Model Predictive Control: Past, Present and Future *

Manfred Morari
Institut für Automatik
ETH - Z/ETL
CH-8092 Zürich Switzerland
morari@aut.ee.ethz.ch

Jay H. Lee
Department of Chemical Engineering
Auburn University
Auburn, AL 36849-5127
jhl@eng.auburn.edu

Abstract

More than 15 years after Model Predictive Control (MPC) appeared in industry as an effective means to deal with multivariable constrained control problems, a theoretical basis for this technique has started to emerge. The issues of feasibility of the on-line optimization, stability and performance are largely understood for systems described by linear models. Much progress has been made on these issues for nonlinear systems but for practical applications many questions remain, including the reliability and efficiency of the on-line computation scheme. To deal with model uncertainty "rigorously" an involved dynamic programming problem must be solved. The approximation techniques proposed for this purpose are largely at a conceptual stage. Among the broader research needs the following areas are identified: multivariable system identification, performance monitoring and diagnostics, nonlinear state estimation, and batch system control. Many practical problems like control objective prioritization and symptom-aided diagnosis can be integrated systematically and effectively into the MPC framework by expanding the problem formulation to include integer variables yielding a Mixed-Integer Quadratic or Linear Program. Very efficient techniques for solving these problems are becoming available.

INTRODUCTION

The intention of this paper is to give an overview of the origins of Model Predictive Control and its glorious present. No attempt is made to categorize and comprehensively review the literature which includes several books (Robert R. Bitmead and Wertz, 1990; Soeterboek, 1992; Martín Sánchez and Rodellar, 1996; Clarke, 1994; Berber, 1995; Camacho and Bordons, 1995) and hundreds of papers (Kwon, 1994). The review should give the novice reader an impression which practical objectives have been pursued, which theoretical problems have been formulated and what progress has been made without undue mathematical complexity. All citations are only exemplary and should point the reader in a direction where more details are available. There is more emphasis on the future of MPC than on its past. MPC brings out new needs in related areas like system identification, state estimation, monitoring and diagnostics, etc. We show that many important practical and theoretical problems can be formulated in the MPC framework. Pursuing them will assure MPC of its stature as a vibrant research area, where theory is seen to support practice more directly than in most other areas of control research.

THE PAST

Though the ideas of receding horizon control and model predictive control can be traced back to the 1960s (García *et al.*, 1989), interest in this field started to surge only in the 1980s after publication of the first paper on Dynamic Matrix Control (DMC) (Cutler and Ramaker, 1979; Cutler and Ramaker, 1980) and the first comprehensive exposition of Generalized Predictive Control (GPC) (Clarke *et al.*, 1987*a*; Clarke *et al.*, 1987*b*). At first sight, the ideas underlying the two methods are similar.

The objectives behind the developments of DMC and GPC were very different, however. DMC was conceived to tackle the multivariable constrained control problems typical for the oil and chemical industries. In the pre-DMC era these problems were handled by single loop controllers augmented by various selectors, overrides, decouplers, time-delay compensators, etc. For the DMC task a time-domain model (finite impulse or step response model) was natural. GPC was intended to offer a new adaptive control alternative. In the tradition of much of the work in adaptive control input/output (transfer function) models were employed. Stochastic aspects played a key role in GPC from the very beginning, while the original DMC formulation was completely deterministic and did not include any explicit disturbance model.

The GPC approach is not suitable or, at the very least, awkward for multivariable constrained systems which are much more commonly encountered in the oil and chemical industries than situations where adaptive control is needed. Essentially all vendors have adopted a DMC-like approach (Qin and Badgwell, 1996). Because of these reasons and because of the type of applications of interest to the readers of this journal GPC will not be discussed any further. The interested reader is referred to several recent books on this subject (Robert R. Bitmead and Wertz, 1990; Soeterboek, 1992; Martín Sánchez and Rodellar, 1996).

DMC had a tremendous impact on industry. There is probably not a single major oil company in the world, where DMC (or a functionally similar product with a different trade name) is not employed in most new installations or revamps. For Japan some statistics are available (Ohshima *et al.*, 1995). The initial research on MPC is characterized by attempts to *un-*

*This paper was presented at the Joint 6th International Symposium on Process Systems Engineering (PSE'97) and 30th European Symposium on Computer Aided Process Engineering (ESCAPE-7), May 25 - 29 1997, Trondheim, Norway

derstand DMC, which seemed to defy a traditional theoretical analysis because it was formulated in a nonconventional manner. One example was the development of Internal Model Control (IMC) (García and Morari, 1982) which failed to shed light on the constrained behavior of DMC but led to some insights on robust control (Morari and Zafiriou, 1989)

THE PRESENT

LINEAR MPC

Nowadays in the research literature MPC is formulated almost always in the state space. The system to be controlled is described by a linear discrete time model.

$$x(k+1) = Ax(k) + Bu(k), \quad x(0) = x_0, \tag{1}$$

where $x(k) \in \Re^n$ and $u(k) \in \Re^m$ denote the state and control, respectively. A receding horizon implementation (García *et al.*, 1989) is typically formulated by introducing the following open-loop optimization problem.

$$J_{(p,m)}(x_0) = \inf_{u(\cdot)} \left[x^T(p)P_0x(p) + \sum_{i=0}^{p-1} x^T(i)Qx(i) + \sum_{i=0}^{m-1} u^T(i)Ru(i) \right] \tag{2}$$

subject to

$$Ex + Fu \le \psi \tag{3}$$

($p \geq m$) where p denotes the length of the *prediction horizon* or *output horizon*, and m denotes the length of the *control horizon* or *input horizon*. (When $p = \infty$, we refer to this as the *infinite horizon problem*, and similarly, when p is finite, we refer to it as a finite horizon problem.)

The expressions (1,2,3) define a quadratic program for which many algorithms and commercial software exist. Let $u^*_{(p,m)}(i|k)$, $i = 0, \cdots, m-1$ be the minimizing control sequence for $J_{(p,m)}(x(k))$ subject to the system dynamics (1). A receding horizon policy proceeds by implementing *only the first* control $u^*_{(p,m)}(0|k)$ to obtain $x(k+1) = Ax(k) + Bu^*_{(p,m)}(0|k)$. The rest of the control sequence $u^*_{(p,m)}(i|k)$ is discarded and $x(k+1)$ is used to update the optimization problem (2) as a new initial condition. This process is repeated, each time using only the first control action to obtain a new initial condition, then shifting the cost ahead one time step and repeating, hence the name receding horizon control. In the special case when $p = m = N$, then $J_{(p,m)} = J_N$ as defined in (2).

We note that as the control horizon and the prediction horizon both approach infinity and when there are no constraints we obtain the standard Linear Quadratic Regulator (LQR) problem, which was studied extensively in the 1960s and 1970s (Kwakernaak and Sivan, 1972). The optimal control sequence is generated by a static state feedback law where the feedback gain matrix is found via the solution of an Algebraic Riccati Equation (ARE). This feedback law has some well known nice properties, in particular, it guarantees closed-loop stability for any positive (semi) definite choice of weighting matrices Q and R.

With constraints an infinite dimensional optimization problem results, which is – at least at first sight – not a very practical proposition. On the other hand, by choosing both the control and the output horizons to be finite, the quadratic program is finite dimensional and can be solved easily on-line at every time step.

Three practical questions are immediate: 1) When is the problem formulated above feasible, so that the algorithm yields a control action which can be implemented? 2) When does the sequence of computed control actions lead to a system which is closed-loop stable? 3) What closed-loop performance results from repeated solution of the specified open-loop optimal control problem?

Feasibility

The constraints stipulated in (3) may render the optimization problem infeasible. Input saturation constraints cannot be exceeded, while constraints involving outputs can be violated, albeit with undesirable consequences for the controlled system. It may happen, for example, because of a disturbance, that the optimization problem posed above becomes infeasible at a particular time step. Obviously a real-time control algorithm must not fail in this trivial fashion. Therefore in all commercial algorithms the hard constraints are softened by introducing slack variables which are kept small by introducing a corresponding penalty term in the objective (Zheng and Morari, 1995*b*). There are many variations on this theme to suit different tastes. At issue are magnitude of the violation versus duration and if the solution of the problem with softened constraints may lead to a constraint violation though a feasible solution without constraint violation exists (Scokaert and Rawlings, 1996*b*; Scokaert and Rawlings, 1996*a*).

If the system is unstable then, in general, the system cannot be stabilized globally, when there are input saturation constraints. Algorithms for precalculating a feasible region were proposed by Zheng and Morari (1995*a*) and Gilbert and Tan (1991).

Finally, it may happen, that the algorithm which minimizes an *open-loop* objective, inadvertently drives the closed-loop system outside the feasible region. This difference between open-loop objective and closed-loop behavior is addressed below.

