

## OPTIMIZATION-BASED CONTROL - SOME CRITICAL ISSUES

Bjarne A. Foss\* Halgeir Ludvigsen\*\* Stein O. Wasbø\*\*\*

\* *Department of Engineering Cybernetics, NTNU, Trondheim,*

*e-mail: Bjarne.Foss@itk.ntnu.no*

\*\* *Cybernetica AS, Trondheim*

\*\*\* *Eramet Norway AS, Trondheim*

Abstract: This paper investigates critical issues in model development for nonlinear MPC. The results are based on experience with industrial application as well as theoretical work on MPC. *Copyright ©2002 IFAC*

Keywords: Optimization-based control, model predictive control, industrial application.

### 1. INTRODUCTION

Optimization-based control refers to a class of control algorithms where the control input is computed as the solutions of an optimization problem on some time-horizon (or prediction horizon) from the current time and onwards. One important class of such algorithms is model predictive control (MPC). In MPC the first part of the computed open-loop control input trajectory is injected to the system and the system is re-optimized at a later stage to calculate a new control input trajectory. Hence, the prediction is moved forwards in time and thereby the alternative name receding horizon control. There exists good survey papers on MPC, see e.g. Qin and Badgwell and Rawlings. The optimization problem being the heart in all MPC applications consists of three parts; an objective function, equality constraints and inequality constraints. The equality constraints are defined by the dynamic model of the system while the inequality constraints are linked to constraints like control input limitations. The term nonlinear MPC (NMPC) refers to problems in which there are nonlinear equality constraints induced by a nonlinear dynamic model. An excellent survey on this is provided by Qin and Badgwell (2000). Following this reference a typical NMPC optimization problem is defined as follows:

$$\begin{aligned} \min_{\{u_k, \dots, u_{k+M}\}} J = & \sum_{j=k+1}^{k+N} \|e_j^y\|_{Q_j}^2 + \sum_{j=k}^{k+M} \|u_j - u_{j-1}\|_{S_j}^2 \quad (1) \\ & + \sum_{j=k}^{k+M} \|e_j^u\|_{R_j}^2 + \|s\|_T^2 \\ & \text{subject to} \\ & x_{j+1} = f(x_j, u_j, v_j) \quad \forall j \in \{k, \dots, k+N-1\} \\ & y_j = g(x_j) \quad \forall j \in \{k+1, \dots, k+N\} \\ & \underline{y}_j - s \leq y_j \leq \bar{y}_j + s \quad \forall j \in \{k+1, \dots, k+N\} \\ & \underline{u}_j \leq u_j \leq \bar{u}_j \quad \forall j \in \{k, \dots, k+M\} \\ & \underline{\Delta u}_j \leq u_j - u_{j-1} \leq \bar{\Delta u}_j \quad \forall j \in \{k, \dots, k+M\} \\ & s \geq 0 \\ & x_j \in R^n, \quad y_j \in R^{m_y}, \quad u_j \in R^{m_u}, \quad v_j \in R^{m_v} \end{aligned}$$

$Q_j, S_j, R_j, T$  refers to positive, symmetric semidefinite matrices that defines (possibly) time-varying norms.  $k$  is the current time and  $N$  is the prediction horizon.

The objective function weights the control error  $e_j^y = y_j - y_j^{ref}$  according to some (possibly) time-varying norm on an  $N$ -step prediction horizon.  $y_j^{ref}$  is a time-varying reference value for the system output. The 2nd term considers control input changes and the 3rd term penalizes deviation from

control input reference values ( $e_j^u = u_j - u_j^{ref}$ ). These terms are defined on the control horizon  $M$  which typically is shorter than the prediction horizon ( $M < N$ ). The fourth term of the objective function weighs output constraint violations.

The model is defined by a discrete nonlinear dynamic model state-space which may include measurable disturbances  $v_j$ . Further, the system output is defined by a nonlinear equation. Finally, note that the inequality constraints are linear.

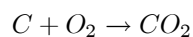
The scope of this paper is to delve into some critical issues when developing an optimization-based control application. We focus on applications that use mechanistic models and are particularly interested in the difference between the typical formulation seen in the open literature, e.g. (1), and formulations that are used in industrial applications. In this paper the dynamic model and the optimization problem will be the focal points.

The reference cases for this paper are applications for batch processes. Batch process optimization can be divided into three categories (Bonvin (1998)): (i) One-time optimization, (ii) run-to-run optimization and (iii) on-line optimization. In run-to-run optimization information from previous runs are used to optimize the current run, e.g. by updating the dynamic model. On-line optimization means re-optimization of the computed control input trajectories as in MPC. The metal refining case is a run-to-run optimization application while the NMPC application for a batch PVC-reactor is an on-line optimization application.

