ELSEVIER

Chemical Engineering Journal 112 (2005) 93–108

Chemical Engineering Journal

www.elsevier.com/locate/cej

Observer-based fault diagnosis in chemical plants

Oscar A.Z. Sotomayor, Darci Odloak*

University of São Paulo, Department of Chemical Engineering, Av. Prof. Luciano Gualberto, trav.3, n.380, 05508-900 São Paulo SP, Brazil

Received 22 March 2005; received in revised form 30 June 2005; accepted 6 July 2005

Abstract

The main focus of this paper is on the study and application of existing methodology on unknown input observer (UIO) for on-line diagnosis of faults on input actuators, output sensors, model parameters and disturbances of complex chemical plants operating under model predictive control (MPC). Two industrial systems are studied by simulation: styrene polymerization reactor and fluid catalytic cracking (FCC) unit. For each of these cases, the development of the method is presented and the design of the fault diagnosis system is discussed. In the first case, a bank of reduced-order UIOs is used for fault diagnosis of process parameters and external disturbances. The design is based on the rigorous first principles model of the polymerization process. In the second case, a single bank of full-order UIOs is used for sensor and actuator fault diagnosis based on an input–output linear model of the FCC unit. In both cases, extensive simulation results are presented and discussed.

Keywords: Fault diagnosis; Unknown input observers; Analytical redundancy; Polymerization process; Fluid catalytic cracking unit; Model predictive control

1. Introduction

The size and complexity of modern industrial plants have led to the adoption of MPC as the standard control solution for most processes. At the same time, the integration of large process plants in a single centralized MPC controller has increased the vulnerability of these systems to abnormal occurrences, which can significantly degrade the performance of the overall control system. Frequently, the MPC system will hide a gradual incident until the control failure becomes unavoidable. In other cases, the multivariable character of the MPC controller can amplify a local incident that is propagated to the whole process system resulting in a premature plant shutdown. Therefore, an important issue to achieve high performance, efficiency, reliability and safety of large scale processes is to supervise the centralized MPC controller, i.e. to diagnose faults in the control system while the plant is still operating in a controllable region.

As any control strategy, MPC is usually vulnerable to malfunctioning of sensors and actuators, which can be considered as additive faults. For example, a biased sensor may lead the process to an operating point far from the optimal one and can cause saturation of a manipulated valve. Moreover, sensor faults make the plant partially unobservable, while actuator faults make the plant partially uncontrollable. The supervisory system may help MPC to deal with these abnormal situations. In this sense, the fault diagnosis system will be seen as part of a larger strategy for optimal process control.

The fault diagnosis area has grown as an active research topic for more than 30 years. Extensive reviews of different fault diagnosis methods can be found in the literature [1–7]. Existing methods can be grouped into three general categories: quantitative model-based methods, qualitative model-based methods and data-driven methods. Our interest is on the class of quantitative model-based techniques, namely observer-based approach, parameter estimation approach and parity-space approach, which have received considerable attention in recent years. These approaches are based on the concept of analytical or functional redundancy, which makes use of a mathematical model of the process to obtain behavior estimates. The inconsistencies between the estimated and actual behavior are symptoms or fault indicators. They are called residuals and may reflect the abnormal situation of the supervised process. Later, the residuals are evaluated aiming at to localize the fault. Although there is a close relationship among the quantitative model-based

* Corresponding author. Tel.: +55 11 30912237; fax: +55 11 38132380.
E-mail addresses: oscar@pqi.ep.usp.br (O.A.Z. Sotomayor), odloak@usp.br (D. Odloak).

1385-8947/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.cej.2005.07.001

techniques, observer-based approach have become the most popular and important method for model-based fault diagnosis [8,9], especially within the automatic control community.

We do not find in the literature a significant number of applications of observers for fault diagnosis in chemical processes of industrial relevance [10] use the unknown input observer (UIO) in a continuous stirred tank reactor (CSTR), while [11,12] apply the extended Kalman filter (EKF) in a two-tank system and a fluid catalytic cracking (FCC) unit, respectively. A common feature in these works is that the fault diagnosis is performed in processes operating in open loop. However, most chemical processes operate in closed-loop, and the control action, due to the feedback, may affect the fault diagnosis system in detriment of its performance. There are a few papers dealing with monitoring of processes operating under conventional regulatory control, e.g. see [13] that use the EKF in a binary distillation column [14] that apply the generalized Luenberger observer (GLO) in a pre-heating tank and [15] that use the EKF in an exothermic CSTR. However, we could not find studies on the use of observer-based fault diagnosis systems in processes operating under advanced control, namely model predictive control (MPC).

The basic idea of this paper is to show the application and to evaluate the practical feasibility of UIO for on-line fault diagnosis in complex chemical plants operating under MPC. Two industrial systems are studied: the styrene polymerization reactor and the FCC reactor–regenerator unit. The application of UIOs to such systems is not trivial and there are some challenges in the design of a diagnosis system that is able to achieve a satisfactory performance. In the first case, a bank of reduced-order UIOs is used for process parameter fault and external disturbance diagnosis based on the linearized model of the polymerization process. In the second case, a single bank of full-order UIOs is used for sensor and actuator fault diagnosis based on a linear model obtained from input–output data of the FCC unit. The paper is organized as follows. Section 2 presents preliminary concepts related to observer-based fault diagnosis systems. Section 3 describes the full and a reduced-order UIO for fault detection purposes. The description of the UIO versions, as presented here, aims at a better understanding and comprehension of this subject by the reader in order to facilitate the practical implementation. Section 4 presents the application of the fault diagnosis to the styrene polymerization reactor and Section 5 discusses a similar application to an industrial FCC unit. Finally, conclusions are reported in Section 6.

2. Observer-based fault diagnosis

Fault diagnosis is usually performed to accomplish one or more of the following tasks: fault detection (or monitoring), to indicate the fault occurrence; fault isolation, to determine the exact location of the fault and fault identification, to estimate the fault magnitude. The design of an observer-based fault diagnosis system is normally carried out following a three-stage procedure.

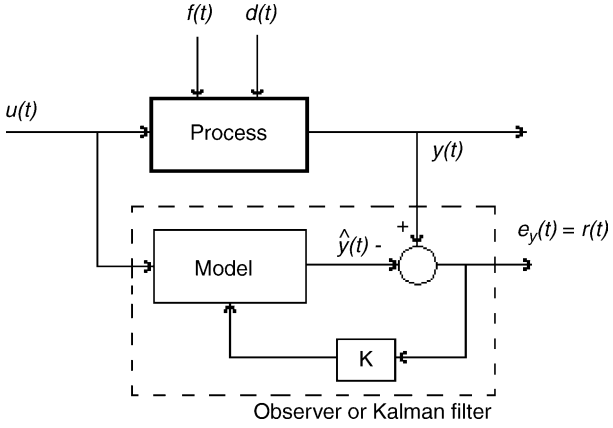

Fig. 1. Scheme of a linear observer for residual generation.

2.1. Residual generation

This is the core element of a fault diagnosis system. It consists in estimating the process output by using either a Luenberger observer in a deterministic setting or a Kalman filter in a stochastic setting. The estimation error (or innovation in the stochastic case) is defined as the residual. Observer-based fault detection makes use of the disturbance decoupling principle, in which the residual is computed assuming the decoupling of the effects of faults on different inputs. For the purpose of fault isolation, it is also assumed that the effect of a fault is decoupled from the effects of other faults. A way to achieve this is by using the so-called unknown input observer (UIO) or unknown input filter (UIF) for the deterministic or stochastic cases, respectively Patton and Chen (1997) [8].

The basic idea of a linear observer-based residual generator is illustrated in Fig. 1. The observer consists of a model parallel to the process with a feedback of the output estimation error, $e_y(t) = y(t) - \hat{y}(t)$.

In Fig. 1, u and y denote the input and output vectors, respectively, f the vector of faults to be detected and d is the vector of unknown inputs, to which the detection system should be insensitive. Variable $\hat{y}$ corresponds to the vector of estimated outputs, r the residual vector and K is the observer gain. To provide useful information for fault diagnosis, the residual should be defined in such a way that

$r(t) = 0$ (or $r(t) \approx 0$), if $f(t) = 0$ (fault-free case),

$r(t) \neq 0$, if $f(t) \neq 0$ (faulty case)

2.2. Residual evaluation

The residual is examined in terms of the likelihood of a fault, and a logical decision-making process is then applied aiming at to decide if the fault has occurred and avoid wrong

decisions, such as false alarm and fault ignored. There are several decision-making methods, but they all end in a simple binary decision variable S_r, where it is used a fixed or adaptive threshold $T(t)$ on a residual evaluation function $J(r(t))$:

$$J(r(t)) \leq T(t), \quad \text{for } f(t) = 0 \text{ then } S_r = 0 \text{ (fault-free case)},$$
$$J(r(t)) > T(t), \quad \text{for } f(t) \neq 0 \text{ then } S_r = 1 \text{ (faulty case)}$$

2.3. Fault analysis

Consists in determining the magnitude and time-variant behavior of the fault. This subject itself has not gained enough research attention, but it can be of crucial importance if control reconfiguration is required.

3. Fault detection with unknown input observers

The unknown input observer (UIO) is a generalization of the Luenberger observer, which is here designated as the unknown input fault detection observer (UIFDO). In this section, two linear versions of UIFDO, the full-order and the reduced-order, are detailed.

Considering that malfunctioning can be caused by faults on plant components as sensors and actuators or by unknown disturbances, we represent the process by a state-space model in the following form:

$$x(k+1) = Ax(k) + Bu(k) + Ef(k), \qquad y(k) = Cx(k) \tag{1}$$

where $x \in \mathbb{R}^{n\times 1}$ is the state, $u \in \mathbb{R}^{m\times 1}$ the input, $y \in \mathbb{R}^{l\times 1}$ the output, $f \in \mathbb{R}^{g\times 1}$ the fault (including disturbances) and k is the discrete sampling instant. $A \in \mathbb{R}^{n\times n}$ is the system matrix, $B \in \mathbb{R}^{n\times m}$ the input matrix, $C \in \mathbb{R}^{l\times m}$ the output matrix and $E \in \mathbb{R}^{n\times g}$ the fault distribution matrix, which is assumed to be known. For the purpose of fault isolation, vector f is partitioned into $f \equiv [f_1 f_2]^T$, where f_1 contains the faults that will not be detected by the fault detector and f_2 contains the faults that will be monitored by the fault detection system. Also, matrix E is partitioned into $E \equiv [E_1 E_2]$.