Closed Loop Stability

In either the infinite or the finite horizon constrained case it is not clear under what conditions the closed loop system is stable. Much recent research on linear MPC has focused on this problem. Two approaches have been proposed to guarantee stability: one based on the original problem (1), (2), and (3) and one where a "contraction constraint" is added (Polak and Yang, 1993*a*; Polak and Yang, 1993*b*). With the contraction constraint the norm of the state is *forced* to decrease with time and stability follows trivially independent of the various parameters in the objective function. Without the contraction constraint the stability problem is more complicated. The most comprehensive and also most compact analysis was presented in Nevistić and Primbs (1997) and Primbs and Nevistić (1997) whose arguments we will sketch here. To simplify the exposition we assume $p = m = N$, then $J_{(p,m)} = J_N$ as defined in (2). The key idea is to use the optimal finite horizon cost J_N, the *value function*, as a Lyapunov function. One wishes to show that $J_N(x(k)) - J_N(x(k+1)) > 0$ for $x \neq 0$. Rewriting $J_N(x(k)) - J_N(x(k+1))$ gives:

$$J_N(x(k)) - J_N(x(k+1)) = [x^T(k)Qx(k) + u_N^{*T}(x(k))Ru_N^*(x(k))] + [J_{N-1}(x(k+1)) - J_N(x(k+1))] \quad (4)$$

If it can be shown that the right hand side of (4) is positive, then stability is proven. Assuming $Q > 0$, the first term $[x^T(k)Qx(k) + u_N^*(x(k))Ru_N^*(x(k)))]$ is positive. In general, it *cannot* be asserted that the second term $[J_{N-1}(x(k+1)) - J_N(x(k+1))]$ is **non-negative**.
Several approaches have been presented to assure that the right hand side of (4) is positive:

- Primbs and Nevistić (1997) showed that the second term approaches zero as $N \to \infty$ and that there exists a finite N^* such that for $N > N^*$ the first term dominates over the second. In general, the solution of a non-convex min-max problem is necessary in order to determine N^*. However, when the system is open-loop stable and when the constraints involve the control inputs only, the solution of a somewhat conservative version of this problem requires eigenvalue computations only, which is quite remarkable.
- When an end constraint $x(k+N) = 0$ (Kwon and Pearson, 1977) is imposed it can be argued in a straight forward manner that J_N is monotonically non–increasing as a function of N, which trivially guarantees stability.
- When P_0 is chosen as the solution to the Lyapunov equation $A^T P_0 A + Q = P_0$ (Rawlings and Muske, 1993), then J_N is again monotonically non–increasing and stability follows. This choice of P_0 amounts to using an infinite output horizon (with input horizon of N). The constraint horizon, however, must be chosen large enough so that satisfying the constraints within the finite horizon imply the same for the infinite horizon. Such a choice of horizon can be determined, for instance, using the concept of the maximal output admissible sets (Gilbert and Tan, 1991).
- In the special case when there are no constraints and N is finite, the FARE conditions for guaranteeing stability (Robert R. Bitmead and Wertz, 1990) follow directly.

Some remarks are in order. Despite the fact that there exist now techniques to test for stability of constrained systems with finite p, this choice is generally not recommended. The system behavior is relatively insensitive to changes in both p and m over a wide range of values, therefore Q and R are the typical tuning parameter to affect performance. Soeterboek (1992) has shown, however, that for a finite p the effect of the control weighting R may be "non-monotonic", i.e., increasing R may lead to instability, which is counter-intuitive. This type of behavior was not observed for the infinite horizon case, though no proof exists.
It is a common practice in the process industries to add integrators to the model by expressing it in terms of differenced inputs δu and outputs and re-integrating the differenced outputs. In this case the objective function (2) includes a penalty term on δu rather than u which effectively adds integral action to the controller. For $p = \infty$ and with the constraints softened as discussed above, systems with integrators can be globally stabilized with MPC (Zheng and Morari, 1995*b*). For FIR systems, using an infinite output horizon is equivalent to setting $p = m + t_s$ and adding the constraint $y(m + t_s) = 0$ (where y represents the output of the FIR system and t_s the number of time steps it takes for the system to settle) (Lee, 1996).
For general state-space systems, integrating modes must be zeroed at the end of the control horizon in order for the infinite horizon cost to be bounded. This gives rise to some technical complications since zeroing of integrating modes is not always possible for two reasons: First, the chosen control horizon may not be sufficiently large. In other words, there may not be enough degrees of freedom available to force the integrating modes to zero at the end of the horizon. Second, one may have hard constraints on the input u, which translate into constraints on the *integrated* δu. Both problems can be overcome by performing a bi-level optimization, i.e., steady-state error minimization followed by dynamic error minimization, as suggested by Lee (1996). The asymptotic stability is preserved, if the constraints on u are such that returning the integrating modes to the origin is possible. It is interesting to note that a similar bi-level optimization has been a standard feature in popular commercial algorithms (in the name of "local optimization" and "ideal input resting values" (Qin and Badgwell, 1996)).

Open-loop Performance Objective vs. Closed Loop Performance

In receding horizon control only the first of the computed control moves is implemented; the remaining ones are discarded. Therefore the sequence of actually implemented control moves may differ significantly from the sequence of control moves calculated at a particular time step. Consequently the finite horizon objective which is minimized may have only a tentative connection with the value of the objective function as it is obtained when the control moves are implemented. As mentioned above, it is even conceivable that the sequence of calculated control moves leads the system outside the feasible region. When both input and output horizons are infinite, there is no difference between the sequence determined at a time step and the implemented sequence. As the control horizon is lengthened we should expect the difference to diminish. A measure introduced by Primbs and Nevistić (1997) quantifies this difference and can be used to decide on the horizon length.
By choosing the output horizon long relative to the input horizon short-sighted control policies are avoided but the mismatch criticized above is not eliminated. Thus, it was proposed to set both horizons to infinity which also reduces the number of tuning parameters to be selected (Scokaert and Rawlings, 1997). The computational effort increases but apparently not unduly.

Research Issues

A major problem is the stability analysis of constrained finite horizon systems. The computations suggested by Primbs and Nevistić (1997) are rather difficult except when the state dimension is low.
It was proven (Scokaert *et al.*, 1997) that if an exponentially converging observer is combined with a stable MPC algorithm where access to all the states is assumed, then this observer-controller system is stable, though the controller is nonlinear and the separation principle obviously does not hold. A Kalman filter could serve as the observer. Guidelines for selecting the noise/tuning parameters and efficient implementation schemes were discussed by Lee *et al.* (1994).
In all these deterministic formulations "certainty equivalence" was assumed tacitly. It has been argued (Rawlings and Muske, 1994) that performance gains could be achieved by accounting more accurately for the characteristics of this nonlinear stochastic system. It is unclear how much could be gained from tackling this difficult theoretical problem.

NONLINEAR MPC

The same receding horizon idea which we discussed in detail above is also the principle underlying nonlinear MPC, with the exception that the model describing the process dynamics is nonlinear. Various model forms (differential equations, differential - algebraic equations, discrete time algebraic descriptions, Wiener models, neural nets, etc.) have been tried and some specific theoretical results for some of them are available (Patwardhan *et al.*, 1990; Li and Biegler, 1988; Bhat and McAvoy, 1990; Koulouris, 1995; Maner *et al.*, 1996; Eskinat *et al.*, 1991; Norquay *et al.*, 1996; Hernandez, 1992; Tulleken, 1993). Also see Bequette (1991) for a review on nonlinear process control, which includes an extensive list of different methods for solving nonlinear model predictive control problems. Not to be led astray by these specifics, we will focus on general issues common to all nonlinear MPC algorithms independent of the model form. We will also not go into a discussion of continuous vs. discrete time which can bring up a wealth of hairy technicalities but no new concepts.
Contrary to the linear case, feasibility, closed-loop stability, and the possible mismatch between the open-loop performance objective and the actual closed loop performance are largely unresolved research issues in nonlinear MPC. An additional difficulty is that the optimization problems to be solved on line are generally nonlinear programs without any redeeming features, which implies that convergence to a global optimum cannot be assured. For the quadratic programs arising in the linear case this is guaranteed. This may appear to be a technicality, but it is not: all stability proofs for the linear case rely critically on the fact that this global optimum is found by the control algorithm.
We will discuss some of the ideas in nonlinear MPC and their implications for the issues listed above. The intention is to summarize, complement and update the excellent survey by Mayne (1995).

