Comparison of model predictive control strategies for a fluidized catalytic cracker
Alexandre Boum, Abderrazak Latifi, Jean Pierre Corriou

To cite this version:

HAL Id: hal-02130571
https://hal.archives-ouvertes.fr/hal-02130571
Submitted on 15 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Comparison of model predictive control strategies for a fluidized catalytic cracker

A.T. Boum 1*, A. Latifi 2 and J.P. Corriou 3

1 Universite de Douala, ENSET B.P 1872 Douala, Cameroon
2 Universite de Lorraine, LRGP, UMR 3349, Nancy, F-54001, France
3 Universite de Lorraine, LRGP, UMR 3349, Nancy, F-54001, France
*Corresponding author E-mail: boumat2002@yahoo.fr

Abstract

A FCC model is used to compare five different Model Predictive Control (MPC) strategies. The FCC process is a complex petrochemical unit with catalyst recycling that makes its behaviour highly nonlinear. The FCC comprises a riser, a separator and a regenerator with important heat coupling due to the endothermic cracking reactions of gas oil in the riser and the exothermic combustion reactions in the regenerator. The riser and the regenerator exhibit fast and slow dynamics respectively. The temperatures at riser top and in the regenerator should be controlled by manipulation of catalyst and air flow rates. All these nonlinear and coupled characteristics render the multivariable control problem difficult and thus the FCC process constitutes a valuable benchmark for comparing control strategies. Here, the performances of Dynamic Matrix Control, Quadratic Dynamic Matrix Control, MPC control with penalty on the outputs, NonLinear MPC control, Observer Based MPC control are compared.

Keywords: multivariable control ; model predictive control ; fluidized catalytic cracker ; step response ; benchmark.

1. Introduction

Catalytic cracking has been used for more than seventy years [1, 2, 3] in petrochemical plants for conversion of heavy gas oil into high-value gasoline and lighter components. It is considered as the main unit in a refinery. Its major economic importance has fostered many technical developments and several versions of Fluidized Catalytic Cracking (FCC) units have been patented and installed, such as Exxon Model IV FCC [4], UOP type FCC [3], Kellog Orthoflow Model F FCC [5].

The simulation of FCCs is performed at many different levels of detail [6, 7, 8, 9, 10, 11, 12], either for the riser or the regenerator. Even CFD models describe the behaviour of the riser [13] or the regenerator. Generally, the riser model complexity depends only on the number of lumps used to describe the cracking kinetics. However, the regenerator model differs much more according to the studies. In some of them, the regenerator is modelled as a Continuously Stirred Tank Reactor (CSTR), whereas in other studies the two-phase model of Kunii-Levenspiel [14]. is used as the basis for describing the dense bed mainly composed of the bubbles and the emulsion phase with a high density of catalyst particles. Even, in this latter case, sometimes the bubbles are considered like the dense phase as a CSTR [15], sometimes as a plug flow system [9, 10].

Many researchers underline the nonstationary behavior of FCCs evidently due to the catalyst recycling which provokes an important heat coupling between the riser and the regenerator. Thus, some authors have signaled multiple steady states [16, 17, 18, 19] though it is not a general case. In our case, we did not find multiple steady states.

Because of its economic importance and the control difficulties encountered due to its multivariable characteristics, large interactions, important nonlinearities and the presence of constraints, FCCs have been the subject of many control studies such as controllability analysis and control structure [20, 21], [12], [22], [23], model predictive control [24, 25, 26, 5, 27], nonlinear model predictive control [4, 28], PI observers [29]. Different control configurations are studied by these authors with various choices for inputs, outputs and their respective numbers. It remains that, owing to the complexity of the FCC dynamics, the multivariable characteristics and the presence of constraints, MPC is often considered as the reference control method.

On-line optimization with large economic benefits is also reported by [30, 31, 32].

In this work, multivariable control of an FCC process by several MPC strategies using two manipulated inputs and two controlled outputs is addressed. The present paper is organized as follows. Section 2 is related to the FCC description together with its constitutive units. Section 3 deals with the different model predictive control strategies followed by their respective simulation results. Finally, the conclusions of the paper are presented in Section 4.

2. FCC model

Basically, the FCC process is a complex petrochemical unit the role of which is to produce lighter components such as gasoline, diesel from gas oil by cracking reactions. These latter take place in the riser...
by contact between vaporized gas oil and the fluidized rising catalyst on the surface of which coke, a reaction product and a poison for the catalyst, is deposited. The FCC process (Figure 1) comprises a riser, a separator and a regenerator with important heat coupling between the endothermic cracking reactions of gas oil in the riser and the exothermic combustion reactions of coke and oxygen in the regenerator. The riser and the regenerator are the main units which exhibit fast and slow dynamics respectively. One of the main features of FCCs is the catalyst recycle that makes its behaviour highly nonlinear.

Cracking reactions occurring in the riser reactor are highly complex due to the nature of feed oil and for that reason, they are modelled using lumps, i.e. groups of chemical components having close chemical properties, such as gasoline. The number of lumps involved in the models of the literature varies a lot and some models, especially those with a large number of lumps, do not provide the heats of reaction which are necessary for describing the energy balance in the riser. Among these cracking models, can be mentioned three lumps [33], four lumps [34, 35], five lumps [36], six lumps [37], seven lumps [38], eight lumps [39, 40], nine lumps [41], ten lumps [42], fourteen lumps [13, 43]. For control purposes, a large number of lumps is not necessary but the energy balance is of prime importance. The role of the regenerator is to regenerate the spent catalyst coming from the separator by allowing the combustion of the coke deposited on the catalyst [26, 44, 45, 46]. It is operated either in full or partial combustion. The regenerated catalyst is returned to the foot of the riser.