3.1. Full-order UIFDO

Following [16], a full-order UIFDO that decouples f_1 from the rest of system (1) has the following form:

$$z(k+1) = Fz(k) + TBu(k) + K_{12}y(k),$$
$$\hat{x}(k) = z(k) + Hy(k) \tag{2}$$

where $z \in \mathbb{R}^{n\times 1}$ is the state of the UIFDO, obtained by the linear transformation $z = Tx$, and $\hat{x}$ is the estimated state vector, whilst F, T, K and H are matrices that will be designed such that the unknown input will be decoupled from other inputs, and certain design requirements will be attended. The state estimation error ($e_x = x - \hat{x}$) and the residual are defined by

$$\begin{aligned} e_x(k+1) = {} & (A - HCA - K_1C)e_x(k) \\ & + ((A - HCA - K_1C) - F)z(k) \\ & + ((I - HC) - T)Bu(k) \\ & + ((A - HCA - K_1C)H - K_2)y(k) \\ & + (I - HC)E_1 f_1(k) + (I - HC)E_2 f_2(k) \end{aligned} \tag{3}$$

$$r(k) = y(k) - C\hat{x}(k) = (I - CH)y(k) - Cz(k) \tag{4}$$

If the following relationships are true:

$$T = I - HC, \qquad TE_1 = 0, \qquad F = TA - K_1C,$$
$$K_2 = FH, \qquad K_{12} = K_1 + K_2 \tag{5}$$

then Eq. (3) becomes

$$e_x(k+1) = Fe_x(k) + TE_2 f_2(k) \tag{6}$$

If the eigenvalues of F are stable and assuming $f_2 = 0$, e_x will approach zero asymptotically, i.e. $\hat{x} \to x$, and, therefore, $r(k) = Ce_x(k)$. This means that, the residual defined in (4) will be insensitive to the faults represented in f_1.

Assuming that the conditions defined in (5) hold true, a particular solution to the fault detection problem in terms of matrix H is given by [16]:

$$H = E_1(CE_1)^+ \tag{7}$$

where $(^+)$ denotes the Moore–Penrose pseudo-inverse. The following theorem states the existence conditions for the full-order UIFDO.

Theorem 3.1. *The necessary and sufficient conditions for the existence of a full-order UIFDO for the system defined by (1) are* [16]:

(i) $rank(CE_1) = rank(E_1)$,
(ii) (C, TA) *is a detectable pair*.

3.2. Reduced-order UIFDO

Full-order observers, as presented in Section 3.1, allow the estimation of the whole state vector. However, in many applications, to know all the components of the state may not be really necessary or not even possible as when condition (ii) in Theorem 3.1 is not attended. This may happen when the linear state space model is obtained by linearizing a complex nonlinear model of large dimension. In this case, if the observation of only a part or a linear combination of the states is sufficient for residual generation, the solution to the observer problem can be obtained under less restrictive existence conditions than those imposed by the full-order observer. Several approaches to the design of the reduced-order UIFDO have been proposed in the literature. Here, we follow the method of Hou and Müller [17], to produce a reduced-order UIFDO to decouple f_1 from other faults in system (1). For this purpose,

consider the non-singular transformation matrix $T = U_1^{\mathrm{T}}$, where U_1 is obtained from the singular value decomposition (SVD) of E_1, i.e. $E_1 = U_1 \begin{bmatrix} \Sigma_1 \\ 0 \end{bmatrix} V_1^{\mathrm{T}}$. Thus, applying the state transformation $z = Tx$ to system (1) results:

$$z(k+1) = TAT^{-1}z(k) + TBu(k) + TE_1 f_1(k) + TE_2 f_2(k),$$
$$y(k) = CT^{-1}z(k) \tag{8}$$

Consider now the following partitions:

$$z(k) = \begin{bmatrix} z_1(k) \\ z_2(k) \end{bmatrix}, \qquad TAT^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

$$TB = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \qquad CT^{-1} = \begin{bmatrix} C_1 & C_2 \end{bmatrix},$$

$$TE_1 = \begin{bmatrix} E_{11} \\ 0 \end{bmatrix}, \qquad TE_2 = \begin{bmatrix} E_{21} \\ E_{22} \end{bmatrix}$$

The transformed system can be separated in two subsystems as follows:

$$z_1(k+1) = A_{11}z_1(k) + A_{12}z_2(k) + B_1u(k) + E_{11}f_1(k) + E_{21}f_2(k) \tag{9}$$

$$z_2(k+1) = A_{21}z_1(k) + A_{22}z_2(k) + B_2u(k) + E_{22}f_2(k) \tag{10}$$

$$y(k) = C_1z_1(k) + C_2z_2(k) \tag{11}$$

From Eq. (9), we observe that subsystem 1 is affected by both f_1 and f_2, while Eq. (10) shows that subsystem 2 is affected only by f_2 and consequently is independent of f_1. Under the assumption that the state of subsystem 1 is computed using the output measurement, we can construct an observer-based residual generator that is insensitive to f_1.

If C_1 is of full rank, $z_1(k)$ can be eliminated from (10) by substituting $z_1(k)$ obtained in (11). Considering $C_1^{\perp}$ the left annihilator of C_1, i.e. $C_1^{\perp}C_1 = 0$, and by applying the output transformation $y_2^* = C_1^{\perp}y$ to (11), a modified subsystem 2 can be obtained. Otherwise, if C_1 is not of full rank, consider a non-singular transformation matrix $T_1 = U_2^{\mathrm{T}}$, such that U_2 is obtained from the SVD of C_1, i.e. $C_1 = U_2 \begin{bmatrix} \Sigma_2 \\ 0 \end{bmatrix} V_2^{\mathrm{T}}$. Applying the output transformation $y^* = T_1y$ to (11) results

$$y^*(k) = U_2^{\mathrm{T}}y(k) = \begin{bmatrix} \Sigma_2 \\ 0 \end{bmatrix} V_2^{\mathrm{T}}z_1(k) + U_2^{\mathrm{T}}C_2z_2(k) \tag{12}$$

Consider now the partitions:

$$y^*(k) = \begin{bmatrix} y_1^*(k) \\ y_2^*(k) \end{bmatrix}, \qquad U_2^{\mathrm{T}}C_2 = \begin{bmatrix} C_{21} \\ C_{22} \end{bmatrix}$$

Then, the transformed output defined in (12) can be written as follows:

$$y_1^*(k) = \Sigma_2 V_2^{\mathrm{T}}z_1(k) + C_{21}z_2(k) \tag{13}$$

$$y_2^*(k) = C_{22}z_2(k) \tag{14}$$

Next, substitute $z_1(k)$ from (13) into (10). The modified subsystem 2 is given represented by

$$z_2(k+1) = (A_{22} - A_{21}(\Sigma_2 V_2^{\mathrm{T}})^{+}C_{21})z_2(k) + B_2u(k) + A_{21}(\Sigma_2 V_2^{\mathrm{T}})^{+}y_1^*(k) + E_{22}f_2 \tag{15}$$

$$y_2^*(k) = C_{22}z_2(k) \tag{16}$$

Finally, with the assistance of a Luenberger observer, a residual generator for (15)–(16) can be written as follows:

$$\hat{z}_2(k+1) = \bar{A}_{22}\hat{z}_2(k) + B_2u(k) + A_{21}(\Sigma_2 V_2^{\mathrm{T}})^{+}y_1^*(k) + K(y_2^*(k) - C_{22}\hat{z}_2(k)) \tag{17}$$

$$r(k) = y_2^*(k) - C_{22}\hat{z}_2(k) \tag{18}$$

where $\bar{A}_{22} = A_{22} - A_{21}(\Sigma_2 V_2^{\mathrm{T}})^{+}C_{21}$.

The observer gain K in Eq. (17) can be obtained following the usual pole placement approach. Notice that the order of the UIFDO described above is $(n - n_{\mathrm{f}})$, where $n_{\mathrm{f}} = \mathrm{rank}(E_1)$.

It is easy to verify that in the design of the reduced-order UIFDO we need to satisfy the following existence conditions.

Theorem 3.2. *Necessary and sufficient conditions for the existence of a reduced-order UIFDO for the system defined by (1) are [17]:*

(i) $\mathrm{rank}(C\bar{E}_1) = \mathrm{rank}(E_1)$,
(ii) $(C_{22}, \bar{A}_{22})$ *is a detectable pair.*

Remark 2.1. In Theorems 3.1 and 3.2, condition (i) is usually assumed in the observer design for linear systems with unknown inputs. It means that the maximum number of unknown inputs that can be decoupled is limited to the number of independent measurements.

Remark 2.2. In Theorems 3.1 and 3.2, condition (ii) is equivalent to the condition that all the non-observable modes of the system are stable. The remaining modes are stabilized by gain K_1, in the full-order case, and K in the reduced-order case, that are chosen to stabilize the system dynamic matrices F and $(\bar{A}_{22} - KC_{22})$, respectively.

Next, we illustrate the application of UIO systems to the styrene polymerization reactor and to the fluid catalytic

cracking (FCC) unit. In both cases, we assume that the process is operating under MPC control. In the polymerization reactor, the objective of the fault diagnosis system is to supervise faults on a process parameter and external disturbances. In the FCC system the objective is to supervise the malfunctioning of actuators (inputs) and sensors (outputs) of the centralized control system.

4. Styrene polymerization reactor

The polymerization reactor is usually the heart of the polymer production process and its operation may be difficult as it involves exothermic reactions, unknown reaction kinetics and high viscosity. Furthermore, in case that the process is not working as it should be, it can be very difficult to tell which factor is responsible for the abnormality. Besides, the impossibility to purify products that are out of specification makes a proper fault diagnosis system very useful to this kind of process [18].

4.1. Dynamic process model

Here, we consider the industrial process described by [19], for free-radical solution polymerization of styrene in a jacketed CSTR. A schematic diagram of this process is shown in Fig. 2.