Infinite horizon / terminal constraint

The idea of using infinite prediction and control horizons or, alternatively, to set up the optimization problem to force the state to zero at the end of the prediction horizon was analyzed by Keerthi and Gilbert (1988) for the discrete time and by Mayne and Michalska (1990) for the continuous time case. Just as outlined for the linear case in the proof the value function is employed as a Lyapunov function. A global optimum must be found at each time step to guarantee stability. When the horizon is infinity, feasibility at a particular time step implies feasibility at all future time steps. Unfortunately, contrary to the linear case, the infinite horizon problem cannot be solved numerically. The optimization problem with terminal constraint can be solved in principle, but equality constraints are computationally very expensive and can only be met asymptotically. In addition, one cannot guarantee convergence to a feasible solution even when a feasible solution exists, a discomforting fact. Furthermore, specifying a terminal constraint which is not met in actual operation is always somewhat artificial. Finally, to reduce the complexity of the optimization problem it is desirable to keep the control horizon small. Thus there may be quite a gap between the open-loop performance objective and the actual closed loop performance.

Variable horizon / hybrid MPC

These techniques were proposed by Michalska and Mayne (1993) to deal with both the global optimality and the feasibility problems, which plague nonlinear MPC with a terminal constraint. Variable horizon MPC also employs a terminal constraint, but the time horizon at the end of which this constraint must be satisfied is itself an optimization variable. In hybrid MPC the terminal constraint is replaced by a "terminal region" which must be reached at the end of a variable horizon. It is assumed that inside this region another controller is employed for which it is somehow known that it asymptotically stabilizes the system. With these modifications a global optimum is no longer needed and feasibility at a particular time step implies feasibility at all future time steps. The terminal constraint is somewhat less artificial here because it may be met in actual operation. However, a variable horizon is inconvenient to handle on-line, an exact end constraint is difficult to satisfy, and the exact determination of the terminal region is all but impossible. In order to show that this region is invariant and that the system is asymptotically stable in this region, usually a global optimization problem needs to be solved.

Contractive MPC

The idea of contractive MPC was mentioned by Yang and Polak (1993), the complete algorithm and stability proof were developed by De Oliveira and Morari (1997). In this approach a constraint is added to the usual formulation which forces the *actual* and not only the *predicted* state to contract at discrete intervals in the future. From this requirement a Lyapunov function can be constructed easily and stability can be established. The stability is independent of the objective function and the convergence of the optimization algorithm as long as a solution is found which satisfies the contraction constraint. Feasibility at future time steps is not necessarily guaranteed unless further assumptions are made. Because the contraction parameter implies a specific speed of convergence, its choice comes natural to the operating personnel.

Quasi-infinite horizon MPC

The technique recently introduced by Chen and Allgöwer (1996) and Chen and Allgöwer (1997*a*) uses an infinite horizon and overcomes both the global optimization and the feasibility problems without making use of artificial terminal constraints, terminal regions and controller switching. Because the infinite horizon costs cannot be evaluated for nonlinear problems, an upper bound is employed, which can be calculated relatively easily and which is minimized by the control algorithm. The open-loop optimal control problem is formulated as

$$\min_{\overline{u}} J\left(x(t), \overline{u}(\cdot)\right)$$

with

$$J\left(x(t), \overline{u}(\cdot)\right) = \int_t^{t+T_p} \left(\|\overline{x}(\tau; x(t), t)\|_Q^2 + \|\overline{u}(\tau)\|_R^2\right) d\tau + \|\overline{x}(t+T_p; x(t), t)\|_P^2$$

subject to

$$\overline{x}(t+T_p; x(t), t) \in \Omega\,, \tag{5}$$

where the penalty term on the final state $\overline{x}(t+T_p)$, the second term in the objective function, is determined to bound the infinite horizon cost:

$$\|\overline{x}(t+T_p; x(t), t)\|_P^2 \geq \int_{t+T_p}^{\infty} \left(\|\overline{x}(\tau; x(t), t)\|_Q^2 + \|\overline{u}(\tau)\|_R^2\right) d\tau \quad \forall \overline{x}(t+T_p; x(t), t) \in \Omega.$$

This bound is established by controlling the nonlinear model fictitiously by linear optimal state feedback within the region Ω after $t+T_p$. The control sequence computed at time k is feasible at all future times and only "improvement" is necessary from time step to time step to guarantee stability.

The method holds much promise. The main unresolved difficulty at this point is the determination of the region Ω which appears to require that some global test is satisfied which again may not be trivial except for academic examples. Recently, a similar technique that removes the need for inequality constraint (5) has been proposed for open-loop stable systems (Chen and Allgower, 1997*b*). The method still requires the Ω region to be defined, however, for determining the terminal weighting matrix and prediction horizon.

MPC with linearization

All the methods discussed so far require a nonlinear program to be solved on-line at each time step. The effort varies somewhat because some methods require only that a feasible (and not necessarily optimal) solution be found or that only an "improvement" be achieved from time step to time step. Nevertheless the effort is usually formidable when compared to the linear case and stopping with a feasible rather than optimal solution can have unpredictable consequences for the performance. The computational effort can be greatly reduced when the system is linearized first in some manner and then the techniques developed for linear systems are employed on-line. Three different approaches have been proposed.

- Nevistić and Morari (1995) apply first feedback linearization and then use MPC in a cascade arrangement for the resulting linear system. The resulting optimization problem is "almost" a Quadratic Program and conditions for global stability can be established. The method is limited to low order systems which fulfill the conditions required for feedback linearization.

- In the first reported industrial approach to nonlinear MPC García (1984) uses at each time step a different *linear* model derived from a local (Jacobian) linearization, and employs standard linear DMC. Lee and Ricker (1994) proposed to add the extended Kalman filter to deal with unstable dynamics and to improve disturbance estimation. De Oliveira (1996) develops this idea further, imposes contraction constraints and derives explicit stability conditions which show the dependence on the quality of the linear approximation and various tuning parameters like the contraction constant.

- Nevistić (1997) shows excellent simulation results when a linear time varying (LTV) system approximation is used which is calculated at each time step over the predicted system trajectory. The time-invariant MPC algorithm can be easily modified to accommodate LTV systems.

Research Issues

This area is wide open for future research and all proposed approaches are little more than initial steps in more or less promising directions. Though the theoretical purists tend to stay away from linearization approaches, linearization is the only method which has found any use in industry beyond demonstration projects. For industry there has to be clear justification for solving nonlinear programs on-line in a dynamic setting and there are no examples to bear that out in a convincing manner. In some sense and with further development quasi-infinite MPC may be "tunable" to use nonlinear MPC only when really needed (far away from equilibrium) and linear MPC otherwise, thus combining the best of the "exact" and the "linearization" methods.

ROBUST MPC

When we say that a control system is robust we mean that stability is maintained and that the performance specifications are met for a specified range of model variations (uncertainty range). To be meaningful, any statement about "robustness" of a particular control algorithm must make reference to a specific uncertainty range as well as specific stability and performance criteria. Although a rich theory has been developed for the robust control of linear systems, very little is known about the robust control of linear systems with constraints.

In the main stream robust control literature "robust performance" is measured by determining the *worst* performance over the specified uncertainty range. In direct extension of this definition it is natural to set up a new "robust" MPC objective where the control action is selected to minimize the worst value the objective function can attain as a function of the uncertain model parameters. This describes the first attempt toward a robust MPC algorithm which was proposed by Campo and Morari (1987). They showed that for FIR models with uncertain coefficients and an ∞-norm objective function the optimization problem which must be solved on-line at each time step is a Linear Program of moderate size. Unfortunately it is well known now that robust stability is not guaranteed with this algorithm (Zheng and Morari, 1993).

Zafiriou (1990) used the contraction principle to derive some necessary and some sufficient conditions for robust stability. The conditions are conservative and difficult to verify. Genceli and Nikolaou (1993) showed how to determine weights such that robust stability can be guaranteed for a set of FIR models. However, weights may not exist even though robust stabilization is possible for a set of FIR models. Also, they assume independent uncertainty bounds on the FIR coefficients which can be very conservative.

The Campo algorithm fails to address the fact that only the first element of the optimal input trajectory is implemented and the whole min-max optimization is repeated at the next time step with a feedback update. In the subsequent optimization, the worst-case parameter values may change because of the feedback update. This is why robust stability cannot be assured as can be easily demonstrated with a counter example.

A true bound on the worst-case cost can be determined when the uncertain parameters are arbitrarily time varying within specified bounds. For this case Lee and Yu (1997) have defined a dynamic programming problem (thus accounting for feedback) to determine the control sequence minimizing the worst case cost. They show that with the horizon set to infinity this procedure guarantees robust stability. However, the approach suffers from the "curse of dimensionality" and the optimization problems at each time step of the dynamic programming problem are usually nonconvex. Thus, in its generality the method is unsuitable for on-line (or even off-line) use except for low order systems with simple uncertainty descriptions.