The paper is structured as follows: First, the metal refining case is presented. The paper's two main sections, on the model and the optimization problem, respectively, succeeds this. Some additional issues are included in a discussion section before the paper ends with some conclusions.

## 2. CASE DESCRIPTION AND APPLICATION

The reference case is a metal refining process for removing carbon from manganese metal. The process is sketched in figure 1. It consists of a ladle which is filled with liquid-phase high-carbon ferromanganese metal. This implies that about 7% of the metal consists of carbon. In addition there is some iron and  $MnO$  in the metal. Carbon is removed by blowing  $O_2$  at a high velocity rate on to the surface of the metal bath.  $O_2$  is supplied from a lance. The main overall reaction is



The refining process produces different products with a carbon content in the range 0.5% – 1.5%. Downstream the refining process the metal is

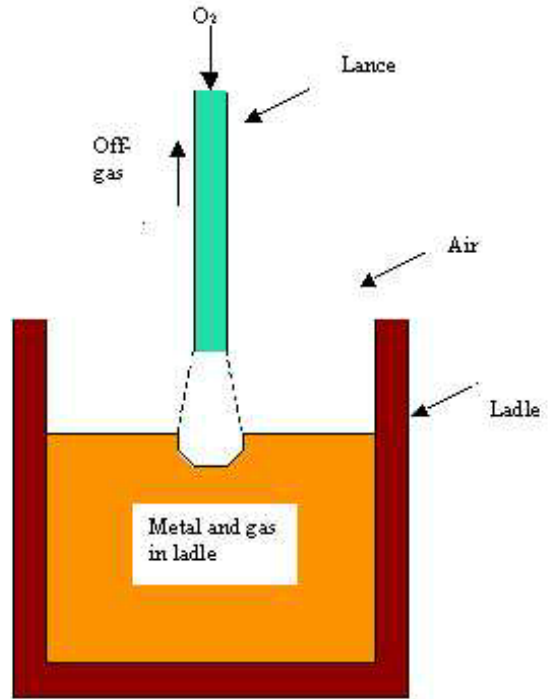
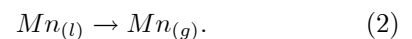


Fig. 1. The figure depicts the metal refining process.

casted, crushed and screened before it is packed and shipped to customers.

In addition to the main reaction there are intermediate reactions as well as side reactions. One important side reaction is evaporation of manganese metal.



Fumes generated during the batch are collected in an off-gas system and routed to a filter-system for removing dust.

The metal refining process is operated as a fed-batch process. The batch sequence is as follows:

- Fill the ladle with high-carbon ferromanganese metal
- Start  $O_2$ -blowing
- Empty the metal in the ladle

More information on  $Mn$  decarburization can be found in Dresler (1989).

The length of a batch sequence is in the order of 1 – 2 hours, and the ladle will typically contain 30–40.000kg of metal. The same ladle is used from one batch to the next. It is replaced when the inner lining becomes too thin. The economic incentive for improved control is minimizing metal loss due to evaporation, while satisfying an upper limit on the concentration of carbon at the end of the batch.

The data from the instrumentation system and the laboratory analysis are of good quality. This includes data of the mass and temperature of the metal before starting  $O_2$ -blowing and at the end of the batch cycle. Further, the precision of the control inputs is high. The online data which is received after  $O_2$ -blowing starts and before the ladle is emptied are less precise. The online data consist of temperature and  $CO_2$ -concentration in the off-gas system. Based on the instrumentation and control input infrastructure it was decided to develop a model-based control strategy to improve operations, or more specifically, an optimization-based control strategy. To elaborate, the lack of good quality online data was one reason for choosing a run-to-run update strategy instead of on-line optimization. If economically interesting an NMPC strategy will be considered at a later stage.

Trial revealed the necessity to use a nonlinear model in the application since nonlinear effects are very pronounced during the course of the batch. A physically-based dynamic model was developed and validated. Further, an optimization problem consistent economic incentive for the project was defined. The application was developed using the CENIT-software developed and marketed by Cybernetica AS, and interfaced to the current data system by adding new tables to an existing online database. CENIT runs on a Windows 2000/NT/XP platform. The system was commissioned in May 2002. Experience so far indicates that the project goals will be met.