Assuming the standard mechanism for free-radical polymerization [19] presented the following model for the styrene reactor:

$$\frac{\mathrm{d}[I]}{\mathrm{d}t} = \frac{(Q_i[I_\mathrm{f}] - Q_\mathrm{t}[I])}{V} - k_i[I] \tag{19}$$

$$\frac{\mathrm{d}[M]}{\mathrm{d}t} = \frac{(Q_\mathrm{m}[M_\mathrm{f}] - Q_\mathrm{t}[M])}{V} - k_\mathrm{p}[M][P] \tag{20}$$

$$\frac{\mathrm{d}[T_\mathrm{e}]}{\mathrm{d}t} = \frac{Q_\mathrm{t}(T_\mathrm{f} - T_\mathrm{e})}{V} + \frac{(-\Delta H_\mathrm{r})}{\rho C_\mathrm{p}} k_\mathrm{p}[M][P] - \frac{hA}{\rho C_\mathrm{p} V}(T_\mathrm{e} - T_\mathrm{c}) \tag{21}$$

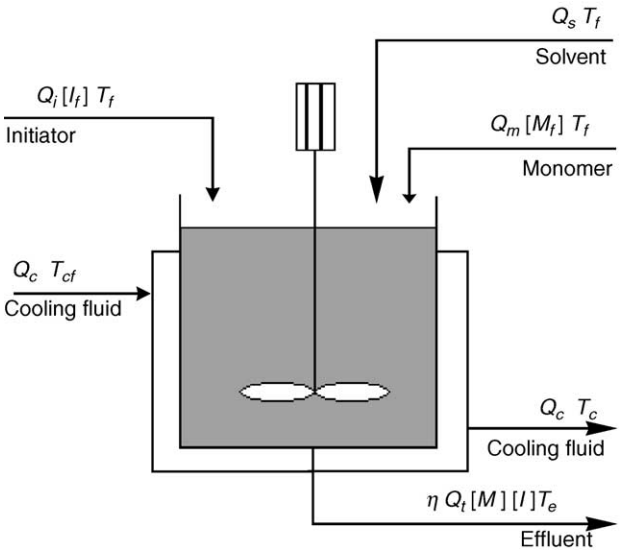

Fig. 2. Process diagram of the styrene polymerization reactor.

$$\frac{\mathrm{d}[T_\mathrm{c}]}{\mathrm{d}t} = \frac{Q_\mathrm{c}(T_\mathrm{cf} - T_\mathrm{c})}{V_\mathrm{c}} + \frac{hA}{\rho_\mathrm{c} C_\mathrm{pc} V_\mathrm{c}}(T_\mathrm{e} - T_\mathrm{c}) \tag{22}$$

$$\frac{\mathrm{d}D_0}{\mathrm{d}t} = 0.5 k_\mathrm{t}[P]^2 - \frac{Q_\mathrm{t} D_0}{V} \tag{23}$$

$$\frac{\mathrm{d}D_1}{\mathrm{d}t} = M_\mathrm{m} k_\mathrm{p}[M][P] - \frac{Q_\mathrm{t} D_1}{V} \tag{24}$$

$$\eta = 0.0012(M_\mathrm{w})^{0.71} \tag{25}$$

where $M_\mathrm{w} = D_1/D_0$, $[P] = [2 f_i k_i [I]/k_t]^{0.5}$, $k_j = A_j \exp(-E_j/T_\mathrm{e})$, $j = i, p, t$, $Q_\mathrm{t} = Q_i + Q_\mathrm{s} + Q_\mathrm{m}$.

Output of Eq. (25) is included in the model to simulate the measurement of the intrinsic viscosity (η) instead of the average molecular weight (M_w), which is rarely available on-line. For the set of parameters presented in Table 1, the reactor has three steady states, but it is designed to operate at the steady state that produces the highest conversion. The operational conditions corresponding to this steady state are listed in Table 2.

Table 1
Process parameters for the polymerization reactor

Variable description	Tag	Value
Frequency factor in Arrhenius equation for initiation reaction (h^{-1})	A_i	2.142×10^{17}
Activation energy for initiation reaction (K)	E_i	14897
Frequency factor in Arrhenius equation for propagation reaction (L/mol h)	A_p	3.816×10^{10}
Activation energy for propagation reaction (K)	E_p	3557
Frequency factor in Arrhenius equation for termination reaction (L/mol h)	A_t	4.50×10^{12}
Activation energy for termination reaction (K)	E_t	843
Initiator efficiency	f_i	0.6
Heat of polymerization reaction (cal/mol)	$-\Delta H_\mathrm{r}$	16700
Monomer molecular weight (g/mol)	M_m	104.14
Overall heat transfer coefficient × heat transfer area of CSTR (cal/K h)	hA	2.52×10^5
Mean density of reactor fluid × mean heat capacity of reactor fluid (cal/K L)	ρC_p	360
Density of cooling jacket fluid × heat capacity of cooling jacket fluid (cal/K L)	$\rho_\mathrm{c} C_\mathrm{pc}$	966.3

Table 2
Steady-state operational condition for the polymerization reactor

Variable description	Tag	Value
Flowrate of solvent (L/h)	Q_s	459
Flowrate of monomer (L/h)	Q_m	378
Reactor volume (L)	V	3000
Volume of cooling jacket (L)	V_c	3312.4
Temperature of reactor feed (K)	T_f	330
Inlet temperature of cooling jacket fluid (K)	T_{cf}	295
Concentration of initiator in feed (mol/L)	$[I_f]$	0.5888
Concentration of monomer in feed (mol/L)	$[M_f]$	8.6981
Concentration of initiator in reactor (mol/L)	$[I]$	6.6832×10^{-2}
Concentration of monomer in reactor (mol/L)	$[M]$	3.3245
Temperature of cooling jacket fluid (K)	T_c	305.17
Molar concentration of the dead polymer chains (mol/L)	D_0	2.7547×10^{-4}
Mass concentration of the dead polymer chains (g/L)	D_1	16.110
Flow rate of initiator (L/h)	$u_1 = Q_i$	108
Flow rate of cooling jacket fluid (L/h)	$u_2 = Q_c$	471.6
Intrinsic viscosity (L/g)	$y_1 = \eta$	2.9091
Temperature of reactor (K)	$y_2 = T_e$	323.56

4.2. The control system

The control objective is to manufacture a polymer with a target molecular weight M_w while regulating the temperature of the reactor for safety and economic considerations. However, as mentioned earlier, on-line measurement of M_w is rarely available and the polymer viscosity η is used instead.

For this process, a 2×2 MIMO control system, based on the infinite-horizon MPC (IHMPC) algorithm, as presented by Rodrigues and Odloak [20], was implemented. In this system $(y_1,y_2) = (\eta,T_e)$ are the controlled variables and $(u_1,u_2) = (Q_i,Q_c)$ are the manipulated variables. The design of the controller incorporates a process model, which is obtained by step response test. Four SISO models are obtained and they are combined in a transfer function matrix model as presented in Table 3.

Some parameters of the IHMPC controller are the sampling time $T = 1$ h and the control horizon $m = 3$. Other tuning parameters are not shown here.

In addition to the MPC control structure, a ratio control law is implemented as [19]:

$$Q_s = 1.5Q_m - Q_i \tag{26}$$

Table 3
Polymerization reactor transfer function matrix model

	u_1	u_2
y_1	$\frac{-45.37}{5.795s+1}e^{-1.547s}$	$\frac{3.66}{9.098s+1}e^{-4.169s}$
y_2	$\frac{121.17}{7.049s+1}e^{-0.771s}$	$\frac{-38.88}{7.206s+1}e^{-1.768s}$

This is done aiming at to maintain a nearly constant volume fraction of solvent in the reactor.

4.3. Fault diagnosis system

We intend to design a fault diagnosis system that will be able to supervise faults (changes) in the process parameters A_t, f_i, M_m and disturbances in the temperature T_f. For this purpose, the fault diagnosis system requires the use of a more detailed model of the reactor than the model presented in Table 3.

Observing Eqs. (19)–(25), it is clear that the process model can be written as a nonlinear state-space of the form:

$$\frac{d\bar{x}(t)}{dt} = g(\bar{x}(t), \bar{u}(t), \bar{f}(t)), \quad \bar{y}(t) = h(\bar{x}(t)) \tag{27}$$

where $\bar{x} = [\,[I] \quad [M] \quad T_e \quad T_c \quad D_0 \quad D_1\,]^T$ is the state vector, $\bar{u} = [\,Q_i \quad Q_c \quad Q_s\,]^T$ the input vector, $\bar{y} = [\eta T_e]^T$ the output vector and $\bar{f} = [\,A_t \quad f_i \quad M_m \quad T_f\,]^T$ is the fault vector. This model can be linearized, around the operating point $(\bar{x}_0(t), \bar{u}_0(t), \bar{f}_0(t))$ shown in Table 2, and the model in deviation form can be transformed into a discrete form similar to Eq. (1), with a sampling time T, where $x = \bar{x} - \bar{x}_0$, $u = \bar{u} - \bar{u}_0$, $y = \bar{y} - \bar{y}_0$ and $f = \bar{f} - f_0$.

Based on this linear model, a bank of four reduced-order UIFDOs (Eqs. (17) and (18)) was implemented in a generalized observer scheme (GOS) [21]. In this bank of observers, each residual is constructed such that it is insensitive to the fault of interest and sensitive to all other faults.

