Issues in the Design of a Multirate Model-Based Controller for a Nonlinear Drug Infusion System

Ravi Gopinath,[†,‡] B. Wayne Bequette,[*,†] R. J. Roy,[§] and H. Kaufman[||]

The Howard P. Isermann Department of Chemical Engineering, Biomedical Engineering Department, and Electrical, Computer and Systems Engineering Department, Rensselaer Polytechnic Institute, Troy, New York 12180-3590

Clem Yu

The BOC Group, Murray Hill, New Jersey 17974

Multivariable controller design for the regulation of mean arterial pressure (MAP) and cardiac output (CO) in congestive heart failure patients is restricted by the limited frequency of CO sampling. Performance criteria for the controller specify maximum allowable transient settling times for both variables, and the design should account for the inherent multirate nature of the process in order to satisfy these criteria. We present a multirate model predictive control (MPC) design for MAP and CO regulation by combined infusion of sodium nitroprusside and dopamine, based on a comprehensive nonlinear model of the system. The multirate MPC algorithm is based on nonlinear quadratic dynamic matrix control. To reduce computation time, we introduce a selective linearization technique that linearizes the model on the basis of trends in the plant–model mismatch. The problem is complicated by restrictions on initial dopamine infusion, prescribed to avoid extremely slow responses. We present a novel rule-based override (RBO) to the MPC controller that uses a set of heuristics to initialize dopamine. The performance of the MPC/RBO controller is illustrated using simulation results.

1. Introduction

The regulation of mean arterial pressure (MAP) and cardiac output (CO) in heart attack patients is an important control problem. In a typical intensive care or operating room scenario, these variables are usually maintained at their desired values by the controlled infusion of drugs. Two drugs that are commonly used in practice are sodium nitroprusside and dopamine. A review of the area of hemodynamic control by selective drug infusion has been provided by Yu (1990). The development of advanced control strategies for the preceding drug delivery system requires a descriptive model, and Yu *et al.* (1990) have presented a nonlinear model of the circulatory system describing the effect of the infused drugs on the controlled variables. The drug delivery control problem is multirate due to the limitations on CO measurement frequency. The performance criteria for the closed-loop system specify the maximum allowable settling times for MAP and CO to reach their steady-state values, and if all measurements are slowed down to correspond to the CO sampling rate, the settling time criteria may not be met. This supports the need for an explicit multirate controller design for this system. We present a multirate MPC design for the drug infusion system, based on the nonlinear model, that meets the specified performance criteria.

For linear systems where some outputs are sampled at a slower rate than the others, with no secondary measurements available, Gopinath and Bequette (1991) presented a multirate MPC formulation for single input–single output (SISO) subsets of a multivariable system, based on a linear convolution model. The outputs are estimated at the intersample points using the linear convolution model and control move history, when measurements are not available. It is now well-known that linear MPC can be represented as a special case of linear quadratic (LQ) control and that the state–space representation provides a framework for disturbance prediction *via* state estimation (Prett and García, 1988; Ricker, 1991). The improvement in the disturbance estimates improves regulatory control and also the estimates of the outputs made at intersample points. However, this method is valid only for systems representable in the standard state–space form. While the nonlinear quadratic dynamic matrix control (NL-QDMC; García, 1984) approach used in this paper does employ a linear model in the prediction phase, the temporal nature of the state variables does not permit the model to be written in the standard state–space form. Hence, we use a multivariable extension of the multirate method based on convolution models. This method does not include any disturbance estimation, but the controller is designed to handle the specific constraints of this system and is seen to adequately meet all the performance criteria.

This paper is divided into the following sections. Section 2 provides a design overview, section 3 is used to describe the nonlinear model and issues involved in designing a model-based controller to regulate the drug infusion rates, and section 4 describes the controller design and implementation issues in utilizing a complex

* Author to whom all correspondence should be addressed. E-mail: bequeb@rpi.edu.
† The Howard P. Isermann Department of Chemical Engineering.
‡ Current affiliation: Tata Research, Development and Design Centre, Pune 411001, India.
§ Biomedical Engineering Department.
|| Electrical, Computer and Systems Engineering Department.

8756-7938/95/3011-0318$09.00/0

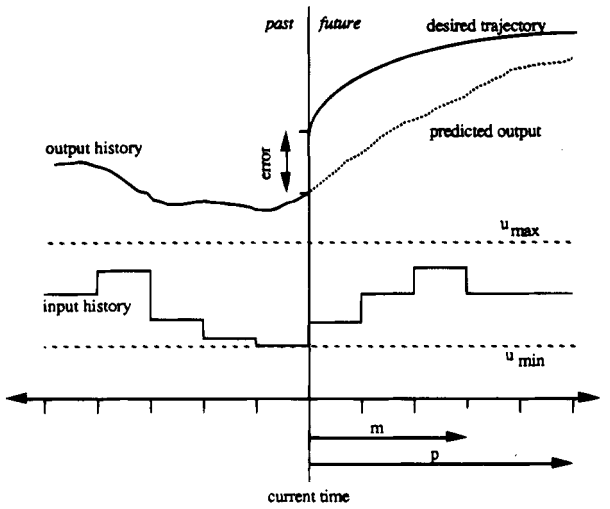

Figure 1. Moving horizon concept in model predictive control.

high-order nonlinear model in the control algorithm. Simulation results and conclusions are presented in sections 5 and 6, respectively. A detailed mathematical description of the nonlinear model is presented in Appendix A, from the point of view of setting up a numerical simulation. Details of the MPC formulation are given in Appendix B.

2. MPC Design Overview

The current MPC class of controllers had their origin in heuristic algorithms developed in the late 1970s (Richalet *et al.*, 1978; Cutler and Ramaker, 1980). These algorithms found widespread application mainly due to their ability to handle multivariable systems with deadtimes and constraints. Recent developments have included a number of modifications to the basic heuristic algorithms, incorporating concepts of LQ design (Prett and García, 1988; Ricker, 1991). However, these methods all share the same design principle, which will be outlined in the following.

MPC design is based on the use of a model of the process in parallel with the plant to compute the predicted output over a certain number of future sample intervals, or "prediction horizon", at each sample point. An optimization problem is then set up at each sample point, where the predicted deviation from the desired output trajectory (error) is minimized over the prediction horizon. The decision variables in the optimization comprise the set of future optimal control moves that could be implemented. A reduced set (usually one) of these moves is then implemented on the plant and the model and the process is repeated, thus setting up a "moving horizon" framework. The MPC concept is shown schematically in Figure 1. The number of future moves computed at each point in the optimization is the "control horizon" and serves as an important tuning parameter.

An important feature of all predictive control algorithms is the ability to handle constraints. The incorporation of constraints in the design is in the optimization step. In dynamic matrix control (DMC), the optimization is a least-squares minimization of the predicted error. García and Morshedi (1986) proposed an extension to this method where a quadratic programming (QP) problem is derived from the least-squares objective function. The main feature of quadratic dynamic matrix control (QDMC), as this method is known, is that process constraints can be incorporated as explicit hard constraints in the QP problem, in contrast to DMC where they are satisfied in a least-squares sense.

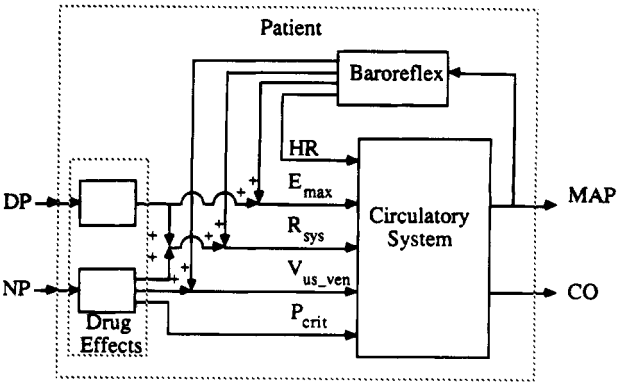

Figure 2. Schematic diagram of the nonlinear drug delivery model.

Since the control move computation (error minimization) is based entirely on the model prediction (if no constraints are active), it is obvious that the accuracy of the model has a significant effect on the ultimate performance of the closed-loop system. Hence, it is relevant to study the open-loop characteristics of the model carefully before designing a controller based on it. The drug delivery system exhibits several interesting characteristics in its modeled response to drug infusions under various initial conditions. In addition, drug infusions for this system are restricted by several constraints that significantly affect the solution to the optimization problem. The impact of these issues on the controller design are discussed in the following section, which describes the model in detail.

3. Model Description

To aid in simulating the effects of sodium nitroprusside and dopamine infusion on heart attack patients, a comprehensive nonlinear model that dynamically describes drug effects on the circulatory system was developed by Yu (1990). A schematic diagram of this model is shown in Figure 2. A broad division of the model yields three sets of descriptive equations: (a) the circulatory system equations, which describe the effect of specific body parameters on the controlled variables MAP and CO; (b) the drug effect relationships, which describe the influence of the infused drugs on the specific body parameters; and (c) the baroreflex model, which describes the effect of the arterial baroreceptors in short-term MAP regulation.

3.1. Modeling Equations. In the following description, note that sodium nitroprusside and nitroprusside are used interchangeably. Yu *et al.* (1990) used an electrical circuit analogy to describe the lumped parameter model of the circulatory system. The forcing function is the time-varying elastance of the heart. The maximum value of this elastance, E_{max}, is used to characterize ventricular contractility. Body compartments and blood vessels are represented as capacitances, and the viscous forces and resistance to blood flow in the systemic and pulmonary vasculature are modeled as resistors. MAP is then the voltage measured after the left ventricle, and CO is the current flow measured at that point in the circuit.