![Figure 1: Schematic diagram of a FCC unit](image)

### 2.1. Riser model

The riser is considered as a plug flow reactor. At the inlet, the feed is assumed to be instantaneously vaporized [12]. In the present model, it is assumed that no slip occurs between the solid phase and the gas phase, i.e. the velocities of both phases are equal. The residence time of catalyst and gas stream in the riser is supposed to be a few seconds. Consequently, the riser is only described by spatial equations and considered as an algebraic system with respect to time, opposite to the separator and regenerator whose equations are time-dependent. The kinetic model is based on a three-lump scheme [33] (Figure 2) to describe the cracking reactions in the riser [47]. The formation of coke is given by a simple algebraic equation at the outlet. In the literature, there exist many different models for the cracking in the riser, with lumps from three components up to fifteen components. Even if the number of components is particularly low in the present model, the main characteristics of the riser are taken into account such as the consumption of gas oil, the formation of gasoline and the endothermicity of the reaction with the decrease of temperature along the riser.

![Figure 2: Three lump kinetics](image)

In the riser model, a dimensionless spatial variable \( z \) is used, i.e. at bottom \( z = 0 \) and at top \( z = 1 \) corresponding to the actual height \( L \). The feed temperature \( T_{\text{ris}}(z=0) \) at riser inlet results from the following heat balance

\[
F_{\text{cat,reg}}C_{\text{pcat}}(T_{\text{reg}} - T_{\text{ris},0}) = F_{\text{feed}}[C_{\text{p,ol}}(T_{\text{boil}} - T_{\text{feed}}) + \Delta H_{\text{vap}} + C_{\text{p,rg}}(T_{\text{ris},0} - T_{\text{boil}})]
\]

that describes the heating of the liquid feed, its vaporization and finally the heating of the gas feed. \( F_{\text{cat,reg}} \) is the catalyst flow rate from the regenerator.

**Mass balance of gas oil**

\[
\frac{d y_{\text{go}}}{d z} = -k_{1}y_{\text{go}}C_{\text{owr}} \theta_{t_{c}} \tag{1}
\]

where \( k_{1} \) is the kinetic constant for gasoil consumption, \( C_{\text{owr}} \) is the ratio of catalyst flow rate to oil flow rate, \( y_{\text{go}} \) is the mass fraction of gasoil in the riser, \( t_{c} \) is the residence time of the catalyst in the riser. \( \theta \) is the deactivation factor of the catalyst due to coke deposition.

**Mass balance of gasoline**

\[
\frac{d y_{g}}{d z} = (\alpha_{2}k_{2}y_{\text{go}} - k_{3}y_{g})C_{\text{owr}} \theta_{t_{c}} \tag{2}
\]

where \( y_{g} \) is the mass fraction of gasoline.

**Energy Balance**

\[
\frac{d T_{\text{ris}}}{d z} = \frac{\Delta H_{\text{cark}}F_{\text{feed}}}{(F_{\text{cat,reg}}C_{\text{pcat}} + F_{\text{feed}}C_{\text{po}} + \lambda F_{\text{feed}}C_{\text{p,steam}})} \frac{d y_{\text{go}}}{d z} \tag{3}
\]

where \( \Delta H_{\text{cark}} \) is the heat of reaction and \( F_{\text{feed}} \) and \( F_{\text{cat,reg}} \) are the respective flow rates of the gasoil and catalyst. The kinetic constants follow Arrhenius law, \( \lambda \) is the weight fraction of steam in the feed stream to the riser. The deactivation of the catalyst by the coke deposition is given as

\[
\phi = (1 - mC_{\text{c,reg}}) \exp(-\alpha t_{c}zC_{\text{owr}}) \tag{4}
\]

Following [48], the produced coke concentration is empirically given by

\[
C_{\text{c,prod}} = k_{1}\sqrt{\frac{T_{\text{ris},1}}{R_{\text{c,irr}}} \exp\left(\frac{E_{\text{w}}}{RT_{\text{ris},1}}\right)} \tag{5}
\]

where \( T_{\text{ris},1} \) is the temperature at the riser outlet. The amount of coke leaving the riser is given by

\[
C_{\text{c,ris},1} = C_{\text{c,reg}} + C_{\text{c,prod}} \tag{6}
\]

### 2.2. Separator model

The residence time of catalyst in the separator is frequently of the order of one minute. This separator can be modelled as a Continuously Stirred Tank Reactor (CSTR).

**Mass balance of coke on catalyst**

\[
\frac{d C_{\text{c,sep}}}{d t} = F_{\text{cat,sep}}(C_{\text{c,ris},1} - C_{\text{c,sep}}) \tag{7}
\]

where \( C_{\text{c,sep}} \) is the coke concentration in the separator, \( F_{\text{cat,sep}} \) is the flow rate of catalyst from the riser, \( m_{\text{cat,sep}} \) is the catalyst holdup.

**Energy balance**

\[
\frac{d T_{\text{sep}}}{d t} = C_{\text{p,cat}}F_{\text{cat,sep}}(T_{\text{ris},1} - T_{\text{sep}}) \tag{8}
\]

where \( T_{\text{sep}} \) is the separator temperature.
2.3. Regenerator model

The regenerator model is inspired from [15]. The regenerator is a fluidized bed where air bubbles cross the dense bed formed by the catalyst. This bed is considered as a CSTR where the residence time of catalyst is frequently between ten and twenty minutes. The air bubbles through the bed could be modelled as a plug flow, however for ease of simulation, they are modelled as a CSTR. The temperature and amount of coke are considered uniform throughout the dense bed as well as the oxygen concentration. An important feature of the FCC is that the reactions in the riser are mainly endothermic whereas those in the regenerator are exothermic, thus the heat released in the regenerator is used by the riser by means of the transported catalyst. As the process involves a recycle, the behavior of a FCC is difficult to simulate correctly in both steady and transient states. In the FCC, partial combustion is assumed, i.e. carbon monoxide leaves the dense bed and is later oxidized to carbon dioxide in the freeboard, not represented here. The model equations are derived from mass and energy balances.