Most other papers in the literature aim at explicitly or implicitly approximating the problem above by simplifying the objective and uncertainty description, and making the on-line effort more manageable, but still guarantee at least robust stability. For example, Lee and Yu (1997) use a 2-norm and Zheng and Morari (1994) an ∞-norm *open-loop* objective function. Both assume FIR models with uncertain coefficients. A similar technique has also been proposed for state-space systems with bounded input matrix (Lee and Cooley, 1997*b*).

These formulations may be conservative for certain problems leading to sluggish behavior because of two reasons. First of all, arbitrarily time-varying uncertain parameters are usually not a good description of the model uncertainty encountered in practice, where the parameters may be either constant or slowly varying but unknown. Second, the computationally simple open-loop formulations neglect the effect of feedback. **Third, the worst-case error minimization itself may be a conservative formulation for most problems.**

Zheng and Morari (1994) and Zheng (1995) propose to optimize *nominal* rather than robust performance and to achieve robust stability by enforcing a *robust contraction constraint*, i.e., requiring the worst-case prediction of the state to contract. With this formulation robust global asymptotic stability can be guaranteed for a set of linear *time-invariant* stable systems. The optimization problem can be cast as a quadratic program of moderate size for a broad class of uncertainty descriptions.

To account for the effect of feedback Kothare and Morari (1996) propose to calculate at each time step not a sequence of control moves but a state feedback gain matrix which is determined to minimize an upper bound on robust performance. For fairly general uncertainty descriptions, the optimization problem can be expressed as a set of Linear Matrix Inequalities for which efficient solution techniques exist.

Lastly, it is possible to adopt a stochastic uncertainty description (instead of a set-based description) and develop an MPC algorithm that minimizes the expected value of a cost function. In general, the same difficulties that plagued the set-based approach are encountered here. One notable exception is that, when the stochastic parameters are independent sequences, the true closed-loop optimal control problem can be solved analytically using dynamic programming (Lee and Cooley, 1997*a*). In many cases, the expected error may be a more meaningful performance measure than the worst-case error. A contraction constraint can be added to guarantee robust stability for a model set corresponding to a given probability level.

FUTURE – WHAT'S NEEDED?

As we saw in the previous section, the theory of MPC has matured considerably. However, practitioners contend (and rightly so) that what limits the performance and applicability of MPC in practice are not the deficiencies of the control algorithm, but issues like modeling difficulties, lack of suitable sensors, insufficient robustness to failures, etc. MPC points out new needs in these areas and also suggest new approaches: For example, in the past, tasks like fault detection were dealt with at the supervisory level in the form of a "fuzzy" or "knowledge-based" decision maker. As we will point out, there exist now new formulations of MPC involving integer variables, which hold promise for a combined approach to control and diagnosis. Similarly, there is the possibility to include *qualitative knowledge* in a systematic manner in the control decision process.

IMPROVED IDENTIFICATION

Model development is by far the most critical and time-consuming step in implementing a model predictive controller. It is estimated that, in a typical commissioning project, modeling efforts can take up to 90% of the cost and time (Andersen and Kummel, 1992). Quite commonly MPC applications in industry involve dozens of inputs and outputs. To determine such a multivariable model from data puts unprecedented demands on model identification techniques. The conventional steps to arrive at models for MPC applications are illustrated in Figure 1. Each of the steps can be improved greatly, as discussed below:

- **Test Protocol Design**
  Conventionally, models used in MPC applica-

tions are identified through a series of step tests. In some cases, PRBS tests instead of step tests are used and impulse response coefficients are fitted through least squares or through ridge regression (Cutler and Yocum, 1991). In most cases, input channels are perturbed *one at a time*, leading to SISO identification. While this practice is simple and easy to implement, it emphasizes the accuracy of individual SISO models and may not yield a multivariable model of required accuracy. One can easily construct an example where the open-loop responses (either step responses or frequency responses) for all the SISO systems are fitted almost perfectly, but the prediction of the multivariable model when several inputs are changed simultaneously is very poor (Li and Lee, 1996*b*). Implementing a controller designed on the basis of such a model can cause closed-loop instability.

One can experience the same problem with MISO/MIMO identification, as long as perturbations introduced to different input channels are independently designed. This is because, in a highly interactive process, gain directionality of the process causes the responses of output channels to exhibit strong correlations to the point of near colinearity. This can lead to problems like poor signal-to-noise ratios (for low-gain directions) and undesirable distribution of model bias (Andersen and Kummel, 1992).

- **Identification Algorithm**
  In most cases, model fitting is done using SISO or MISO methods. Because the model for each output is fitted separately in these methods, correlations that exist among different outputs cannot be captured or exploited. A true MIMO identification algorithm fits a single model for all outputs simultaneously and accounts for existing correlations. Not only can this improve identification of the deterministic part, but the correlations captured in the form of a stochastic model can also be used in prediction. This can be particularly useful in designing a model predictive control system for quality control, as most quality variables cannot be measured online and must be inferred from secondary process measurements (see Amirthalingam and Lee (1997) for an example application).

- **Model Validation**
  Model validations in most cases amount to examining the prediction errors of individual SISO models with some additional data. As we mentioned earlier, this can lead to misleading conclusions about model quality. SISO models that are very accurate can together constitute a very poor MIMO model. What is needed is a more rigorous model analysis scheme that quantifies the achievable closed-loop performance.

There are results in the literature that provide partial solutions to the above discussed problems. For instance, proposed remedies against the gain directionality problem include: correlated design based on the SVD analysis (Koung and MacGregor, 1994), closed-loop identification (Li and Lee, 1996*a*; Li and Lee, 1996*b*; Jacobsen, 1994), and iterative / adaptive input design (Cooley and Lee, 1996). The recently introduced subspace identification method (Van Overschee and Moor, 1994) may fill the need for a practical MIMO identification algorithm. In addition, several investigators have developed methods to obtain frequency-domain uncertainty bounds, albeit mostly in the SISO context (Goodwin *et al.*, 1992; Wahlberg and Ljung, 1992; Cooley and Lee, 1997). These tools pave the way for *integrated identification and control*, which is depicted in Figure 2 (Cooley and Lee, 1997). The integrated methodology we envision includes: (1) optimal test signal generation based on the collected plant information, closed-loop objectives and plant constraints, (2) quantification of model uncertainty, and (3) rigorous analysis of stability and achievable performance on the basis of the model and its uncertainty. The tools and theories discussed above represent merely a few pieces of the whole puzzle. To realize the concept, new ideas need to be carved out and put together with the existing ones.

PERFORMANCE MONITORING AND DIAGNOSIS

It has been noted by several practitioners that many model predictive controllers perform well when first commissioned, but their performance deteriorates over time leading to eventual shut-downs (Studebaker, 1995). In an industrial setting, maintainability of control systems in the face of various adversities like instrumentation malfunctioning, nonlinearity, parameter variations, etc. is key to long-term success. In order to sustain the intended benefits of model predictive controllers over a long period of time, a mechanism to detect an abnormality and diagnose its root cause is needed. The results can be communicated to engineers and can also be used to adapt control parameters.

Recent publicity of the maintenance problem for industrial control loops has stimulated the research in the area of control system performance monitoring. Thus far, most researchers have concentrated on developing performance measures for existing loops (Stanfelj *et al.*, 1993; Kozub, 1996; Tyler and Morari, 1996*a*; Harris *et al.*, 1995). Very few researchers have examined the problem specifically for model-based control systems. In the model-based control system context, Kesavan and Lee (1997) proposed to monitor the prediction error to detect an abnormality and run a few simple diagnostic tests to gain insights into the source of the abnormal trend.

The problem of fault diagnosis in the model-based setting has been studied by researchers in many disciplines and there is a wealth of literature on the subject (Wilsky, 1976; Isermann, 1984). For instance, with fault states created in the model, it can be viewed as a state estimation problem. It is, however, an unconventional kind in that joint-Gaussian statistics poorly describe the characteristics of most fault signals. Better statistics can be assigned to them using Gaussian-sum models, leading to multiple filter estimation (Tugnait and Haddad, 1979; Kesavan and Lee, 1997).

Some MPC vendors have recognized the importance of self-managing abnormal situations and have launched major R & D efforts on the subject. The next generation of commercial MPC algorithms is sure to be equipped with self-diagnostic features and schemes to manage abnormal situations in an autonomous fashion. However, there is yet to be a consensus on what specific approaches are to be taken. Many believe that a synergistically combined variety of tools (e.g., analytical redundancy, pattern recognition, hardware redundancy) will be needed.

PRACTICAL EXTENSIONS TO NONLINEAR SYSTEMS

In most applications, it is neither technically nor economically feasible to develop detailed first principles models. One of the important factors for MPC's success in industries has been the ability of engineers to construct required models from simple plant tests. Unlike the linear case, however, there is no established method to construct a nonlinear model from input/output data. Recognition of the need has made empirical modeling of nonlinear systems a focal research topic within the research community. Many model forms have been proposed and studied, including simple extensions of a linear FIR model like the Volterra kernel and novel connection structures like the artificial neural networks.