All the circulatory system elements are described in terms of the following (time-varying) body parameters: (a) heart rate (HR) affects the contraction time of the ventricle, which in turn affects the time-varying elastance; (b) maximum elastance (E_{max}), which is used to characterize ventricular contractility; (c) unstressed venous volume ($V_{us\text{-}ven}$), which provides a measure of resistance

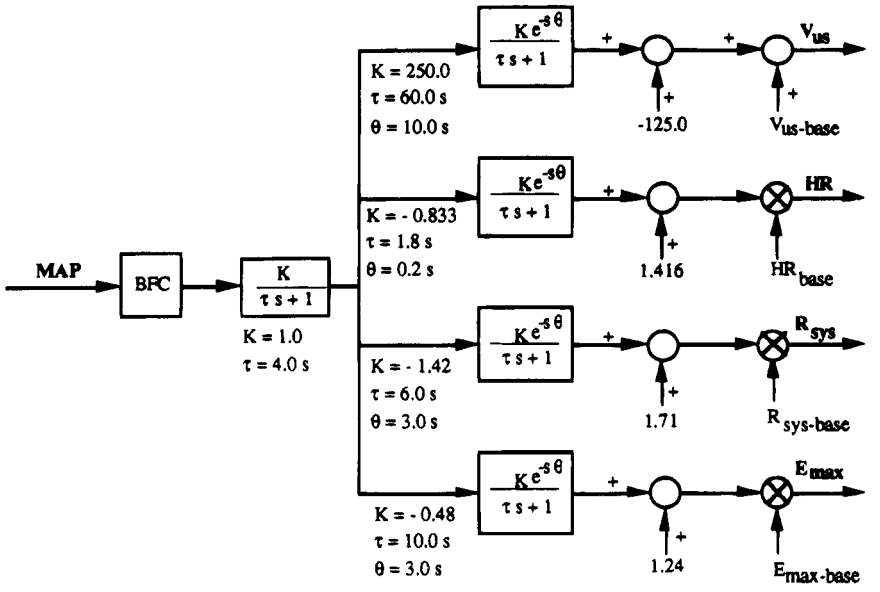

Figure 3. Schematic diagram of the baroreflex model.

to blood flow through the veins; (d) systemic resistance (R_{sys}), which is the resistance to blood flow through the smaller blood vessels; and (e) critical closing pressure (P_{crit}), which is the minimum pressure required to prevent the collapse of blood vessels in pulmonary circulation.

In congestive heart failure (CHF), there are usually a number of factors that cause the symptomatic drop in CO and MAP. One of the causes is a reduction in the effective contractility of the heart. Sodium nitroprusside and dopamine are chosen to increase ventricular contractility and to reduce resistance to blood flow. Dopamine is inotropic in nature, i.e., it increases ventricular contractility and primarily affects E_{max}. Nitroprusside is a vasodilator and reduces resistance to blood flow by reducing R_{sys} and increasing $V_{us\text{-}ven}$. This drug also reduces P_{crit}, increasing the number of blood vessels in the pulmonary circulation and facilitating blood flow through the lungs at lower pressures. The drug infusion thus affects the controlled variables MAP and CO through these body parameters. An increase in dopamine infusion increases MAP and CO, and an increase in nitroprusside infusion reduces MAP but increases CO. The baroreflex model developed by Wesseling *et al.* (1983) describes the effect of the arterial baroreceptors on short-term MAP regulation. The model described here uses a modified version of the baroreflex and is shown schematically in Figure 3. The baroreflex model uses MAP as the input to modify $V_{us\text{-}ven}$, R_{sys}, E_{max}, and HR. Congestive heart failure is modeled by a reduction in E_{max} by 50–70% in the left ventricle. Right ventricular contractility is assumed to be unaffected. The baroreflex gains and time constants are also assumed to remain constant during reduced LV contractility.

The nonlinear model is represented by 38 differential equations. The first set of seven equations is given by flow relationships for each descriptive vessel in the body and are of the form

$$\left(\frac{dV}{dt}\right)_i = Q_{i-} - Q_{i+} \tag{1}$$

where i is the vessel being considered. Q_{i-} represents flow in from the previous vessel and Q_{i+} represents flow out to the next vessel. There are seven different descriptive vessels: the left ventricle, the large arteries, the small arteries, the venous system, the right ventricle, the pulmonary artery, and the pulmonary vein. In addition, the following equation describes the pressure–flow relationships for the large arteries:

$$\left(\frac{dQ}{dt}\right)_{\text{large arteries}} = \frac{P_{\text{large arteries}} - P_{\text{small arteries}}}{L} \tag{2}$$

where L is a constant inertance element. The next set of 20 equations describes the time-dependent concentration of the drugs in the descriptive vessels:

$$\left(\frac{dm}{dt}\right)_{j,i} = (C_d Q)_{j,i-} - (C_d Q)_{j,i+} - (m/\tau_{1/2})_{j,i} \tag{3}$$

where j refers to the drug and i represents the vessel being considered. m is the mass of the drug in the vessel being considered. C_d is the drug concentration in the vessel, and V is the vessel volume ($C_d = (m/V)$. $\tau_{1/2}$ is the half-life of the drug in the vessel. The descriptive vessels are the same as for the flow relationships, but the venous compartments are split into four virtual venous volumes for this computation to induce a delay in the effect of the drugs. The largest blood volume is in the venous chamber, and if the instantaneous mixing assumption is applied to the venous volume considered as one unit, the model predicts an unrealistically fast drug effect.

The third group of five equations is the drug effect relationships, given by

$$\frac{d(\text{Eff})}{dt} = k_1 C_d{}^p(\text{Eff}_{max} - \text{Eff}) - k_2\text{Eff} \tag{4}$$

Eff is the quantitative measure of the effect of a drug on its affected parameter, in the chamber where the effect is assumed to be concentrated. The drug effects on the variables and the relevant chambers in which the drug is assumed to act are given in Table 1. The remaining equations are the five differential equations that describe the first-order transfer functions of the Hammerstein baroreflex model, as shown in Figure 3. The gains and time constants in the baroreflex model are assumed to be time-invariant. A description of the modeling equations and the procedure for solving them are given in Appendix A.

3.2. Dynamic Behavior. The open-loop behavior of the nonlinear model plays an important role in the design

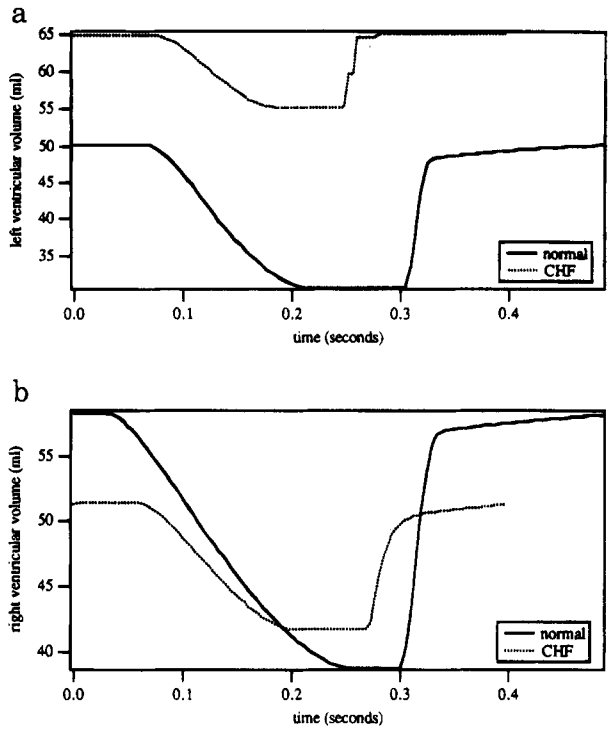

Figure 4. Comparison of ventricular volume variation over one heartbeat under normal conditions and CHF. The systolic volume is significantly lower during CHF, which is indicative of reduced baseline contractility. (a) Left ventricular volume; (b) right ventricular volume.

Table 1. Drug Effects

drug	affected parameters	area of effect
NP	R_{sys}, P_{crit}	small arteries, pulmonary artery
NP	$V_{us\text{-}ven}$	veins
DP	E_{max}	large arteries
DP	R_{sys}	small arteries

and implementation of a model-based controller. This section describes the model dynamics.

3.2.1. Hemodynamics. The 38 state variables can be divided into two sets on the basis of their temporal characteristics: (a) variables that remain constant over a heartbeat (these include the drug masses and the drug effect variables) and (b) variables that change over a heartbeat (these include volumes and blood flow rates in the circulatory compartments and the baroreflex state variables). The drug-related variables are assumed to remain constant over a heartbeat within their respective compartments since the transport is *via* the circulation. Changes in values are assumed to take place with the fresh in flow of blood into the compartment, which in turn coincides with the heartbeat. The volumes and flow rates within the circulatory compartments, however, reflect the changes within a heartbeat. This is illustrated in Figures 4 and 5, which show the open-loop variation in ventricular volume and blood flow rate through the aorta over one heartbeat. The variation is shown for a normal patient and for a patient with CHF. The systolic and diastolic regions of the heartbeat can be seen clearly in the ventricular volume profiles in Figure 4a,b. The onset of CHF is simulated by reducing the baseline contractility of the left ventricle by 60%. Right ventricle contractility is assumed to be unchanged but the effect is seen in both LV and RV profiles as a higher end systolic volume. In the RV profile, the end diastolic volume is actually lower in CHF than in the normal case, but this is due to the modeling approximation that only LV contractility is

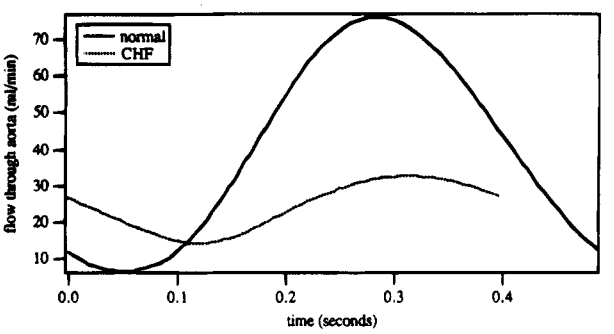

Figure 5. Flow through the aorta over one heartbeat under normal conditions and CHF. The amplitude of the pulse is lower during CHF. The heartbeat is shorter during CHF, mainly due to the action of the baroreceptors.

lowered during CHF. The effect on flow rate through the aorta is also pronounced, and the pulse of blood has a much lower amplitude during CHF, as seen in Figure 5. The volume and flow rate plots span a shorter time period as the heart rate increases during CHF, reducing the duration of each heart beat. The primary symptoms of CHF are a reduction in CO and MAP, and the control objective for this system is to maintain MAP and CO at specified set points and to return the controlled variables to as close to normal stable conditions in the presence of any external disturbance, such as a change in contractility. Controller performance strongly depends on how each drug affects the controlled variables, and this will be discussed in the description of the pharmacodynamics.

3.2.2. Pharmacodynamics. Drug effects on the controlled variables are important in the design and tuning of a controller. Sodium nitroprusside effects are primarily a reduction in MAP and an increase in CO. The dynamics under normal and CHF conditions is slightly different with respect to settling times, with the CHF condition exhibiting the slower response. The dopamine response is more complex and strongly depends on the drug dosage. At higher doses, there is usually some overshoot at high contractilities, and the settling time is in the 10–20 min range. At low infusion rates, the time constants are much larger (50–60 min settling time), and the system exhibits low gain and significant delay. Gingrich and Roy (1991) have analyzed the dopamine hemodynamic response under CHF, and the model predictions for dopamine response closely match their experimental results.

