PERGAMON

Computers and Chemical Engineering 23 (1998) 187–202

Computers & Chemical Engineering

Nonlinear model predictive control: current status and future directions

Michael A. Henson*

Department of Chemical Engineering, Louisiana State University, Baton Rouge, LA 70803-7303, U.S.A.

Received 13 July 1998

Abstract

Linear model predictive control (LMPC) is well established as the industry standard for controlling constrained multivariable processes. A major limitation of LMPC is that plant behavior is described by linear dynamic models. As a result, LMPC is inadequate for highly nonlinear processes and moderately nonlinear processes which have large operating regimes. This shortcoming coupled with increasingly stringent demands on throughput and product quality has spurred the development of nonlinear model predictive control (NMPC). NMPC is conceptually similar to its linear counterpart except that nonlinear dynamic models are used for process prediction and optimization. The purpose of this paper is to provide an overview of current NMPC technology and applications, as well as to propose topics for future research and development. The review demonstrates that NMPC is well suited for controlling multivariable nonlinear processes with constraints, but several theoretical and practical issues must be resolved before widespread industrial acceptance is achieved.

1. Introduction

Model predictive control (MPC) refers to a class of control algorithms in which a dynamic process model is used to predict and optimize process performance. The first MPC techniques were developed in the 1970s because conventional single-loop controllers were unable to satisfy increasingly stringent performance requirements (Qin and Badgwell, 1997). MPC is well suited for high performance control of constrained multivariable processes because explicit pairing of input and output variables is not required and constraints can be incorporated directly into the associated open-loop optimal control problem. The current generation of commercially available MPC technology is based on linear dynamic models, and therefore is referenced by the generic term linear model predictive control (LMPC). Although often unjustified, the assumption of process linearity greatly simplifies model development and controller design.

Many processes are sufficiently nonlinear to preclude the successful application of LMPC technology. Such processes include highly nonlinear processes that operate near a fixed operating point (e.g. high-purity distillation columns) and moderately nonlinear process with large operating regimes (e.g. multi-grade polymer reactors). This has led to the development of nonlinear model predictive control (NMPC) in which a more accurate nonlinear model is used for process prediction and optimization. While NMPC offers the potential for improved process operation, it offers theoretical and practical problems which are considerably more challenging than those associated with LMPC. Many of these problems are associated with the nonlinear program which must be solved on-line at each sampling period to generate the control moves.

The goal of this paper is to present the current status of NMPC technology and to outline directions for future research. The critically important issue of nonlinear process modeling is discussed in Section 2. In Section 3, the NMPC problem is described with particular emphasis on a prototypical formulation. Computational issues associated with on-line solution of the nonlinear program are discussed in Section 4. In Section 5, process applications of NMPC are summarized. Finally, some topics for future research are presented in Section 6.

*Tel.: + 1-504-388-3690; fax: + 1-504-388-1476; e-mail: henson@nlc.che.lsu.edu

0098-1354/98/$ – see front matter
PII: S0098-1354(98)00260-9

2. Nonlinear process models

The industrial success of LMPC is largely attributable to the availability of commercial software packages which can be used to develop linear dynamic models directly from process data (Qin and Badgwell, 1997). These linear empirical models are used by the LMPC controller to predict and optimize process performance. NMPC requires the availability of a suitable nonlinear dynamic model of the process. Consequently, the development of nonlinear process models is of paramount importance. Due to the complexity of nonlinear systems, it is not possible to develop nonlinear system identification techniques by straightforward extension of the linear theory (Pearson and Ogunnaike, 1997). As an alternative, the NMPC controller may be based on a fundamental model which is derived from basic conservation laws and constitutive relations. These two general classes of nonlinear models are described below. Also discussed are hybrid nonlinear models, which are developed by a combination of the fundamental and empirical modeling approaches.

2.1. Fundamental models

Fundamental dynamic models are derived by applying transient mass, energy, and momentum balances to the process (Ogunnaike and Ray, 1994). In the absence of spatial variations, the resulting models have the general form

$$\dot{x} = f(x, u), \tag{1}$$

$$0 = g(x, u), \tag{2}$$

$$y = h(x, u), \tag{3}$$

where x is a n-dimensional vector of state variables, u is a m-dimensional vector of manipulated input variables, and y is a p-dimensional vector of controlled output variables. The ordinary differential equations (1) and the algebraic equations (2) are derived from conservation laws and various constitutive relations, while the output equations (3) are chosen by the control system designer. Because NMPC is most naturally formulated in discrete time (see Section 3), it is necessary to discretize the continuous-time differential equations (1). As discussed in Section 4, this usually is achieved using orthogonal collocation on finite elements (Meadows and Rawlings, 1997).

Fundamental models have several advantages as compared to nonlinear empirical models. Because fundamental models are highly constrained with respect to their structure and parameters, less process data is required for their development. In particular, model parameters may be estimated from laboratory experiments and routine operating data instead of time-consuming plant tests. As long as the underlying assumptions remain valid, fundamental models can be expected to extrapolate to operating regions which are not represented in the data set used for model development (Meadows and Rawlings, 1997). This property is particularly important when a process operates over a wide range of conditions, such as a polymer reactor which produces several different product grades. Some disadvantages of the fundamental modeling approach are discussed below.

A large number of NMPC studies based on fundamental models have been reported. In most cases, the process consists of a single unit operation, and the nonlinear dynamic model is relatively simple. With the possible exception of the paper by Ricker and Lee (1995), studies which reflect the large-scale nature of typical industrial plants are not available in the open literature. This probably is attributable to the inherent difficulties involved in deriving fundamental dynamic models for large-scale processes. An alternative approach for developing rigorous nonlinear models is to utilize a commercial dynamic simulator. A number of vendors offer such products, including ABB Industrial Systems (iGES), Aspen Technology (SPEEDUP), and Hyprotech (HYSYS). The use of commercial simulators for NMPC has not been reported in the open literature, probably because the dynamic model equations are not available to the control system designer.

A potential disadvantage of the fundamental modeling approach is that the resulting dynamic model may be too complex to be useful for NMPC design. This motivates the development of modeling techniques which are constrained by the underlying physics, but yield dynamic models with significantly fewer equations. Reduction techniques such as singular perturbations may be applied to the rigorous model to derive a simplified model which retains the basic dynamic behavior of the full-scale model. This approach has been sucessfully applied to chemical reactors (Duchene and Rouchon, 1996) and distillations columns (Levine and Rouchon, 1991). NMPC design also may be facilitated by developing simplified dynamic models that are capable of describing the most important process characteristics. Applications of this approach to distillation column modeling are described in Benallou et al. (1986) and Hwang (1991).

Before being used for NMPC design, a fundamental model should be validated with plant data which represent typical operating conditions. While more systematic methods are available (Pearson and Ogunnaike, 1997), some measure of model accuracy can be obtained by placing the model on-line in predictive mode. Large deviations between the plant measurements and the model predictions may warrant further modeling effort. Once a suitable model is developed and utilized for NMPC design, long-term maintanence and support by the model developers is required. Otherwise, the NMPC controller will not be used by plant personnel because the nonlinear model will fail to provide satisfactory predictions as the plant is modified and/or operating conditions are changed.

2.2. Empirical models

In many applications, lack of process knowledge and/or a suitable dynamic simulator precludes the derivation of a fundamental model. This necessitates the development of empirical nonlinear models from dynamic plant data. This process is known as nonlinear system identification. Unfortunately, a well developed theory for nonlinear system identification is not currently available (Pearson and Ogunnaike, 1997). Furthermore, few applications of nonlinear system identification techniques to actual plant data have been reported in the literature.

A fundamental difficulty associated with the empirical modeling approach is the selection of a suitable model form. Discrete-time models are most appropriate because plant data is available at discrete time instants and NMPC is most naturally formulated in discrete time. The types of discrete-time nonlinear models utilized for NMPC include (Pearson and Ogunnaike, 1997)

- Hammerstein and Wiener models,
- Volterra models,
- polynomial autoregressive moving average model with exogenus inputs (polynomial ARMAX),
- artificial neural network models.

Each of these models can be represented as a nonlinear autoregressive moving average model with exogenus inputs (NARMAX). The NARMAX form for a single-input, single-output system is

$$y(k) = F[y(k-1), \ldots, y(k-n_y), u(k-1), \ldots, u(k-n_u), e(k), \ldots, e(k-n_e+1)], \tag{4}$$

where F is a nonlinear mapping, k represents the time instant $k\Delta t$, where Δt is the sampling time, y is the controlled output, u is the manipulated input, e is the noise input, n_y is the number of past outputs used, n_u is the number of past inputs used, and n_e is the number of current and past noise inputs used. If necessary, a state-space representation of the input-output model (4) is easily derived (Hernandez and Arkun, 1992; Nahas, Henson and Seborg, 1992). Multivariable systems are handled by building a multiple-input, single-output model for each output variable (Pearson and Ogunnaike, 1997).