**Mass Balance of coke on the catalyst**

\[
\frac{d C_{\text{coked,reg}}}{dt} = \frac{(F_{\text{cat,spent}} C_{\text{coked,reg}} - F_{\text{cat,reg}} C_{\text{coked,reg}})}{m_{\text{cat,reg}}} - \dot{R}_{\text{cb}} \tag{9}
\]

where \( C_{\text{coked,reg}} \) is the coke concentration on catalyst, \( m_{\text{cat,reg}} \) is the catalyst inventory, \( F_{\text{cat,reg}} \) is the catalyst flow rate leaving the regenerator. The rate of coke combustion is given by

\[
\dot{R}_{\text{cb}} = k_{\text{cb}} \exp(-\frac{E_{\text{cb}}}{RT_{\text{reg}}}) x_{\text{O2}} C_{\text{coked,reg}} m_{\text{cat,reg}} \tag{10}
\]

**Energy balance**

\[
\frac{dT_{\text{reg}}}{dt} = \frac{1}{(m_{\text{cat,reg}} C_{\text{p,cat}})(C_{\text{p,cat}}(F_{\text{cat,spent}} T_{\text{sep}} - F_{\text{cat,reg}} T_{\text{reg}}))} \left( F_{\text{air,reg}} C_{\text{p,air}}(T_{\text{air}} - T_{\text{reg}}) - \Delta H_{\text{cb}} \frac{\dot{R}_{\text{cb}}}{M_{\text{coked}}} \right) \tag{11}
\]

where \( T_{\text{reg}} \) is the regenerator temperature, \( \Delta H_{\text{cb}} \) is the enthalpy of coke combustion.

**Mass balance of oxygen in the dense bed**

\[
\frac{d x_{\text{O2}}}{dt} = \frac{1}{m_{\text{sep,reg}}} \left( F_{\text{air,reg}} / M_{\text{w,air}} (x_{\text{O2,in}} - x_{\text{O2,reg}}) \right) - ((1 + \sigma) n_{\text{CH}} + 2 + 4 \sigma) / (4(1 + \sigma)) \dot{R}_{\text{cb}} / M_{\text{coked}} \tag{12}
\]

where \( \sigma \) is the molar ratio of CO2 to CO in the dense bed, \( n_{\text{CH}} \) is the number of moles of hydrogen per mole of carbon in the coke. The signification of parameters as well as their values for the FCC are given in Table 1. Typical steady-state values of the states and manipulated inputs are given in Table 2.

The code representing the FCC behaviour and all the Model Predictive Control codes were written in Fortran90. Efficient numerical and optimization routines were used [49]. Isoda [50] for integration of differential equations, NLPLQ [51] and BFGS [52] for nonlinear optimization. In the particular case of quadratic optimization, the only routines necessary for its solution were extracted from NLPLQ code.

### Table 1: FCC parameters, variables and initial data

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C_{\text{p,air}} )</td>
<td>Heat capacity of air ((\text{J} \cdot \text{kg}^{-1} \cdot \text{K}^{-1}))</td>
<td>1074</td>
</tr>
<tr>
<td>( C_{\text{p,cat}} )</td>
<td>Heat capacity of catalyst ((\text{J} \cdot \text{kg}^{-1} \cdot \text{K}^{-1}))</td>
<td>1005</td>
</tr>
<tr>
<td>( C_{\text{p,steam}} )</td>
<td>Heat capacity of steam ((\text{J} \cdot \text{kg}^{-1} \cdot \text{K}^{-1}))</td>
<td>1015.10^3</td>
</tr>
<tr>
<td>( E_{\text{af}} )</td>
<td>Activation energy for coke formation ((\text{J} \cdot \text{mol}^{-1}))</td>
<td>20.10^3</td>
</tr>
<tr>
<td>( E_{\text{f}} )</td>
<td>Activation energy for cracking of gas oil ((\text{J} \cdot \text{mol}^{-1}))</td>
<td>112.610^3</td>
</tr>
<tr>
<td>( f_{\text{air}} )</td>
<td>Mass flow rate of air to regenerator ((\text{kg} \cdot \text{s}^{-1}))</td>
<td>25.378</td>
</tr>
<tr>
<td>( f_{\text{cat,reg}} )</td>
<td>Mass flow rate of catalyst ((\text{kg} \cdot \text{s}^{-1}))</td>
<td>294</td>
</tr>
<tr>
<td>( m_{\text{air,reg}} )</td>
<td>Mass flow rate of feed ((\text{kg} \cdot \text{s}^{-1}))</td>
<td>40.63</td>
</tr>
<tr>
<td>( m_{\text{cat,reg}} )</td>
<td>Holdup of air in the regenerator ((\text{mol}))</td>
<td>20000</td>
</tr>
<tr>
<td>( m_{\text{sep,reg}} )</td>
<td>Holdup of catalyst in regenerator ((\text{kg}))</td>
<td>175738</td>
</tr>
<tr>
<td>( m_{\text{w,coke}} )</td>
<td>Molecular weight of coke ((\text{kg} \cdot \text{mol}^{-1}))</td>
<td>1074.3</td>
</tr>
<tr>
<td>( n_{\text{CH}} )</td>
<td>Number of moles of hydrogen per mole of carbon in the coke</td>
<td>2</td>
</tr>
<tr>
<td>( T_{\text{cat}} )</td>
<td>Temperature in the regenerator ((\text{K}))</td>
<td>360</td>
</tr>
<tr>
<td>( T_{\text{air}} )</td>
<td>Boiling temperature of the feed ((\text{K}))</td>
<td>700</td>
</tr>
<tr>
<td>( t_{\text{rise}} )</td>
<td>Residence time in the riser ((\text{s}))</td>
<td>9.6</td>
</tr>
<tr>
<td>( T_{\text{f,feed}} )</td>
<td>Feed temperature ((\text{K}))</td>
<td>434.63</td>
</tr>
<tr>
<td>( \alpha_{\text{cat}} )</td>
<td>Catalyst decay rate constant ((\text{s}^{-1}))</td>
<td>0.12</td>
</tr>
<tr>
<td>( \alpha_{\text{f,cb}} )</td>
<td>Fraction of gas oil which cracks to gasoline</td>
<td>0.75</td>
</tr>
<tr>
<td>( \Delta H_{\text{feed,vap}} )</td>
<td>Heat of feed vaporization ((\text{J} \cdot \text{kg}^{-1}))</td>
<td>1.56.10^3</td>
</tr>
<tr>
<td>( \Delta H_{\text{f,cb}} )</td>
<td>Heat of cracking ((\text{J} \cdot \text{kg}^{-1}))</td>
<td>506.210^3</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>Weight fraction of steam in feed to riser</td>
<td>0.035</td>
</tr>
<tr>
<td>( \rho )</td>
<td>Molar ratio of CO2 to CO in the dense bed</td>
<td>0.383</td>
</tr>
</tbody>
</table>