In addition to the reference case we will also include experience from an NMPC application for a batch PVC-reactor reported in Schei, Ludvigsen, Singstad and Foss (2001).

### 3. MODEL DEVELOPMENT

In this section we discuss critical issues in model development. These include the choice of system boundaries, model states, model smoothness, data and information collection, and model implementation.

#### 3.1 System boundaries

The choice of system boundaries is important in all system-oriented methods. The system boundaries may be defined on different levels. One level of particular interest to us is the system defined by the model in the optimization problem. Referring to the reference case shown in figure 1 it may be divided into several parts:

- The liquid and gas within the ladle
- The ladle

- The lance which supplies  $O_2$
- The off-gas collection system
- Process instrumentation

It is not immediate which parts to include in the model. The goal of the application, however, directs focus towards the liquid and gas within the ladle. The model must capture the main kinetics to be able to predict decarburization, i.e. removal of  $C$ , and  $Mn$  evaporation accurately. The ladle need not be included in the dynamic model since its states can be regarded as constant during a batch.

Parts of the off-gas collection system must be modelled by a dynamic model as a means to utilize the temperature and  $CO_2$ -concentration data from the instrumentation located in the off-gas collection system.

Other instrumentation for measuring the mass, concentration and temperature of the metal before starting  $O_2$ -blowing and at the end of the batch cycle is not included in the model. These provide either initial states for the model or measurements for updating the model between two batches.

The current state is required in (1). In run-to-run optimization this means the initial state, i.e. the model states before starting  $O_2$ -blowing. The system boundaries for the dynamic model do not coincide with the measurements that are available before starting  $O_2$ -blowing. To overcome this a static model is used to transform input data to coincide with input data for the dynamic model. To be more specific (i) weight measurements of the ladle before and after filling it with metal for the current batch, (ii) the ladle history, (iii) the concentration data for the input metal, and (iv) the time lag since the last batch, are used to estimate the initial state of the dynamic model.

#### 3.2 Model structure and validation

The rationale for choosing a physically-based model was fivefold.

- A mechanistic model is able to capture fundamental knowledge about the system. In the present case knowledge on the flow pattern in the ladle, species, reaction kinetics and thermodynamics were used to choose the model structure.
- The refining system itself changes from time-to-time. This includes process changes, e.g. changes to the instrumentation system, and new products. It is particularly important to be able to integrate future online measurements, e.g. improved measurements in the off-gas collection system, into the model. This will lead to limited changes in the mechanistic model for the refining process.

- The application is used by metallurgists. To gain confidence in the application it is important to include the fundamental chemical and physics knowledge into the model as opposed to presenting a "black box" model.
- The company Cybernetica AS specializes in applications where the prediction properties of the nonlinear model is critical for success.
- Early tests supported the hypothesis that a physically-based model could predict decarburization and *Mn*-evaporation accurately.

The term physically-based model should be qualified since there is one part of the model which is purely empirical. This is the model of the  $O_2$ -blowing system and its coupling to the metal and gas in the ladle. A mechanistic model of this would require substantial resources. Hence, a nonlinear empirical model structure was selected.

Selected model parameters were estimated on the basis of data series describing the most important operating condition for the refining process. This included different initial conditions and different products. The number of data series was about 20. Afterwards, the model was validated on more than 200 data series. In practice parameter estimation and model validation is no linear process in the sense that several iterations are necessary.

Model validation revealed the need to update some parameters on a batch-to-batch basis to achieve adequate prediction accuracy. This is done using a sequential quadratic programming method which includes bounds on the allowable parameter set.

It is quite important to note that data series do not contain sufficient information for model identification. It is always necessary to supplement this with knowledge provided by the operational staff. This may include detailed information on the instrumentation (what is actually measured, are the data filtered, location of the sensor), and disturbances that are not measured directly. The latter may include operational problems in the upstream *Mn*-furnace, maintenance work that affects production, and different shift teams.

### 3.3 States

The dynamic model includes three types of states. First, there are the masses of the species in the ladle liquid phase; e.g. *Mn*, *Fe*, and *C*; temperatures in the reactor and off-gas collection system; and sensor dynamics in the off-gas collection system. The total number of these states is 12.