Infusion effects can be clearly analyzed from steady-state plots. The MAP–dopamine and CO–dopamine steady-state plots are shown in Figure 6 under varying degrees of baseline contractility (between 7.0 and 1.5 mmHg/mL). Each of the curves shows the presence of a dead zone at the low infusions. This is reflected in the low gains of the step responses under low dopamine infusions. At the higher contractilities, the MAP and CO steady-state curves exhibit input multiplicities, with maxima near a DP infusion of 3.5 μg/kg/min. This can be explained from a physiological standpoint. Dopamine is modeled as a pure inotropic agent and increases baseline contractility. If the baseline contractility is already high, then the result is an increase in MAP to values above normal, once infusions are greater than those corresponding to the dead zone. As the steady-state MAP increases due to infusion, the arterial baroreceptors cause increased excitation of the vagal center and further inhibit the vasomotor system, which normally causes vasoconstriction. This results in a lowered heart rate and some vasodilation in the systemic vasculature, causing MAP and CO to decrease. At higher drug doses, the

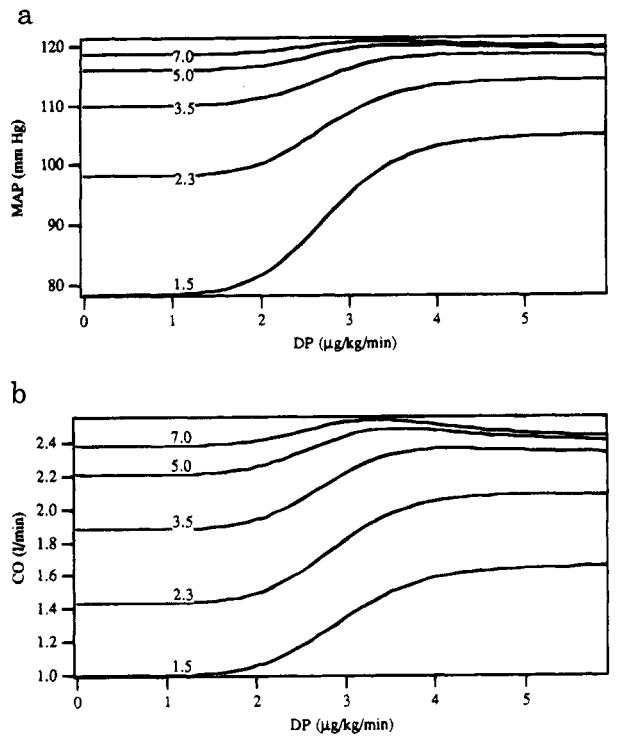

Figure 6. Steady-state response curves for DP infusion as a function of baseline contractility. The numbers on each curve indicate the baseline contractility for the left ventricle in mmHg/mL. (a) MAP steady-state response; (b) CO steady-state response.

baroreflex response is stronger and results in lower steady-state MAP and CO. Once the baseline contractility is low, infusion does not increase MAP above normal. This reduces the effect of the baroreflex, and the steady-state curves at lower contractilities do not exhibit any multiplicities. It should also be noted here that the effect of dopamine is much more pronounced at low contractilities. This is expected, since the drug is inotropic and any effect due to its infusion under normal conditions will be countered by the action of the arterial baroreceptors. The steady-state responses shown here are supported by the experimental results of Gingrich and Roy (1991).

The response to dopamine infusion plays a significant role in the design of the control system, primarily due to the variation in speed of response and the dead zone at low infusions. This factor would play an important role in meeting the controller performance criteria given in section 3.3.1. The high degree of nonlinearity of the system is also seen clearly in the wide range of gains and time constants in the steady-state and transient responses. This underlines the significance of using a rigorous nonlinear model to design the controller. The next section summarizes the control-relevant issues arising from the hemodynamics and pharmacodynamics of the nonlinear model.

3.3. Control-Relevant Issues. There are several issues that affect controller design for this process. Primary among these are the performance criteria. They determine the manner in which the control objectives must be met and govern the tuning of the MPC controller. In addition, there are model characteristics that affect implementation. The effects of low dopamine infusion were introduced earlier and point toward the need for some scheduling. Computation time also has an impact on the applicability of the proposed design. The use of a complex nonlinear model on-line in predictive control is a significant computational load that affects the feasibility of real-time implementation. This is tackled by using a linearized model in the prediction phase. Issues in linearization are discussed in this section.

3.3.1. Performance Criteria. Due to the current technology in CO measurement, which consists of invasive techniques such as the Fick method or the indicator–dilution method, the drug delivery system is multirate in practical application. CO measurements typically are available every 2–3 min, while MAP can be measured as often as every 30 s. The need for an explicit multirate control scheme arises from the strict performance criteria for the controller. The drug delivery controller is responsible for tracking MAP and CO set points that are input by the physician. Since MAP and CO are critical physiological variables, the performance criteria for the controller are based on response time to these set points. If control move implementation is slowed down to match the CO sampling frequency, the performance criteria, summarized here, may not be met: (1) The transient performance criterion for the closed-loop system is a maximum allowable settling time of approximately 10 min for MAP and 20–25 min for CO. (2) The desired transient response must be achieved subject to the following constraints on the manipulated variables (SNP, sodium nitroprusside, DP, dopamine).

$$0 \le \mathrm{DP}(k) \le 7\ \mu\mathrm{g/kg/min} \tag{5a}$$

$$0 \le \mathrm{SNP}(k) \le 10\ \mu\mathrm{g/kg/min} \tag{5b}$$

$$0 \le \Delta\mathrm{SNP}(k) = \Delta\mathrm{DP}(k) \le 0.2\ \mu\mathrm{g/kg/min} \tag{5c}$$

The manipulated variable bounds and velocity constraints are imposed to avoid drug toxicity. In addition, the complications due to low dopamine dosage, as discussed in section 3.2.2, impose additional constraints on the infusion. These constraints and the manner in which they are imposed are discussed later.

3.3.2. Linearizing the Nonlinear Model. In a model predictive control strategy, the model used for control is of primary importance in determining the nature of the output prediction (Bequette, 1991). Use of the full nonlinear model involves integrating the model several times at each time step. For a complex, high-order model such as the one described here, this can be time-consuming, and it is common practice in nonlinear systems to assume that a linear model adequately describes the plant over all future points (García, 1984). In implementation, this involves linearizing the nonlinear model at every sampling instant. Sistu *et al.* (1993) have discussed the computational issues in using a full nonlinear model compared to using a linearized model in the prediction phase. For the drug delivery model, this is not straightforward because of the different time scales associated with the hemodynamic and pharmacodynamic states.

The dynamic state of the model is defined from heartbeat to heartbeat, and steady state is when the volumes and flow rates at the end of each diastole remain constant. Linearization for control is normally performed at every sampling instant. Since there are assumed to be an integral number of heartbeats per sample time, this corresponds to the end of diastole. A linear model obtained in this fashion would adequately reflect the transition of the 38 states from heartbeat to heartbeat and their response to drug infusion. However, this would yield no information on the MAP and CO, since these variables are obtained as functions of the volumes and flow rates, respectively, *integrated over the entire heartbeat*. To describe the outputs in terms of a linear model, the nonlinear model should be linearized at every sub-

interval of the heartbeat. With roughly 60 heartbeats over every sample interval (we use the normal 120 bpm canine heart rate as a measure) and 200 integration intervals per heartbeat, such a linearization procedure may be as time-consuming as using the full nonlinear model for prediction. An alternative is to compute the step response coefficients at each time step. This provides complete input–output information, which is all that is needed in using a linear model for prediction, and solves the problem of obtaining a 38-state linear model.

3.3.3. Scheduling Dopamine Dosage. As seen in section 3.2.2, the dopamine dosage plays a significant role in shaping the response of the model, particularly at low dosages. Accurate scheduling of drug infusion is especially critical when MAP needs to be increased. In the absence of any other drug, MAP can be increased only by the inotropic agent, and it is essential that dopamine infusion be greater than that corresponding to the dead zone. At low infusions, the response may be extremely slow due to very large drug time constants, and the settling time performance criterion may not be met. As suggested by Held and Roy (1993), this could be countered by specifying a given region as a forbidden zone for dopamine infusion. This is also common operating practice, and various heuristics are followed by physicians to determine the span of the forbidden zone. The main factors that govern these heuristics in actual practice are as follows: (1) Current condition of the patient. A measure of the patient's state that is relevant to this problem is the offset of MAP and CO from the desired stable values. (2) Patient sensitivity to the drug. Different patients respond differently at low dopamine dosages, and the physician uses this information to estimate the best dosage at any time. (3) Need for dopaminergic action. Dopamine acts as a vasodilator in the renal vasculature at very low doses. In some patients where the fluid output from the body is very low, the drug may be injected at low doses to increase renal function. The removal of fluids in this fashion may serve to relieve edema in some cases.

The model we use here does not reflect the dopaminergic action, and response time constants are assumed to be the primary factors in scheduling dopamine. In this application, we choose the region between 0 and 4 μg/kg/min as the forbidden zone for initial infusion. Normally the drugs are initialized to zero infusion at startup. The desired dopamine infusion profile is then to step up to 4 μg/kg/min whenever the first infusion is requested. After this, the subsequent infusion can be subject to the other constraints as given in eq 5. The implication of this for the MPC solution and the method for implementation are discussed later. The next section describes the multirate MPC formulation.

4. Controller Design

This section is used to provide an overview of the steps in the MPC algorithm and the implementation issues arising from each step. The details of the MPC formulation are provided in Appendix B.

The MPC algorithm consists of two steps. In the prediction step, the model of the process is propagated over the prediction horizon. The unknowns in this propagation are the future values of the manipulated variables. These values are obtained at each time step through the optimization, where the future moves that minimize the predicted error are computed. In the presence of constraints, the minimization of the predicted error may require several optimization iterations. With a complex nonlinear model, the computational load then becomes significant since the nonlinear model is integrated over a number of intervals at each time step, with each new guess for the decision variable vector (set of future moves). For real-time feasibility, it is important to ensure that the control calculations are completed within the specified sampling interval. To ensure this in the drug delivery system, we use a modification of the nonlinear MPC procedure, as proposed by García (1984). In this method, the system is assumed to be linear at all future points. The prediction then is split into two parts. In the first part the nonlinear model is propagated over the prediction horizon prior to the optimization, with the current input values held constant. In the second part, the nonlinear model is linearized around the current states and inputs. Then the linear model is used in the optimization iterations, with the nonlinear contribution simply added on since it is computed *a priori*. This circumvents the repeated integrations and saves computation time. As discussed in section 3.3.2, linearization of the drug delivery model is nontrivial due to the two time scales, and the determination of a method for obtaining the linear model during control calculations is an important implementation issue.