In the design of the reduced-order observers it was noted that for the styrene reactor the pair $(C_{22}, \bar{A}_{22})$ is only detectable, which is the minimal condition for an observer to be implemented. In this case, the alternative design procedure from [16] can be used. However, because of the nonlinear features of the reactor, to reduce process/model mismatch, we prefer to use the observers with adaptive gain, in the form of a Kalman filter. This feature converts the UIFDO in unknown input fault detection filter (UIFDF), where the gain K is updated as follows:

$$K(k) = (\bar{A}_{22}P(k-1)C_{22}^T)(I + C_{22}P(k-1)C_{22}^T)^{-1} \tag{28}$$

$$P(k) = (\bar{A}_{22} - K(k)C_{22})P(k-1)\bar{A}_{22}^T \tag{29}$$

where P is the prediction error covariance matrix. In the styrene system, the residuals were computed using a cumulative average of the residual, with a weighting factor $N = 10$ and exponential forgetting factor $\lambda = 0.1$, in the form:

$$J(r(t)) = N(1-\lambda)\sum_{i=0}^{\infty}\lambda^i r(t-i) \tag{30}$$

The performance of the fault diagnosis system is illustrated for two fault scenarios as shown below. For each of the other possible fault scenarios, the performance of the UIFDO system was similar to the ones presented here and does not

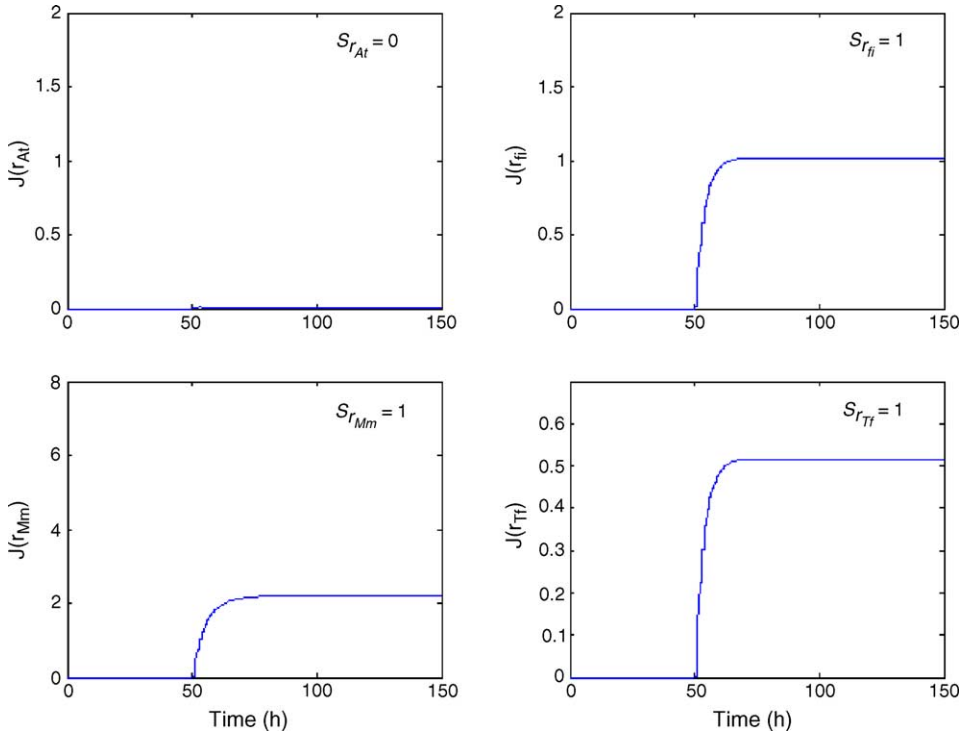

Fig. 3. Residual responses for fault f_{A_t} (−5% in A_t).

bring any new insight into the application of the method and so was not included here. In all the cases presented here, the process was assumed to be initially at steady state.

4.4. Fault scenario 1

A sudden decrease of 5% in parameter A_t (f_{A_t}) occurs at $t = 50$ h. As can be seen in Fig. 3, the fault is isolated perfectly, as it is alarmed by residuals r_{fi}, r_{M_m} and r_{T_f} and not by residual r_{A_t}. The fault magnitude is estimated using state z_1 inferred with (13) and substituting the result into (9). The estimation of the fault, presented in Fig. 4, is obtained from the insensitive observer as

$$\hat{f}_1(k) = (E_{11})^{+}[\hat{z}_1(k+1) - A_{11}\hat{z}_1(k) - A_{12}\hat{z}_2(k) - B_1 u(k)] \tag{31}$$

It can be seem from Eq. (31) that the estimation of the fault magnitude at instant k depends on the inferred state z_1 at instant $k+1$. To avoid this problem, the computation of the fault estimation is delayed one sampling period.

This fault scenario corresponds to a change in the termination rate constant k_{t}, which is the sum of the effects of reaction disproportionation and combination. These contributions are not easily estimated as they vary with temperature and composition, causing an uncertainty in the overall constant k_{t}. In addition, k_{t} presents a phenomenon known as gel or Trommsdorff effect, when its value falls due to strong diffusion limitations at higher monomer conversions.

It is shown in Fig. 4 that detection and isolation of the fault f_{A_t} is achieved in approximately 2 h and its estimation in 15 h, after the sudden fault occurrence. We can also observe that the MPC control system included in the simulation can reject this particular fault without major consequences to the styrene reactor as shown in Fig. 5. The output variables y_1 and y_2 increase as a consequence of the fault and to correct this situation, the IHMPC controller increases the control signals, mainly input u_2. This causes the ratio control to diminish the

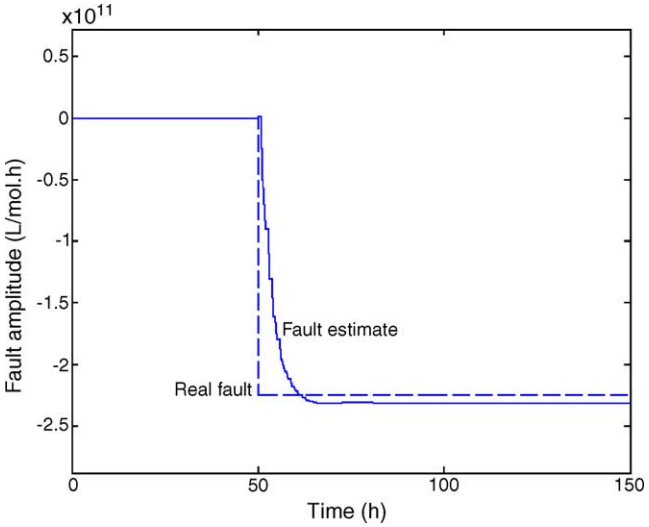

Fig. 4. Estimation of fault f_{A_t} (−5% in A_t).

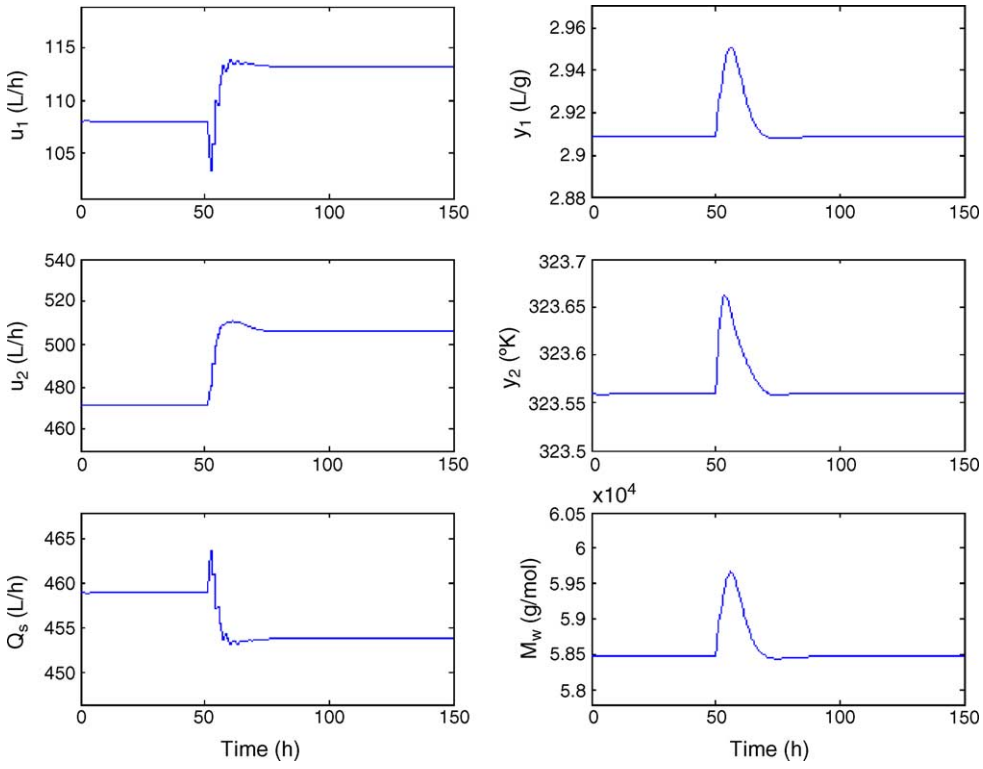

Fig. 5. Manipulated and controlled variables for fault f_{A_t} (−5% in A_t).

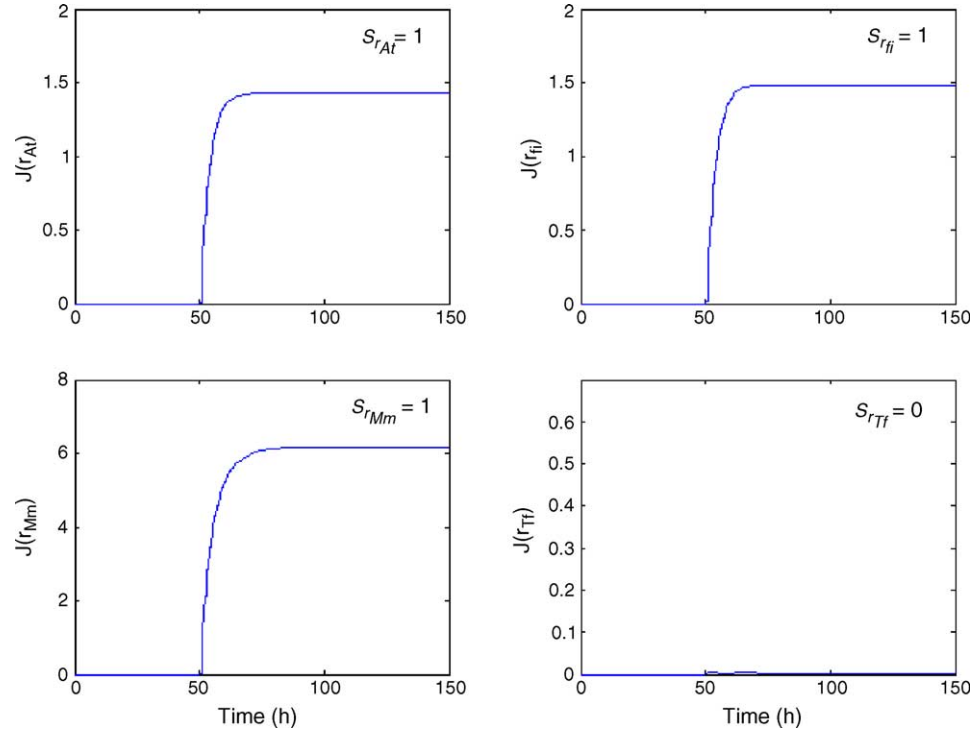

Fig. 6. Residual responses for fault f_{t_f} (+0.5 K in T_f).

flow rate of solvent (Q_s). The controlled secondary variable M_w is also presented in Fig. 5.