Table 2: GCC steady-state values of the states for \( F_{\text{air,reg}} = 25.571 \text{kg} \cdot \text{s}^{-1} \) and \( f_{\text{cat,reg}} = 295.03 \text{kg} \cdot \text{s}^{-1} \) at a given temperature set point

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C_{\text{coked,r1}} )</td>
<td>Coke concentration at top of riser ((\text{kg coke/kg cat}))</td>
<td>0.00974</td>
</tr>
<tr>
<td>( C_{\text{coked,r2}} )</td>
<td>Coke concentration in the regenerator ((\text{kg coke/kg cat}))</td>
<td>0.0305</td>
</tr>
<tr>
<td>( T_{\text{ris},1} )</td>
<td>Temperature in the riser ((\text{K}))</td>
<td>805.70</td>
</tr>
<tr>
<td>( T_{\text{sep}} )</td>
<td>Temperature in the separator ((\text{K}))</td>
<td>780.00</td>
</tr>
<tr>
<td>( T_{\text{reg}} )</td>
<td>Temperature in the regenerator ((\text{K}))</td>
<td>972.00</td>
</tr>
<tr>
<td>( x_{\text{O2}} )</td>
<td>Oxygen mole fraction in regenerator</td>
<td>0.0593</td>
</tr>
<tr>
<td>( Y_{\text{f,sep}} )</td>
<td>Mass fraction of gas oil at top of riser</td>
<td>0.458</td>
</tr>
<tr>
<td>( Y_{\text{g,sep}} )</td>
<td>Mass fraction of gasoline at top of riser</td>
<td>0.383</td>
</tr>
</tbody>
</table>

3. Comparison of Model Predictive Control Strategies

Several Model Predictive Control (MPC) strategies will be compared on the FCC process. Thus the simple model of the FCC can be considered as a valuable benchmark to perform a comparison of multivariable control schemes. The intention is not to provide the best control tuning for each control strategy, but rather to demonstrate how, by means of FCC model, it is possible to highlight the differences between different strategies and choices of control parameters. Table 3 shows the two manipulated inputs and the two controlled outputs of the FCC. Table 4 gives a brief summary of the different strategies which are compared in the present paper. All these strategies are very well detailed in [53, 54].

### Table 3: Manipulated inputs and controlled outputs

<table>
<thead>
<tr>
<th>Manipulated input</th>
<th>Manipulated output</th>
<th>Controlled output</th>
</tr>
</thead>
<tbody>
<tr>
<td>( u_{1} )</td>
<td>( u_{2} )</td>
<td>( y_{1} )</td>
</tr>
<tr>
<td>Mass flow rate of catalyst ( F_{\text{cat,reg}} )</td>
<td>Mass flow rate of air ( F_{\text{air,reg}} )</td>
<td>Temperature at top of riser ( T_{\text{ris},1} )</td>
</tr>
<tr>
<td>Temperature in the regenerator ( T_{\text{reg}} )</td>
<td>Temperature in the regenerator ( T_{\text{reg}} )</td>
<td></td>
</tr>
</tbody>
</table>

3.1. Identification by open loop responses

In many cases, open loop responses of the system are necessary to identify a model of the system which is later used for closed loop control. The sampling period is taken equal to 250s. In all cases, the model horizon was equal to 60 sampling time. In several of the MPC strategies (DMC, QDMC, MPC-penalty), the linear model of the
Table 4: Summary of control strategies

<table>
<thead>
<tr>
<th>Control strategy (Symbol)</th>
<th>Basis (Type of optimization)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic Matrix Control (DMC)</td>
<td>Linear model using step response coefficients, without constraints (analytical solution)</td>
</tr>
<tr>
<td>Quadratic Dynamic Matrix Control (QDMC)</td>
<td>Linear model using step response coefficients, with hard constraints on $u$ and $\Delta u$ (quadratic programming)</td>
</tr>
<tr>
<td>Model Predictive Control (MPC-penalty)</td>
<td>Linear model using step response coefficients, with hard constraints on $u$, $\Delta u$ and soft constraints on $y$ (nonlinear optimization with penalty function)</td>
</tr>
<tr>
<td>NonLinear Model Predictive Control (NLMPC)</td>
<td>Nonlinear model of the process, with hard constraints on $u$, $\Delta u$ (nonlinear optimization)</td>
</tr>
<tr>
<td>Observer Based Model Predictive Control (OBMPC)</td>
<td>State-space linear model using step response coefficients, with hard constraints on $u$, $\Delta u$, Kalman filter (quadratic programming)</td>
</tr>
</tbody>
</table>

The plant is represented by a stable step response, i.e. the step response of the studied controlled outputs with respect to all the manipulated inputs. Even, OBMPC uses this information, but in a different way. Thus, step variations of mass flow rate of catalyst (5%) and mass flow rate of air (0.5%) are imposed to the process around a steady state (Figure 3). The responses of the controlled outputs with respect to the mass flow rate of catalyst are shown in Figure 4 as they are particularly interesting. The normalized time is indicated in numbers of sampling periods. When the step on $u_1$ is imposed at normalized time 81, immediately the temperature at the top of the riser increases abruptly. This is due to the algebraic character of the riser model equations, i.e. its dynamics is neglected with respect to that of the rest of the system. After that, the riser temperature decreases and then increases, thus showing an inverse response characteristics. The first decrease is due to the fact that the endothermic heat of reaction overcomes the sensible heat brought by the catalyst, but later the regenerator temperature will increase and finally cause increase of bottom riser temperature and consequently top riser temperature. The regenerator temperature is influenced by the riser temperature and will vary in a relatively similar way to the step of mass flow rate of catalyst. The responses of the top riser temperature and regenerator temperature to steps of mass flow rate of air are simpler and close to first order responses. Of course, the step response coefficients (Figures 5 and 6) display the same dynamical tendencies that were previously described.