In spite of vigorous research, many fundamental issues remain unresolved in the nonlinear system identification area. One outstanding issue is the model structure determination. The questions regarding the structure determination include: (1) What are the intrinsic differences between various structures like NARX, NARMAX, NMA, Hammerstein, Wiener, etc. and what prior knowledge and plant tests are needed to determine the correct structure? (2) How many delayed input and / or output terms should be included in the model? (3) What basis functions and connection structure should be used? Although general solutions do not appear to be within reach any time soon, there are some promising directions, for example on item (2) (Rhodes and Morari, 1997).

Another difficult issue is the test signal design. Unlike the linear case, conditions for parameter convergence have not been established, except in some special cases. In addition, the need to integrate the closed-loop robustness considerations into the experiment design is even more compelling than in the linear case, since nonlinear system dynamics are much more general and the characteristics of the resulting model are very much shaped by those of the data. A similar approach to the one discussed earlier for linear system identification can be envisioned for nonlinear system identification as well.

Finally, since nonlinear models derived from input-output data will inevitably contain significant bias and variance, the uncertainties need to be quantified and used in the controller design and analysis. The theory for doing this is still at the developmental stage, even for linear systems. However, the need for systematic tools to deal with them is clear in the nonlinear case as insights and heuristics developed for linear controllers do not apply to nonlinear controllers in general.

In terms of practical applications, two approaches seem to be best developed or most in line with the current industrial practice. The first is MPC based on the Volterra kernel, which can be viewed as an immediate high-order extension of the current commercial algorithms. Identification of the Volterra kernel has been well studied and conditions on the input test signals for asymptotic convergence of the parameters under prediction error minimization have been established (Koh and Powers, 1985; Pearson *et al.*, 1993; Pearson *et al.*, 1996). MPC algorithms using second-order Volterra models have also been derived and the properties have been investigated (Doyle III *et al.*, 1995; Maner *et al.*, 1996). A stumbling block for embracing this model type as the choice for general nonlinear control problems is the large number of parameters which explodes with the system's input dimension. Volterra models beyond second order seem impractical. In addition, one must address the problem of large parameter variances, for instance, by quantifying them and accounting for them in the control computation (Genceli and Nikolau, 1994; Chikkula *et al.*, 1993).

The second is the scheduling of multiple linear models within MPC algorithms. The model-scheduling can be done either statically or dynamically, and can be viewed as a form of the popular industrial practice of gain-scheduling (In the model-based control context, it is more appropriate to schedule the model parameters rather than the controller parameters). Model development and scheduling can be performed in a variety of ways, but one systematic way is to identify a *piece-wise linear* model from input output data, for instance, by fitting so called *hinging-hyperplanes* (Breiman, 1993). This model has a nice local linear interpretation and is conducive to *dynamic* scheduling of linear models (Chikkula and Lee, 1995). An approach related to this is to linearly interpolate several a priori constructed models in the state space (Johansen and Foss, 1994; Arkun *et al.*, 1995). The interpolation parameters can be determined *a priori* on the basis of off-line data and prior knowledge (Johansen and Foss, 1994) or can be estimated on-line (Arkun *et al.*, 1995). Kothare *et al.* (1997) "space" the linear models to minimize some measure of modeling error and show that scheduled MPC results in much smoother behavior of the level of a steam generator in a nuclear power plant than many other schemes which have been tried over the years on this problem. At this point, no theory exist, however, which shows under what conditions such scheduled schemes are stable.

CONTROL OF BATCH PROCESSES

Control problems in batch processes are usually posed as tracking problems for time-varying reference trajectories defined over a finite time interval. During the course of a typical batch, process variables swing over wide ranges and process dynamics go through significant changes due to the nonlinearity, making the task of finding an accurate process model very difficult. Because of this, a conventional model-based control system is likely to lead to significant tracking errors. This may explain why there have been so few applications of MPC to batch processes.

A unique aspect of batch operations that must be exploited for tight control is that they are repetitive. Hence, errors in one batch are likely to repeat in the subsequent batches. A framework to use the past batch data along with the real-time data is clearly needed. As a step toward this, Lee and coworkers (Lee and Lee, 1995; Lee and Lee, 1997) took the idea of iterative learning control (popular in robot arm training) and developed an MPC algorithm tailored to the specific needs and characteristics of the batch process control problem. The model used correlates the error trajectory of one batch to the next and includes stochastic components. Previous batches are remembered through state estimation and used in the predictive control computation. The method can also be applied to processes that undergo the same transitions repeatedly. It should be mentioned that the idea of run-to-run learning has also be used in the context of batch optimization (Zafiriou and Zhu, 1990; Zafiriou *et al.*, 1995).

Another largely unexplored aspect of batch system control is quality management. Quality variables can be controlled in a cascade control fashion, i.e., by adjusting the reference trajectories fed to the tracking controllers. However, feedback-based on-line adjustments are often infeasible as most quality variables cannot be measured on-line. The standard industrial practice is to use the statistical monitoring charts (for off-line quality measurements available after the batches) to make adjustments only when significant and prolonged deviations are observed. Not only is this approach ineffective in reducing often-significant batch-to-batch variations, it also results in large amounts of off-spec products due to the delay. A more promising approach is to build a statistical model between the process variables and the quality variables and control the quality variables in an inferential manner. Such an approach has been found to be extremely effective in quality control systems for a pulp digester and a Nylon autoclave (Kesavan *et al.*, 1997).

The above-mentioned concepts and methods need to be tested on practical problems. After some refinements on the basis of practical trials, a separate general software package could be built for batch systems.

MOVING HORIZON ESTIMATION

In most practical problems, states of the system are not directly accessible and must be estimated. The quality of state estimates has important bearings on the overall performance of a model predictive controller, especially of one based on a nonlinear model. Unlike the linear case, however, there is no established method for nonlinear state estimation. The most popular method is the Extended Kalman filter, which simply relinearizes the nonlinear model at each time step and updates the gain matrix and the covariance matrix on the basis of linear filtering theory.

Motivated by the success of MPC, a similar optimization-based state estimation technique has been studied by several investigators recently (Robertson *et al.*, 1994; Michalska and Mayne, 1992). The idea is to formulate the estimation problem within a finite moving window and to find the values of the unknown sequences (e.g., initial condition, state noise, measurement noise) in some least squares sense. Once the unknowns are estimated, the states can be reconstructed using the model. In the linear case with no constraints, it can be shown that moving horizon estimation is equivalent to the Kalman filter for certain choices of weighting matrices (Robertson *et al.*, 1994). A statistical interpretation also exists for the nonlinear case, which suggests the choice of the weighting matrices. The advantage of the formulation is two-fold: First, a nonlinear model can be used directly, at least within the estimation window, which should improve the estimation. Second, constraints can be imposed. Robertson (1996) shows how the constraints can be used to alter the assumed distributions of the unknown sequences, when the strategy is viewed as a Maximum Likelihood or Bayesian estimation method. Michalska and Mayne (1992) establishes the stability of a very restrictive form of moving horizon estimator. In the presence of constraints it appears that, in general, additional somewhat artificial assumptions need to be made to guarantee stability (Tyler and Morari, 1996*c*).

In many nonlinear systems, the conventional certainty equivalence control approach delivers only limited performance. In this sense, it is desirable to compute the probability distribution of the states, albeit approximately, and not just the best estimate. This information should be useful in computing the control input, even though the computation of an optimal input may be intractable. In general, the interaction between estimation and control cannot be ignored and hence the two must be viewed as one problem.

IMPROVED OPTIMIZATION

A demanding feature of most model predictive controllers is that an optimization must be solved online. Depending on the nature of the model and the performance specification, this may be an LP, QP or NLP. Though LPs and QPs are thought to be easy to solve, they can still be computationally demanding for large-scale problems. The NLP is solved in most cases using the Sequential Quadratic Programming (SQP) technique, which is computationally very expensive and comes with no guarantee of convergence to global optimum. For efficiently, many vendors currently solve the QP and LP in a heuristic manner, for example, by using dynamic weighting matrices.

Recently, the so called interior-point methods for solving LPs have been drawing much attention. Originally developed about 15 years ago, reliable public-domain and commercial codes are becoming available nowadays. A remarkable feature of these methods is that, though not proven, they all seem to converge within 5–50 Newton iterations regardless of the problem size (Boyd, 1997), a very attractive feature for on-line use. Moreover, these methods are readily extendible to QPs and SQPs (Wright, 1996; Biegler, 1997). These developments are expected to have major bearings on the future practice of MPC since they will enable the user to solve large-scale problems very efficiently and reliably (without resorting to heuristics and fudge factors which may or may not work).