4.5. Fault scenario 2

In this case we simulate a small increase of 0.5 K in the feed temperature $T_f(f_{t_f})$, occurring at $t = 50$ h. Fig. 6 shows that, the residuals r_{A_t}, r_{fi} and r_{M_m} are sensitive and residual r_{T_f} is insensitive to this fault. The fault is isolated correctly and its magnitude is well estimated by means of Eq. (31), as illustrated in Fig. 7. For this disturbance, complete detection and isolation is obtained in approximately 2 h after its occurrence and its estimation in achieved after 7 h of its abrupt occurrence.

The occurrence of this fault increases the temperature profile of the reactor (y_2), which produces an increase in the intrinsic viscosity (y_1) and, consequently, the molecular weight (M_w) also increases. As expected, this disturbance does not causes major problems to the control system. The MPC control results in good regulation of the outputs as shown in Fig. 8, where it is observed that input u_2 was more demanding, which is rather obviously justified. The profiles of Q_s and M_w are also presented.

A fault diagnosis system based on the full-order UIFDO was also implemented for this process. The performance of the full-order system was similar to the performance of reduced-order system, with respect to fault detection and isolation. However, the full-order system did not converge when estimating the size of fault. This can be explained by observing that for the system represented in (1), the fault estimation using full-order UIFDO can be obtained by

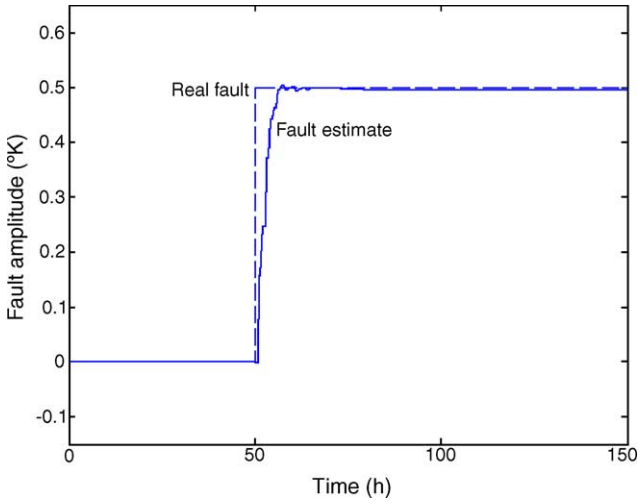

Fig. 7. Estimation of fault f_{t_f} (+0.5 K in T_f).

$$\hat{f}_1(k) = (E_1)^+[\hat{x}(k+1) - A\hat{x}(k) - Bu(k)] \quad (32)$$

where $\hat{x} = T^{-1}\hat{z}$, and $\hat{z}$ is the estimated state of the UIFDO. Thus, the full-order system requires the state transformation using the inverse of matrix T, which cannot be computed adequately if T is ill-conditioned. This problem appears in the polymerization process, when the linearized process model is computed around the middle operating point. On the other

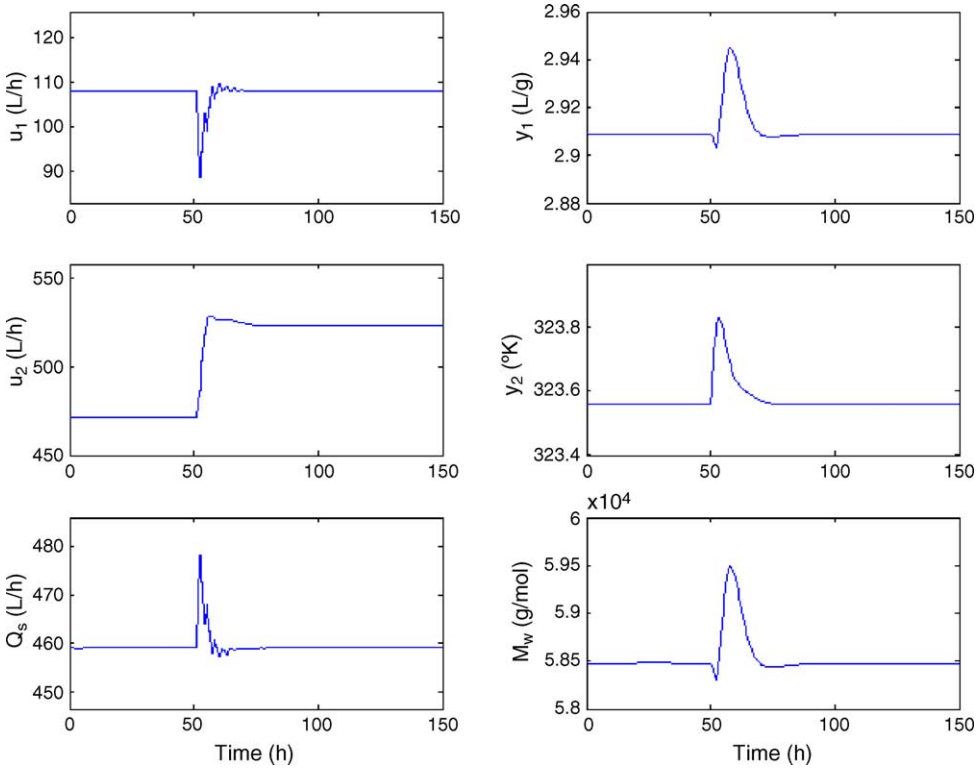

Fig. 8. Manipulated and controlled variables for fault f_{t_f} (+0.5 K in T_f).

hand, by using the reduced-order UIFDO, the SVD approach improves in a natural way the condition number of the estimation problem, besides not requiring the transformation of states to estimate the fault.

5. Industrial fluid catalytic cracking (FCC) unit

In a modern oil refinery, the most critical component, is the fluid catalytic cracking (FCC) unit. The FCC unit involves a nonlinear and multivariable process, with many cross-couplings, subject to a number of operational constraints and working at high temperatures. The extent of potential losses due to unexpected faults in a FCC unit is enormous. As the FCC unit involves high temperatures and the flow of solid catalyst at high velocities, sticking of control valves and erosion in measuring instruments are more likely to occur than in most of the process units of the oil refinery. Thus, unexpected faults in the MPC control system of the FCC unit are also more likely to occur than in the control systems of other process units.

5.1. Dynamic process model and control system

A simplified schematic diagram of this process is shown in Fig. 9. Moro and Odloak [22] provide an extensive mathematical model of this process, validated with industrial data, which has become a standard for validation of FCC control structures in Petrobras refineries [23]. The nominal steady-state operational conditions are shown in Table 4. The model equations and process parameters are here omitted but they can be found in Moro and Odloak [22].

The performance of a FCC process highly depends on the selected control structure. In this paper, the control structure is based on the original version of the MPC system proposed for the FCC unit of the REVAP refinery. From this scheme we consider only a 4×4 MIMO control system, where the inputs or manipulated variables are: u_1, u_3, u_4 and u_9. The outputs or controlled variables are: y_1, y_2, y_3 and y_4. The control objective is stability rather economic, i.e. to minimize disturbance and interaction effects and to keep the operation point as close as possible to its nominal values, where, usually, optimal operation conditions lie.

The controller is also the IHMPC of Rodrigues and Odloak [20]. It is based on the transfer function model presented in Table 5. In the implementation of the IHMPC controller simulated here, the sampling time is $T = 3$ min and the control horizon is $m = 4$. Other controller tuning parameters are omitted.

5.2. Fault diagnosis system

Several fault diagnosis systems have been applied to the FCC unit. However, these diagnosis systems are mainly based on data-driven approaches. Pranatyasto and Qin [24] have developed a fault detection system for the FCC sensors

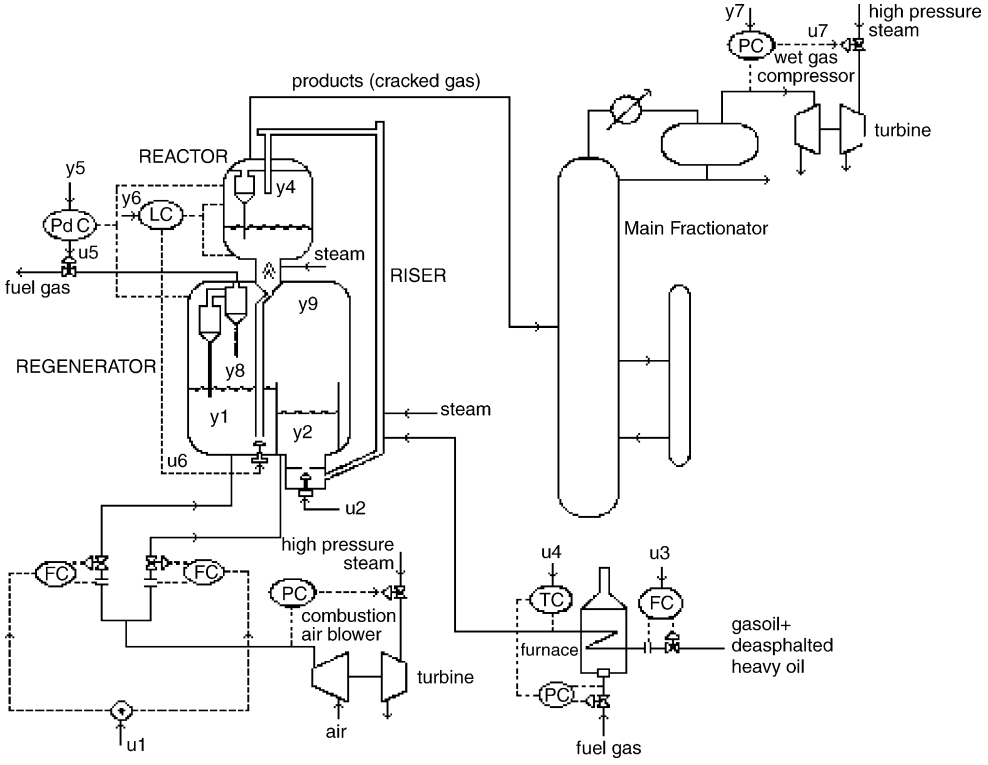

Fig. 9. FCC Reactor–regenerator system.