Nonlinear system identification involves the following tasks:

1. Stucture selection — parameterization of the nonlinear mapping F and selection of the model parameters n_y, n_u, and n_e.
2. Input sequence design — determination of the input sequence $u(k)$ which is injected into the plant to generate the output sequence $y(k)$.
3. Noise modeling — determination of the dynamic model which generates the noise input $e(k)$.
4. Parameter estimation — estimation of the remaining model parameters from the dynamic plant data $u(k)$ and $y(k)$ and the noise input $e(k)$.
5. Model validation — comparison of plant data and model predictions for data not used in model development.

Each task represents a very challenging theoretical and practical problem. A general theory is not available, but results have been presented for specific classes of nonlinear models (Pearson and Ogunnaike, 1997).

As compared to fundamental models, empirical nonlinear models offer several important advantages. First and foremost, detailed process understanding is not required for empirical model development. This is an important consideration for complex industrial process, such as polymerization reactors, which are difficult to model from fundamental principles. Because NMPC requires on-line solution of a nonlinear programming problem, computational overhead and reliability is intimately connected with the complexity of the nonlinear model. An advantage of empirical models is that the nonlinear model form can be chosen to restrict model complexity (Pearson and Ogunnaike, 1997). This is considerably more difficult to achieve with the fundamental modeling approach.

The use of empirical models for NMPC has been studied by several investigators. Artificial neural networks are the most popular framework for empirical model development (Su and McAvoy, 1997), although techniques based on Hammerstein and Wiener models (Chu and Seborg, 1994), Volterra models (Maner, Doyle, Ogunnaike and Pearson, 1996), and polynomial ARMAX models (Sriniwas and Arkun, 1997) also have been presented. For the most part, studies have been restricted to small-scale simulated processes which are not indicative of industrial systems. The development of empirical nonlinear models for large-scale processes is a very challenging problem which requires the availability of suitable software tools. Neural network modeling packages are offered by several vendors, the most popular of which is Pavilion Technologies' Process Insights. With the notable exception of published studies by Pavilion (Martin, 1997), applications of these commercial packages for NMPC are not available in the open literature.

2.3. Hybrid models

Hybrid nonlinear models are developed by combining the fundamental and empirical modeling approaches. This allows the advantages of each modeling approach to be exploited. A common method for developing hybrid models is to use empirical models to estimate unknown functions in the fundamental model; e.g. reaction rates in a chemical reactor model (Pottmann and Henson, 1997). In this case, steady-state empirical models usually are sufficient. Another possible approach is to utilize a fundamental model to capture the basic process characteristics, and then to describe the residual between the plant and the model using a nonlinear empirical model. Both techniques allow the nonlinear model to be constrained by the underlying physics, but they do not require a complete rigorous model of the plant. While hybrid models hold great promise, their use for NMPC design has not been explored.

3. Problem formulation

NMPC is an optimization-based control strategy which is well suited for constrained, multivariable processes. A sequence of control moves is computed to minimize an objective function which includes predicted future values of the controlled outputs. The predictions are obtained from a nonlinear process model, hence the terminology nonlinear model predictive control (NMPC). The optimization problem is solved subject to constraints on input and output variables, as well as constraints imposed by the nonlinear model equations. This formulation yields an open-loop optimal controller. Feedback is included by implementing only the manipulated inputs computed for the present time step, then moving the prediction horizon forward one step and resolving the problem using new process measurements. For this reason, NMPC often is called nonlinear receding horizon control (Mayne and Michalska, 1991a). Calculation of the manipulated input sequence requires the on-line solution of a nonlinear programming problem.

The basic formulation of the NMPC problem is discussed below. The presentation focuses on a prototypical discrete-time formulation which is representative of techniques proposed in the literature. It is important to note that a wide variety of alternative discrete-time formulations (Badgwell, 1997; Zheng, 1997; Maner et al., 1996; Pottmann and Seborg, 1997) and continuous-time formulations (Chen and Allgower, 1997; Mayne and Michalska, 1990, 1991b; Peterson et al., 1992) are available. In addition to the constrained optimization problem, issues such as disturbance and state estimation, stability, and controller tuning are discussed. Computational issues associated with on-line solution of the nonlinear program are discussed in the following section.

3.1. Optimization problem

The nonlinear process model is assumed to have the following discrete-time representation,

$$x(k+1) = F[x(k), u(k)], \tag{5}$$

$$y(k) = h[x(k)], \tag{6}$$

where x is a n-dimensional vector of state variables, u is a m-dimensional vector of manipulated input variables, and y is a p-dimensional vector of controlled output variables. Such a model can be obtained by discretizing a continuous-time, state-space model or by deriving a state-space realization of a discrete-time, input–output model. Several discretization methods for nonlinear systems, such as Taylor-based linearization (Kazantzis and Kravaris, 1994), have been proposed. It is important to note that time delays can be handled by augmenting the state vector such that the resulting state-space model has no delays.

The optimization problem for the prototypical NMPC formulation is (Meadows and Rawlings, 1997)

$$\min_{u(k|k),u(k+1|k),\ldots,u(k+M-1|k)} J = \phi[y(k+P|k)] + \sum_{j=0}^{P-1} L[y(k+j|k), u(k+j|k), \Delta u(k+j|k)], \tag{7}$$

where $u(k+j|k)$ is the input $u(k+j)$ calculated from information available at time k, $y(k+j|k)$ is the output $y(k+j)$ calculated from information available at time k, $\Delta u(k+j|k) = u(k+j|k) - u(k+j-1|k)$, M is the control horizon; P is the prediction horizon and ϕ and L are (possibly) nonlinear functions of their arguments. The optimization problem is solved subject to the constraints discussed below. The functions ϕ and L can be chosen to satisfy a wide variety of objectives, including minimization of overall process cost. However, economic optimization may be performed by a higher-level system which determines appropriate setpoints for the NMPC controller. In this case, it is meaningful to consider quadratic functions of the following form:

$$L = [y(k+j|k) - y_s(k)]^{\mathrm{T}} Q [y(k+j|k) - y_s(k)] + [u(k+j|k) - u_s(k)]^{\mathrm{T}} R [u(k+j|k) - u_s(k)] + \Delta u^{\mathrm{T}}(k+j|k)\, S \Delta u(k+j|k), \tag{8}$$

$$\phi = [y(k+P|k) - y_s(k)]^{\mathrm{T}} Q [y(k+P|k) - y_s(k)], \tag{9}$$

where $u_s(k)$ and $y_s(k)$ are steady-state targets for u and y, respectively (see Section 3.3), and Q, R, and S are positive-definite weighting matrices. The principal controller tuning parameters are M, P, Q, R, S, and the sampling period Δt.

The predicted outputs are obtained from the nonlinear model (5)–(6). Successive iterations of the model equations yield

$$y(k+1|k) = h[x(k+1|k)] = h[F[x(k|k), u(k|k)]],$$

$$\equiv G_1[x(k), u(k|k)],$$

$$y(k+2|k) = G_1[x(k+1|k), u(k+1|k)],$$

$$= G_1[F[x(k|k), u(k|k)], u(k+1|k)],$$

$$\equiv G_2[x(k), u(k|k), u(k+1|k)],$$

$$\vdots$$

$$y(k+j|k) = G_j[x(k), u(k|k), u(k+1|k), \ldots, u(k+j-1|k)],$$

where $x(k|k) = x(k)$ is a vector of current state variables. If the control horizon (M) is less than the prediction horizon (P), the output predictions are generated by setting inputs beyond the control horizon equal to the last computed value: $u(k+j|k) = u(k+M-1|k)$, $M \leqslant j \leqslant P$. Note that the prediction $y(k+j|k)$ depends on the current state variables, as well as the calculated input sequence. Therefore, NMPC requires measurements or estimates of the state variables. This is discussed in more detail below.

Solution of the NMPC problem yields the input sequence $\{u(k|k), u(k+1|k) \ldots, u(k+M-1|k)\}$. Only the first input vector in the sequence is actually implemented: $u(k) = u(k|k)$. Then the prediction horizon is moved forward one time step, and the problem is resolved using new process measurements. This receding horizon formulation yields improved closed-loop performance in the presence of unmeasured disturbances and modeling errors.

3.2. Process constraints

An important characteristic of process control problems is the presence of constraints on input and output variables. Input constraints arise due to actuator limitations such as saturation and rate-of-change restrictions. Such constraints take the form,

$$u_{min} \leqslant u \leqslant u_{max}, \tag{10}$$

$$\Delta u_{min} \leqslant \Delta u \leqslant \Delta u_{max}, \tag{11}$$

where u_{min} and u_{max} are the minimum and maximum values of the inputs, respectively, and Δu_{min} and Δu_{max} are the minimum and maximum values of the rate-of-change of the inputs, respectively. Output constraints usually are associated with operational limitations such as equipment specifications and safety considerations. These constraints can be posed as

$$y_{min} \leqslant y \leqslant y_{max}, \tag{12}$$

where y_{min} and y_{max} are the minimum and maximum values of the outputs, respectively. Constraints on the state variables also may be specified if appropriate.