Another way to increase the efficiency and reliability is to exploit the structure of the problem. The Hessian and gradient of the QPs are highly structured and exploiting this fact has been shown to speed up the computation by orders of magnitude (Biegler, 1997). This may be the key to solving NLPs and large-scale QPs reliably and efficiently. Similar efforts are also under way for highly structured, large-scale LPs (Doyle III *et al.*, 1997).

IMPROVED INTERFACE WITH DIFFERENT LAYERS OF AUTOMATION HIERARCHY

Model predictive controllers are intended to work within the plant operation hierarchy which includes the plant optimizer, low-level PID loops and monitoring schemes. Although the individual components may be well developed, issues regarding their interface have not been investigated much. For instance, in implementing a model predictive controller, one is faced with the option of breaking the low-level loops (and manipulating the valves directly) or retaining them (and manipulating their set points). Both have merits and drawbacks. In addition, the previously discussed performance monitoring / diagnosis scheme must be coordinated with the model-based control system and the process identification scheme to obtain a truly self-sustaining ("adaptive") control system. Finally, there are many issues regarding the interface between the model predictive controller and the plant optimizer. For example, what information from the MPC layer is needed by the optimizer? How (in what form and how often) should the results from the plant optimization be passed onto the MPC layer? Some standards need to be established first, however, before meaningful research can take place in this area.

NEW OPPORTUNITIES BY INCLUDING INTEGER DECISION VARIABLES IN MPC

Integer variables and linear constraints can be used to represent heuristic process knowledge. Any relationship which can be expressed as propositional logic can be translated into this framework (Raman and Grossman, 1992). Apparently, it was not recognized that many possible applications of this approach exist in the area of control and detection (Tyler and Morari, 1996*b*).

In the area of control, by including integer variables representing logic propositions, it is possible to combine logic based control decisions within the MPC framework. This allows innovative control strategies which are capable of prioritizing constraints as well as altering the control objective depending upon the positions of control inputs. By implementing such a strategy, controller performance can be improved. For example, for multivariable systems wherein saturation of one of the manipulated variables prevents all objectives from being met, integer constraints can be used to improve performance and prioritize the objectives.

Integer variables can be used in detection problems to represent the occurrence of symptoms which are indicative of classes of failures. In applications where uncertain models must be used, false alarms due to uncertainty can be reduced by combining quantitative fault estimation with symptom based fault estimation. When residuals are primarily due to modeling uncertainty, the use of logic variables corresponding to symptoms will prevent erroneous fault alarms.

CONCLUSIONS

Over the last decade a mathematically clean formulation of MPC emerged which allows researchers to address problems like feasibility, stability and performance in a rigorous manner. In the nonlinear area a variety of issues remain which are technically complex but have potentially significant practical implications for stability and performance and the computational complexity necessary to achieve them. There have been several innovative proposals how to achieve robustness guarantees but no procedure suitable for an industrial implementation has emerged. While a resolution of the aforementioned issues will undoubtedly change our understanding of MPC and be of high scientific and educational value, it is doubtful that it will have more than a minor effect on the *practice* of MPC. Seemingly peripheral issues like model identification and monitoring and diagnostics will be decisive factors if MPC will or will not be used for a certain application. By generalizing the on-line MPC problem to include integer variables it will be possible to address a number of practical engineering problems directly which may lead to a qualitative change in the type of problems for which MPC is used in industry.

Acknowledgment: JHL gratefully acknowledges the financial support from the NSF NYI Program (CTS #9357827).

REFERENCES

Amirthalingam, R. and J. H. Lee (1997). Subspace identification based inferential control of a continous pulp digester. In: *Proc. PSE'97 / ESCAPE -7*. Trondheim, Norway.

Andersen, H. W. and M. Kummel (1992). Evaluating estimation of gain directionality. part 2: A case study of binary distillation. *Journal of Process Control* **2**(2), 67–86.

Arkun, Y., B. A. Ogunnaike, A. Banarjee and R.K. Pearson (1995). Robust multiple model based control of nonlinear systems. In: *AIChE Annual Meeting*. Miami Beach, FL.

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

Berber, R., Ed.) (1995). *Methods of Model Based Process Control*. Vol. 293 of *NATO ASI Series E: Applied Sciences*. Kluwer Academic Publications. Dortrecht, Netherlands.

Bhat, N. and T.J. McAvoy (1990). Use of neural networks for dynamic modeling and control of chemical process systems. *Comp. and Chem. Eng.* **14**, 573–582.

Biegler, L. T. (1997). Advances in nonlinear programming concepts for process control. In: *IFAC Adchem Conference*. Banff, Canada. pp. 587–598.

Boyd, S. (1997). New advances in convex optimization and control applications. In: *IFAC Adchem Conference*. Banff, Canada.

Breiman, L. (1993). Hinging hyperplanes for regression, classification and function approximation. *IEEE Trans. Information Theory* **39**, 999–1013.

Camacho, E.F. and C. Bordons (1995). *Model Predictive Control in the Process Industry*. Advances in Industrial Control. Springer Verlag.

Campo, P. J. and M. Morari (1987). Robust model predictive control. In: *Proceedings of the American Control Conference*. pp. 1021–1026.

Chen, H. and F. Allgöwer (1996). A quasi-infinite horizon predictive control scheme for constrained nonlinear systems. In: *Proceedings of the 16th Chinese Control Conference*. Qindao, China. pp. 309–316.

Chen, H. and F. Allgöwer (1997*a*). A quasi-infinite horizon nonlinear model predictive control scheme with guaranteed stability. *Automatica*. (in press).

Chen, H. and F. Allgower (1997*b*). A quasi-infinite horizon nonlinear predictive control scheme for stable systems. In: *IFAC Adchem Conference*. Banff, Canada. pp. 471–476.

Chikkula, Y. and J. H. Lee (1995). Local linear model based identification and control of batch processes. In: *AIChE Annual Meeting*. Miami Beach, FL.

Chikkula, Y., J. H. Lee and B. A. Ogunnaike (1993). Robust model predictive control of nonlinear systems using input-output models. In: *Proc. American Control Conf.*. Seattle, WA. pp. 2205–2209.

Clarke, D. W., C. Mohtadi and P. S. Tuffs (1987*a*). Generalized predictive control–I. The basic algorithm. *Automatica* **23**, 137–148.

Clarke, D. W., C. Mohtadi and P. S. Tuffs (1987*b*). Generalized predictive control–II. Extensions and interpretations. *Automatica* **23**, 149–160.

Clarke, D.W., Ed.) (1994). *Advances in Model-Based Predictive Control*. Oxford University Press.

Cooley, B. L. and J. H. Lee (1996). Experimental design for control-relevant multivariable system identification. In: *AIChE Annual Meeting*. Chicago, IL.

Cooley, B. L. and J. H. Lee (1997). Integrated identification and control. In: *IFAC Adchem Conference*. Banff, Canada. pp. 43–48.

Cutler, C. R. and B. L. Ramaker (1979). Dynamic matrix control– A computer control algorithm. In: *AIChE 86th National Meeting*. Houston, TX.

Cutler, C. R. and B. L. Ramaker (1980). Dynamic matrix control– A computer control algorithm. In: *Joint Automatic Control Conf.*. San Francisco, California.

Cutler, C. R. and F. H. Yocum (1991). Experience with the DMC inverse for identification. In: *Conf. Chemical Process Control (CPC-IV)* (Y. Arkun and W.H. Ray, Eds.). South Padre Island, Texas. pp. 297–318.

De Oliveira, S. L. (1996). Model Predictive Control (MPC) for Constrained Nonlinear Systems. PhD thesis. California Institute of Technology. Pasadena, CA.

De Oliveira, S. L. and M. Morari (1997). Contractive model predictive control for constrained nonlinear systems. *IEEE Transactions on Automatic Control*. (in press).

Doyle III, F. J., B. A. Ogunnaike and R. K. Pearson (1995). Nonlinear model based control using second order Volterra models. *Automatica* **31**, 697–714.

Doyle III, F. J., J. F. Pekny, P. Dave and S. Bose (1997). Specialized programming methods in the model predictive control of large-scale systems. *Comp. and Chem. Engr.* **21**, S847–852.

Eskinat, E., S.H. Johnson and W.L. Luyben (1991). Use of hammerstein models in identification of nonlinear systems. *AIChE J.* **37**, 255–268.

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

García, C. E. and M. Morari (1982). Internal model control 1. A unifying review and some new results. *Ind. Eng. Chem. Process Des. & Dev.* **21**, 308–232.

García, C. E., D. M. Prett and M. Morari (1989). Model predictive control: Theory and practice – A survey. *Automatica* **25**(3), 335–348.