The second issue in controller design involves the multirate nature of the system. In the MPC procedure, the output prediction has to be corrected at each time step using feedback information. The correction term is the observed mismatch between the plant and model outputs (disturbance variable). In conventional MPC formulations such as DMC, this term is assumed to remain constant over the prediction horizon, since no future measurements are available. In recent formulations of MPC, the variation of the plant–model mismatch over the prediction horizon is estimated by characterizing it as a stochastic variable and computing the Kalman filter gains over the prediction horizon. In a multirate system, the values of all outputs may not be available at a given sample point. The best information available to correct the prediction then is the disturbance term computed at the previous sampled point for a given output. In a DMC-type formulation, the disturbance term at an intersample point for a given output is assumed to be the same as that obtained at the previous sampled point. If the disturbance term is estimated over the prediction horizon as described earlier, then the estimate for the current point can be computed by the same procedure. Once the disturbance term is estimated at the current point, the prediction step is analogous to the single-rate case. The Kalman filter design procedure is valid for systems cast in the standard state–space form. While the NL-QDMC approach used here ensures a linear model over future points at every time step, the linear model used cannot be cast in the standard state–space form, as explained in section 3.3.2. The disturbance estimation procedure thus cannot be used, and we use the DMC assumption of constant additive disturbance to update the CO plant–model mismatch term at its intersample points. This procedure yields the desired closed-loop response specified in eq 5 for various simulation tests. In the most general case, however, it is expected that disturbance estimation would significantly improve regulatory performance.

4.1. Implementation Issues. We now describe the linearization methods used to generate the on-line model and the rule-based scheduling of dopamine infusion to avoid the low infusion effects.

4.1.1. Linearization. In using a nonlinear model, the most rigorous approach is to use the full model in prediction. This necessitates calling the model from the optimization routine, with every guess for the future

control moves, at every iteration. With a multivariable nonlinear model with large prediction horizons, this is computationally intensive. On a test run with the drug delivery system, use of the full nonlinear model took roughly 30 min per MPC iteration. As discussed earlier, an approach based on using the linear model in prediction should be used. Due to the problems in linearization to a state−space form, as discussed in section 3.3.2, a multivariable step response model is generated for use in prediction. Generation of the step response at every time step is computationally expensive as well, since the coefficients would have to be generated by integrating the nonlinear model at every time step over a large enough number of intervals to correspond to the open-loop settling time. We propose, as an alternative, the following procedure of selective linearization that is based on trends in the plant−model mismatch: (1) At every step, the change in the plant−model mismatch term is noted. (2) If the difference in this term grows over more than a specified number of intervals or changes in one step by an amount greater than a specified tolerance, it is assumed that this is either due to the plant moving to a different region based on the current input or due to an unmodeled disturbance entering the system. A flag is set to indicate this. (3) If the flag in (2) is set, the step responses are computed.

This procedure thus calls for generation of the step response model only when the current step response model is seen to be inadequate in tracking the plant. The computational load is reduced significantly as a result. In all the simulations reported here, this method was used, and the maximum computational time in an iteration (including the linearization step) was 27 s. This is within real time, since the base sampling interval for the drug delivery system is chosen to be 30 s.

4.1.2. Rule-Based Override Option. To achieve the desired scheduling of dopamine infusion around the forbidden zone, we propose a rule-based override (RBO) to the MPC controller that alters the infusion on the basis of the current action requested by the MPC controller. It must be emphasized here that the override is *not* a controller, since its outputs are not influenced by the current values of the controlled variables. In future work, this option could be extended into a complete fuzzy logic controller that is activated only in the forbidden zone.

At the beginning of the simulation, both drugs are initialized to zero. When set point changes are requested, the MPC controller requests the infusion of dopamine, if MAP or CO has to be increased. Velocity constraints are imposed as in eq 5, so that the first requested action will usually be in the forbidden zone. Even if the velocity constraints are removed within the forbidden zone, the requested action depends on the tuning of the MPC controller, and it requires extremely tight tuning to make a large initial step out of this region. As is well-known, such tuning adversely affects the robustness of the controller. Instead, the initial infusion for dopamine is determined by the rule-based override as follows: (1) Waiting periods are specified for control actions in the up and down directions. These are denoted as iu and id. The initial requests for dopamine infusion are monitored, and if these exceed iu in number, dopamine is stepped up to 4 μg/kg/min. (2) Dopamine is held at an infusion of 4 μg/kg/min for at least id intervals. After this, control is passed completely to the MPC controller.

The concept of a forbidden zone suggests that *any* infusion in this range must be avoided. It must be recognized here, however, that the limits of this zone are determined rather arbitrarily. The steady-state solution to a given set point change or disturbance rejection may actually require a dopamine infusion that is within this region, and constraining it to be either zero or greater than 4.0 may result in some steady states being unreachable. The rule-based override thus is used only to avoid initial low doses that may result in extremely slow response. The important tuning parameters for the RBO are the waiting periods iu and id. These should be adjusted on the basis of the observed response and/or the requested set point. If the measurements are very noisy, it might be necessary to desensitize the RBO to noise effects to reduce chattering in the response. Once the dopamine infusion has been initialized at 4.0 μg/kg/min, the subsequent control actions are determined by the MPC controller and are affected only by the MPC tuning and the constraints (eq 5).

4.2. Controller Tuning. (1) The prediction horizon p is chosen on the basis of the open-loop settling time. The open-loop response of the output to each input is generated, and p is set to the number of intervals corresponding to the settling time for the slowest input−output transient. (2) The control horizon m is used to tighten or detune the controller. In general, larger values of m for an input will result in more aggressive action. This yields faster response, but the closed-loop system is less robust to disturbances. m is chosen on the basis of the allowed trade-off. (3) The output weighting matrix $\mathbf{Q}$ is a diagonal matrix used to assign weights to the components of the error function, corresponding to each output, in the optimization step. A larger weight for an output will result in tighter control. (4) The input penalty matrix $\mathbf{R}$ is also a detuning parameter and is used to penalize control action in the objective function. This parameter is especially useful when a large m is used. The effects of these tuning parameters on closed-loop response have been discussed at length in the MPC literature (Maurath *et al.*, 1988; García *et al.*, 1989). For the multirate drug delivery system, the following tuning parameters were used.

controlled variable	prediction horizon	weight
MAP	15	1.0
CO	20	20.0
manipulated variable	**control horizon**	**penalty**
SNP	3	100.0
DP	5	20.0

In addition, the application-specific tuning parameters used were in the selective linearization and in the RBO. In selective linearization, the number of intervals over which the plant−model mismatch is allowed to grow or reduce before the model is relinearized was 5. The tolerance for maximum allowable change in the plant−model mismatch over one interval was 10mmHg for MAP and 0.25 min^{-1} for CO. For rule-based override, the waiting period in the up direction (iu) was 3. The waiting period in the down direction was varied from case to case and is reported in the discussion of simulation results.

5. Simulation Results and Discussion

The simulation results in this section illustrate the performance of the multirate MPC controller in servo and regulatory mode. Simulations were performed on an IBM RS/6000 Model 550 with 256M memory, running AIX Version 3.2.4. The maximum CPU time per MPC iteration was 27 s. For similar applications reported in the literature, this lies within real time [Voss *et al.* (1987) used a 40−60 s sample interval, while Yu *et al.* (1990) used a 30 s sample time]. The GRG2 package (Lasdon

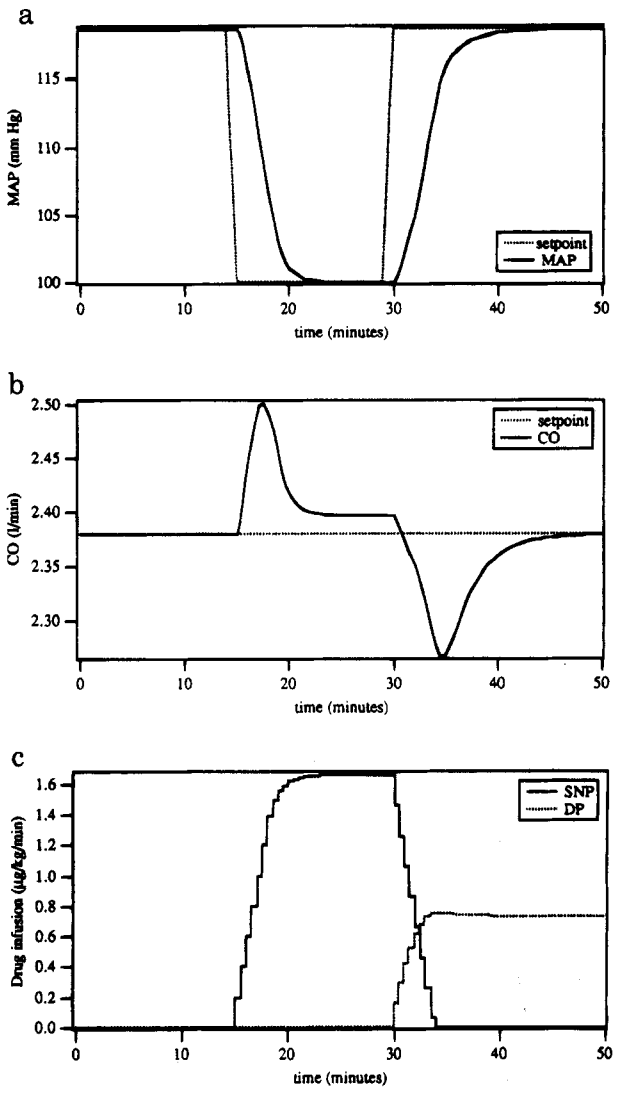

Figure 7. Closed-loop response to MAP set point changes under normal conditions. The MAP and CO responses are seen to meet the settling time performance criteria. The intermediate offset in CO is because DP is constrained at its lower bound. (a) MAP profile; (b) CO profile; (c) drug infusion profile.

et al., 1978) was used to solve the optimization problem. In the multirate framework, the MAP sampling rate is assumed to be 5 times that of CO. All control moves are made at the base sampling rate. The following cases were simulated: (1) Normal patient, set point changes in MAP only; (2) patient with CHF, set point change in CO with a change in the plant operating region; and (3) patient with CHF, set point changes in MAP and CO. The model was initialized with the appropriate data sets for each case (C. M. Held, personal communication, 1993), as described in Appendix A. Unless stated otherwise, the model description of the plant is assumed to be perfect.

