, Newcastle, Australia.
- Genceli, H., M. Nikolaou (1996). New approach to constrained predictive control with simultaneous model identification. *AIChE Journal*, **42**(10): 2857-2868.
- Godfrey, K.R. (1993). *Perturbation signals for system identification*. Prentice Hall: Hertfordshire, UK.
- Gopaluni, R.B., R.S. Patwardhan, S.L. Shah (2003). The nature of data pre-filters in MPC relevant identification – Open and closed-loop issues. *Automatica*, **39**(9): 1617-1626.
- Gustavsson, I., L. Lung, T. Söderström (1977). Identification of processes in closed loop – Identifiability and accuracy aspects. *Automatica*, **13**(1): 59-75.
- Hjalmarsson, H., M. Gevers, F. De Bruyne (1996). For model-based control design, closed-loop identification gives better performance. *Automatica*, **32**(12): 1659-1673.
- Kline, S., R. Street, R.S. Patwardhan, A. Bae (2006). Closed loop identification for multivariable model predictive controller on a catalytic gasoline splitter. In: *7th International Conference on Chemical Process Control (CPC 7)*, Lake Louise – Alberta, Canada.
- Landau, I.D. (2001). Identification in closed loop: a powerful design tool (better design models, simpler controllers). *Control Engineering Practice*, **9**(1): 51-65.
- Li, W., J.H. Lee (1996). Frequency-domain closed-loop identification of multivariable systems for feedback control. *AIChE Journal*, **42**(10): 2813-2827.
- Ljung, L. (1999a). *System identification – Theory for the user*. 2<sup>nd</sup> Ed., Prentice Hall: Upper Saddle River, NJ.
- Ljung, L. (1999b). Model validation and model error modeling. Technical Report N°LiTH-ISY-R-2125, Linköping University, Sweden.
- Maciejowski, J.M. (2002). *Predictive control with constraints*. Pearson Education Limited: Harlow, England.
- Porfirio, C.R., E.A. Neto, D. Odloak (2003). Multi-model predictive control of an industrial C3/C4 splitter. *Control Engineering Practice*, **11**(7): 765-779.
- Qin, S.J., T.A. Badgwell (2003). A survey of industrial model predictive control technology. *Control Engineering Practice*, **11**(7): 733-764.
- Rivera, D.E., M.E. Flores (1999). "Plant-friendly" closed-loop identification using model predictive control. In: *1999 AIChE Annual Meeting*, paper 227d, Dallas – TX.



19. Saffer, D.R., F.J. Doyle (2002). Closed-loop identification with MPC for an industrial scale CD-control problem. *IEE Proceedings – Control Theory and Applications*, **149**(5): 448-456.
20. Sogias, D., M. Nikolaou (2001). Closed-loop identification of processes operating under model predictive control without external excitation. In: *2001 AIChE Annual Meeting*, paper 279i, Reno – NV.
21. Skogestad, S., I. Postlethwaite (1996). *Multivariable feedback control – Analysis and design*. John Wiley & Sons Ltd.: Chichester, England.
22. Sotomayor, O.A.Z., D. Odloak (2006). Performance assessment of model predictive control systems. In: *8th IFAC International Symposium on Advanced Control of Chemical Processes (ADCHEM 2006)*, CD Rom, Gramado-RS, Brazil.
23. Thornhill, N.F., M. Oettinger, P. Fedenczuk (1999). Refinery-wide control loop performance assessment. *Journal of Process Control*, **9**(2): 109-124.
24. Treiber, S., R. Stelmaier, M. Starling (2003). Sustaining benefits of process automation. In: *2003 AIChE Spring National Meeting*, New Orleans – LA.
25. Tulleken, H.J.A.F. (1990). Generalized binary noise test-signal concept for improved identification-experiment design. *Automatica*, **26**(1): 37-49.
26. Zhu, Y. (2001). *Multivariable system identification for process control*. Elsevier Science Ltd.: Oxford, UK.
27. Zhu, Y., F. Butoyi (2002). Case studies on closed-loop identification for MPC. *Control Engineering Practice*, **10**(4): 403-417.