Genceli, H. and M. Nikolaou (1993). Robust stability analysis of constrained L_1−norm model predictive control. *AIChE Journal* **39**(12), 1954–1965.

Genceli, H. and M. Nikolau (1994). Design of robust constrained model predictive controllers with Volterra series. In: *AIChE Annual Meeting*. San Francisco, CA.

Gilbert, E. G. and K. T. Tan (1991). Linear systems with state and control constraints: The theory and application of maximal output admissible sets. *IEEE Trans. Aut. Control* **36**(9), 1008–1020.

Goodwin, G. C., M. Gevers and B. Ninness (1992). Quantifying the error in estimated transfer functions with application to model order selection. *IEEE Transactions on Automatic Control* **37**(7), 913–928.

Harris, T. J., F. Boudreau and J. F. MacGregor (1995). Performance assessment of multivariable feedback controllers. In: *AIChE Annual Meeting*. Miami Beach, FL.

Hernandez, E. (1992). Control of nonlinear systems using input-output information. In: *Ph.D. Thesis*. GIT, Atlanta, GA.

Isermann, R. (1984). Process fault detection based on modeling and estimation methods - a survey. *Automatica* **20**, 387–404.

Jacobsen, E. W. (1994). Identification for control of strongly interactive plants. *AIChE Annual Meeting*, St. Louis, MO.

Johansen, T.A. and B.A. Foss (1994). Identification of nonlinear system structure and parameters using regime decomposition. *Automatica* **30**, 321–326.

Keerthi, S. and E. Gilbert (1988). Optimal infinite-horizon feedback laws for a general class of constrained discrete-time systems: Stability and moving-horizon approximations. *Journal of Optimization Theory and Applications* pp. 265–293.

Kesavan, P. and J. H. Lee (1997). Diagnostic tools for multivariable model-based control systems. *Ind. Eng. Chem. Res.*

Kesavan, P., S. Russell and J. H. Lee (1997). PLS based monitoring and control of product quality in batch processes. In: *To be presented at AIChE Annual Meeting*. Los Angeles, CA.

Koh, T. and E. J. Powers (1985). Second-order Volterra filtering and its speech and signal processing. *IEEE Trans. Acoustics, Speech and Signal Processing* **33**, 1445–1455.

Kothare, K.V., V. Balakrishnan and M. Morari (1996). Robust constrained model predictive control using linear matrix inequalities. *Automatica* **32**(10), 1361–1379.

Kothare, M. V., B. Mettler, M. Morari, P. Bendotti and C.-M. Falinower (1997). Linear parameter varying model predictive control for steam generator level control. *Comp. and Chem. Eng.* **21S**, S861–S866.

Koulouris, A. (1995). Multiresolution learning in nonlinear dynamic process modeling and control. In: *Ph.D. Thesis*. MIT, Cambridge, MA.

Koung, C. W. and J. F. MacGregor (1994). Identification for robust multivariable control: the design of experiments. *Automatica* **30**(10), 1541–1554.

Kozub, D. J. (1996). Monitoring and diagnosis of chemical processes with automated process control. In: *Conf. Chemical Process Control (CPC-V)*. Lake Tahoe, CA.

Kwakernaak, H. and R. Sivan (1972). *Linear Optimal Control Systems*. Wiley-Interscience. New York.

Kwon, W. H. (1994). Advances in predictive control: Theory and application. In: *1st Asian Control Conf.*. Tokyo. (updated in October, 1995).

Kwon, W.H. and A.E. Pearson (1977). A modified quadratic cost problem and feedback stabilization of a linear system. *IEEE Trans. Aut. Control* **22**(5), 838–842.

Lee, J. H. and B. L. Cooley (1997*a*). Optimal feedback control strategies for state-space systems with stochastic parameters. *IEEE Transactions on Automatic Control*. in press.

Lee, J. H. and B. L. Cooley (1997*b*). Stable min-max control for state-space systems with bounded input matrix. In: *Proc. American Control Conf.*. Vol. 5. Alberquerque, NM. pp. 2945–2949.

Lee, J. H. and N. L. Ricker (1994). Extended Kalman Filter Based Nonlinear Model Predictive Control. *Ind. Eng. Chem. Res.* **33**, 1530–1541.

Lee, J. H. and Z. Yu (1997). Worst-case formulations of model predictive control for systems with bounded parameters. *Automatica* **33**(5), 763–781.

Lee, J. H., M. Morari and C. E. García (1994). State space interpretation of model predictive control. *Automatica* **30**(4), 707–717.

Lee, Jay H. (1996). Recent advances in model predictive control and other related areas. In: *Conf. Chemical Process Control (CPC-V)*. Tahoe City, CA.

Lee, K. S. and J. H. Lee (1995). Quadratic performance based iterative learning control of batch and transition systems. In: *AIChE Annual Meeting*. Chicago, IL.

Lee, K. S. and J. H. Lee (1997). Model predictive control for nonlinear batch processes with asymptotically perfect tracking. *Comp. and Chem. Eng.* **21S**, S873–S879.

Li, W. and J. H. Lee (1996*a*). Control relevant identification of ill-conditioned processes. *Comp. and Chem. Engr.* **20**, 1023–1042.

Li, W. and J. H. Lee (1996*b*). Frequency-domain closed-loop identification of multivariable systems for feedback control. *AIChE J.* **20**, 2813–2827.

Li, W. Ch. and L. Biegler (1988). Process control strategies for constrained nonlinear systems. *Ind. Eng. Chem. Res.* **27**, 1421–1433.

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

Martín Sánchez, J.M. and J. Rodellar (1996). *Adaptive Predictive Control*. International Series in Systems and Control Engineering. Prentice Hall.

Mayne, D. and H. Michalska (1990). Receding horizon control of nonlinear systems. *IEEE Trans. Aut. Control* **35**, 814–824.

Mayne, D. Q. (1995). Optimization in model based control. In: 4^{th} *IFAC Symposium on Dynamics and Control of Chemical Reactors, Distillation Columns, and Batch Processes (DYCORD+ '95)* (J. B. Rawlings, Ed.). Danish Automation Society. pp. 229–242.

Michalska, H. and D. Q. Mayne (1992). Moving horizon observers. In: *Proceedings IFAC Symposium on Nonlinear Control Systems Design*. Bordeaux, France. pp. 576–581.

Michalska, H. and D.Q. Mayne (1993). Robust receding horizon control of constrained nonlinear systems. *IEEE Trans. Aut. Control* **38**(11), 1623–1633.

Morari, M. and E. Zafiriou (1989). *Robust Process Control*. Prentice-Hall, Inc.. Englewood Cliffs, N.J.

Nevistić, V. (1997). Constrained control of nonlinear systems. PhD thesis. ETH - Swiss Federal Institute of Technology Zürich.

Nevistić, V. and J. A. Primbs (1997). Finite receding horizon linear quadratic control: A unifying theory for stability and performance analysis. Technical Report CIT-CDS 97-001. California Institute of Technology. Pasadena, CA 91125.

Nevistić, V. and M. Morari (1995). Constrained control of feedback-linearizable systems. In: *Proceedings of the European Control Conference*. Rome, Italy. pp. 1726–1731.

Norquay, S.J., A. Palazoglu and J.A. Ramagnoli (1996). Nonlinear model predictive control of ph neutralization using weiner models. In: *Preprints* 13^{th} *World Congress of IFAC*. Vol. M. San Francisco, CA. pp. 31–36.

Ohshima, M., H. Ohno and I. Hashimoto (1995). Model Predictive Control - Experiences in the university-industry joint projects and stastistics on MPC appliciation in Japan -. In: *Int. Workshop on Predictive and Receding Horizon Control, Seoul, Korea*. Seoul National University. pp. 1–16.

Patwardhan, A.A., J.B. Rawlings and T.F. Edgar (1990). Nonlinear model predictive control. *Chemical Engineering Communications* **87**, 123–.

Pearson, R. K., B. A. Ogunnaike and F. J. Doyle III (1993). Identification of nonlinear input / output models using non-Gaussian input sequences. In: *Proc. American Control Conf.*. San Francisco, CA.

Pearson, R. K., B. A. Ogunnaike and F. J. Doyle III (1996). Identification of structurally constrained second-order Volterra models. *IEEE Trans. Acoustics, Speech and Signal Processing* **44**, 2837–2846.

Polak, E. and T. H. Yang (1993*a*). Moving horizon control of linear systems with input saturation and plant uncertainty–Part 1. Robustness. *International Journal of Control* **58**(3), 613–638.

Polak, E. and T. H. Yang (1993*b*). Moving horizon control of linear systems with input saturation and plant uncertainty–Part 2: Disturbance rejection and tracking. *International Journal of Control* **58**(3), 639–663.