5.1. Case 1. Normal Patient, MAP Set Point Changes. The closed-loop response under multirate MPC without the rule-based override is shown in Figure 7. The plant and model are initialized with normal condition parameters, and the drug concentrations are initialized to zero. At $t = 0$, the nonlinear model is linearized by computing the open-loop step response. A dopamine step of 4.0 is used to compute the step response, and the coefficients are normalized according to this. Further linearization would follow the logic in section 4.1.1. A set point reduction is requested in MAP at $t = 15$. This is possible only by nitroprusside infusion, and the expected response is seen. MAP settles to the set point within 10 min, which satisfies the settling time performance criterion. CO is seen to settle with offset, since dopamine is constrained at its lower bound. At $t = 30$, the set point is increased to the initial value. Two control actions accompany this: nitroprusside infusion decreases, which lowers CO below its steady-state value. This is countered by dopamine infusion that serves to increase MAP and CO back to the set point. The settling time in both profiles satisfies the performance criterion. An important observation is the steady-state value of dopamine (0.72 μg/kg/min). This is in the low end of the forbidden zone under normal conditions, but still results in the desired response. This illustrates that the final steady-state infusion for a given set point change may actually lie within the 0−4 μg/kg/min infusion range. We use this as a basis for handing control back to MPC after the waiting period in the implementation of the RBO.

5.2. Case 2. Patient with CHF, CO Set Point Change. This case is used to study the disturbance rejection properties of the controller in regulating MAP and also its performance, with model uncertainty introduced. Transient responses are shown in Figure 8. The plant and model are initialized with CHF parameters, and drug infusions are set to zero initially. We compare three cases here: single-rate MPC (SR), multirate MPC (MR), and multirate MPC with the rule-based override (RBO). The control objective is to maintain MAP at its set point of 98.2 mmHg, while a set point change is requested in CO from 1.43 to 2.00 min^{-1}. Single-rate MPC was implemented with the common sampling period now set to 2.5 min (to correspond to CO sampling frequency). The rule-based override was implemented with iu = 3. id = 3 was used for this case, on the basis of the observed distances of MAP and CO from their respective set points. In the extension of this concept to a complete fuzzy logic override controller, the selection of the waiting periods could be determined by the controller, on the basis of current offset. For the single-rate case, the horizons were based on the larger sampling time, and adjusted to represent roughly the same prediction time as the multirate case. The prediction and control horizons used were $p(\text{MAP}) = 3$, $p(\text{CO}) = 4$, $m(\text{SNP}) = 1$, and $m(\text{DP}) = 3$. The CO weight function was increased and the dopamine penalty was reduced from the multirate values to get the fastest single-rate response to the requested set point change.

A CO set point change is requested at $t = 5.0$. The MPC controller requests an increase in the infusion of both drugs. The infusion in the single-rate case is lower for both drugs, which is expected as a result of the detuning effect of a larger sample time. The MR-MPC control action is subject to the velocity constraints on infusion and several steps are made in the forbidden zone. With the RBO, the MPC output is held at 0 for three intervals and then stepped up to 4.0 μg/kg/min. The effect of this is seen immediately in the MAP response time. The single-rate response is extremely slow and is entirely unsatisfactory. The multirate response is faster, but still does not satisfy the settling time criterion. The lower initial doses result in a smaller peak, but much slower settling. The RBO response satisfies the settling time criterion. The slight overshoot is due to the chosen value of id. With a smaller value, the control would have switched to MPC at an earlier point, eliminating the overshoot. The CO responses show the same trend. The single-rate response is extremely slow, while the multirate MPC still shows some deviation from set point 1 h after the set point change. The benefits of using the RBO to initialize dopamine infusion are obvious from this simulation.

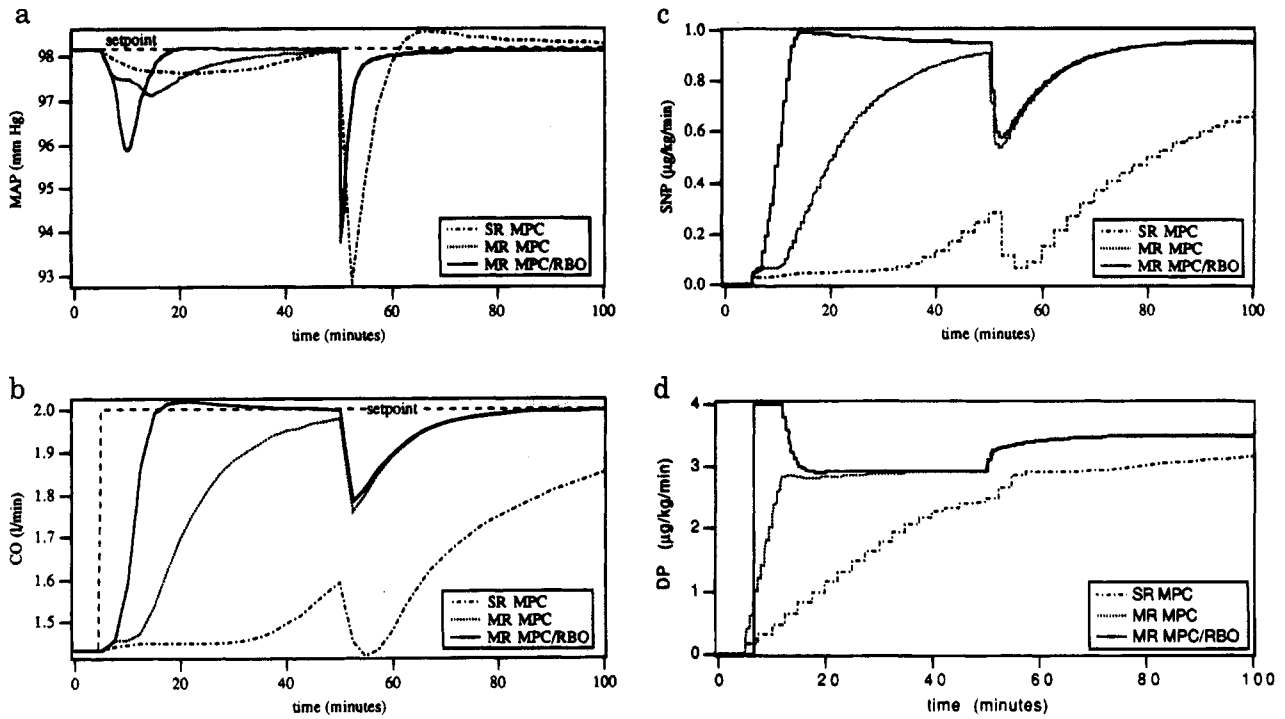

Figure 8. Comparison of various MPC implementation methods. Transient responses to CO set point change under CHF with a sudden change in plant parameters introduced at $t = 50$. (a) MAP profile; (b) CO profile; (c) SNP infusion profile; (d) DP infusion profile.

At $t = 50$, an uncertainty is introduced by reducing the LV baseline contractility, in the plant only, from 2.33 to 1.9. This results in a sudden drop in MAP and CO. Nitroprusside infusion is immediately reduced to bring MAP rapidly back to set point. Further, dopamine infusion also increases, which raises MAP and CO back to set point. The responses of the MR-MPC and the RBO are similar in terms of response time, since the RBO had handed control back to MPC and the drug infusions had nearly stabilized, at the time of the reduction in plant $E_{\max\ base}$. The response times for both MAP and CO satisfy the settling time criterion. Again, the significant difference is with the single-rate strategy, which shows the expected slow response time. These results underline the importance of designing an effective multirate control system for combined MAP and CO regulation, and the proposed design illustrates a feasible method, depending on the availability of a reliable model.

5.3. Case 3. Patient with CHF, MAP and CO Set Point Changes. The previous simulation established the performance deterioration as a result of slowing down all sampling frequencies to correspond to that of CO. We do not show further single-rate results, and conclusions regarding the controller design will be based on the performance of the multirate controller. The next simulation compares the MR-MPC with MR/RBO for set point changes in MAP and CO. The closed-loop response is shown in Figure 9. A perfect model is assumed in this case, and the CHF parameters are used to initialize the plant and model, with zero drug infusion. At $t = 5.0$, a set point change in MAP occurred from 98.2 to 110 mmHg and a change in CO from 1.43 to 2.1 min^{-1} occurred at the same time. The RBO was implemented with id = 10, since the final desired value of CO is higher. After an initial wait of three intervals, the RBO increases dopamine infusion to 4 μg/kg/min and holds it there until a further increase to the final value is requested by the MPC controller. The nitroprusside infusion in the earlier stage serves to increase CO. The effect on MAP is countered by the dopamine infusion. Once CO has been brought reasonably close to its set point, the nitroprusside infusion levels off. The final increase in both MAP and CO to their set points is brought about by dopamine infusion. The improvement in the speed of response using the RBO is clearly seen in the transients. Differences of roughly 15 min in MAP and roughly 30 min in CO settling time are seen. The reason is that the entire range of infusion that corresponds to a slow response is avoided (small steps from zero infusion). Once the RBO initializes dopamine to the specified level (4.0 μg/kg/min in this case), further adjustments can be made on the basis of the MPC output.

6. Conclusions

Control system design plays a very important role in the successful implementation of an automatic drug infusion system, the objective being to achieve tight control, as determined by various performance criteria. We have presented a control system design for the combined mean arterial pressure–cardiac output regulation problem that satisfactorily meets the specified control objectives and performance criteria. The multirate controller design allows MAP and CO to be sampled at different rates, as in actual practice, clearly demonstrating the need for explicit multirate controller design when compared with the single-rate case. The success of any model-based controller design critically depends on the accuracy of the descriptive model. The rigorous nonlinear model used in this work has previously been shown to accurately predict system response (Yu *et al.*, 1990), as determined by the limits on the drug infusions. Further aspects of the open-loop behavior of the modeled system, from both the hemodynamic and pharmacodynamic points of view, were discussed in this paper. Open-loop response is important from the point of view of controller design, as it provides a valuable guideline in tuning the controller.