The industrial success of LMPC is largely attributable to its explicit constraint handling capability (Qin and Badgwell, 1997). By including constraints in the optimization problem, the controller is able to predict future constraint violations and respond accordingly. A major advantage of NMPC as compared to other nonlinear control strategies is that it provides the same constraint handling capability. This is achieved by solving the nonlinear optimization problem (7) subject to the following inequality constraints:

$$u_{min} \leqslant u(k+j|k) \leqslant u_{max}, \quad 0 \leqslant j \leqslant M-1, \tag{13}$$

$$\Delta u_{min} \leqslant \Delta u(k+j|k) \leqslant \Delta u_{max}, \quad 0 \leqslant j \leqslant M-1, \tag{14}$$

$$y_{min} \leqslant y(k+j|k) \leqslant y_{max}, \quad 1 \leqslant j \leqslant P. \tag{15}$$

In addition, the nonlinear model (5)–(6) is posed as a set of equality constraints,

$$x(k+j+1|k) = F[x(k+j|k), u(k+j|k)], \quad 0 \leqslant j \leqslant P-1, \tag{16}$$

$$y(k+j|k) = h[x(k+j|k)], \quad 1 \leqslant j \leqslant P, \tag{17}$$

where $x(k|k) = x(k)$ if the state variables are measured.

It is important to note that input constraints are hard constraints in the sense that they must be satisfied. Conversely, output constraints can be viewed as soft constraints because their violation may be necessary to obtain a feasible optimization problem. More specifically, the output constraints can be relaxed during part of the prediction horizon,

$$y_{min} \leqslant y(k+j|k) \leqslant y_{max}, \quad j_1 \leqslant j \leqslant P, \tag{18}$$

where j_1 represents the lower limit for output constraint enforcement. In LMPC, it is possible to determine a priori the limit j_1 such that feasibility is guaranteed if there is no plant/model mismatch (Rawlings and Muske, 1993). This may require $j_1 > P$, in which case the output constraints are relaxed over the entire prediction horizon. An analogous result is not available for NMPC, therefore feasibility is more difficult to establish in the nonlinear case.

3.3. Disturbance and state estimation

The goal of the NMPC controller is to drive the process inputs and outputs to their target values in an optimal manner. If the target values u_s and y_s in Eqs. (8)–(9) are not chosen properly, the controller can exhibit steady-state offset in the presence of unmeasured disturbances and modeling errors (Meadows and Rawlings, 1997). As in LMPC, the offset problem can be handled by designing a disturbance estimator which gives the controller implicit integral action. Recall from Section 3.1 that current values of the state variables are required to compute the output predictions. In the absence of full-state feedback, it becomes necessary to design a nonlinear observer to generate estimates of the unmeasured state variables.

The simplest method for incorporating integral action is to generate the output targets (y_s) by shifting the setpoints with the disturbance estimates. In this method, the penalty on the inputs is eliminated ($R = 0$) such that the quadratic function L becomes:

$$L = [y(k+j|k) - y_s]^T Q [y(k+j|k) - y_s] + \Delta u^T(k+j|k) S \Delta u(k+j|k). \quad (19)$$

The output targets are computed as follows:

$$y_s(k) = y_{sp}(k) - \hat{d}(k), \quad (20)$$

$$\hat{d}(k) = y(k) - y(k|k), \quad (21)$$

where $y_{sp}(k)$ are setpoints for the output variables, $y(k)$ are the plant outputs, $y(k|k)$ are estimated outputs obtained from the nonlinear model (5)–(6); and $\hat{d}(k)$ are the estimated disturbances. This disturbance model assumes plant/model mismatch is attributable to a step disturbance in the output and the disturbance remains constant over the prediction horizon. While these assumptions rarely hold in practice, the disturbance model does eliminate offset for asympotically constant setpoints under most conditions (Meadows and Rawlings, 1997). A more sophisticated method for incorporating integral action based on steady-state target optimization is described in (Meadows and Rawlings, 1997).

Simultaneous state and disturbance estimation can be performed using the augmented state-space model,

$$x(k+1) = F[x(k), u(k)] \quad (22)$$

$$d(k+1) = d(k) \quad (23)$$

$$y(k) = h[x(k)] + d(k) \quad (24)$$

where $d(k)$ is a constant output disturbance. The augmented system can be used to design a nonlinear observer which generates state estimates $\hat{x}(k)$ and disturbance estimates $\hat{d}(k)$ from the measured process variables (Muske and Edgar, 1997). Unfortunately, a well-developed theory for nonlinear observers is not available. Existing design methods are applicable only to specific classes of nonlinear systems and restrictive assumptions are required to ensure stability. Nevertheless, some promising state estimation techniques have been proposed for both continuous-time and discrete-time nonlinear systems. These methods include extended Kalman filters (Muske and Edgar, 1997), extended Luenberger observers (Soroush, 1997), receding horizon observers (Michalska and Mayne, 1995), adaptive observers (Bastin and Gevers, 1988), and inversion-based observers (Moraal and Grizzle, 1995). Due to the complication associated with nonlinear state estimation, input–output models are preferred to state-space models when full-state feedback is not available.

3.4. Stability

NMPC is a nonlinear open-loop optimal control technique where feedback is incorporated via the receding horizon formulation and the disturbance estimator. From a theoretical perspective, the minimum requirement of a model-based feedback controller is that it yields a stable closed-loop system if a perfect model of the plant is available. This is known as nominal closed-loop stability. It is important to note that techniques such as PID control and linear model predictive control most often applied in the process industries do not meet this requirement. While this has not precluded their successful application, the development of rigorous stability theories is desirable for several reasons. Stability theory provides a systematic framework for the derivation and refinement of control algorithms. Without such tools, the potential usefulness of a particular control strategy can be evaluated only via simulation and experimental studies. Furthermore, such analysis facilitates the determination of tuning parameter values which result in closed-loop stability. This allows the controller to be tuned to improve performance rather than to establish stability.

The objective of this section is to present a representative NMPC stability result. This requires a slightly different formulation of the NMPC problem as Eqs. (7)–(9) is not a stabilizing formulation unless certain modifications are introduced. The cost function (7) is modified as follows:

$$J = x^T(k+P|k) Q x(k+P|k) + \sum_{j=0}^{P-1} x^T(k+j|k) Q x(k+j|k) + u^T(k+j|k)\, Qu(k+j|k). \quad (25)$$

This differs from the prototypical formulation in that: (i) the control horizon is equal to the prediction horizon; (ii) the state variables are penalized rather than the output variables; (iii) a penalty on the rate-of-change of the input is not included; and (iv) and the steady-state target values are zero. The first and third modifications are not very limiting, while the fourth modification can be achieved simply by moving the target values to the origin via a change of coordinates. By contrast, the second modification represents a fundamental change of the prototypical formulation.

The primary tool for NMPC stability analysis is Lyapunov theory (Khalil, 1992). Available theorems are applicable to the NMPC problem only if the prediction horizon is infinite ($P \rightarrow \infty$) or a terminal state constraint is imposed:

$$x(k+P|k) = 0 \quad (26)$$

In the linear case, the infinite horizon controller can be reformulated as a finite horizon controller with a terminal state penalty (Muske and Rawlings, 1993). This is not possible in the nonlinear case. Consequently, the terminal constraint (26) is utilized to derive the result presented below. If there exists a prediction horizon (P)

such that constraint (26) is satisfied, the nonlinear system is said to be constrained null controllable.

The following result is derived in Meadows and Rawlings (1997).

Theorem 1. *If the NMPC objective function* (25) *is continuous, then the origin of the nonlinear system* (5) *is an asymptotically stable fixed point with a region of attraction that includes all initial conditions which are constrained null controllable.*

The assumption that the objective function is continuous is not particularly restrictive, although it can be difficult to verify. The major limitation of this result is the assumption that the system is constrained null controllable. It is important to note that this is a very strong assumption. In fact, there does not exist general results which ensure the existence of a control horizon such that Eq. (26) is satisfied even for the unconstrained case. While this condition often holds in practice, it is very difficult to verify a priori. Additional nominal stability results for discrete-time systems (Alamir and Bornard, 1994; Keerthi and Gilbert, 1988; Meadows et al., 1995; Meadows, 1997; Nicolao et al., 1997) and continuous-time systems (Chen and Shaw, 1982; Mayne and Michalska, 1990; Sistu and Bequette, 1996) are available.