Table 4
Nominal steady-state operational conditions for the FCC unit

Variable description	Tag	Value
Input variables		
Total combustion air flow rate (t/h)	u_1	221
Regenerated catalyst valve position (%)	u_2	82
Total oil feed flowrate (gasoil + deasphalted heavy oil) (m^3/d)	u_3	9700
Temperature of the total oil feed (°C)	u_4	235
Flue gas flow rate valve position (%)	u_5	63
Spent catalyst valve position (%)	u_6	73
Flow of steam to the gas compressor Turbine driver valve position (%)	u_7	74
Combustion air temperature (°C)	u_8	190
First stage/feed ratio of air flow rate (%)	u_9	90.72
Output variables		
Temperature of the regenerator first stage dense phase (°C)	y_1	670.14
Temperature of the regenerator second stage dense phase (°C)	y_2	700.88
Estimated cracking reaction severity (%)	y_3	77.5
Riser (reactor) temperature (°C)	y_4	542.2
Δp between reactor and regenerator (kgf/cm^2)	y_5	0.65
Catalyst inventory in the reactor (t)	y_6	90
Pressure at the wet gas compressor (kgf/cm^2)	y_7	1.0
Temperature of the regenerator first stage dilute phase (°C)	y_8	704.39
Temperature of the regenerator general dilute phase (°C)	y_9	697.53

following the multivariate statistical principal component analysis (PCA) method. Vedam and Venkatasubramaniam [25] introduced PCA-SDG (signed digraphs) fault diagnosis approach and [26] developed a recursive partial least squares (PLS) method to detect faults in a FCC unit. Sebzalli and Wang [27] applied PCA-Fuzzy clustering fault diagnosis approach to the FCC main fractionator. The main disadvantage of these methods is that they need a large amount of plant data that are collected along a quite large window of operating time and are used to construct a statistical model of the process. This new model is not the same as the one used in the MPC controller to predict the system output. Here, the same model is used both in the MPC controller and in the fault diagnosis system.

The transfer function matrix model shown in Table 5 is transformed into a modified state-space model with the following structure:

$$\tilde{x}(k+1) = \tilde{A}\tilde{x}(k) + \tilde{B}u(k) + R_1 f,$$
$$y(k) = \tilde{C}\tilde{x}(k) + R_2 f \tag{33}$$

where $\tilde{x} \in \mathbb{R}^{n\times 1}$, $f = \begin{bmatrix} f_a \\ f_s \end{bmatrix} \in \mathbb{R}^{p\times 1}\ (p = m + l)$, $f_a \in \mathbb{R}^{m\times 1}$ denotes the vector of actuator faults, $f_s \in \mathbb{R}^{l\times 1}$ denotes the vector of sensor faults, $R_1 = [\tilde{B}0_{n\times l}] \in \mathbb{R}^{n\times p}$ and $R_2 = [0_{l\times m} I_l] \in \mathbb{R}^{l\times p}$ are the actuator and sensor fault distribution matrices respectively.

Here, we assume that the probability of two or more faults to occur at the same time is negligible. Thus, the model defined in (33) can be written for the case of actuator faults as follows:

$$\tilde{x}(k+1) = \tilde{A}\tilde{x}(k) + \tilde{B}u(k) + \tilde{B}_i f_i + [\,\bar{B} \quad 0_{n\times l}\,]\bar{f},$$
$$y(k) = \tilde{C}\tilde{x}(k) \tag{34}$$

where $\tilde{B}_i \in \mathbb{R}^{n\times 1}$ is the ith column of matrix $\tilde{B}$, $\bar{B} \in \mathbb{R}^{n\times(m-1)}$ corresponds to matrix $\tilde{B}$ without the ith column, $f_i \in \mathbb{R}$ is the ith component of vector f and $\bar{f} \in \mathbb{R}^{(p-1)\times 1}$ corresponds to vector f without the ith element.

Analogously, for sensor faults f_{m+j} $(1 \le j \le l)$ system (33) can be written as follows:

$$\tilde{x}(k+1) = \tilde{A}\tilde{x}(k) + \tilde{B}u(k) + [\,\tilde{B} \quad 0_{n\times(l-1)}\,]\bar{f},$$
$$y(k) = \tilde{C}\tilde{x}(k) + I_j f_{m+j} + [\,0_{l\times m} \quad \bar{I}\,]\bar{f} \tag{35}$$

where $I_j \in \mathbb{R}^{l\times 1}$ is the jth column of the identity matrix I_l, $\bar{I} \in \mathbb{R}^{l\times(l-1)}$ the identity matrix without column j, $f_{m+j} \in \mathbb{R}^{1\times 1}$ the $(m+j)$th component of vector f and $\bar{f} \in \mathbb{R}^{(p-1)\times 1}$ is obtained from vector f by removing the $(m+j)$th component.

As discussed by Park et al. [28], sensor faults can be also represented by actuator faults. For this purpose, assume that a random walk model describes the dynamics of the sensor fault:

$$f_{m+j}(k) = f_{m+j}(k) + T\xi(k)$$

Table 5
FCC unit transfer function matrix model

	u_1	u_3	u_4	u_9
y_1	$\frac{0.0661s + 0.0018}{s^2 + 0.0422s + 0.0019}$	$\frac{-10^{-3}(0.7857s + 0.0425)}{s^2 + 0.0434s + 0.0022}$	$\frac{0.0271s + 0.016}{s^2 + 0.0561s + 0.0033}$	$\frac{45.3896s + 7.9639}{s^2 + 1.4619s + 0.0757}$
y_2	$\frac{0.0730s + 0.0020}{s^2 + 0.0421s + 0.0019}$	$\frac{-10^{-3}(0.6939s + 0.0430)}{s^2 + 0.0393s + 0.0023}$	$\frac{0.0256s + 0.015}{s^2 + 0.0504s + 0.0032}$	$\frac{-44.3671s - 0.2308}{s^2 + 0.3213s + 0.0195}$
y_3	$\frac{0.0123s + 0.0025}{s^2 + 0.6546s + 0.0231}$	$\frac{-1.5 \times 10^{-3}(0.1082s + 0.0032)}{s^2 + 0.0339s + 0.0009}$	$\frac{0.0181s + 0.0039}{s^2 + 0.3455s + 0.0409}$	$\frac{-0.2655s - 0.0086}{s^2 + 0.0376s + 0.0028}$
y_4	$\frac{0.0219s + 0.0008}{s^2 + 0.0400s + 0.0016}$	$\frac{-1.25 \times 10^{-3}(0.6931s + 0.0136)}{s^2 + 0.0339s + 0.0009}$	$\frac{0.0653s + 0.0096}{s^2 + 0.4273s + 0.0188}$	$\frac{-3.1350s - 0.0568}{s^2 + 0.0724s + 0.0043}$

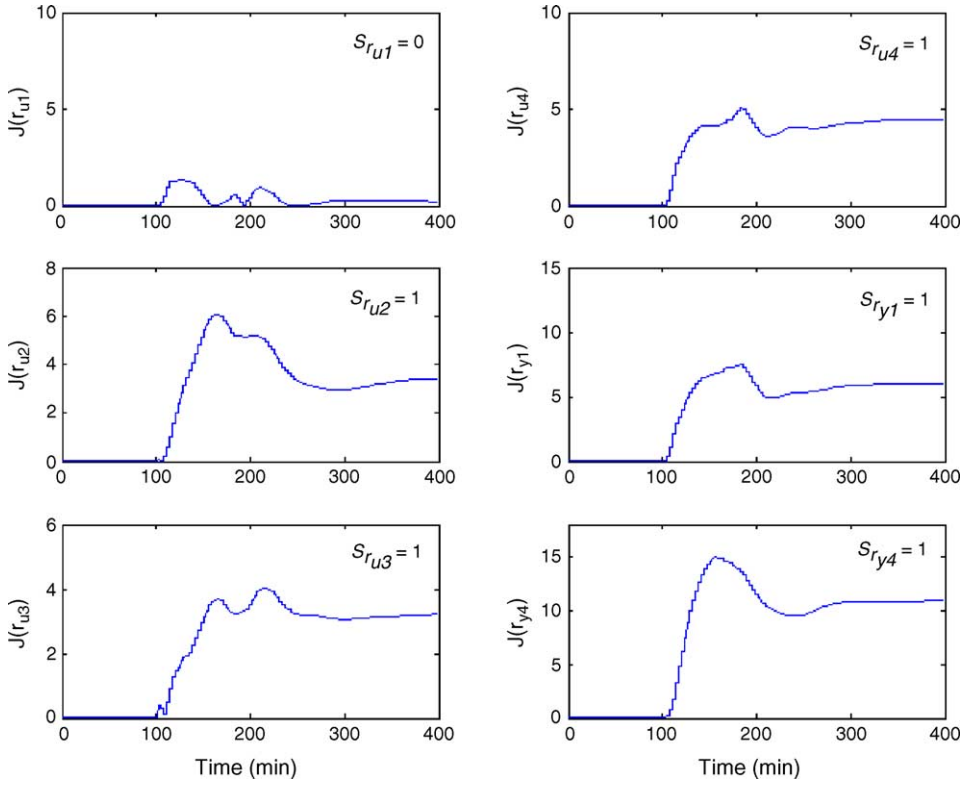

Fig. 10. Residual responses for fault f_{u_1} (−5% in actuator 1).

where ξ is the sensor error. Therefore, for a sensor fault, an augmented system is obtained as follows:

$$\begin{bmatrix} \tilde{x}(k+1) \\ f_{m+j}(k+1) \end{bmatrix} = \begin{bmatrix} \tilde{A} & 0_{n\times 1} \\ 0_{1\times n} & 1 \end{bmatrix} \begin{bmatrix} \tilde{x}(k) \\ f_{m+j}(k) \end{bmatrix} + \begin{bmatrix} \tilde{B} \\ 0_{1\times m} \end{bmatrix} u(k) + \begin{bmatrix} 0_{n\times 1} \\ T \end{bmatrix} \xi(k) + \begin{bmatrix} \tilde{B} & 0_{n\times(l-1)} \\ 0_{1\times m} & 0_{1\times(l-1)} \end{bmatrix} \bar{f},$$

$$y(k) = \begin{bmatrix} \tilde{C} & I_j \end{bmatrix} \begin{bmatrix} \tilde{x}(k) \\ f_{m+j}(k) \end{bmatrix} \tag{36}$$

Observe that, for actuator or sensor fault, the faulty process models (34)–(36) have the same state-space structure as Eq. (1).