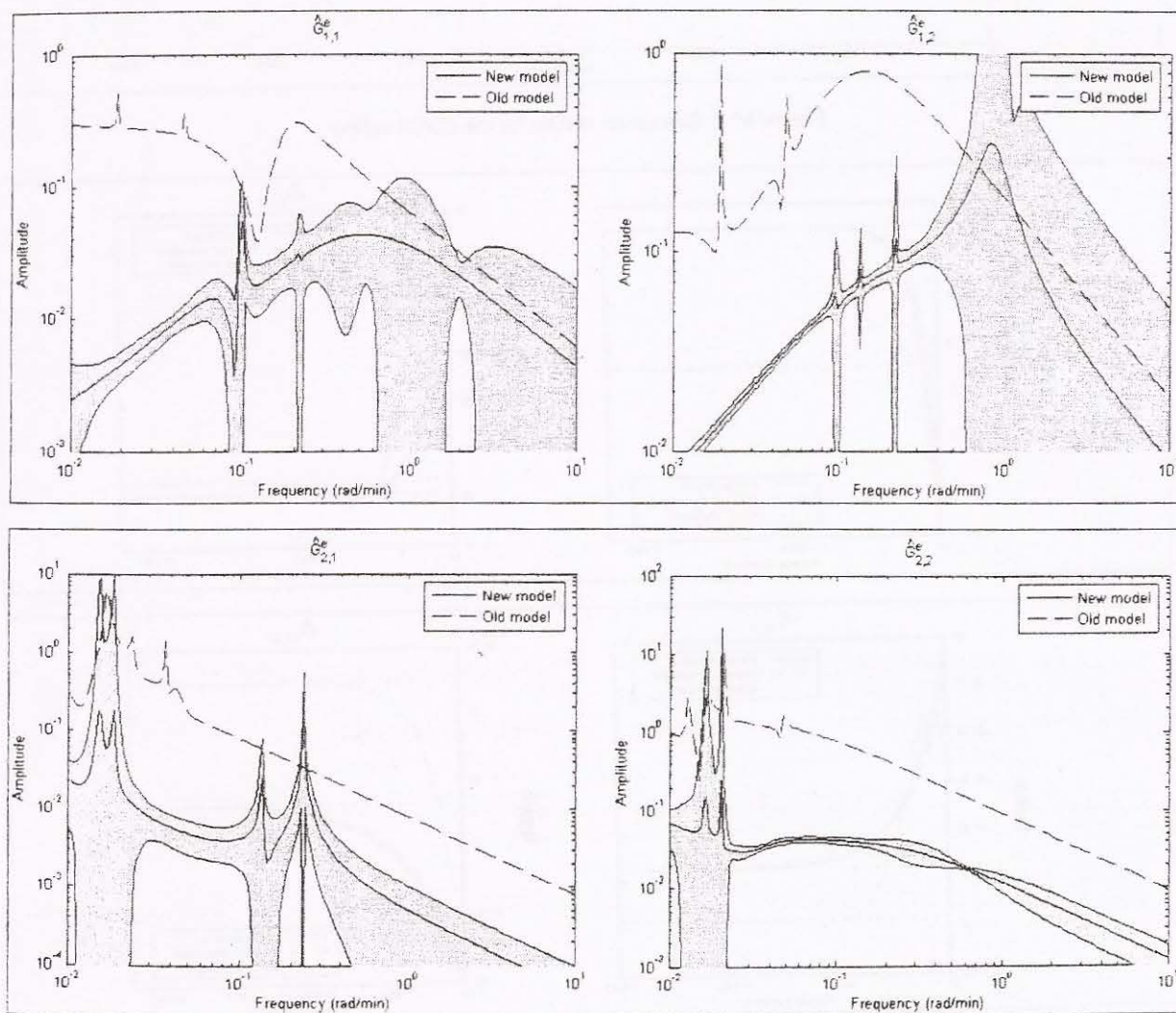


Figure N° 6. C3/C4 model error models and its confidence regions