3.5. Tuning

For the quadratic NMPC formulation (7)–(9), the primary controller tuning parameters are the sampling period (Δt), the control horizon (M), the prediction horizon (P), and the weighting matrices (Q, R, S). A limitation of NMPC is that the effect of these parameters on closed-loop performance is difficult to predict a priori. To date, parameter values which ensure nominal closed-loop stability have been determined only if the prediction horizon is infinite or a terminal state constraint is imposed. Because these conditions are rarely satisfied in practice, it is important to develop heuristic tuning guidelines. The following results are summarized from Meadows and Rawlings (1997).

For stable nonlinear systems, the sampling interval (Δt) should be chosen to provide a compromise between closed-loop performance and on-line computation. Small values generally improve performance but require a longer prediction horizon to adequately capture the process dynamics. This leads to more finite elements in the nonlinear program (see Section 4.1) and increased computation. Large sampling intervals reduce on-line computation, but they can result in poor performance such as ringing. The choice of sampling interval can have a dramatic effect on robustness when the nonlinear system is unstable. There is an inverse relationship between Δt and the allowable modeling error. As modeling error increases, more frequent feedback of process measurements (i.e. small Δt) is required to indicate the onset of unstable behavior.

The effect of the control horizon (M) on NMPC performance is similar to that observed in the linear case. For a fixed prediction horizon, smaller control horizons yield more sluggish output responses and more conservative input moves. Large control horizons have the opposite effect on performance. In addition, large values of M lead to increased on-line computation as M is linearly related to the number of decision variables in the nonlinear programming problem. In practice, M often must be chosen to provide a balance between performance and computation. The prediction horizon (P) has similar effects as the control horizon. Large prediction horizons result in more aggressive control and increased computation.

The weighting matrices (Q, R, S) can be the most difficult tuning parameters to select because their values depend on the scaling of the problem. Typically, they are chosen to be diagonal matrices with positive elements. The magnitude of the diagonal elements depends both on the scaling and the relative importance of the variables. For problems in which all the variables are scaled similarly, it is suggested in Meadows and Rawlings (1997) that the output penalties (Q) be chosen in the range 1–100 and the input penalties (R, S) be chosen in the range 1–10. The final parameter values can be obtained by fine tuning via simulation study.

4. Computational issues

The prototypical NMPC formulation described in the previous section requires that a nonlinear programming problem be solved on-line at each time step to determine the manipulated inputs. The optimization problem generally is nonconvex because the model equations are nonlinear. Consequently, the major practical challenge associated with NMPC is on-line solution of the nonlinear program (NLP). Efficient and reliable NLP solution techniques are required to make NMPC a viable control technique. In addition, it may be necessary to derive alternative formulations of the NMPC problem with improved computational properties. Below some general characteristics of the NLP problem are discussed, and the most widely studied solution algorithms are reviewed. Alternative NMPC formulations with improved computational characteristics also are described.

4.1. Nonlinear programming problem

The prototypical NMPC formulation is based on a discrete-time state-space model (1)–(3) of the nonlinear process. Such a model can be derived by performing state-space realization on a discrete-time input–output model obtained via nonlinear system identification. In many applications of practical interest, an inherently continuous-time nonlinear model derived via fundamental modeling is available for NMPC design. In this case, a discrete-time nonlinear model can be obtained by explicitly discretizing the fundamental model. However, discretization usually is performed implicitly as part of the NLP solution using a numerical technique such as orthogonal collocation. Consequently, continuous-time models should be considered when discussing computational issues.

For a continuous-time model, the NMPC problem can be represented as

$$\min_{U(k)} J = \Psi[Y(k), U(k)], \tag{27}$$

$$\dot{x}(t^*) = f[x(t^*), u(t^*)],\ t \leqslant t^* \leqslant t + P\Delta t, \tag{28}$$

$$u(t^*) = u[t + (M-1)\,\Delta t],$$

$$t + (M-1)\Delta t \leqslant t^* \leqslant t + P\Delta t, \tag{29}$$

$$0 = \Phi[X(k), Y(k), U(k)], \tag{30}$$

$$0 \geqslant DU(k), \tag{31}$$

$$0 \geqslant DY(k), \tag{32}$$

where

$$U(k) \equiv \begin{bmatrix} u(k|k) \\ u(k+1|k) \\ \vdots \\ u(k+M-1|k) \end{bmatrix}, \quad Y(k) \equiv \begin{bmatrix} y(k|k) \\ y(k+1|k) \\ y(k+2|k) \\ \vdots \\ y(k+P|k) \end{bmatrix},$$

$$X(k) \equiv \begin{bmatrix} x(k|k) \\ x(k+1|k) \\ x(k+2|k) \\ \vdots \\ x(k+P|k) \end{bmatrix}.$$

The constraint (28) corresponds to satisfaction of the continuous-time model equations over the prediction horizon, while (29) enforces the requirement that all inputs beyond the control horizon are held constant. The constraint (30) represents the algebraic equations in the model, while Eqs. (31) and (32) correspond to the constraints on the input variables and output variables, respectively. As discussed below, NMPC solution techniques differ primarily according to the method used to handle the model Eqs. (28) and (30).

The development of efficient and reliable solution methods for the NLP problem (27)–(32) is a challenging problem. The most obvious difficulty is that the optimization problem is nonlinear and has a potentially large number of decision variables. The NLP solution can be too computationally intensive for on-line implementation using conventional process control computers. An equally important problem is that the model constraints (28) and (30) generally yield a nonconvex optimization problem (Mayne, 1997). As a result, standard NLP techniques such as successive quadratic programming (SQP) cannot be expected to find the global minimum. Furthermore, there is no theoretical guarantee that any feasible solution can be determined in the presence of nonconvex constraints. Some widely studied solution algorithms for the NLP problem are presented below.

4.2. Successive linearization of model equations

The simplest way to deal with the model equations (28) and (30) is to perform Jacobian linearization about a nominal operating point and discretize the resulting linear model. This yields linear model predictive control if the objective function is quadratic. Local linearization allows the optimization problem to be solved with simple quadratic programming (QP) techniques, but it provides no compensation for process nonlinearities. A straightforward extension of this idea is to use the current operating point to linearize the model before each execution of the NMPC controller (Bequette, 1991). The primary advantage of successive model linearization is that the NMPC problem is reduced to a LMPC problem at each time step. However, this approach only provides indirect compensation for process nonlinearities.

NMPC techniques based on successive model linearization has been proposed by a number of investigators. Typically, the linearized model is used to predict future process behavior, while the original nonlinear model is used to compute the effect of past input moves (Garcia, 1984). The accuracy of the linear model can be improved by relinearizing the model equations several times over the sampling period (Brengel and Seider, 1989) or by linearizing the model along the computed system trajectory (Li and Biegler, 1989). In the event that the current operating point cannot be determined directly from available process measurements, it becomes necessary to perform the linearization using an estimate of the state variables (Gattu and Zafiriou, 1992; Lee and Ricker, 1994). A related approach is to perform on-linear updating of the linear model using the difference between the linear and nonlinear model responses (Peterson et al., 1992).

4.3. Sequential model solution and optimization

Improved closed-loop performance can be expected if the nonlinear model is used directly in the NMPC calculations. However, standard NLP codes are not designed to handle ODE constraints. This limitation can be overcome using a two-stage solution procedure in which a standard NLP solver is used to compute the manipulated inputs and an ODE solver is used to integrate the nonlinear model equations. This is known as sequential solution because the optimization and integration problems are solved iteratively until the desired accuracy is obtained (Bequette, 1991). As compared to the simultaneous solution method discussed below, an important advantage of the sequential approach is that the manipulated inputs are the only decision variables. Disadvantages of the sequential approach include difficulty in incorporating state/output constraints and poor reliability for large problems (Meadows and Rawlings, 1997).

Several investigators have proposed NMPC solution techniques based on sequential solution of the NLP problem and the model equations. Gradients of the objective function are obtained via simultaneous integration of the model and sensitivity differential equations (Jang et al., 1987; Morshedi, 1986). The model solution phase can be simplified by discretizing the differential equations (28). While other methods are available, the most popular discretization technique is orthogonal collocation on finite elements (Bequette, 1991). This procedure yields nonlinear algebraic model equations of the form

$$WX = \Gamma(X, U), \tag{33}$$

where X is a matrix of state values at the collocation points, U is a vector of inputs which change over the finite elements, W is a matrix of collocation weights, and Γ is a matrix of nonlinear functions derived from the model function f. More details about this numerical procedure are presented in Meadows and Rawlings (1997).

4.4. Simultaneous model solution and optimization

An alternative to the sequential solution approach is to solve the optimization problem and the model equations simultaneously (Meadows and Rawlings, 1997). The simultaneous solution method requires the model equations to be discretized as in Eq. (33) since ODEs cannot be handled by standard NLP solvers. The decision variables are the inputs on each finite element and the state variables at each collocation point. Therefore, the number of decision variables increases as: (i) the sampling period is decreased and/or the prediction horizon is increased (both increase the number of finite elements); or (ii) the number of collocation points on each finite element is increased. The simultaneous approach is best suited for large NLP problems with state/output constraints (Eaton and Rawlings, 1990; Patwardhan et al., 1990).