In this case study, the fault diagnosis system supervises inputs u_1, u_3, u_4 and u_9, and outputs y_1 and y_4 of the FCC MPC system. A bank of six full-order UIFDOs (Eqs. (2) and (4)) is designed in a GOS scheme. The first four residuals, for actuator fault detection, are based on model (34) and the last two residuals, for sensor fault detection, are based on model (33). The static gain of the full-order observers has the form $\mathbf{K_1} = \begin{bmatrix} I_l & \mathbf{0}_{n\times l} \end{bmatrix}^{\mathrm{T}}$. The residuals are evaluated according Eq. (30), with $N = 1$ and $\lambda = 0.1$. The performance of the fault diagnosis algorithm is illustrated for two scenarios as described below. The algorithm was also tested for all the other possible scenarios corresponding to a single fault and the obtained performances were similar to the ones presented here.

5.2.1. *Fault scenario 1*

The fault is represented by a 5% decrease in actuator 1 (f_{u_1}), occurring at $t = 100$ min. In this case, residual r_{u_1} is

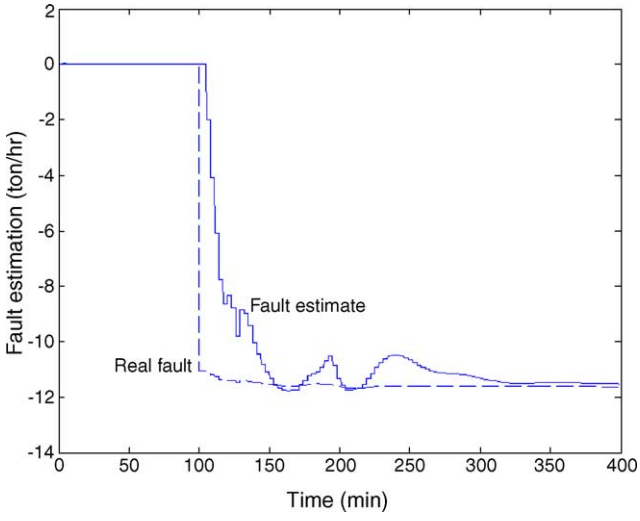

Fig. 11. Estimation of fault f_{u_1} (−5% in actuator 1).

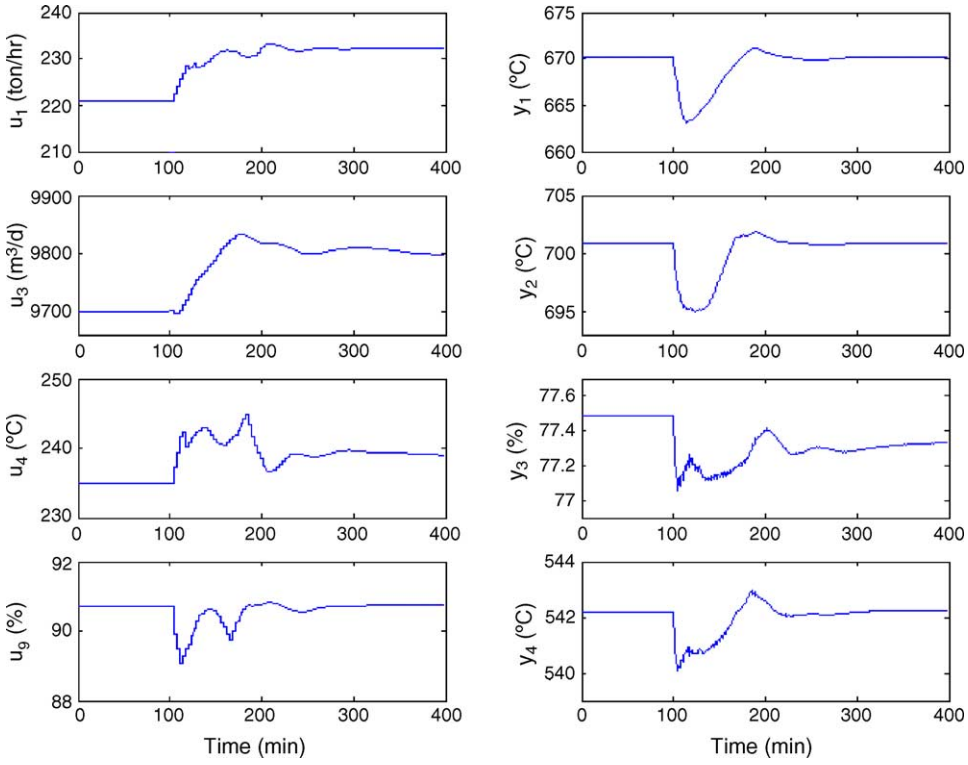

Fig. 12. Inputs and controlled outputs for fault f_{u_1} (−5% in actuator 1).

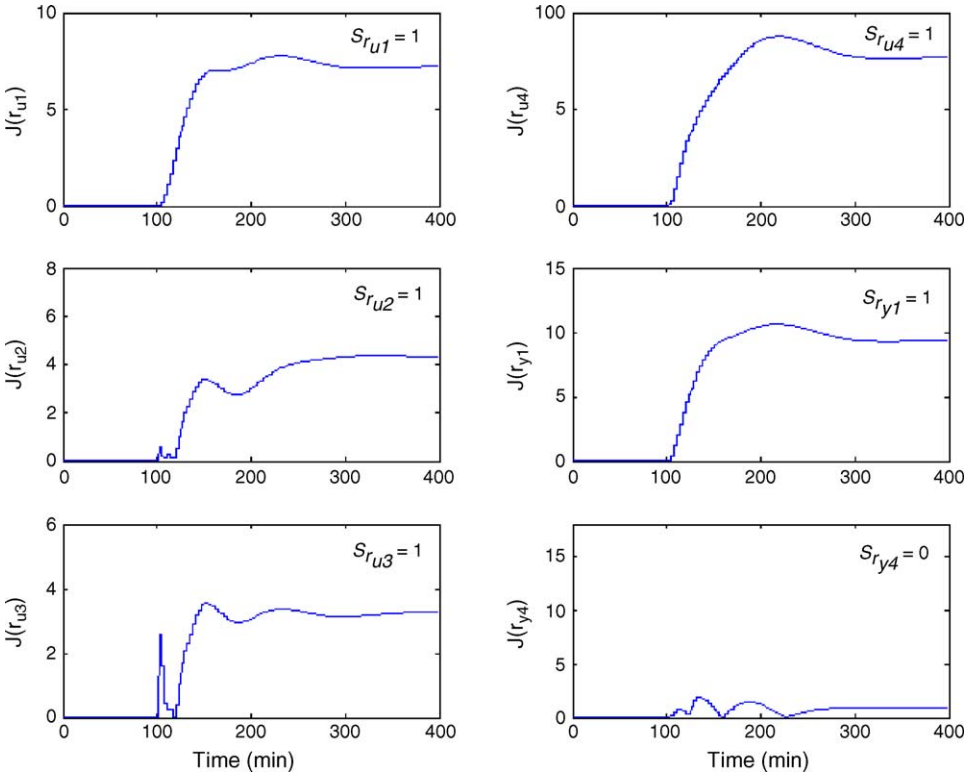

Fig. 13. Residual responses for fault f_{y_4} (+5% in sensor 4).

insensitive to the fault and the other residuals are sensitive. Fig. 10 shows that the proposed approach is able to detect and isolate this fault satisfactorily. The fault magnitude is adequately estimated as shown in Fig. 11 and this is done by substituting the state estimation $\hat{z}$ provided by the observer of the faulty input, into the faulty model defined in (35), and rearranging the resulting equation as in (32).

The effects of this fault on the process are shown in Fig. 12, where it can be seen that the temperature in the regenerator section is decreased. To compensate this undesirable effect the controller supplies a larger amount of combustion air, which is not attractive economically. The controller increases the control signals u_1, u_3 and u_4 trying to keep the output variables at their desired values.

In the case presented here, actuator fault detection and isolation is achieved in approximately 10 min and its estimation in 200 min, after the fault occurrence.

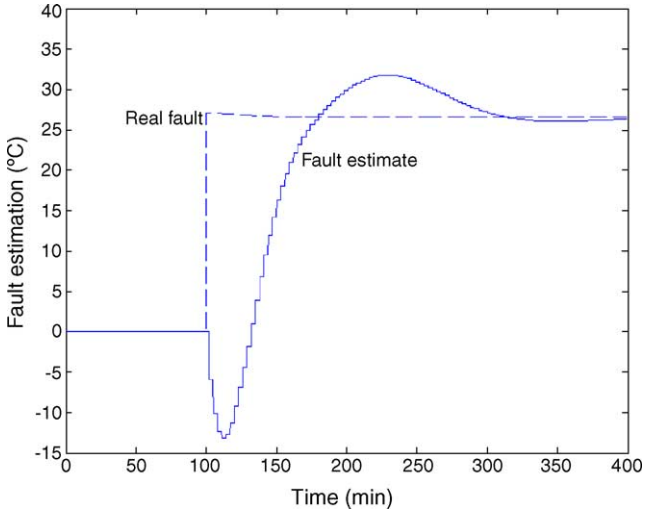

Fig. 14. Estimation of fault f_{y_4} (+5% in sensor 4).

5.2.2. Fault scenario 2

The fault corresponds to a 5% increase in the measurement of sensor 4 (f_{y_4}), occurring at $t = 100$ min. Fig. 13 shows that residual r_{y_4} is insensitive to this fault while the other residuals are sensitive. The magnitude of the fault is estimated using a dedicated Luenberger observer. The fault estimation is illustrated in Fig. 14 where it is indicated that the estimation tends to the correct value of the fault.