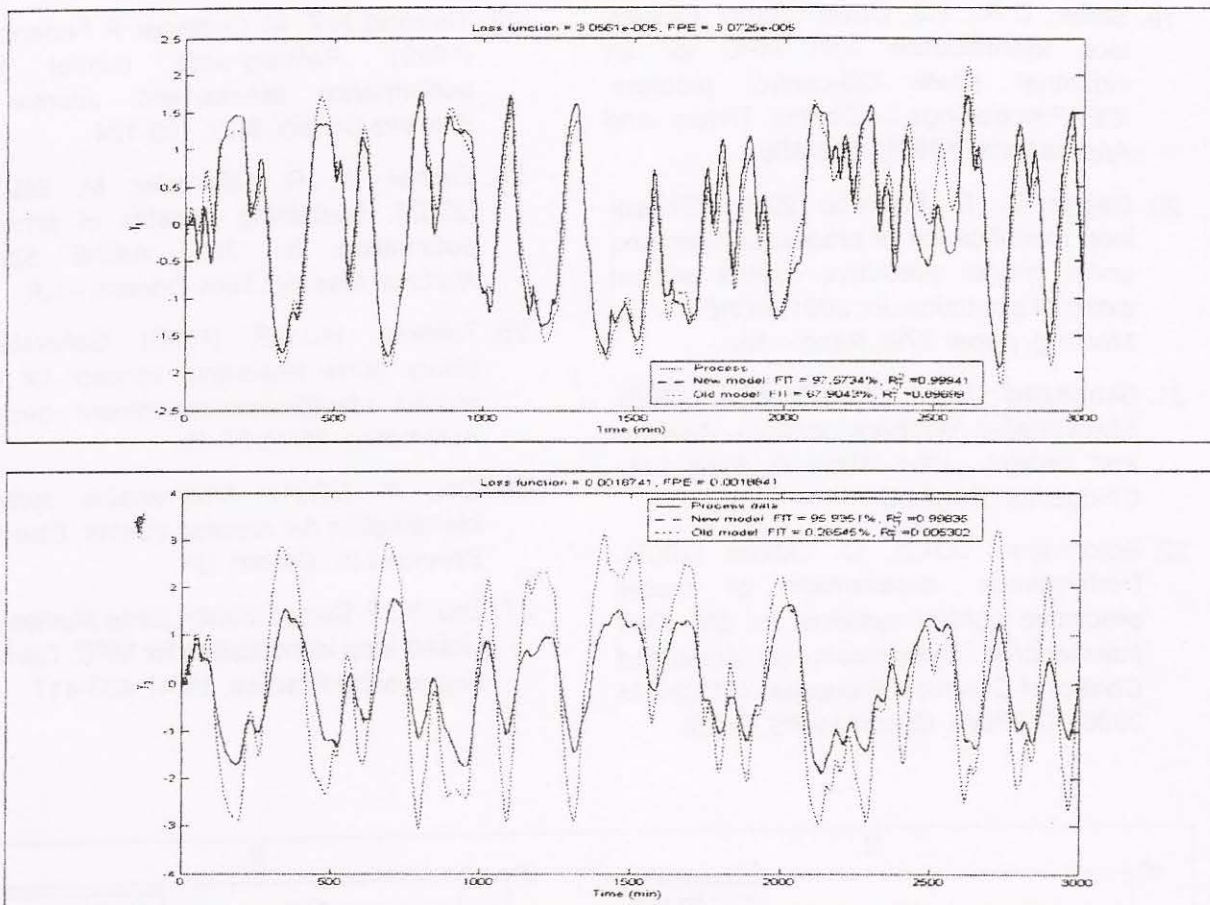


Figure N° 5. Simulation results for the C3/C4 splitter

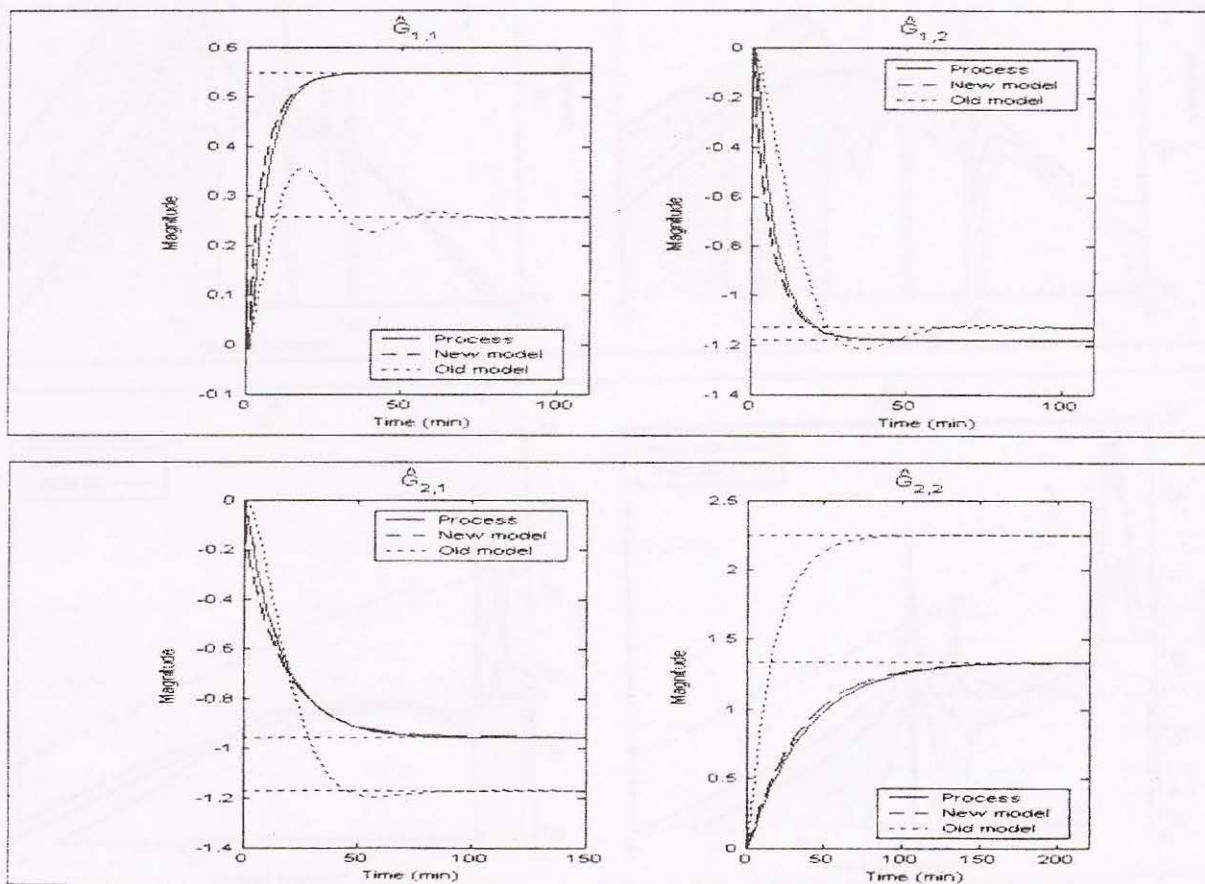