4.5. Alternative NMPC formulations

The algorithms discussed above are developed to provide efficient and robust solution of the standard NMPC problem. An alternative approach for enhancing the applicability of NMPC is to derive alternative problem formulations with inherently better computational properties. A number of these methods have been proposed during the past few years. An obvious approach to improve computational efficiency is to reduce the number of decision variables in the NMPC problem. The NMPC formulation proposed by Zheng (1997) achieves this dimensionality reduction by allowing only the current inputs to be decision variables. The remaining inputs in the control horizon are computed with unconstrained LMPC controllers designed for embedded operating regions in the state space.

The structure of certain nonlinear empirical models allows the NMPC optimization problem to be solved more efficiently than is possible with other model forms. Second-order Volterra models with linear autoregressive terms yield an NLP problem whose complexity does not depend on the prediction horizon (Maner and Doyle, 1997). Consequently, the optimization problem is easier to solve. Polynomial ARMAX models allow the NMPC problem to be reformulated via change of coordinates as a convex NLP (Sriniwas and Arkun, 1997). This allows the global optimum to be determined using recently developed algorithms for convex NLPs.

Several investigators have proposed NMPC formulations which ensure nominal closed-loop stability by relaxing the infinite prediction horizon or terminal state constraint requirements. An algorithm in which the terminal state constraint is satisfied approximately is proposed in (Mayne, 1997). The hybrid approach involves a local controller which stabilizes the nonlinear system near the desired operating point and a NMPC controller which forces (in finite time) the system to enter the domain of attraction of the linear controller. A related approach in which the terminal state constraint is replaced by a terminal state penalty in the objective function has been proposed for stable nonlinear systems (Chen and Allgower, 1997). A terminal penalty matrix and prediction horizon which ensure stability are determined off-line. An NMPC technique in which the prediction horizon is allowed to be a decision variable is proposed in Mayne (1997).

Another approach to reduce on-line computation is to transform the NMPC problem into a LMPC problem. This idea has been pursued by several investigators (Kurtz and Henson, 1997; Morningred et al., 1992; Nevistic and Morari, 1995). The unconstrained nonlinear system is transformed into a linear system using a feedback linearizing control law. The input constraints are mapped into constraints on the manipulated input of the transformed system, and the resulting constrained linear system is controlled using LMPC. The primary challenge associated with this method is mapping of the input constraints. In practical implementations, the future input constraints must be mapped approximately because future values of the state vector are unknown. Because the predictive control problem is posed in the transformed coordinates, these methods are most accurately classified as feedback linearizing control techniques instead of NMPC techniques.

5. Process applications

NMPC has been applied to a wide variety of simulated and experimental process systems. Such studies are critical for evaluating alternative NMPC formulations and solution algorithms. The vast majority of application studies are restricted to small-scale simulated processes. However, applications to more complex multivariable processes have been investigated, and additional plant-wide control studies should appear in the future. NNMPC has been applied to several experimental processes, but studies focusing on larger scale systems are needed. Perhaps the most exciting development from an applications perspective is the availability of commercial NMPC packages which are being applied in the process industries. Applications of NMPC to simulated, experimental, and commercial processes are reviewed below.

5.1. Simulation studies

Computer simulation represents a very powerful tool to evaluate the efficacy and limitations of NMPC techniques. Simulation studies often provide the only means to evaluate closed-loop stability and performance for situations which are not amenable to theoretical analysis. However, it is important to emphasize that simulation is not a substitute for experimental evaluation. Table 1 contains a summary of some representative NMPC simulation studies. For each reference, the table contains the process considered, the type of model used for controller design, the size of the control problem (number inputs $\times$ number outputs), and the general type of solution algorithm employed. In some papers, the solution algorithm is not explained in much detail and the table contains the most accurate characterization possible given the available information. Additional NMPC application studies are discussed in (Bequette, 1991; Meadows and Rawlings, 1997).

Several general trends can be observed from Table 1. Because of their highly nonlinear behavior, chemical reactors are the most common application in NMPC simulation studies. A constant volume reactor with a single irreversible reaction has become a benchmark problem because the nonlinear model can exhibit a wide range of steady-state and dynamic behaviors (Uppal, Ray and Poore, 1974). Slightly more complex chemical reactor models can exhibit non-minimum phase behavior, as well as a change in the sign of the steady-state gain. These type of reactors are very difficult to operate effectively with conventional linear controllers. Batch, semi-batch, and continuous reactors have been investigated. NMPC also has been applied to polymerization reactors, biochemical reactors, crystallizers, distillation columns, and paper machines.

The most common model form used for NMPC studies is ordinary differential equations (ODEs) derived from basic conservation laws. The derivation of such models is feasible because most of the processes considered are sufficiently simple for fundamental modeling. Several empirical model forms, including Volterra models, polynomial ARMAX models, and neural network models, also have been investigated. To date, a detailed comparison of fundamental models and empirical models for NMPC has not been conducted. Most simulation studies are restricted to single-input, single-output (SISO) processes. Although several multivariable control problems have been investigated, in most studies the number of input and output variables do not approach that encountered in typical industrial applications. A notable exception is Ricker and Lee (1995), where the NMPC controller has 10 inputs and 23 outputs.

A wide variety of solution techniques have been investigated in NMPC simulation studies. Unfortunately, some papers do not contain a sufficiently detailed description to allow the solution algorithm to be adequately characterized. In this case, the solution technique is classified as a generic nonlinear program (NLP). The most widely studied methods involve successive linearization of the model equations (SL) and simultaneous model solution and optimization (SIM). NLP techniques based on sequential solution and optimization (SEQ) do not appear to be as popular. A case study comparing these solution techniques is presented in Sistu et al. (1991).

5.2. Experimental studies

Table 2 contains a summary of experimental NMPC studies. While the number of papers is surprisingly low, it is likely that a few additional experimental applications have appeared in the literature. The processes considered include two pH neutralization reactors, a fixed-bed water-gas shift reactor, and a packed distillation column. It is interesting to note that a wide range of model forms, including fundamental models with ODEs and partial differential equations (PDEs) and empirical models based on the Hammerstein representation and radial basis function networks, have been utilized. Most of the studies focus on SISO control problems, with the exception of Patwardhan and Edgar (1993) where dual composition control of a distillation column is considered. Several solution techniques have been investigated, including an off-line method in which a radial basis function network is trained to emulate the NMPC controller (Pottmann and Seborg, 1997). While computationally appealing, this method probably is impractical for

Table 1
Representative NMPC simulation studies

Reference	Process	Model	Size	Solution
Bequette (1990)	Biochemical reactor	ODE	1 × 1	SEQ
Brengel and Seider (1989)	Biochemical reactor	ODE	1 × 2	SL
Chen and Allgower (1997)	Chemical reactor	ODE	2 × 4	NLP
Chen et al. (1995)	Chemical reactor	ODE	2 × 1	NLP
Eaton and Rawlings (1990)	Batch chemical reactor	ODE	1 × 1	SIM
	Batch crystallizer	PDE	1 × 1	SM
Hernandez and Arkun (1990)	Chemical reactor	Neural network	1 × 1	SL
Hernandez and Arkun (1993)	Chemical reactor	Polynomial ARMAX	1 × 1	SL
Lee and Ricker (1994)	Paper machine	ODE	2 × 2	SL
Maner and Doyle (1997)	Polymerization reactor	Volterra	1 × 1	NLP
	Copolymerization reactor	Volterra	3 × 3	NLP
Maner et al. (1996)	Polymerization reactor	Volterra	1 × 1	SL
	Polymerization reactor	Volterra	2 × 2	NLP
Marco et al. (1997)	Terpolymerization reactor	ODE	3 × 3	SL
Meadows and Rawlings (1997)	Chemical reactor	ODE	1 × 1	SIM
	Fluidized bed reactor	ODE	1 × 1	SIM
Ogunnaike et al. (1993)	Distillation column	Polynomial ARMAX	2 × 2	SL
Patwardhan et al. (1990)	Chemical reactor	ODE	1 × 1	SIM
Peterson et al. (1992)	Semibatch polymerization reactor	ODE	2 × 2	SL
Rawlings et al. (1989)	Semibatch chemical reactor	ODE	1 × 1	NLP
Ricker and Lee (1995)	Chemical reactor/separators	ODE	10 × 23	SL
Sistu and Bequette (1996)	Chemical reactor	ODE	1 × 1	NLP
Wright et al. (1993)	Chemical reactor	ODE	1 × 1	SIM
Zheng (1997)	Distillation column	ODE	2 × 2	NLP

NLP = nonlinear program, SEQ = simultaneous method, SIM = simultaneous method, SL = successive linear.