Second, there are the masses of the species in the gas phase; e.g.  $O_2$ , and *CO*. The gas phase appears both in the ladle and in the off-gas system. The dynamic modes of the gas phase are much faster (in the range of seconds) than the

dynamic modes that are important for predicting decarburization and *Mn* evaporation (minutes and upwards). Hence, they could have been converted to a static model. In the metal refining case artificial gas phase dynamics were included. These dynamics are slower than the actual dynamics of the gas phase, but at the same time significantly faster than the minute-scale dynamics. By this, the solution satisfied two requirements: (i) The gas phase dynamics do not interfere significantly with the slower dynamics that are important for the prediction properties. (ii) The dynamic model has limited stiffness which is important for fast numerical integration. The total number of gas-related states is 18.

Third, the dynamic model include auxiliary states. These are states that are used in the optimization problem to be discussed later, or states that are used by the operators for surveillance purposes. These states are all (simple) integrators. Examples are the consumption of  $O_2$  and the total evaporation of *Mn* after starting  $O_2$ -blowing. The total number of these states is 4.

### 3.4 Smoothness

In addition to prediction accuracy a model used in optimization-based control should be smooth to facilitate the search algorithm for solving the optimization problem. Hence, it is not adequate to develop a model with good prediction accuracy. The model must also be smooth with respect to the control inputs eligible for optimization. This has been a key issue to obtain robust and computationally efficient performance of the optimization algorithm both in the reference case on metal refining and in the PVC-application reported in Schei et al. (2001).

To elaborate, in the reference case several kinetics models are non-smooth. Non-smooth functions were transformed by using sigmoid-functions. To illustrate, assume the following quite simple kinetic model for the reaction rate  $r$  for the reaction  $A \rightarrow B$ .

$$r = \begin{cases} a(p_B - p_{equil}) & \text{if } p_B > p_{equil} \\ 0 & \text{if } p_B \leq p_{equil} \end{cases}$$

$a > 0$  is some constant,  $p_B$  is the partial pressure of (gas) component  $B$ , and  $p_{equil}$  is the equilibrium partial pressure. A smooth approximate model for the reaction rate is

$$r = h(p_B, p_{equil}) \cdot [a(p_B - p_{equil})]$$

where

$$h(p_B, p_{equil}) = \frac{1}{1 + e^{-\alpha(p_B - p_{equil})}}, \quad \alpha > 0$$

Note that this approximation does allow negative reaction rates.

#### 4. THE OPTIMIZATION PROBLEM

In this section we discuss critical issues in defining the optimization problem. These include the choice of the objective function, inequality constraints, control input parametrisation and the infeasibility problem.

##### 4.1 Prediction horizon

In the reference case the prediction horizon was not included as a free variable in the optimization. The total batch time is precalculated as a function of the initial metal weight and the product in question. More metal requires longer batch time. So does a (final) product with lower  $C$  concentration than a product with higher  $C$ . There were three reasons for this strategy. First, the metal refining system does not limit total production, i.e. it is no bottleneck in the production chain. Second, the precalculated batch time is equal to the batch times used in the former control system. It is reasonable to waive changes that do not improve performance. Third, the complexity of the optimization algorithm is limited by the computational resources available. Including batch time in the optimization problem increases computational load.

As a contrast to the reference case, in the PVC-application reported in Schei et al. (2001) the minimization of the batch time was important to increase production capacity.

##### 4.2 Objective function

Two alternative approaches to select the objective function is the use of economically inspired terms or terms directly linked to physical variables. The latter approach would imply that the objective function should minimize the evaporation of  $Mn$ , typically

$$J_1 = \sum_{j=k+1}^{k+N} \|r_j\|_Q^2 \quad (3)$$

where  $r_j$  is the reaction rate of (2) at time-step  $j$ . An economically based objective function reflects variable costs and revenues, typically

$$J_2 = a_1 \cdot x_{Mn,k+N} - \sum_{j=k+1}^{k+N} b^T \cdot u_j \quad (4)$$

where  $a$  is the sales price for  $Mn$ .  $b$  includes the costs for the control inputs.  $x_{Mn,k+N}$  is the mass of  $Mn$  in the ladle at the end of the batch ( $k+N$ ). The latter approach provides the user with the highest flexibility in the sense that it is easy to vary the objective function in the wake of changing sales prices and cost.

Weighting control input changes, cf. the term  $\sum_{j=k}^{k+M} \|u_j - u_{j-1}\|_{S_j}^2$  in (1), was added to (4). This is important to prevent the optimization algorithm from computing excessive changes in the control input that provide minute improvements to (4). Further, it can be noted that the term  $\sum_{j=k}^{k+M} \|u_j - u_{j-1}\|_{S_j}^2$  contributes with a positive definite matrix to the Hessian of the Lagrangian. Hence, in practice the term improves convergence speed of the search algorithm.