Figure N° 7. Step response for the C3/C4 splitter



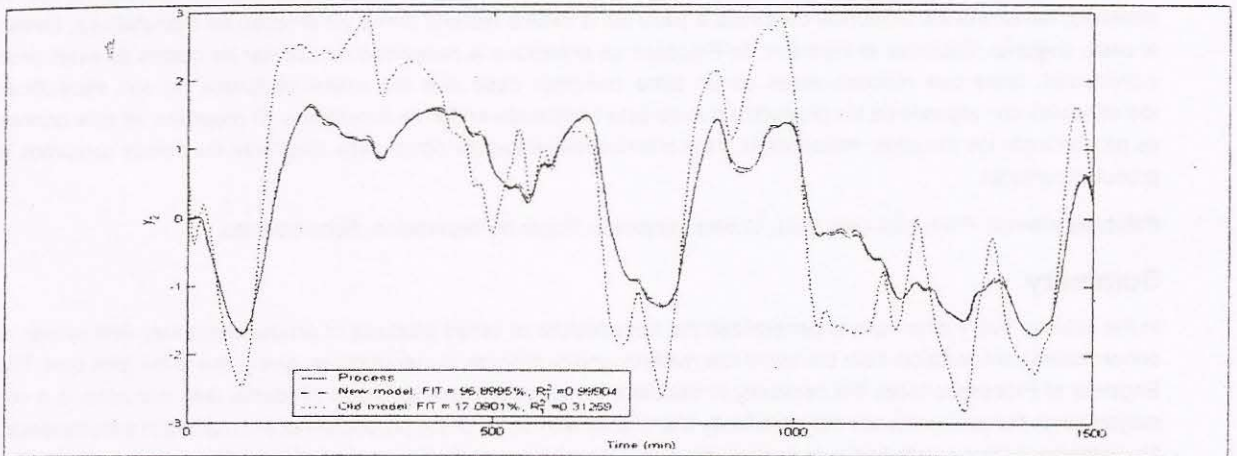
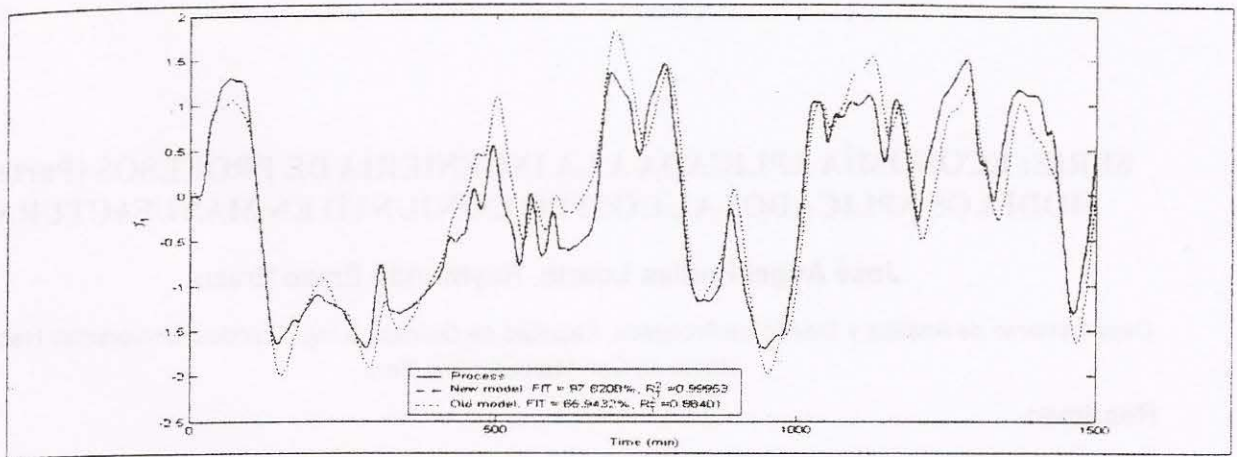