Table 2
NMPC experimental studies

Reference	Process	Model	Size	Solution
Patwardhan and Edgar (1993)	Distillation column	ODE	2 × 2	SEQ
Pottmann and Seborg (1997)	pH neutralization reactor	Radial basis function	1 × 1	Off-line
Wright and Edgar (1994)	Chemical reactor	PDE	1 × 1	SIM
Zhu et al. (1991)	pH neutralization reactor	Hammerstein	1 × 1	NEW

SEQ = simultaneous method, SIM = simultaneous method, NEW = Newton search.

industrial problems where retuning of a large controller is necessary.

5.3. Commerical applications

During the past few years, several vendors have introduced commercial nonlinear control products which are being applied in the process industries. Available products include the NOVA nonlinear controller (Dynamic Optimization Technology Product, 1996) from Dynamic Optimization Technology (DOT) Products and the Process Perfector (Martin, 1997) from Pavilion Technologies. Detailed information concerning the nonlinear control algorithms is not available in the open literature. However, information in the public domain strongly indicates that these products are based on NMPC techology.

The NOVA controller has been installed on several polyethylene and polypropylene processes worldwide. Existing applications utilize fundamental models, but empirical nonlinear models can be accomodated. The product includes utilities for fundamental model development and parameter estimation. It appears that the NOVA control algorithm is a nonlinear extension of the IDCOM technique developed for linear process models (Qin and Badgwell, 1997). The optimization problem considers both the desired closed-loop behavior and an economic objective function. Process constraints are treated as controlled variables. The NOVA optimization engine is used to solve the resulting NMPC problem. The controller is tuned by specifying parameters which determine the desired closed-loop response. According to Dynamic Optimization Technology Products (1996), the controller is capable of providing high performance control even during polymer grade changes in which the process gain changes by several orders of magnitude.

As of March, 1997, the Process Perfector had been applied to 15 refining and polymerization processes. With the exception of a hydrocracker, all the installations involve SISO or small multivariable control problems. The technique combines steady-state and dynamic neural network models. The steady-state model is used for steady-state optimization and to calculate updated process gains for the dynamic model. The solution of the steady-state optimization problem determines the target values for the dynamic optimization, similar to the hierarchical optimization strategy used in linear DMC (Qin and Badgwell, 1997). The target values and the updated dynamic model are utilized for dynamic optimization. According to Martin (1997), the controller is capable of reducing polymer grade transition times by a factor of three.

6. Future research directions

During the past decade, considerable process has been made in the theory and practice of nonlinear model predictive control (NMPC). NMPC is most appropriate for constrained multivariable processes which are sufficiently nonlinear that conventional linear control techniques are inadequate. The commercial market for such a technology is large, as exemplified by the recent development of NMPC products by several vendors. The future of NMPC depends on continued development of effective modeling techniques, alternative problem formulations, efficient solution strategies, and novel process applications. Below some specific topics which warrant further investigation by both the academic and industrial communities are discussed.

6.1. Nonlinear modeling

The NMPC approach assumes the availability of a suitable nonlinear dynamic model of the controlled process. It is well known that model development is the most time consuming activity in linear model predictive control (LMPC) projects. In most application studies of NMPC, the nonlinear model is readily obtained due to the simplicity of the process considered. The nonlinear modeling problem is considerably more challenging in an industrial setting due to the large-scale nature of manufacturing plants. Consequently, the development of nonlinear modeling tools is of paramount importance to the continued evolution of NMPC.

The fundamental modeling approach could be enhanced by the integration of NMPC with sophisticated dynamic simulators. This would reduce the time and effort required to derive fundamental models for large-scale processes. The utility of the resulting models could be enhanced by the development of suitable nonlinear model reduction techniques (Levine and Rouchon, 1991). Another promising research direction is the development of low-order modeling strategies which provide a more reasonable compromise between model complexity and prediction accuracy (Benallou et al., 1986). More systematic techniques for nonlinear model validation also are needed. An important area for theoretical research is analysis of the necessary smoothness properties the nonlinear model should possess to be effectively utilized with various optimization codes. A very important practical issue is the development of computer-aided tools for maintaining the resulting nonlinear models.

The empirical modeling approach is promising because detailed process knowledge is not required and the model form can be chosen to reduce on-line computation. The industrial success of LMPC largely is attributable to the development of well established guidelines for dynamic data collection and analysis. Many theoretical and practical problems must be resolved before empirical nonlinear modeling is viable for large-scale industrial processes. A particularly important topic is the design of input sequences which provide sufficient excitation yet are acceptable in industrial situations (Pearson and Ogunnaike, 1997). A related issue is characterization of the amount and type of process data required to build nonlinear empirical models with satisfactory predictive capability. Another topic for future research is the identification of nonlinear model structures which are capable of capturing a wide variety of process behaviors and are amenable to NMPC optimization (Sriniwas and Arkun, 1997). It may be necessary to perform online model adaptation when the process deviates significantly from the operating conditions used for empirical model development. This requires the development of recursive parameter estimation techniques for nonlinearly parameterized models (Marino and Tomei, 1993). A very important topic which has received little attention is the development of hybrid modeling approaches which allow fundamental and empirical process knowledge to be integrated.

6.2. Nonlinearity measures

This review is based on the underlying premise that the potential benefits of nonlinear control are sufficient to warrant the additional effort required to develop, implement, and maintain NMPC control systems. In practice, the assessment of potential benefits is a critical issue which should be addressed in the initial stages of a control project. This requires tools for assessing process nonlinearity. Recently, there has been considerable research on the development of nonlinearity measures. Initial work focused on characterizing open-loop nonlinearity (Allgower and Gilles, 1992; Guay et al., 1995). The results obtained using these methods can be misleading because open-loop nonlinearity measures do not necessarily provide an accurate assessment of the inherent limitations associated with linear control techniques. It is well known that some nonlinear processes can be optimally controlled with a linear controller (Stack and Doyle, 1997). This has led to the development of closed-loop nonlinearity measures which consider performance objectives as well as the controlled process (Guay et al., 1997; Stack and Doyle, 1997). These methods yield measures of control-relevant nonlinearity.

The continued development of nonlinearity measure theory is needed to provide quantitative tools for evaluating the potential benefits of NMPC. An important problem is the determination of information required to adequately characterize control-relevant nonlinearity. Such information may include the nonlinear process model, the desired operating regime, the performance objective, and the control structure. Also needed is the derivation of metrics which provide measures of control-relevant nonlinearity as well as the potential benefits of nonlinear control. The ultimate goal is to utilize nonlinearity measures to determine the most appropriate control technology given the process characteristics and the performance objectives.

6.3. *Problem formulation*

The NMPC problem involves on-line computation of a sequence of manipulated inputs which optimize an objective function and satisfy process constraints. A wide variety of problem formulations have been proposed. Most techniques are developed with an emphasis on computational efficiency, although some methods focus on stability properties. Comparative studies which reveal the relative advantages and disadvantages of these different formulations would be useful. Of particular interest would be a characterization of the class of nonlinear systems which are well suited for a particular formulation.

The development of new NMPC formulations also is an important topic for future research. There has been some exciting progress in this direction during the past few years (see Section 4.4) and further work is needed. There should be a particular emphasis on formulations which are suited for large-scale systems. With the exception of Ricker and Lee (1995), this topic largely has been ignored by academic researchers. The development of NMPC techniques for large-scale systems may require problem formulations which exploit the specific structure of the nonlinear model. The ultimate goal should be the development of computationally efficient NMPC techniques with guaranteed stability properties.

6.4. *Algorithm development*

LMPC techniques require a linear or quadratic programming problem to be solved at every sampling period. The industrial success of LMPC is partially attributable to the widespread availability of efficient and robust solution techniques for these types of optimization problems. On the other hand, NMPC requires on-line solution of a nonlinear program (NLP) at each execution of the controller. The solution of such NLP problems can be very time consuming, especially for large-scale systems. An additional complication is that the optimization problem generally is nonconvex because the nonlinear model equations are posed as constraints. Consequently, NLP solvers designed for convex problems may converge to a local minima or even diverge.

Future research should focus on the development of efficient and reliable algorithms for on-line solution of NLP problems arising from various NMPC formulations. This will require the development of improved solution algorithms for nonconvex NLP problems. Studies which extend the class of nonlinear models and NMPC formulations which yield convex NLPs also should be pursued (Sriniwas and Arkun, 1997). Particular emphasis should be placed on large-scale optimization problems arising from plant-wide applications. This may require a delicate interplay between the nonlinear model form, the NMPC problem formulation, and the NLP solution algorithm. Recent work (Maner and Doyle, 1997) has demonstrated the potential advantages of such an integrated modeling/optimization approach. Another important area for future research is the development of NLP solution strategies which use available computing resources more effectively. For instance, the NLP solver might be stopped near convergence since the last few iterations usually result in small changes in the solution. Then the remaining time could be used to precompute information required for the next execution of the controller.