Figure 3: Mass flow rate of catalyst and mass flow rate of air imposed for open loop identification

Figure 4: Responses of top riser temperature and regenerator temperature used for open loop identification

Figure 5: Unit step response coefficients for both outputs with respect to manipulated mass flow rate of catalyst

Figure 6: Unit step response coefficients for both outputs with respect to manipulated mass flow rate of air

3.2. Dynamic Matrix Control

3.2.1. General presentation

To simplify the presentation, DMC is first presented in a SISO framework [55, 56, 53, 54]. At time $k$, the following quadratic criterion taking into account the differences between the estimated outputs $\hat{y}(k+i|k)$ and the references on the prediction horizon $H_p$

$$J = \sum_{i=1}^{H_p} \| \hat{y}(k+i|k) - y^{ref}(k+i) \|^2$$

is minimized, in the absence of constraints, with respect to the variations of $\Delta u(k)$ of the manipulated input considered over a control horizon $H_c$.

The plant is represented by a stable step response, i.e. the step response coefficients. The prediction of the output based on past and future inputs is given by

$$\hat{y}(k+l|k) = y_{ss} + \sum_{i=1}^{H_u-1} h_i \Delta u(k+l-i) + h_M(u(k+l-M) - u_{ss}) + \sum_{i=1}^{H_d} h_i \Delta u(k+l-i) + d(k+l|k)$$

where $H_d$ is the model horizon which must be larger than or equal to the prediction horizon $H_p$. The influence of unmodelled dynamics and disturbances is globally taken into account by the disturbance term

$$d(k) = y(k) - y^*(k|k)$$
Given the measured output $y^m(k)$, the future disturbances are estimated as

$$d(k+l|k) = d(k|k) = y^m(k) - y^*(k|k)$$

(16)
i.e. all the future predicted disturbances are equal to the present disturbance estimation.

The output prediction based on past inputs is defined as

$$y^*(k+l|k) = y_{sa} + \sum_{i=1}^{M-1} h_i\Delta u(k+l-i) + h_M(u(k-1) - u_{sa})$$

(17)
The vector of output predictions $\mathbf{y}(k+l|k)$ is related to the vector of output predictions $y^*(k+l|k)$ based on past inputs, to the vector of inputs $\Delta u(k)$ and to the vector of predicted disturbances as

$$\mathbf{y}(k+l|k) = \begin{bmatrix} y^*(k+1|k) \\ \vdots \\ y^*(k+H_p|k) \end{bmatrix} + sA \begin{bmatrix} \Delta u(k) \\ \vdots \\ \Delta u(k+Hc-1) \end{bmatrix} + \begin{bmatrix} d(k+1|k) \\ \vdots \\ d(k+H_p|k) \end{bmatrix}$$

where $sA$ is the dynamic matrix made of step response coefficients $h_i$ of the plant outputs to the manipulated inputs.

$$sA = \begin{bmatrix} h_1 & 0 & \ldots & 0 \\ h_2 & h_1 & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ h_M & h_{M-1} & \ldots & h_1 \\ \vdots & \vdots & \ddots & \vdots \\ h_M & h_{M-1} & \ldots & h_{M-Hc+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_M & h_M & \ldots & h_M \end{bmatrix}$$

(18)

For a multivariable system of dimension $n_u \times n_y$, the dynamic matrix is simply composed of submatrices as

$$sA = \begin{bmatrix} sA_{11} & \cdots & sA_{1n_u} \\ \vdots & \ddots & \vdots \\ sA_{n_u,1} & \cdots & sA_{nn_u} \end{bmatrix}$$

(19)

According to past equations, the vector of future input moves is given as

$$\Delta u(k) = [\Delta u_1(k) \ldots \Delta u_{n_u}(k)]^T$$

(20)
which is the least-squares solution of the following linear system

$$\begin{bmatrix} y^{ref}(k+1) - y^*(k+1|k) - d(k|k) = e(k+1) \\ \vdots \\ y^{ref}(k+H_p) - y^*(k+H_p|k) - d(k|k) = e(k+H_p) \end{bmatrix} = \begin{bmatrix} e(k+1) \\ \vdots \\ e(k+H_p) \end{bmatrix}$$

(21)

In the absence of constraints, the least-squares solution is

$$\Delta u(k) = (sA^TsA)^{-1}sA^Te(k+1)$$

(22)

Only the first component of $u(k)$ is implemented at time $k$.

In fact, the modified criterion (24) of QDMC will be used even for DMC to rigorously use the same criterion.

The different MPC strategies which are presented were developed in Fortran90 and they can be applied for a stable system having any number of inputs and outputs. The values of the main parameters which influence the studied MPC strategies are given in Table 5.

### 3.2.2. Numerical results

In the open loop responses, it has been observed that when a step on the catalyst flow rate was imposed, the outputs exhibited an inverse response behaviour. It will be shown in the following that it has important consequences. The inverse response is even more difficult to handle than a pure delay to which it could be compared. For that reason, a modified criterion with a lower $H_{p,low}$ bound of the prediction horizon which can be varied is implemented as

$$J = \sum_{i=H_{p,low}}^{H_p} (\hat{y}(k+i|k) - y^{ref}(k+i))^2$$

(23)
The lower bound is used in MPC to avoid problems arising from the inverse response of processes and in our case we have a case of an inverse response. The set points have been chosen as positive and negative steps occurring at different times to emphasize the interaction effects that can influence the closed-loop control of such a process. Thus, with respect to the steady state, the set point of the riser temperature undergoes an increase of 5K at $t = 20000s$ and a decrease of 5K at $t = 40000s$ whereas the set point of the regenerator temperature undergoes an increase of 5K at $t = 60000s$ and a decrease of 5K at $t = 80000s$, the total simulation time being equal to 100000s.

**DMC case with $H_{p,low} = 1$.**

The first choice of parameters was $H_p = 20$, $H_{p,low} = 1$, $H_c = 3$. When the original criterion (13) is used (Figures 7 and 8), i.e. with $H_{p,low} = 1$, strong oscillations of the riser temperature are visible at each set point change corresponding to strong oscillations of the catalyst flow rate and also sharp variations of the air flow rate. The oscillations are due to the inverse response behaviour of the temperatures to variations of the catalyst flow rate. The set points are well tracked for the riser and regenerator temperatures.