Figure N° 8 Cross-validation results for the C3/C4 splitter

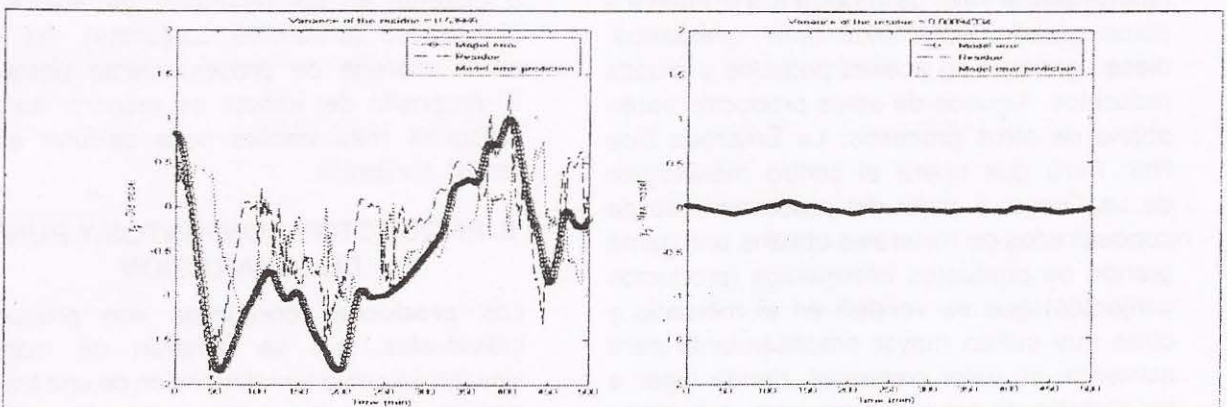
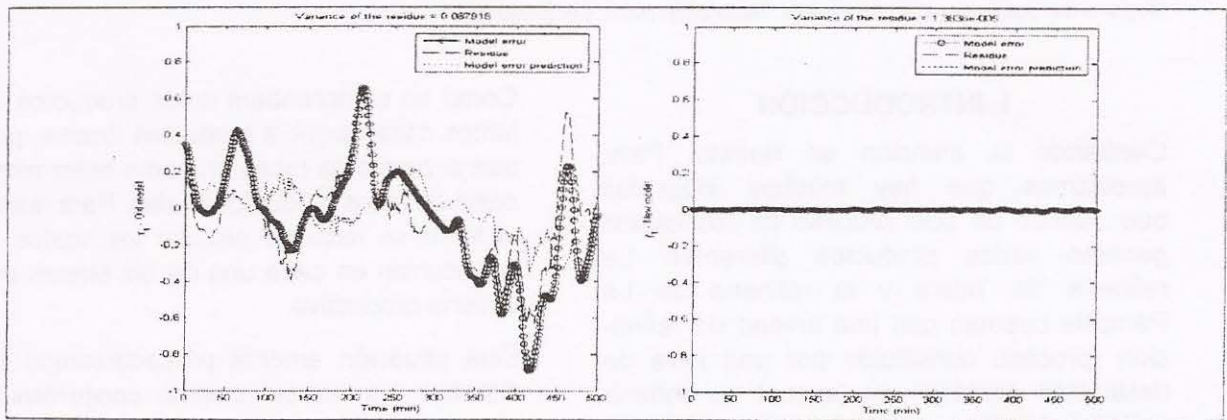


Figure N° 10. Predictability of the C3/C4 model errors