6.5. *Stability and robustness analysis*

Nominal stability results are available for NMPC when the prediction horizon is infinite or a terminal state constraint is imposed. Unfortunately, both these conditions are problematic from an implementation perspective. It is not possible to maintain the stabilizing properties of NMPC by reformulating the infinite horizon problem as a finite horizon problem with a simple terminal state penalty in the objective function. The assumption that the state vector can be driven to the origin in the presence of constraints is known as constrained null controllability. The terminal state constraint is limiting because it is very difficult to ensure the existence of a control horizon such that a general nonlinear system is constrained null controllable (Meadows, 1997). Consequently, it is necessary to derive stabilizing NMPC formulations which are more suitable for on-line implementation. Some initial results in this direction are summarized in Section 4.5.

The primary reason for including feedback in NMPC is to account for mismatch between the actual plant and the process model. Simulation and experimental studies demonstrate that NMPC has some degree of robustness to modeling errors. Nevertheless, it is important to develop a rigorous theory which allows the robustness of different NMPC formulations to be analyzed and facilitates the derivation of new formulations with improved robustness properties. Some interesting robustness results are presented by Badgwell (1997), Mayne and Michalska (1991c), and Nicolao et al. (1996). Another source of plant/model mismatch is attributable to the actual state variables being replaced by state estimates obtained from a nonlinear observer. An important problem is the derivation of conditions under which a stabilizing NMPC controller and a stable observer can be combined to yield a stable closed-loop system. Some promising results for discrete-time NMPC are presented by Scokaert et al. (1997).

6.6. *Process applications*

NMPC has been applied to a wide variety of simulated processes. Most application studies consider benchmark processes with a small number of input and output variables. While these problems are useful for evaluating new NMPC formulations, future studies should focus on novel process applications and large-scale systems. In particular, there is a need to develop NMPC formulations which are amenable to real-time implementation on large-scale processes. Some interesting results in this direction are presented in Ricker and Lee (1995), and further advances would be facilitated by the development of new industrial challenge problems. In applications where only particular process units are strongly nonlinear, it may be unnecessary to apply NMPC to the entire plant. This motivates the development of theoretical tools and practical guidelines which facilitate integration of NMPC and conventional linear control technologies. The number of experimental applications of NMPC is surprisingly low. More experimental studies are needed to obtain realistic comparisons of the relative performance of NMPC and conventional linear control techniques.

Commercial application of NMPC requires the development of software tools that provide integrated dynamic modeling and real-time optimization capabilities. Several vendors currently offer such software products, and it is likely that process economics will spur the development of new products. Also needed are software support tools for long-term maintenance of nonlinear models and controllers. This toolkit would enable the control engineer to analyze model accuracy and evaluate closed-loop performance. The deployment of NMPC technology would be advanced by establishing hardware and software standards. This would allow components from different products to be combined to achieve the best solution for a particular application.

Acknowledgements

The author would like to acknowledge financial support from an NSF Career Development Award (CTS-9501368). The author also would like to acknowledge members of the nonlinear model predictive control breakout group at the NSF/NIST Workshop on Process Measurements and Control who contributed to many of the ideas in this paper.

References

Alamir, M., & Bornard, G. (1994). On the stability of receding horizon control of nonlinear discrete-time systems. *Systems Control Lett.*, *23*, 291–296.

Allgower, F., & Gilles, E. D. (1992). Approximate input/output linearization of nonlinear systems. *A.I.Ch.E. Annual Mtg.*, Miami.

Badgwell, T. A. (1997). A robust model predictive control algorithm for stable nonlinear plants. *Proc. IFAC Symposium on Advanced Control of Chemical Processes*, Banff, Canada, pp. 477–481.

Bastin, G., & Gevers, M. R. (1988). Stable adaptive observers for nonlinear time-varying systems. *IEEE Trans. Automat. Control*, *AC-33*, 650–658.

Benallou, A., Seborg, D. E., & Mellichamp, D. A. (1986). Dynamic compartmental models for separation processes. *A.I.Ch.E. J.*, *32*, 1067–1078.

Bequette, B. W. (1990). Process control using nonlinear programming techniques. Lecture Notes in Control and Information Sciences. pp. 57–66. New York, NY: Springer.

Bequette, B. W. (1991). Nonlinear control of chemical processes: a review. *Ind. Engng Chem. Res.*, *30*, 1391–1413.

Brengel, D. M., & Seider, W. D. (1989). Multistep nonlinear predictive controller. *Ind. Engng Chem. Res.*, *28*, 1812–1822.

Chen, C. C., & Shaw, L. (1982). On receding horizon feedback control. *Automatica*, *18*, 349–352.

Chen, H., & Allgower, F. (1997). A quasi-infinite horizon nonlinear predictive control scheme for stable systems: application to a CSTR. *Proc. IFAC Symp. on Advanced Control of Chemical Processes*, Banff, Canada, pp. 471–476.

Chen, H., Kremling, A., & Allgower, F. (1995). Nonlinear predictive control of a benchmark CSTR. *Proc. European Control Conf.*, Rome, Italy, pp. 3247–3252.

Chu, X., & Seborg, D. E. (1994). Nonlinear model predictive control based on Hammerstein models. *Proc. Int. Symp. on Process Systems Engineering*, Seoul, Korea, p. 995.

Duchene, P., & Rouchon, P. (1996). Kinetic scheme reduction via geometric singular perturbation techniques. *Chem. Engng Sci.*, *51*, 4661–4672.

Dynamic Optimization Technology Products (1996). Technical description of NOVA nonlinear controller. *Hydrocarbon Processing*, Issue on Advanced Control and Information Systems.

Eaton, J. W., & Rawlings, J. B. (1990). Feedback control of chemical processes using on-line optimization techniques. *Comput. Chem. Engng*, *14*, 469–479.

Garcia, C. E. (1984). Quadratic dynamic matrix control of nonlinear processes: An application to a batch reactor process. AIChE Annual Mtg., San Francisco.

Gattu, G., & Zafiriou, E. (1992). Nonlinear quadratic dynamic matrix control with state estimation. *Ind. Engng Chem. Res.*, *31*, 1096–1104.

Guay, M., McLellan, P. J., & Bacon, D. W. (1995). Measurement of nonlinearity in chemical process control systems: the steady-state map. *Canadian J. Chem. Engng*, *73*, 868–882.

Guay, M., McLellan, P. J., & Bacon, D. W. (1997). Measure of closed-loop nonlinearity and interaction for nonlinear chemical processes. *A.I.Ch.E. J.*, *43*, 2261–2278.

Hernandez, E., & Arkun, Y. (1990). Neural network modeling and an extended DMC algorithm to control nonlinear systems. *Proc. American Control Conf.*, San Diego, CA, pp. 2454–2459.

Hernandez, E., & Arkun, Y. (1992). A study of the control relevant properties of backpropagation neural net models of nonlinear dynamical systems. *Comput. Chem. Engng*, *16*, 227–240.

Hernandez, E., & Arkun, Y. (1993). Control of nonlinear systems using polynomial ARMA models. *A.I.Ch.E. J.*, *39*, 446–460.

Hwang, Y. L. (1991). Nonlinear wave theory for dynamics of binary distillation column. *A.I.Ch.E. J.*, *37*, 705–723.

Jang, S., Joseph, B., & Mukai, H. (1987). Control of constrained multivariable nonlinear processes using a two-phase approach. *Ind. Engng Chem. Res.*, *26*, 2106–2114.

Kazantzis, N., & Kravaris, C. (1994). System-theoretic properties of sampled-data representations of nonlinear systems obtained via Taylor methods. *Proc. American Control Conf.*, Baltimore, MD, pp. 1354–1360.

Keerthi, S. S., & Gilbert, E. G. (1988). Optimal infinite-horizon feedback laws for a general class of constrained discrete-time systems: stability and moving-horizon approximations. *J. Optimiz. Theory Applic.*, *57*, 265–292.

Khalil, H. (1992). *Nonlinear systems*. Londom: Macmillan.

Kurtz, M. J., & Henson, M. A. (1997). Input–output linearizing control of constrained nonlinear processes. *J. Process Control*, *7*, 3–17.

Lee, J. H., & Ricker, N. L. (1994). Extended Kalman filter based nonlinear model predictive control. *Ind. Engng Chem. Res.*, *33*, 1530–1541.

Levine, J., & Rouchon, P. (1991). Quality control of binary distillation columns via nonlinear aggregated models. *Automatica*, *27*, 463–480.

Li, W. C., & Biegler, L. T. (1989). Multistep, Newton-type control strategies for constrained nonlinear systems. *Chem. Engng Res. Dev.*, *67*, 562–577.

Maner, B. R., & Doyle, F. J. (1997) Polymerization reactor control using autoregressive-plus Volterra-based MPC. *A.I.Ch.E. J.*, *43*, 1763–1784.

Maner, B. R., Doyle, F. J., Ogunnaike, B. A., & Pearson, R. K. (1996). Nonlinear model predictive control of a simulated multivariable polymerization reactor using second-order Volterra models. *Automatica*, *32*, 1285–1301.