**DMC case with $H_{p,low} = 10$.**

To avoid the oscillations visible with $H_{p,low} = 1$, a larger value of $H_{p,low}$ equal to 10 is considered. It must be noticed that the duration of the inverse response (Figure 4) is about 15 sampling periods so that the value of $H_{p,low}$ is close to that duration. In these conditions, compared to the previous case where $H_{p,low} = 1$, the behaviour of the closed loop system is much improved with a smoother tracking of the set points (Figure 9). However, the interaction effects are
Table 5: Conditions of Model Predictive Control Strategies

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Prediction horizon ( H_p )</th>
<th>Lower bound of prediction horizon ( H_{p,low} )</th>
<th>Control horizon ( H_c )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic Matrix Control, first case</td>
<td>20</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Dynamic Matrix Control, second case</td>
<td>20</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>Quadratic Dynamic Matrix Control</td>
<td>20</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>Model Predictive Control with nonlinear optimization by NLPQL</td>
<td>20</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Nonlinear Model Predictive Control</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Observer Based Model Predictive Control, first case</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Observer Based Model Predictive Control, second case</td>
<td>0.01</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

3.3. Quadratic Dynamic Matrix Control

In order to take into account the constraints, quadratic dynamic matrix control (QDMC) is used instead of DMC. Furthermore, a modification of the quadratic criterion as the sum of a performance term and an energy term is introduced in QDMC [56] as

\[
J = \sum_{i=1}^{H_c} \gamma_i ((y(k+i) - \tilde{y}_c)^2 + \sum_{i=1}^{H_c} \lambda_i^2 (\Delta u(k+i))^2) \tag{24}
\]

where the weights \( \gamma \) for the outputs are the elements of a diagonal matrix \( \Gamma \), and the weights \( \lambda \) for the inputs are the elements of a diagonal matrix \( \Lambda \). Hard constraints with respect to the manipulated variables and their moves are taken into account

\[
\begin{align*}
    u_{min} \leq u & \leq u_{max} \\
    \Delta u_{min} \leq \Delta u & \leq \Delta u_{max}
\end{align*}
\tag{25}
\]

These constraints can be gathered as a system of linear inequalities incorporating the dynamic information concerning the projection of constraints [57]

\[
B \Delta u(k) \leq c(k+1)
\tag{26}
\]

where \( B \) is a matrix and \( c \) a vector.

In the presence of constraints (25), the QDMC problem can be formulated as a quadratic programming problem [51] such as

\[
\begin{align*}
    \min_{\Delta u(k)} & \quad \frac{1}{2} \gamma_i ((y(k+i) - \tilde{y}_c)^2 + \lambda_i^2 (\Delta u(k+i))^2) \\
    \text{subject to constraints} & \quad (25), H \text{ is the Hessian matrix which is equal to} \\
    & \quad H = a^T \Gamma^T \Gamma a' + \Lambda^T \Lambda
\end{align*}
\tag{27}
\]

where \( a' \) is the dynamic matrix, \( g \) is the gradient vector equal to \( g(k+1) = a^T \Gamma^T \Gamma e(k+1) \)

\[
\tag{28}
\]

In all cases, the weights \( \Gamma \) and \( \Lambda \) were diagonal matrices with respective values equal to 10 and 1, putting more weight on the controlled outputs. The constraints on \( u \) were chosen in such a way that they were never reached and no constraints on \( \Delta u \) were considered. The choice of parameters \( H_p = 20, H_{p,low} = 10, H_c = 3 \), resulted in the same results as DMC with the same set of parameters. For that reason, the Figures are not shown. This is due to the absence of constraints so that the problems are exactly of the same nature and the numerical solution by optimizer NLPQL obtained for QDMC is the same as the analytical solution as DMC provided the same criterion (24) is minimized.

3.4. Model Predictive Control with nonlinear optimization by NLPQL

When soft constraints on the controlled outputs \( y \) are considered, the nature of the optimization problem changes with respect to QDMC. These soft constraints are taken into account by addition of a penalty term to the quadratic criterion (24) used in QDMC. To be able to compare with DMC and QDMC, the soft constraints on \( y \) were taken large so that they were never reached, making the penalty term equal to zero. In spite of this, the code was required to perform the non-linear optimization as if soft constraints were present. Thus, the control problem results in a general problem of non-linear optimization which is solved by NLPQL optimizer [51]. The same values of prediction horizon \( H_p = 20 \) and control horizon \( H_c = 3 \) were taken. However, between the two values of \( H_{p,low} \) equal to 1 and 10, the value \( H_{p,low} = 1 \) gave much better results which are presented in Figures 11 and 12. This \( H_{p,low} = 1 \) is due to the difficulty of the

![Figure 9: DMC control with \( H_{p,low} = 10 \). Left: Riser temperature (K). Right: Regenerator temperature (K).](image)

![Figure 10: DMC control with \( H_{p,low} = 10 \). Left: Catalyst flow rate (kg/s). Right: Air flow rate (kg/s).](image)
nonlinear optimization for the FCC. The set point tracking is excellent and the interaction effects at the set point changes are relatively small. The difference between these results and those of QDMC may be explained by the numerical approximations of the Lagrangian function and its gradient that are performed during the nonlinear minimization by a quasi-Newton method and which would result in a filtering and smoothing of the oscillations previously present in QDMC.

Figure 11: MPC control with nonlinear optimization by NLPQL and \( H_p, \text{low} = 1 \). Left: Riser temperature (K). Right: Regenerator temperature (K).

Figure 12: MPC control with nonlinear optimization by NLPQL and \( H_p, \text{low} = 1 \). Left: Catalyst flow rate (kg/s). Right: Air flow rate (kg/s).