Primbs, J. and V. Nevistić (1997). Constrained finite receding horizon linear quadratic control. Technical Report CIT-CDS 97-002. California Institute of Technology. Pasadena, CA 91125.

Qin, S. Joe and T.A. Badgwell (1996). An overview of industrial predictive control technology. In: *Chemiccal Process Control - V, Assessment and New Directions for Research*. Tahoe City, CA.

Raman, R. and I. E. Grossman (1992). Integration of logic and heuristic knowledge in minlp optimization for process synthesis. *Comp. and Chem. Eng.* **16**(3), 155–171.

Rawlings, J. B. and K. R. Muske (1993). The stability of constrained receding horizon control. *IEEE Trans. Aut. Control* **38**(10), 1512–1516.

Rawlings, J.B., E.S. Meadows and K.R. Muske (1994). Nonlinear model predictive control: A tutorial and survey. In: *Advanced Control of Chemical Processes*. IFAC Symposium. IFAC. Kyoto, Japan. pp. 203–214.

Rhodes, C. and M. Morari (1997). Determining the model order of nonlinear input/output systems. *AIChE Journal*. (in press).

Robert R. Bitmead, Michel Gevers and Vincent Wertz (1990). *Adaptive Optimal Control. The Thinking Man's GPC*. International Series in Systems and Control Engineering. Prentice Hall.

Robertson, D. (1996). Development and statistical interpretation of tools for nonlinear estimation. In: *Ph.D. Thesis*. Auburn University, Auburn AL.

Robertson, D., J. H. Lee and J. Rawlings (1994). A moving horizon based approach for least squares estimation. *AIChE Journal* **42**, 2209–2224.

Scokaert, P. O. M. and J. B. Rawlings (1996*a*). Infinite horizon linear quadratic control with constraints. In: *IFAC '96 World Congress*. San Francisco. pp. 109–114.

Scokaert, P.O.M. and J.B. Rawlings (1996*b*). On infeasibilities in model predictive control. In: *Chemical Process Control - V, Assessment and New Directions for Research*. Tahoe City, CA, USA.

Scokaert, P.O.M. and J.B. Rawlings (1997). Constrained linear quadratic regulation. *IEEE Transactions on Automatic Control*. (in press).

Scokaert, P.O.M., J. B. Rawlings and E. S. Meadows (1997). Discrete time stability with perturbations: Applications to model predictive control. *Automatica* **33**(3), 463–470.

Soeterboek, R. (1992). *Predictive Control - A Unified Approach*. International Series in Systems and Control Engineering. Prentice Hall.

Stanfelj, N., T. E. Marlin and J. F. MacGregor (1993). Monitoring and diagnosis of process control performance. *Ind. Eng. Chem. Res.* **32**, 301–314.

Studebaker, P. (1995). Staying on top of advanced controls. *Control Magazine* **Dec.**, 46–48.

Tugnait, J. K. and A. H. Haddad (1979). A detection-estimation scheme for state estimation in switching environments. *Automatica* **15**, 477–481.

Tulleken, H.J.A.F. (1993). Grey-box modeling and identification using physical knowledge and bayesian techniques. *Automatica* **29**, 285–308.

Tyler, M. L. and M. Morari (1996*a*). Performance monitoring of control systems using likelihood methods. *Automatica* **32**(8), 1145–1162.

Tyler, M.L. and M. Morari (1996*b*). Propositional logic in control and monitoring problems. Technical Report AUT96–15. Automatic Control Lab, ETH - Zurich.

Tyler, M.L. and M. Morari (1996*c*). Stability of constrained moving horizon estimation schemes. Technical Report AUT96–18. Automatic Control Lab, ETH-Zurich.

Van Overschee, P. and B. De Moor (1994). N4SID: subspace algorithm for the identification of combined deterministic-stochastic systems. *Automatica*.

Wahlberg, B. and L. Ljung (1992). Hard frequency domain model error bounds from least-squares like identification techniques. *IEEE Transactions on Automatic Control* **37**(7), 900–912.

Wilsky, A. S. (1976). A survey of design methods for failure detection in dynamical systems. *Automatica* **12**, 601–611.

Wright, S. (1996). Applying new optimization algorithms to model predictive control. In: *Conf. Chemical Process Control (CPC-V)* (C. Garcia and J. Kantor, Eds.). Tahoe City, CA.

Yang, T. H. and E. Polak (1993). Moving horizon control of nonlinear systems with input saturation, disturbances and plant uncertainty. *Int. J. Control* **58**(4), 875–903.

Zafiriou, E. (1990). Robust model predictive control of processes with hard constraints. *Computers & Chemical Engineering* **14**(4/5), 359–371.

Zafiriou, E. and J. M. Zhu (1990). Optimal control of semi-batch processes in the presence of modeling error. In: *Proc. American Control Conf.*. San Diego, CA.

Zafiriou, E., H. W. Chiou and R. A. Adomaitis (1995). Nonlinear model based run-to-run control for rapid thermal processing with unmeasured variable estimation. In: *Proc. Symposium on Process Control, Diagnoistics and Modeling of Semiconductor Manufacturing*. San Diego, CA.

Zheng, A. and M. Morari (1994). Robust control of linear time-varying systems with constraints. In: *Proceedings of the 1994 American Control Conference, Baltimore, MD*. Vol. 3. pp. 2416–2420.

Zheng, A. and M. Morari (1995*a*). Control of linear unstable systems with constraints. In: *Proceedings of the American Control Conference*. Seattle. pp. 3704–3708.

Zheng, Z. Q. (1995). Robust Control of Systems Subject to Constraints. PhD thesis. California Institute of Technology. Pasadena, California.

Zheng, Z. Q. and M. Morari (1993). Robust stability of constrained model predictive control. In: *Proceedings of the 1993 American Control Conference*. Vol. 1. San Francisco, CA. pp. 379–383.

Zheng, Z. Q. and M. Morari (1995*b*). Stability of model predictive control with mixed constraints. *IEEE Trans. Aut. Control* **40**(10), 1818–1823.

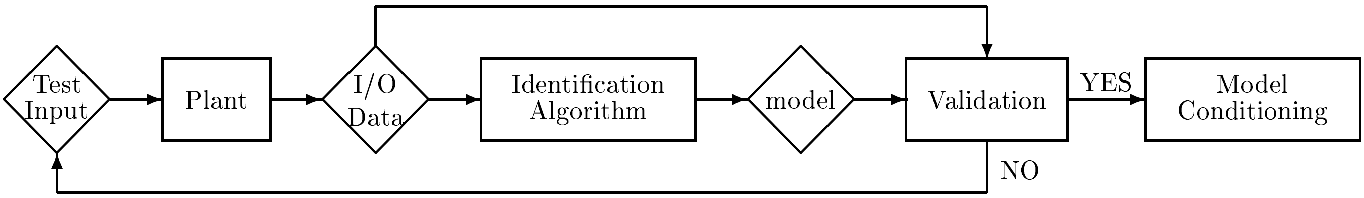

Figure 1: Conventional model identification practice

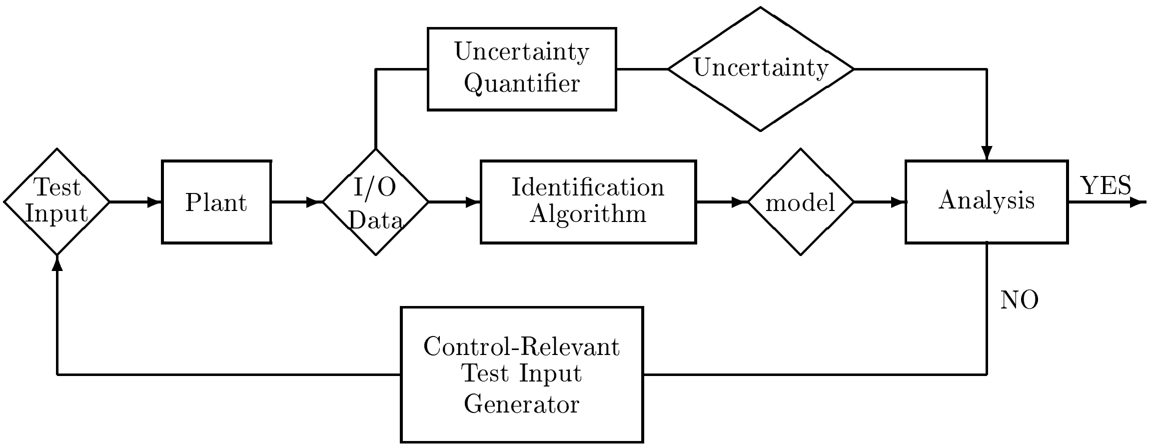

Figure 2: Integrated identification and control methodology