This faulty scenario is unsafe to the FCC operation, as it may cause an undesirable breakdown of the production. To revert this situation the temperature of the inflow feed of the riser has to be rapidly reduced. Fig. 15 presents the responses of the manipulated and controlled variables associated with the fault in sensor 4. As it can be seen, this fault causes saturation in u_3 and u_4. Moreover, an acceptable control is never achieved with MPC since bias appears in all the outputs. For this fault, the complete detection and isolation is obtained in approximately 20 min, while its estimation occurs only about 200 min, after the fault occurrence.

It is interesting to mention that, for the FCC process example, fault diagnosis systems based on the full-order UIFDO

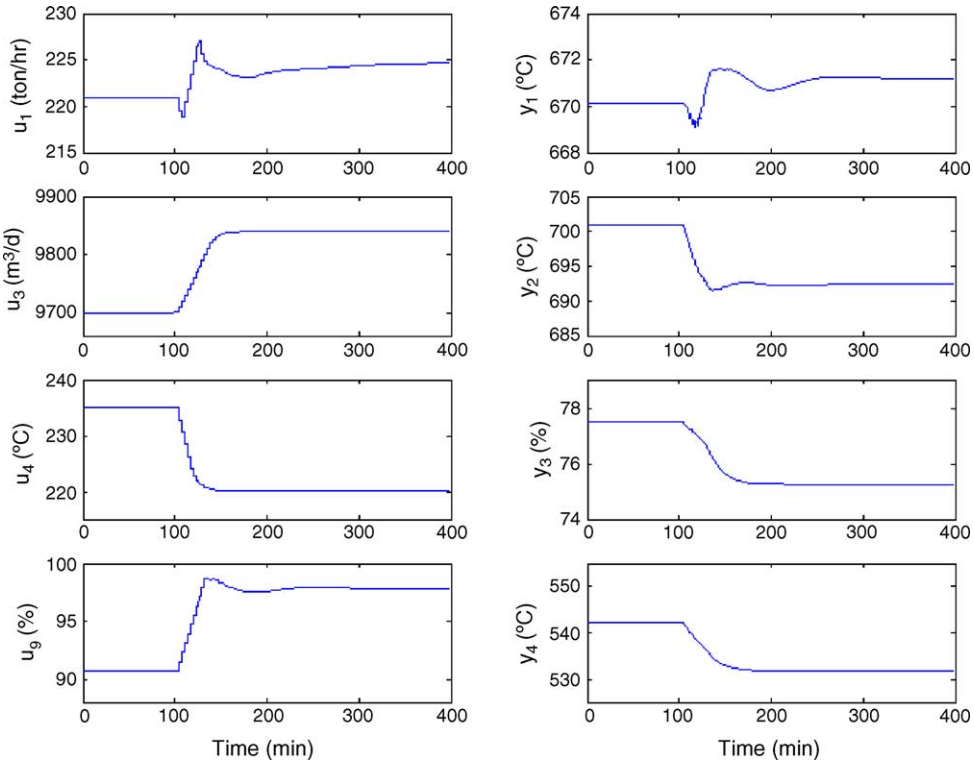

Fig. 15. Inputs and controlled outputs responses for fault f_{y_4} (+5% in sensor 4).

or on the reduced order UIFDO produced similar performances. This is so because the state space model utilized in the fault detection system was obtained via reliable linear model identification methods, which produce minimal order fully observable models. Thus, model reduction was not necessary to be implemented in the fault detection system in order to produce a correct estimation of the magnitude of the fault.

As seen in the presented cases, there are time delays between the fault occurrence and its detection, isolation and identification. In the cases studied, the delay related to the fault identification is larger than the others. This is due to the lack of sufficient input–output measurement data in the early stages of the identification step. Gradually, as more data become available from the monitored system, the fault estimate is refined, i.e. it becomes more accurate and the estimate error decreases until it reaches a minimum and the fault can be identified. Moreover, the final estimate of the fault cannot be expected to match perfectly the true value of the fault, due to model uncertainties. Therefore, imperfections of the fault analysis system, in terms of delays and model uncertainties, should be taken into account if control reconfiguration is to be considered.

6. Conclusions

In this paper, unknown input observer (UIO) methods have been utilized to design fault diagnosis systems for two complex process systems of the chemical and oil refining industries. Although the systems considered in this study are essentially nonlinear, the diagnose systems used here, are based on linear models obtained at the most probable operating condition. The practical application of the diagnose systems developed here, was tested by simulation the styrene polymerization reactor and the FCC reactor–regenerator system. The results of this study shows that the method is very promising for future practical implementation, mainly when the system remains close to the operating point in which the model, that is used in the observer, was obtained. The application of the approach followed here to large-scale systems does not seem to increase significantly the complexity of the fault detection algorithm, as long as the hypothesis of occurrence of a single fault is maintained. However, future research should focus on the robustness of this approach for large-scale systems as well as for nonlinear systems that operate in a broader range of operating conditions.

Acknowledgments

Support for this work was provided by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) under Grants 02/08119-2 and 01/06657-4 and by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) under Grant 300860/97-9.

References

[1] A.S. Willsky, A survey of design methods for failure detection in dynamic systems, Automatica 12 (6) (1976) 601–611.
[2] J.J. Gertler, A survey of model-based failure detection and isolation in complex plants, IEEE Contr. Syst. Mag. 8 (6) (1988) 3–11.
[3] P.M. Frank, Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy—a survey and some new results, Automatica 26 (3) (1990) 459–474.
[4] R.J. Patton, Robustness in model-based fault diagnosis: the 1995 situation, Annu. Rev. Contr. 21 (1) (1997) 103–123.
[5] R. Isermann, Supervision, fault-detection and fault-diagnosis methods—an introduction, Contr. Eng. Pract. 5 (5) (1997) 639–652.
[6] P.M. Frank, S.X. Ding, T. Marcu, Model-based fault diagnosis in technical processes, Trans. Inst. Measure. Contr. 22 (1) (2000) 57–101.
[7] V. Venkatasubramanian, R. Rengaswamy, K. Yin, S.N. Kavuri, A review of process fault detection and diagnosis. Part I. Quantitative model-based methods, Comput. Chem. Eng. 27 (3) (2003) 293–311.
[8] R.J. Patton, J. Chen, Observer-based fault detection and isolation: robustness and applications, Contr. Eng. Pract. 5 (5) (1997) 671–682.
[9] P.M. Frank, X. Ding, Survey of robust residual generation and evaluation methods in observer-based fault detection systems, J. Process Contr. 7 (6) (1997) 403–424.
[10] S.K. Dash, R. Rengaswamy, V. Venkatasubramanian, Fault diagnosis in a nonlinear CSTR using observers, in: 2001 Annual AIChE Meeting, Paper 282i, Reno, NV, 2001.
[11] C.-T. Chang, J.-W. Chen, Implementation issues concerning the EKF-based fault diagnosis techniques, Chem. Eng. Sci. 50 (18) (1995) 2861–2882.
[12] Y. Huang, G.V. Reklaitis, V. Venkatasubramanian, A heuristic extended Kalman filter based estimator for fault identification in a fluid catalytic cracking unit, Ind. Eng. Chem. Res. 42 (14) (2003) 3361–3371.
[13] R. Li, J.H. Olson, Fault detection and diagnosis in a closed-loop nonlinear distillation process: application of extended Kalman filter, Ind. Eng. Chem. Res. 30 (5) (1991) 898–908.
[14] R. Tarantino, F. Szigeti, E. Colina-Morles, Generalized Luenberger observer-based fault-detection filter design: an industrial application, Contr. Eng. Pract. 8 (6) (2000) 665–671.
[15] Y. Chetouani, N. Mouhab, J.M. Cosmao, L. Estel, Application of extended Kalman filtering to chemical reactor fault detection, Chem. Eng. Commun. 189 (9) (2002) 1222–1241.
[16] J. Chen, R.J. Patton, Robust model-based fault diagnosis for dynamic systems, Kluwer Academic Publishers, Massachusetts, 1999.
[17] M. Hou, P.C. Müller, Fault detection and isolation observers, Int. J. Contr. 60 (5) (1994) 827–846.
[18] P. Kaboré, S. Othman, T.F. McKenna, H. Hammouri, Observer-based fault diagnosis for a class of non-linear systems—application to a free radical copolymerization reaction, Int. J. Contr. 73 (9) (2000) 787–803.
[19] B.R. Maner, F.J. Doyle III, B.A. Ogunnaike, R.K. Pearson, Nonlinear model predictive control of a simulated multivariable polymerization reactor using second-order Volterra models, Automatica 32 (9) (1996) 1285–1301.
[20] M.A. Rodrigues, D. Odloak, MPC for stable linear systems with model uncertaintly, Automatica 39 (4) (2003) 569–583.
[21] J. Wünnenberg. Observer-based fault detection in dynamic systems. PhD Thesis, Gerhard-Mercator-University of Duisburg, Germany, 1990.
[22] L.F.L. Moro, D. Odloak, Constrained multivariable control of fluid catalytic cracking converters, J. Process Contr. 5 (1) (1995) 29–39.
[23] A.C. Zanin, M. Tvrzská de Gouvêa, D. Odloak, Integrating real-time optimization into the model predictive control of the FCC system, Contr. Eng. Pract. 10 (8) (2002) 819–831.

[24] T.N. Pranatyasto, S.J. Qin, Sensor validation and process fault diagnosis for FCC under MPC feedback, Contr. Eng. Pract. 9 (8) (2001) 877–888.
[25] H. Vedam, V. Venkatasubramanian, PCA-SDG based process monitoring and fault diagnosis, Contr. Eng. Pract. 7 (7) (1999) 903–917.
[26] X. Wang, U. Kruger, B. Lennox, Recursive partial least squares algorithms for monitoring complex industrial processes, Contr. Eng. Pract. 11 (6) (2003) 613–632.
[27] J. Park, G. Rizzoni, W.B. Ribbens, On the representation of sensor faults in fault-detection filters, Automatica 30 (11) (1994) 1793–1795.
[28] Y.M. Sebzalli, X.Z. Wang, Knowledge discovery from process operational data using PCA and fuzzy clustering, Eng. Appl. Artific. Intell. 14 (5) (2001) 607–616.