3.5. Nonlinear Model Predictive Control with prediction from the nonlinear model described by ordinary differential equations

Nonlinear Model Predictive Control exists under several different forms [53, 58]. In particular, it can be performed by local linearization of the nonlinear state space model \([59, 60]\) at each sampling instant. In the three previously described versions of MPC (DMC, QDMC, MPC with penalty on \( y \)), the outputs were predicted using the linear model based on the step response coefficients. In the nonlinear MPC described in this section, the nonlinear model

\[
\begin{align*}
\dot{x} &= f(x, u) \\
y &= h(x)
\end{align*}
\]  

constituted by the equations of the riser, the separator and the regenerator, is directly used to obtain the predicted states and consequently the predicted outputs as follows. At each sampling instant \( k \), the nonlinear model (30) is integrated from \( k \) until \( k + H_p \), i.e. over the prediction horizon by the integrator lsoda \([50]\) well suited to stiff systems. From the calculated values of the states \( x(k + i), 1 \leq i \leq H_p \), considered as the predicted states \( \hat{x}(k + i) \), the predicted outputs \( \hat{y}(k + i) \) are deduced. They are directly used in the calculation of the same criterion (24) as QDMC, but now the criterion depends in a nonlinear way on the manipulated inputs, opposite to the two first MPC control strategies (DMC, QDMC). Consequently, the optimization to be performed with respect to the future variations of the manipulated inputs \( \Delta u(k + i) \) is nonlinear. Thus, the NLPQL optimizer \([51]\) and BFGS routine \([52]\) are used in the Fortran90 code to perform the nonlinear optimization. In the optimization process, it appears that it is more difficult to achieve a successful optimization at each sampling instant than in the other cases. This may be due to the highly nonlinear character of the FCC. Thus, the large value of the prediction horizon \( H_p = 20 \) resulted in a failure of optimization. Considering \( H_p = 20 \) and \( H_p, \text{low} = 10 \) was no more successful. The best choices which were found corresponded to small values of prediction horizon \( H_p \) and very small control horizon \( H_c \), typically \( H_p = 3 \) and \( H_c = 1 \). Moreover, at each sampling instant \( k \), to find the optimum solution with respect to the manipulated input variations, the optimizer calculates the criterion a large number of times until some tolerance is reached, and for each calculation of the criterion, the nonlinear equations must be integrated. This resulted in a very long computation time which would not be adapted for on-line optimization and control. An important drawback of such a nonlinear optimization problem to be solved for control and on-line implementation is that it is required to guarantee successful optimization at each sampling instant. Other forms of nonlinear control are presented in the literature such as by \([59, 60]\), but the local linearization which is proposed by these latter authors make it closer to linear MPC than real Nonlinear MPC.

Figure 13 shows that the regulation and tracking are well performed, but that the way to reach the setpoints and the interactions shows significant improvement with respect to the linear MPC strategies as the outputs exhibit lower deviations with respect to their respective set points around the set point changes. Figure 14 shows that, in general the variations of the manipulated inputs are smooth, especially the flow rate of catalyst. However, the set point tracking is less satisfactory.

Figure 13: Nonlinear MPC control with \( H_p = 3, H_c = 1 \). Left: Riser temperature (K). Right: Regenerator temperature (K).

Figure 14: Nonlinear MPC control with \( H_p = 3, H_c = 1 \). Left: Catalyst flow rate (kg/s). Right: Air flow rate (kg/s).

3.6. Observer Based Model Predictive Control

3.6.1. General presentation

Many state-space approaches exist for MPC. They were first developed for linear MPC \([61, 62, 63]\) and some of them have been later extended to nonlinear MPC \([58]\). As well as for DMC and QDMC, a predictor of the output is first built, then a state observer is developed to estimate the states. This observer-based model predictive control (OBMPC) developed by \([61, 62]\), is briefly presented \([49, 54]\) in the following.

A fundamental idea \([64]\) is to represent the entire trajectory for a
SISO system as a sequence of states
\[ x_1(k) = x_2(k - 1) + h_1 \Delta u(k - 1) \]
\[ \vdots \]
\[ x_i(k) = x_{i+1}(k - 1) + h_i \Delta u(k - 1), \quad i = 1, M \]
where \( h_i \) are the step response coefficients. These equations are then used to obtain the predicted output corresponding to the influence of past variations
\[ y^*(k + i - 1) = y^*(k + i - 1|k - 1) + h_i \Delta u(k - 1) \]
(32)

In the case of a MIMO system with \( n_u \) inputs and \( n_y \) outputs, the matrix \( S_k \) is defined at each instant \( k \)
\[ S_k = \begin{bmatrix}
  h_{1,1,k} & h_{1,2,k} & \cdots & h_{1,n_u,k} \\
  h_{2,1,k} & h_{2,2,k} & \cdots & h_{2,n_u,k} \\
  \vdots & \vdots & \ddots & \vdots \\
  h_{n_u,1,k} & h_{n_u,2,k} & \cdots & h_{n_u,n_u,k}
\end{bmatrix} \]
(33)
where \( h_{j,i,k} \) is the step coefficient at instant \( k \) of output \( j \) with step input \( i \).

At time \( k \), the future control actions \( \Delta u(k+1) \) are to be determined over the control horizon.