Marco, R. D., Semino, D., & Brambilla, A. (1997). From linear to nonlinear model predictive control: Comparison of different algorithms. *Ind. Engng Chem. Res.*, *36*, 1708–1716.

Marino, R., & Tomei, P. (1993). Global adaptive output-feedback control of nonlinear systems: Part II: nonlinear paramertization. *IEEE Trans. Automat. Control*, *AC-38*, 33–48.

Martin, G. (1997). Nonlinear model predictive control with integrated steady-state model-based optimization. AIChE National Mtg., Houston.

Mayne, D. Q. (1997). Nonlinear model predictive control: an assessment. In: J. C. Kantor, C. E. Garcia & B. Carnahan (Eds). *Chemical Process Control V*, pp. 217–231. CACHE, AIChE, Austin, TX.

Mayne, D. Q., & Michalska, H. (1990). Receding horizon control of nonlinear systems. *IEEE Trans. Automat. Control*, *AC-35*, 814–824.

Mayne, D. Q., & Michalska, H. (1991a). Model predictive control of nonlinear systems. *Proc. American Control Conf.*, (pp. 2343–2348). Pittsburgh, PA.

Mayne, D. Q., & Michalska, H. (1991b). Receding horizon control of constrained nonlinear systems. *European Control Conf.*, Grenoble, France, pp. 2037–2042.

Mayne, D. Q., & Michalska, H. (1991c). Robust receding horizon control. *Proc. IEEE Conf. on Decision and Control*, (pp. 64–69). Brighton, England.

Meadows, E. S. (1997). Dynamic programming and model predictive control. *Proc. American Control Conf.*, Albuquerque, NM, pp. 1635–1639.

Meadows, E. S., Henson, M. A., Eaton, J. W., & Rawlings, J. B. (1995). Receding horizon control and discontinuous state feedback stabilization. *Int. J. Control*, *62*, 1217–1229.

Meadows, E. S., & Rawlings, J. B. (1997). Model predictive control. In M. A. Henson & D. E. Seborg (Eds). *Nonlinear process control*, Chapter 5, (pp. 233–310). Englewood Cliffs, NJ: Prentice-Hall.

Michalska, H., & Mayne, D. Q. (1995) Moving horizon observers and observer-based control. *IEEE Trans. Automat. Control*, *40*, 995–1006.

Moraal, P. E., & Grizzle, J. W. (1995). Observer design for nonlinear systems with discrete-time measurements. *IEEE Trans. Automat. Control*, *AC-40*, 395–404.

Morningred, J. D., Paden, B. E., Seborg, D.E., & Mellichamp, D.A. (1992). An adaptive nonlinear predictive controller. *Chem. Engng Sci.*, *47*, 755–762.

Morshedi, A. M. (1986). Universal dynamic matrix control. In: Morari, M., & McAvoy, T. J. (Eds). *Chemical process control III*, pp. 547–577. CACHE, AIChE, Austin, TX.

Muske, K. R., & Edgar, T. F. (1997) Nonlinear state estimation. In: Henson, M. A., & Seborg, D. E. (Eds). *Nonlinear Process Control*, Chapter 6 (pp. 311–370). Englewood Cliffs, NJ: Prentice-Hall.

Muske, K. R., & Rawlings, J. B. (1993). Model predictive control with linear models. *A.I.Ch.E. J.*, *39*, 262–287.

Nahas, E. P., Henson, M. A., & Seborg, D. E. (1992). Nonlinear internal model control for neural network models. *Computers Chem. Engng*, *16*, 1039–1057.

Nevistic, V., & Morari, M. (1995). Constrained control of feedback-linearizable systems. *Proc. European Control Conf.*, Rome, Italy, pp. 1726–1731.

Nicolao, G. D., Magni, L., & Scattolini, R. (1996). On the robustness of receding-horizon control with terminal constraints. *IEEE Trans. Automat. Control.*, *41*, 451–453.

Nicolao, G. D., Magni, L., & Scattolini, R. (1997). Stabilizing predictive control of nonlinear ARX models. *Automatica.*, *33*, 1691–1697.

Ogunnaike, B. A., Chien, I.-L., & Arkun, Y. (1993). Nonlinear model predictive control of high-purity distillation columns using polynomial ARMA models. *Proc. European Control Conf.*, Groningen, Netherlands.

Ogunnaike, B. A., & Ray, W. H. (1994). *Process dynamics, modeling, and control*. New York, NY: Oxford Press.

Patwardhan, A. A., & Edgar, T. F. (1993). Nonlinear model predictive control of a packed distillation column. *Ind. Engng Chem. Res.*, *32*, 2345–2356.

Patwardhan, A. A., Rawlings, J. B., & Edgar, T. F. (1990). Nonlinear model predictive control. *Chem. Engng Comm.*, *87*, 123–141.

Pearson, R. K., & Ogunnaike, B. A. (1997). Nonlinear process identification. In: Henson, M. A., & Seborg, D. E. (Eds). *Nonlinear process control*, Chapter 2 (pp. 11–110). Englewood Cliffs, NJ: Prentice-Hall.

Peterson, T., Hernandez, E., Arkun, Y., & Schork, F. J. (1992). A nonlinear DMC algorithm and its application to a semibatch polymerization reactor. *Chem. Engng Sci.*, *47*, 737–753.

Pottmann, M., & Henson, M.A. (1997). Compactly supported radial basis functions for adaptive process control. *J. Process Control*, *7*, 345–356.

Pottmann, M., & Seborg, D. E. (1997). A nonlinear predictive control strategy based on radial basis function models. *Comput. Chem. Engng*, *21*, 965–980.

Qin, S. J., & Badgwell, T. A. (1997). An overview of industrial model predictive control technology. In: Kantor, J.C. Garcia, C.E., & Carnahan, B. (Eds). *Chemical process control V* (pp. 232–256). CACHE, AIChE, Austin, TX.

Rawlings, J. B., Jerome, N. F., Hamer, J. W., & Bruemmer, T. M. (1989). End-point control in semi-batch chemical reactors. *Proc. IFAC Symposium DYCORD* + 89, pp. 323–328. Maastricht, Netherlands.

Rawlings, J. B., & Muske, K. R. (1993). The stability of constrained receding horizon control. *IEEE Trans. Automat. Control, AC-38*, 1512–1516.

Ricker, N. L., & Lee, J. H. (1995). Nonlinear model predictive control of the Tennessee Eastman challenge process. *Comput. Chem. Engng, 19*, 961–981.

Scokaert, P. O. M., Rawlings, J. B., & Meadows, E. S. (1997). Discrete-time stability with perturbations: Application to model predictive control. *Automatica, 33*, 463–470.

Sistu, P. B., & Bequette, B. W. (1996). Nonlinear model-predictive control: closed-loop stabi lity analysis. *A.I.Ch.E. J., 42*, 3388–3402.

Sistu, P. B., Gopinath, R. S., & Bequette, B. W. (1991). Implementation issues in nonlinear predictive control. *Proc. American Control Conf.*, Pittsburgh, PA, pp. 2361–2366.

Soroush, M. (1997). Nonlinear state-observer design with application to reactors. *Chem. Engng Sci., 52*, 387–404.

Sriniwas, G. R., & Arkun, Y. (1997) A global solution to the nonlinear model predictive control algorithms using polynomial ARX models. *Comput. Chem. Engng, 21*, 431–439.

Stack, A. J., & Doyle, F. J. (1997). Application of a control-law nonlinearity measure to the chemical reactor analysis. *A.I.Ch.E. J., 43*, 425–439.

Su, H. T., & McAvoy, T. J. (1997). Artificial neural networks for nonlinear process identification and control. In: Henson, M.A., & Seborg, D.E. (Eds). *Nonlinear process control*, Chapter 7 (pp. 371–428). Englewood Cliffs, NJ: Prentice-Hall.

Uppal, A., Ray, W. H., & Poore, A.B. (1974). On the dynamic behavior of continuous stirred tank reactors. *Chem. Engng Sci., 29*, 967.

Wright, G. T., Breedijk, T., & Edgar, T.F. (1991). On-line parameter estimation and adaptation in nonlinear model-based control. *Proc. American Control Conf.*, Pittsburgh, PA, pp. 2782–2787.

Wright, G. T., & Edgar, T.F. (1994). Nonlinear model predictive control of a fixed-bed water-gas shift reactor: An experimental study. *Comput. Chem. Engng, 18*, 83–102.

Zheng, A. (1997). A computationally efficient nonlinear MPC algorithm. *Proc. American Control Conf.*, Albuquerque, NM, pp. 1623–1627.

Zhu, Q. M., Warwick, K., & Douce, J. L. (1991). Adaptive general predictive controller for nonlinear systems. *IEE Proc. – Pt. D, 138*, 33–40.