An important implementation issue was addressed as a direct consequence of the low-dose open-loop response predicted by the model. Dopamine infusion is often

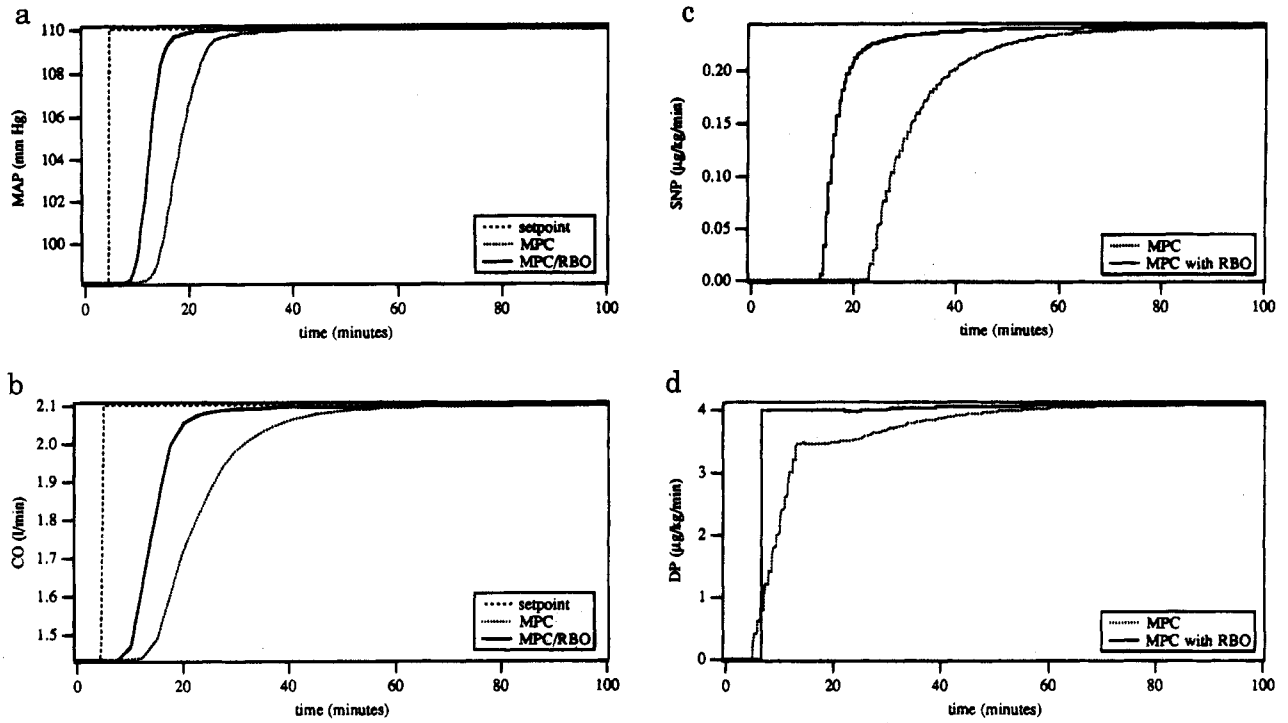

Figure 9. Transint response to MAP and CO set point changes under CHF. A perfect model is used throughout the simulation. The waiting periods in the RBO are adjusted to get the best initial DP infusion. (a) MAP profile; (b) CO profile; (c) SNP infusion profile; (d) DP infusion profile.

scheduled in actual practice, but the scheduling methods vary widely from one practitioner to the next. A novel method for dopamine scheduling that results in the desired drug response was presented. The scheduling was achieved *via* a rule-based override on the MPC controller output, formulated to handle initialization from zero infusion. The initialization avoids initial infusions in a specified forbidden zone, thus preventing the slow responses that characterize low infusion rate response during CHF. This results in an overall superior response, as determined by the settling time performance criteria. The parameters in the RBO were seen to play an important role in the performance of the combined MR/RBO controller, and it is expected that these would be important tuning parameters for a fuzzy logic controller based on the RBO concept. Future work in this area includes the development of a fuzzy logic controller specifically for dopamine initialization that takes into account current-controlled variable offset, in addition to the MPC output which is used by the RBO.

Computational issues are important, especially from the point of view of real-time implementation of the design. The simulations showed that the use of a selectively linearized model in the prediction phase results in significant computational savings. Further efficiency may be achieved by reformulating the current model as a lower order system. We feel that a compartmental model with a lower number of compartments could be used to describe the drug concentrations and drug effect changes and would still predict the dynamics adequately. The resulting reduced order model will reduce computation time considerably. Further, the reduced order may permit the use of the full nonlinear model even during prediction, with the added accuracy resulting in superior controller performance. This option will be addressed in future work. In the current design, however, the multirate MPC controller using the full nonlinear model with a selectively linearized prediction model is seen to run within real time, based on commonly used sample times.

Table 2. Initial Data—Nonlinear Drug Delivery Model

parameter	normal	CHF
$E_{\max \text{ base LV}}$ (mmHg/mL)	7.0	2.333
V_{LV} (mL)	49.813	64.881
$V_{\text{large arteries}}$ (mL)	120.022	112.503
$V_{\text{small arteries}}$ (mL)	125.329	124.30
V_{ven} (mL)	908.138	813.213
V_{RV} (mL)	58.18	50.167
V_{pa} (mL)	24.428	40.172
V_{pv} (mL)	64.097	144.779
$Q_{\text{large arteries}}$ (mL/min)	12.77	25.687
bfc	0.4961	0.22146
MAP (mmHg)	119.6026	99.926
CO (min^{-1})	2.3853	1.4442
MPAP (mmHg)	13.038	26.489
MPVP (mmHg)	4.626	23.262
SV_{LV} (mL)	19.6511	9.7688
SV_{RV} (mL)	19.6511	9.7687

Acknowledgment

This work was partly carried out under NSF Grant BCS-9005678. We thank Claudio Held for initial data used in the simulations shown here.

Appendix A: Nonlinear Modeling Equations and Solution Procedure

The nonlinear modeling equations for the drug delivery system are given below. The equations are presented sequentially and describe the steps to be followed in setting up a time-domain simulation. Data for this section may be found in Tables 2–5.

I. Initialization. As described in section 3.1, the circulatory system is divided into seven compartments. Blood is assumed to flow from one vessel to another in the following sequence: left ventricle → large arteries → small arteries → veins → right ventricle → pulmonary arteries → pulmonary veins → left ventricle. At the beginning of the simulation, the following variables are initialized: volumes in each of the seven compartments, V; flow in the large arteries (from the left ventricle),

Table 3. Ventricular Parameters (Values for an 18 kg Dog)

parameter	left ventricle	right ventricle
A (mL^{-1})	0.145	0.145
P_M (mmHg)	2.0	2.5
E_D (mmHg/mL)	0.09817	0.06544
V_d (mL)	5.0	12.0
R_{src} (mL)	0.0015	0.0045

Table 4. Circulatory Parameters–Nominal Values (Values for an 18 kg Dog)

parameter	nominal value
R_{sysc} (mmHg/(mL·s)$^{-1}$)	0.125
R_{ven} (mmHg/(mL·s^{-1})	0.00125
R_{pv} (mmHg/(mL·s)$^{-1}$)	0.0025
R_{pul} (mmHg/(mL·s)$^{-1}$)	0.013406
$C_{large\ arteries}$ (mL/mmHg)	0.6
$C_{small\ arteries}$ (mL/mmHg)	0.0787
C_{ven} (mL/mmHg)	30.0
C_{pa} (mL/mmHg)	1.0
C_{pv} (mL/mmHg)	4.2

$Q_{large\ arteries}$; BFC parameter to the baroreflex, BFC; initial LV contractility, $E_{max\ base\ LV}$; initial pressures in the systemic and pulmonary circulations, mean arterial pressure (MAP), mean pulmonary arterial pressure (MPAP), and mean pulmonary venous pressure MPVP; initial cardiac output, CO; initial stroke volumes for each ventricle, SV.

Values for these parameters for normal and CHF conditions are given in Table 2. The following ventricular parameters are also initialized. Typical values for an 18 kg dog are given in Table 3. Stiffness parameter, A; limit of linear region in the ventricular P–V relationship, P_M; linear elastance parameter, E_D; dead volume of the ventricle, V_d; characteristic resistance of the ventricle, R_{src}.

In addition, the following ventricular parameters are computed from the preceding ones:

$$V_M = V_{us} + P_M/E_D \tag{A1}$$

$$C_v = \frac{E_D}{A \exp(AV_M)} \tag{A2}$$

$$B_v = C_v \exp(AV_M) - P_M \tag{A3}$$

V_{us} is the unstressed ventricular volume (read in initial data) and V_M is the volume corresponding to P_M. C_v and B_v are derived constants.

Next the circulatory parameters are initialized. These constitute characteristic resistances of vessels directly entering and leaving the heart and the compliances of the systemic and pulmonary vasculature. Values are given in Table 4. Characteristic resistance for flow leaving LV, R_{sysc}; characteristic resistance for flow entering RV, R_{ven}; characterisic resistance for flow leaving RV, R_{pulc}; characterisic resistance for flow entering LV, R_{pv}; compliance of each systemic and pulmonary vessel, C.

The drug parameters are initialized next. For sodium nitroprusside and dopamine, typical values are given in Table 5. Maximum effect of the drug, Eff_{max}; power to which concentration is raised in the drug effect equation, P; drug reaction rate constant, k_2; infusion corresponding to 50% drug effect, i50; drug half-life, $\tau_{1/2}$.

The drug time constant is defined as

$$\tau = \frac{\tau_{1/2}}{\ln(2.0)} \tag{A4}$$

Then the second reaction rate constant is computed as follows. For NP → R_{sys} and NP → $V_{us\text{-}ven}$,

$$x = (i50)\tau/85.0 \tag{A5}$$

For the other effects

$$x = \exp\{P \ln(y)\} \tag{A6}$$

where

$$y = (i50)\tau/85.0$$

Then

$$k_1 = k_2/x \tag{A7}$$

In the preceding equations, 85.0 mL of blood/kg of body weight is assumed. After all the parameters are initialized, initial drug concentrations and infusion rates are computed in the body, for each drug.

$$u_{mass} = uW \tag{A8}$$

$$m_{body} = \sum m_i \tag{A9}$$

$$C_d = m/V \quad \text{in each vessel} \tag{A10}$$

where u is the drug infusion rate in μg/kg/min, u_{mass} is the mass infusion rate, and W is the subject body weight in kg. m_i represents the drug mass in the ith compartment, and m_{body} is the total drug mass in the body.

II. Solution Procedure.

1. The baroreflex input is computed:

$$bfc = \frac{e^{c(MAP-MAP0)}}{1 + e^{c(MAP-MAP0)}} \tag{A11}$$

where MAP0 is the nominal value of MAP (120 mmHg) and c is an empirical constant equal to 0.062 63.