In the absence of disturbances, the state-space form corresponding to the step response model can then be written as
\[ Y(k) = \Phi Y(k - 1) + S \Delta u(k - 1) \]
\[ y^*(k|k) = \Psi Y(k) \]
(34)
with
\[ Y(k) = [y^*(k|k)^T \ y^*(k+1|k)^T \ \ldots \ y^*(k+M-1|k)^T \ x_u(k)^T \ x_w(k)^T]^T \]
\[ y^*(k+i|k) = [y^*_1(k+i|k) \ \ldots \ y^*_n_y(k+i|k)]^T \]
\[ \Delta u(k) = [\Delta u_1(k) \ \ldots \ \Delta u_{n_u}(k)]^T \]
(35)
Similarly to the the LQG state-space model, the state-space model (34) can be extended [62] to take into account disturbances \( w \) and output noise \( v \) as
\[ Y(k) = \Phi Y(k - 1) + S \Delta u(k - 1) + T \Delta w(k - 1) \]
\[ y(k) = y^*(k|k) + v(k) \]
\[ y^*(k|k) = \Psi Y(k) \]
(36)
with the matrix \( \Phi \) equal to
\[ \Phi = \begin{bmatrix}
  0 & I_{n_u} & 0 & \cdots & 0 & 0 & 0 \\
  0 & 0 & I_{n_u} & \cdots & 0 & 0 & \vdots \\
  \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\
  0 & 0 & \cdots & I_{n_u} & 0 & 0 & 0 \\
  0 & 0 & \cdots & I_{n_u} & C_u & C_w & 0 \\
  0 & 0 & \cdots & 0 & A_w & 0 & 0 \\
  0 & 0 & \cdots & 0 & 0 & A_w & 0
\end{bmatrix} \]
(37)
the matrix \( S \) equal to
\[ S = [S_1 \ \ldots \ S_M \ B_u \ 0]^T \]
(38)
the matrix \( \Psi \) equal to
\[ \Psi = [I_{n_u} \ 0 \ \ldots \ 0] \]
(39)
and the matrix \( T \) equal to
\[ T = \begin{bmatrix}
  0 & 0 & \cdots & B_{u,w}
\end{bmatrix}^T \]
(40)
Residual plant dynamics and disturbance dynamics are considered under a state-space form using matrices \( A_u, A_w, B_{u,w}, \) and \( C_u, C_w \) [54, 62].

The future outputs are predicted by means of a state observer such as the optimal linear Kalman filter of matrix gain \( K \), operating in two stages, first, the model prediction
\[ \hat{Y}(k+1|k) = \Phi \hat{Y}(k|k) + S \Delta u(k) \]
(41)
then, the correction based on measurements
\[ \hat{Y}(k|k) = \hat{Y}(k|k - 1) + K [y(k) - \hat{Y}(k|k - 1)] \]
(42)
with
\[ \hat{Y}(k|k - 1) = [\hat{y}^*(k|k - 1)^T \ \hat{y}^*(k+1|k - 1)^T \ \ldots \ \hat{y}^*(k+M-1|k)^T]^T \]
\[ y^*(k+1|k) = \Psi \hat{Y}(k|k) \]
(33)

After state estimation, the optimization is performed as a quadratic programming problem like QDMC [56]. The objective function to be minimized with respect to \( \Delta w(k) \) is
\[ J = || \Gamma [\Delta w(k+1|k) - \Delta w(k+1|k)] ||^2 + || \Delta \Delta w(k|k) ||^2 \]
(44)
In the absence of constraints, the least-squares solution of OB MPC expressed by criterion (44) is
\[ \Delta \Delta w(k|k) = \left[ S_{H_p} \Gamma \Delta \Gamma + \Delta \Gamma \right]^{-1} S_{H_p} \Gamma \Delta \Delta w(k+1|k) - \Phi_{H_p} \hat{Y}(k|k) \]
(45)
where \( S_{H_p} \) is composed of matrices \( S_j \) [54, 62].
Thus, opposite to DMC, OB MPC uses an optimal state observer, the Kalman filter, for the predictor part. According to [62], the different decomposition of OB MPC as a predictor-optimizer avoids limitations of DMC-type algorithms, such as excessive number of step response coefficients, poor behaviour when submitted to ramp-like disturbances or in presence of strong interactions and the necessary stable behaviour of the plant.

3.6.2. Numerical results

The general parameters were always \( H_p = 20, H_{p,low} = 10, H_c = 1 \), as they gave very satisfactory results for tracking and regulation.

OB MPC with \( \sigma_v = 0.01 \)

The case where the standard deviation \( \sigma_v = 0.01 \), i.e. close to zero, for both measured outputs, is first considered to compare this strategy based on a state space approach with the strategies based on step responses such as QDMC. In Figures 15 and 16, the set points are very well tracked; the multivariable interactions are much less important than with DMC and QDMC; the manipulated inputs vary very smoothly. Thus, globally the control is well improved.
OBMPC with \( \sigma_y = 0.5 \)

When a realistic noise of standard deviation \( \sigma_y = 0.5 \) is added to simulate noise on temperature measurement, the tendencies observed in the absence of measurement noise are preserved, i.e. outputs (Figure 17) follow very well the set points and variations of the manipulated inputs are smooth (Figure 18). Compared to all the previous MPC strategies, OBMPC presents the best performance. Obviously, the strong interactions present in the plant are better handled than by DMC-type algorithms, as claimed by [62].

![Figure 16](image1.png)

**Figure 16**: Observer Based Model Predictive Control with \( H_p = 20, H_p,low = 1, \sigma_y = 0.01 \). Left: Catalyst flow rate (kg/s). Right: Air flow rate (kg/s).

![Figure 17](image2.png)

**Figure 17**: Observer Based Model Predictive Control with \( H_p = 20, H_p,low = 1, \sigma_y = 0.5 \). Left: Riser temperature (K). Right: Regenerator temperature (K).

![Figure 18](image3.png)

**Figure 18**: Observer Based Model Predictive Control with \( H_p = 20, H_p,low = 1, \sigma_y = 0.5 \). Left: Catalyst flow rate (kg/s). Right: Air flow rate (kg/s).

4. Conclusion

A nonlinear model of a Fluidized Catalytic Cracker (FCC) is presented. It can serve as a benchmark for Model Predictive Control (MPC) as several multivariable control structures can be tested. Furthermore, it presents important nonlinearities, interactions between the outputs and the outputs display an inverse response behaviour with respect to the catalyst flow rate that is one of the manipulated inputs. Several MPC strategies are presented and tested on this FCC model. DMC with analytical solution, QDMC with quadratic optimization, linear MPC with a penalty function and nonlinear optimization, Nonlinear MPC with model integration over the prediction horizon and nonlinear optimization, Observer Based MPC with a state-space model, an observer and quadratic optimization. The best performance was obtained by OBMPC, but QDMC and MPC-penalty gave also very good results. The most difficult control strategy was Nonlinear MPC which was not easy to tune and is not suited for on-line control, at least under the presented form. Hard constraints on the manipulated inputs and their moves and soft constraints on the controlled outputs could be added to this problem, however, for comparison purpose, it was judged not necessary to introduce them.

Acknowledgement

I am greatful to the university of Douala and Buea for all the help.

References