##### 4.3 Inequality constraints

Inequality constraints limit the search space for the control inputs and may also include limitations on the output or state-space. In the reference case they play an important role in transferring operational experience to the optimization problem. To force one of the control inputs to 0 during certain time periods inequalities are used. One might imagine that this would be taken care of by the objective function and the dynamic model. In the reference case, however, there are both external factor and unmodelled phenomena which necessitates the use of inequalities to set it to 0.

The identified dynamic model is valid for certain operating conditions. In particular there are limitations on the liquid and gas temperature range in which the dynamic model is valid. This is one reason to include an upper bound on the maximum liquid and gas temperature in the ladle.

##### 4.4 Time-varying optimization problem

The optimization problem need not be static with time meaning that external circumstances may trigger changes in the optimization problem. In the reference case it is necessary to change the optimization problem in two operational cases. First, when an old lance is replaced with a new one the dynamic model's prediction accuracy decreases for some runs until the model is tuned. In this case the end time constraint for the  $C$  concentration is reduced to make sure that the real  $C$  concentration is met. Second, when the ladle is changed it is important to account for the fact that the new ladle

is cooler than the normal operating temperature for the ladle. This again necessitates changes in the constraints of the optimization problem.

## 5. DISCUSSION

The sections on model development and on the optimization problem have brought forward some critical issues in optimization-based control, some of which usually are neglected in the open literature. Understanding these issues will direct future academic advancements in a direction that meets the requirements on the implementation side. Some are eligible for further research. We will point to two issues that deserve further attention: These are model smoothness and the dynamic optimization problem.

The requirement on model smoothness springs from the use of the dynamic model. It is used in an optimization search algorithm. As opposed to this the requirement would not be valid if the intended use was in a dynamic simulator. Model smoothness raise two issues. It can be integrated into the model which was done in the reference case. Another approach is the inclusion of a term in the parameter estimation problem that penalizes non-smoothness of the model. This necessitates the design of a measure for non-smoothness of the model.

Application experience pinpoints the need for a dynamic optimization problem in the sense that it varies with time. The objective function may change to accomodate changing priorities. This may be caused by external factors like price and cost changes for products, and raw material and energy, respectively. It may also be caused by maintenance work, e.g. a control actuator may be deactivated. The latter case may alternatively be taken care of by changing the inequality constraints.

The model that defines the equality constraints may be time-varying. In the reference case on metal refining the model is updated from one run to the next. Hence, it is time-varying.

The inequality constraints may vary for several reasons. Changing products may require different operating conditions, e.g. limits on the maximum temperature in a reactor. Wear of equipment may invoke operational conservatism. When the ladle lining becomes thin the user may refrain from producing products that require high temperatures. Uncertainty may lead to the choice of conservative targets to make sure that hard product targets are met.

The time-varying optimization problem complicates analysis, e.g. NMPC stability analysis for continuous processes. The simplest case arises

when the changes are invoked by exogeneous signals. If the changes are based on an adaption loop, however, this feedback loop is added to the control loop. A further complication arises if the adaption loop includes discrete events. The system must then be treated as a hybrid system. The latter case occurs if some (discrete) logic discriminates between different objective functions.

The sections above focus on the development phase of an optimization-based control application. Little is said about the commissioning phase. There are crucial issues related in this phase. They include the application's functionality, graphical user interface, software debugging to make the application robust, and user motivation and training courses.

## 6. CONCLUSIONS

Based on industrial experience this paper highlights some issues that are critical when developing an optimization-based control applications. We argue that model smoothness is important in optimization-based control, and that industrial applications necessitates a time-varying optimization problem.

## 7. ACKNOWLEDGEMENTS

The authors thank Peter Cowx, Ph.D. at Eramet Norway AS for valuable discussions.

## 8. REFERENCES

- Bonvin, Dominique (1998): Optimal operation of batch reactors - a personal view. *J. of Process Control*, **8**, pp.355-368.
- Dresler, W. (1989): Limitations to oxygen decarburization of high-carbon ferromanganese. In: *Proceedings 1989 Steelmaking conference*.
- Qin, S. Joe and Badgwell, Thomas A. (2000). An Overview of Nonlinear Model Predictive Control Applications. In: *Nonlinear Model Predictive Control*, Birkhauser, Basel.
- Schei, Tor S., Ludvigsen, Halgeir, Singstad, Peter and Foss, Bjarne A. (2001): Operation Support System for Optimization of Suspension PVC Batch Polymerization Reactors. In: *Preprints European Control Conference*, Porto, Portugal.