2. Parameters that remain constant over one heartbeat are calculated:

$$HR = HR_{base} - (Y_{HR})HR_{base} \tag{A12}$$

$$R_{sys} = R_{sys} - (Y_{R_{sys}})R_{sys\ base} \tag{A13}$$

$$V_{us\text{-}ven} = V_{us\text{-}ven\ base} + Y_{V_{us\text{-}ven}} \tag{A14}$$

$$E_{max} = E_{max\ base} - (Y_{E_{max\ base}})E_{max\ base} \quad \text{(LV,RV)} \tag{A15}$$

Equations A12–A15 represent the modifications due to the baroreflex. The base values in the nominal case are as follows: HR_{base} = 120 bpm, $R_{sys\ base}$ = 2.81 mmHg/(mL·s)$^{-1}$, $E_{max\ base\ RV}$ = 1.5 mmHg/mL. The other base

Table 5. Drug Parameters

parameters	NP → R_{sys} NP → P_{crit}	NP → $V_{us\text{-}ven}$	DP → E_{max}	DP → R_{sys}
Eff_{max}	0.635	225.0 mL	1.3	0.5
P	1.0	1.0	6.11	1.45298
k_2	0.025 s^{-1}	0.00625 s^{-1}	0.0011316 s^{-1}	0.0125 s^{-1}
i50	1.706 μg/kg/min	0.936 μg/kg/min	4.0 μg/kg/min	92.261 μg/kg/min
$\tau_{1/2}$	15 s	15 s	120 s	120 s

values are read in the initial data. All initial volumes are assumed to be the base unstressed volumes for each vessel. Y in each case represents the baroreflex output at the end of the previous heartbeat. The integration time over each heartbeat is determined by the heart rate (HR). The modified value is used to update the various time steps.

$$T_{beat} = 1/HR \tag{A16}$$

$$T_{step} = 1/n(HR) \tag{A17}$$

$$T_{c'} = 1 + 0.004(HR - HR_{base})\frac{0.65}{HR} \tag{A18}$$

$$T_c = 1.0 - 0.2933\left\{\frac{E_{max\ LV} - E_{max\ base\ LV}}{E_{max\ base\ LV}}\right\}T_{c'} \tag{A19}$$

T_{beat} is the time duration of one heartbeat. T_{step} is the integration step size within one heartbeat. n is the specified number of steps per heartbeat (a constant). T_c is the contraction time of the ventricle (duration of systole).

Each parameter except heart rate in eqs A12–A15 is further modified by the drug effects.

$$R_{sys} = R_{sys} - (Eff_{NP\text{-}R_{sys}} + Eff_{DP\text{-}R_{sys}})R_{sys} \tag{A20}$$

$$E_{max} = E_{max} + (Eff_{DP\text{-}E_{max}})E_{max} \tag{A21}$$

$$V_{us\text{-}ven} = V_{us\text{-}ven} + Eff_{NP\text{-}V_{us\text{-}ven}} \tag{A22}$$

$$P_{crit} = P_{crit} - (Eff_{NP\text{-}P_{crit}})P_{crit} \tag{A23}$$

The drug effects are updated by integrating the drug effect equation.

$$\frac{d(Eff)}{dt} = k_1 C_d^{\ p}[Eff_{max} - Eff] - k_2 Eff \tag{A24}$$

3. Parameters are integrated over one heartbeat. The variable Q_{beat}, which denotes flow out from the heart over one heartbeat, is first set to zero. The following steps are repeated over n intervals.

(a) The baroreflex equations are integrated using the time domain representation of the transfer functions shown in Figure 2. The output from each first-order transfer function is Y.

(b) The heart is in a relaxed state at the beginning of the beat. The diastolic pressure and time-varying elastance are computed for each ventricle.

$$P_D = E_D(V - V_{us}) \tag{A25}$$

In the nonlinear region ($P_D > P_M$), diastolic pressure is given by

$$P_D = (E_D/A)[\exp(A\{V_D - V_M\}) - 1] + P_M \tag{A26}$$

The relaxed-state ventricular elastance is given by

$$E = P_D/(V - V_d) \tag{A27}$$

If the time within the beat t is less than T_c (systole), then

$$E = E + 0.5(E_{max} - E)(1.0 - \cos(2pt/T_c)) \tag{A28}$$

(c) The pressures in the rest of the compartments are computed.

$$P = (V - V_{us})/C \quad V > V_{us} \tag{A29}$$

$$P = 0 \quad V \le V_{us} \tag{A30}$$

where C is the compliance parameter for the compartment being considered.

(d) The pressure–flow relationships in the ventricles are computed as follows.

Flow out

if $P \le P_{art}$

$$Q_{i+} = 0.0 \tag{A31}$$

if $P > P_{art}$

$$P_1 = (V_{ED} - V_d)E \tag{A32}$$

$$Q_{i+} = \frac{P - P_{art}}{R_{src}P_1} \tag{A33}$$

$$P_{art} = P_{art} + Q_{i+}R_{out} \tag{A34}$$

$$P = P_{art} \tag{A35}$$

In the preceding equations, P_{art} is the blood pressure in the artery leaving the heart. R_{out} is the characteristic systemic or pulmonary resistance to blood flow from the heart.

Flow in

if $P \ge P_{ven}$

$$Q_{i-} = 0 \tag{A36}$$

if $P < P_{ven}$

$$Q_{i-} = (P_{ven} - P)/R_{in} \tag{A37}$$

In the preceding equations, P_{ven} is the blood pressure in the vein entering the ventricle, and R_{in} is the characteristic systemic or pulmonary resistance to blood flow into the ventricle.

(e) Blood flow rate in the small arteries is computed:

$$Q_{small\ arteries} = \frac{P_{small\ arteries} - P_{ven}}{R_{sys}} \tag{A38}$$

(f) Pulmonary flow rates are computed as follows. The following terms are defined.

If $P_{pa} \ge P_{pv}$,

$$P_{in} = P_{pa} + H_{ves}$$

$$P_{out} = P_{pv} + H_{ves}$$

If $P_{pa} < P_{pv}$,

$$P_{in} = P_{pv} + H_{ves}$$

$$P_{out} = P_{pa} + H_{ves}$$

If $P_{in} \le P_{crit}$ (minimum pressure necessary to prevent pulmonary vessel collapse),

$$Q = 0$$

If $P_{in} > P_{crit}$,

$$H_1 = P_{out} - P_{crit}$$

$$H_1 = 0 \qquad P_{out} < P_{crit}$$

$$H_2 = P_{in} - P_{crit}$$

$$H_2 = H_T \qquad P_{in} > P_{crit} + H_T$$

then

$$Q = \frac{(P_{in} - P_{out})H_1 + (P_{in} - P_{crit})(H_2 - H_1) - (H_2^2 - H_1^2)/2}{R_{pul}H_T} \tag{A39}$$

In the preceding equations, H_T is the pressure equivalent to that of a column of blood of height equal to that of the lung (15.0 mmHg nominally). H_{ves} is a constant (3.75 mmHg in the nominal case).

(g) The volume–flow differential equations are integrated:

$$\left(\frac{dV}{dt}\right)_i = Q_{i-} - Q_{i+} \tag{A40}$$

$$\left(\frac{dQ}{dt}\right)_{\text{large arteries}} = \frac{P_{\text{large arteries}} - P_{\text{small arteries}}}{L} \tag{A41}$$

In the notation used in this paper, $i-$ refers to the previous vessel and $i+$ refers to the next vessel in sequence. In eq A40, the integration is performed for the seven assigned body compartments.

(h) The flow per beat and pressures are updated:

$$Q_{beat} = Q_{beat} + Q \quad \text{(for each vessel)} \tag{A42}$$

$$\text{MAP} = \text{MAP} + P_{\text{large arteries}} \tag{A43}$$

$$\text{MPAP} = \text{MPAP} + P_{pa} \tag{A44}$$

$$\text{MPVP} = \text{MPVP} + P_{pv} \tag{A45}$$

This completes the integration within a heart beat. Then for each ventricle, the stroke volume is computed from

$$\text{SV} = Q_{beat}T_{step} \tag{A46}$$

The mean values over the heartbeat are given by

$$Q_{beat} = Q_{beat}/n \tag{A47}$$

$$\text{MAP} = \text{MAP}/n \tag{A48}$$

$$\text{MPAP} = \text{MPAP}/n \tag{A49}$$

$$\text{MPVP} = \text{MPVP}/n \tag{A50}$$

Equation A48 is the mean arterial pressure at the end of the heartbeat.

4. Cardiac output is computed from

$$\text{CO} = (\text{SV})(\text{HR}) \tag{A51}$$

5. The end diastolic volumes are updated using the ventricular volumes at the end of the heartbeat. Pressures are updated using eqs A25 and A29.

6. The drug concentration equations are integrated with the current infusion in each body compartment. For drug concentrations, the venous chamber is subdivided into four subchambers to introduce sufficient delay in drug effect and elimination. The new concentrations for each drug for the next cycle are obtained from

$$\left(\frac{dm}{dt}\right)_i = (C_dQ)_{i-} - (C_dQ)_{i+} - (m/\tau_{1/2})_i \tag{A52}$$

where i is the compartment for which the concentration is being computed. The new concentrations and mass infusion rates are computed from eqs A8–A10.

This sequence of steps is repeated over each simulation interval.

Appendix B: Nonlinear Multirate MPC Formulation

For multirate systems, such as the drug delivery system where CO is sampled at a lower rate, the following terms are defined, consistent with the notation in existing literature on multirate control (e.g., Berg *et al.*, 1988). The sample time for the ith output variable is denoted as an integer multiple N_i of a base sample period T_s, corresponding to the fastest sampling rate. The sample times associated with each output are then defined by the following.

Definition 1: Shortest Time Period τ_s. This is the shortest sampling period for an output variable. Assume that all inputs are manipulated at this rate.

$$\tau_s = \text{gcd}(N_1, N_2, \ldots, N_{n_y})T_s \tag{B1}$$

Definition 2: Basic Time Period τ_b. Assuming that at $t = 0$ all variables were sampled synchronously, τ_b is the number of intervals after which the samples coincide. The multirate system is then periodically time varying (PTV). τ_b represents the period of the system as a multiple of T_s:

$$\tau_b = \text{lcm}(N_1, N_2, \ldots, N_{n_y})T_s \tag{B2}$$

We also define

$$\nu = \tau_b/\tau_s \tag{B3}$$

$k\nu$ hence defines the synchronous sample times, where all output measurements are available. In the multirate MPC formulation, we denote a general point $k\nu + j$ that is j steps from the synchronous point, where $k = 0, 1, 2, \ldots$ and $j = 0, 1, 2, \ldots, \nu - 1$.

I. Prediction Step. Consider a general nonlinear model of the s input–r output multivariable system:

$$\dot{x} = \mathbf{f}\{x,u,t\} \tag{B4}$$

$$y = \mathbf{g}\{x,u,t\} \tag{B5}$$

$\mathbf{f}$ and $\mathbf{g}$ are vector functions and $x \in R^n$, $u \in R^s$, and $y \in R^r$. In the NL-QDMC prediction step, the contribution to the predicted output is split into two parts. The first part is computed by integrating the nonlinear model (eq B4) over the prediction horizon holding the current input constant. Then the ith element of the nonlinear component of the predicted output vector, computed at time $k\nu + j$, is given by

$$\hat{y}^{NL}(k\nu + j + i) = \mathbf{g}(\hat{x}(k\nu + j + i), u(k\nu + j), t) \tag{B6}$$

where $\hat{x}(k\nu + j + i)$ is obtained by integrating eq B1 i steps into the future, with input $u(k\nu + j)$. The nonlinear component is computed prior to the optimization step.

If the current point $k\nu + j$ is a linearization step, then the linear multivariable convolution model for each input–output pair is generated by making a unit step change in the input, keeping all other inputs constant at their respective values in the input vector $u(k\nu + j)$. The step response for the (z,q) input–output pair is represented by $\{a_{zq}(i)|i = 1 \rightarrow N_{zq}\}$, where N_{zq} is the truncation order for the (z,q) step response and roughly corresponds to the settling time of the open-loop transient. The linear component of the prediction for the qth

output is then given by

$$\hat{y}_q^{L}(kv+j+i) = \sum_{z=1}^{s}\sum_{t=1}^{\mu} a_{zq}(i-t+1)\Delta u_z(kv+j+t-1) \quad \text{(B7)}$$

$$\mu = \min(i,m)$$

$$i = 1, 2, ..., p$$

where p is the prediction horizon and m is the control horizon. The predicted output is computed by adding the linear and nonlinear contributions for each output over its prediction horizon.

The prediction computed thus far is open-loop since available feedback information has not been used. This is now corrected with the disturbance term for each output, which is the difference between model prediction and the measured value of the output at the current point. For the qth output, if the measurement is available at $kv + j$, the feedback term is updated. If the measurement is not available, then an estimate of the output based on the previous measurement is used (Gopinath and Bequette, 1991). This is equivalent to holding the disturbance term for the qth output constant over all its intersample points. The estimate is given by

$$y_q^{e}(kv+j) = y_q(kv) + \sum_{z=1}^{s}\sum_{i=1}^{N_{zq}}\sum_{t=1}^{j} h_{zq}(i)\Delta u_z(kv-i+t) \quad \text{(B8)}$$

where $h_{zq}(i)$ is the ith (z,q) impulse response coefficient, given by

$$h_{zq}(i) = a_{zq}(i) - a_{zq}(i-1) \quad \text{(B9)}$$

The estimation equation (eq B8) is written by assuming that the qth output is sampled only at the synchronous points. In general, the previous sampled point for the output may be at any $kv + j^*$, where $j^* < v$. For the drug delivery system however, eq B8 is valid since the synchronous points are defined as those where CO measurements are available. The corrected prediction for the qth output at the ith future point, based on information at $kv + j$, is then given by

$$\hat{y}_q(kv+j+i) = \hat{y}_q^{NL}(kv+j+i) + \hat{y}_q^{L}(kv+j+i) + \{\delta_q y_q(kv+j) + [1-\delta_q]y_q^{e}(kv+j) - y_q^{m}(kv+j)\} \quad \text{(B10)}$$

where y_q^{m} is the model prediction at the current point and is given by

$$y_q^{m}(kv+j) = \sum_{z=1}^{s}\sum_{i=1}^{N_{zq}} a_{zq}(i)\Delta u(kv+j-i) \quad \text{(B11)}$$

In eq B10, δ_q is used to represent the measurements available at $kv + j$.

$\delta_q = 1$ if the qth output measurement is available

$\delta_q = 0$ otherwise

II. Optimization Step. Once the output prediction has been computed for the multirate system, the optimization step is the same as in single-rate MPC, since it is assumed that control moves will be updated at the base sampling rate. The optimization step consists of finding the m future optimal control moves that minimize the predicted error in a least-squares sense over the prediction horizon, subject to the satisfaction of the process constraints. In the usual notation, let $\mathbf{e}$ denote the predicted error vector. The jth component of the error vector for the qth output at any (sampled or intersample) point k is defined by

$$\mathbf{e}_q(k+j) = y_q^{d}(k+j) - \hat{y}_q(k+j) \quad \text{(B12)}$$

where y^d is the desired output trajectory over the prediction horizon. The objective function for the minimization is given by

$$\min \Psi(\Delta\mathbf{u}) = \mathbf{e}^{T}\mathbf{Q}\mathbf{e} + \Delta\mathbf{u}^{T}\mathbf{R}\Delta\mathbf{u} \quad \text{(B13)}$$

subject to the specified constraints. In eq B13, $\Delta\mathbf{u}$ is the vector of future control moves and $\mathbf{e}$ is the predicted error vector, constructed by stacking the error vectors for all outputs. Process constraints on manipulated and output variables can be written in terms of constraints on the future control moves (García and Morshedi, 1986). $\mathbf{R}$ and $\mathbf{Q}$ are weighting matrices for the control moves and predicted errors, respectively. Selection of the horizons and determination of the weighting matrices constitute the important tuning steps in MPC, and tuning rules for their selection have been discussed by a number of researchers (Maurath *et al.*, 1988; García *et al.*, 1989).

One of the main advantages of the MPC formulation is the ability to explicitly include constraints in the controller design. This is done by formulating a quadratic programming (QP) subproblem with output and manipulated variable constraints stated as hard constraints on the decision variables:

$$\min \mathbf{q}(\Delta\mathbf{u}) = {}^1/_2\Delta\mathbf{u}^{T}\mathbf{G}\Delta\mathbf{u} + \mathbf{g}^{T}\Delta\mathbf{u} \quad \text{(B14)}$$

$$\text{such that } \mathbf{C}_c^{T}\Delta\mathbf{u} \geq \mathbf{b}_c \quad \text{(B15)}$$

$\mathbf{G}$ is the Hessian of the objective function (eq B13), and $\mathbf{g}$ is the objective function gradient. Equation B15 represents the system of linear equality and inequality constraints. For the unconstrained case, an analytical solution to the QP exists (Fletcher, 1981) and is given by

$$\Delta\mathbf{u}^* = -\mathbf{G}^{-1}\mathbf{g} \quad \text{(B16)}$$

where the Hessian and the gradient of the objective function are obtained by the vector differentiation of the objective function. Equation B16 is the linear unconstrained MPC control law. In the presence of constraints, the solution is not straightforward since the analytical expression (eq B16) is no longer valid and the constrained MPC control law is highly nonlinear even for linear models. However, the constrained QP solution can be written in terms of active sets (Fletcher, 1981), resulting in a piecewise linear MPC control law. The computed solution $\Delta\mathbf{u}^*$ is now implemented on the plant and model, and the prediction and optimization steps are repeated at the next time step.

Literature Cited

Bequette, B. W. Nonlinear Control of Chemical Processes: A Review. *Ind. Eng. Chem. Res.* **1991**, *30*(7), 1391–1413.

Berg, M. C.; Amit, N.; Powell, J. D. Multirate Digital Control System Design. *IEEE Trans. Autom. Control* **1988**, *AC-33* (12), 1139–1150.

Cutler, C. R.; Ramaker, B. L. Dynamic Matrix Control—A Computer Control Algorithm. *Proc. Am. Control Conf.*, San Francisco, CA, 1980, Paper WP5-B.

Fletcher, R., *Practical Methods of Optimization, Vol. 2: Constrained Optimization*; Wiley: New York, 1981.

García, C. E. Quadratic Dynamic Matrix Control of Nonlinear Processes: An Application to a Batch Reaction Process. AIChE Annual Meeting, San Francisco, CA, 1984.

García, C. E.; Morshedi, A. M. Quadratic Programming Solution of Dynamic Matrix Control (QDMC). *Chem. Eng. Commun.* **1986**, *46*, 73–87.

García, C. E.; Prett, D. M.; Morari, M. Model Predictive Control: Theory and Practice—A Survey. *Automatica* **1989**, *23* (3), 335–348.

Gingrich, K. J.; Roy, R. J. Modeling the Hemodynamic Response to Dopamine in Acute Heart Failure. *IEEE Trans. Biomed. Eng.* **1991**, *18* (3), 267–272.

Gopinath, R. S.; Bequette, B. W. Multirate Model Predictive Control: An Analysis for Single Input-Single Output Systems. *Proceedings of the 4th International Symposium on Process Systems Engineering—PSE 91*, Montebello, Québec, Canada, 1991.

Held, C. M.; Roy, R. J. Multiple Drug Hemodynamic Control by Means of a Supervisory-Fuzzy Rule-Based Adaptive Control System: Validation on a Model. *IEEE Trans. Autom. Control* **1993**, submitted for publication.

Lasdon, L. S.; Waren, A. D.; Jain, A.; Ratner, M. Design and Testing of a GRG code for Nonlinear Programming. *ACM Trans. Math. Software* **1978**, *4*, 34–50.

Maurath, P. R.; Mellichamp, D. A.; Seborg, D. E. Predictive Controller Design for Single-Input Single-Output (SISO) Systems. *Ind. Eng. Chem. Res.* **1988**, *27*, 956–963.

Prett, D. M.; García, C. E. *Fundamental Process Control*; Butterworths: Stoneham, MA, 1988.

Richalet, J. A.; Rault, A.; Testud, J. D.; Papon, J. Model Predictive Heuristic Control: Applications to Industrial Processes. *Automatica* **1978**, *14*, 413–428.

Ricker, N. L. Model Predictive Control: State of the Art. *CPC IV—Proceedings of the Fourth International Conference on Chemical Process Control*, Padre Island, TX, February 1991, CACHE, pp 271–296.

Sistu, P. B.; Gopinath, R. S.; Bequette, B. W. Computational Issues in Nonlinear Predictive Control. *Comput. Chem. Eng.* **1993**, *17* (4), 361–366.

Voss, G. I.; Katona, P. G.; Chizek, H. J. Adaptive Multivariable Drug Delivery: Control of Arterial Pressure and Cardiac Output in Anesthetized Dogs. *IEEE Trans. Biomed. Eng.* **1987**, *34*, 617–623.

Wesseling, K. H.; Settels, J. J.; Walstra, H. G.; van Esch, H. J.; Donders, J. H. Baromodulation as the Cause of Short Term Blood Pressure Variability? In *Proceedings of the International Conference on Application of Physics to Medicine and Biology*; Alberi, G., Bajzer, Z., Baxa, P., Eds.; World Scientific Publishing: Singapore, 1983; pp 247–277.

Yu, C. L. Adaptive Control of Cardiac Output and Mean Arterial Pressure Using Multiple Drug Infusion. Ph.D. Thesis, Rensselaer Polytechnic Institute, Troy, NY, 1990.

Yu, C. L.; Roy, R. J.; Kaufman, H. A Circulatory Model for Combined Nitroprusside-Dopamine Therapy in Acute Heart Failure. *Med. Prog. Technol.* **1990**, *16*, 77–88.

Accepted December 20, 1994.[⊗]

BP940107O

[⊗] Abstract published in *Advance ACS Abstracts,* February 15, 1995.