On-line Multivariable Identification and Control of Chemical Processes Using Canonical Variate Analysis

Andreas H. Kemna, Wallace E. Larimore[†],
Dale. E. Seborg[*], and Duncan A. Mellichamp[*]

Weyerhaeuser Company, Tacoma, WA 98477; akemna@wdni.com
† Adaptics, Inc, 40 Fairchild Drive, Reading, MA 01867; adaptics@mcimail.com
*Department of Chemical and Nuclear Engineering
University of California, Santa Barbara, CA 93106
seborg@engrhub.ucsb.edu, dmell@halibut.ucsb.edu

1 Abstract

In this paper, an on-line identification and control strategy is developed for chemical processes. The procedure used for system identification is canonical variate analysis (CVA), a method that operates on input-output data sequences batch-wise to provide an optimized state-space model; the resulting models are then used in a multivariable predictive control (MPC) procedure to yield excellent overall closed-loop results. The resulting adaptive control procedure is demonstrated to have excellent properties for control of a continuous stirred tank reactor (CSTR) with inverse dynamics, delays and nonlinearities.

2 System Identification and Adaptation to Change

The procedure used for system identification is the canonical variate analysis (CVA) method (Larimore, 1990). The CVA method involves computation of the singular value decomposition (SVD) to determine the linear combinations of the past data with predictive information for the future, i.e. states of the system. The computations are entirely automatic and have been shown to approach the theoretical optimum accuracy of maximum likelihood (ML) while avoiding a nonlinear optimization to determine parameter estimates typical of ML methods (Larimore, 1994). CVA has been extensively studied for the identification of chemical processes (Schaper et al, 1994). The CVA method is implemented in the ADAPTx Automated System Identification software package currently running in the MATRIXx/Matlab software environments on Workstations and IBM/PC computers (Larimore, 1992).

A fundamental problem in on-line identification and adaptive control is the adaptation of the identified model to the changes in the actual process dynamics. Previous approaches to the adaptation problem have been largely heuristic using such procedures as exponential forgetting. Only recently has a more fundamental approach been taken using concepts of information. A recent study (Larimore, 1986) approaches the general problem of decisions involving model state order, structure, rate of adaptation, and detection of abrupt changes using an entropy or information measure in a predictive inference framework. The entropy measure has been given a fundamental statistical inference justification in terms of the basic statistical principles of sufficiency and repeated sampling in a predictive setting (Larimore, 1983). This approach provides a sound theoretical basis for the development of adaptive identification procedures.

Consider a piece-wise constant model with constant dynamics governing y_t over the j-th time interval $T_j = \{t_j \leq t < t_{j+1}\}$ where y_{t_j} is the first point of the j-th interval and Y_j is the corresponding set of observations $Y_j = (y_{t_j}, \ldots, y_{t_{j+1}-1})$ Thus consider the model of the form (1) and (2) for $t \in T_j$ with piece-wise constant parameters θ^j

$$x_{t+1} = \Phi^j x_t + G^j u_t + w_t \qquad (1)$$

$$y_t = H^j x_t + D^j u_t + B^j w_t + v_t \qquad (2)$$

The coefficient matrices Φ^j, D^j, B^j, G^j, and H^j and noise covariances Q^j and R^j are functions of the unknown parameters θ^j which are constant over an interval of time T_j and change from one time interval to another.

The problem of adaptation to slow variations is primarily that of determining the lengths of time intervals to use in the time varying model (1) and (2).The predictive inference and entropy methods provide a means of objectively comparing the vast multitude of such hypotheses entailed in the selection of interval length for the adaptive time series analysis problem. Consider the division of a section of data into h subintervals Y_j for $1 \leq j \leq h$. The set $S = \{Y_1, \ldots, Y_h\}$ of subintervals can be associated with the set of times $(t_1, \ldots, t_h)$ where t_j specifies the first point of the interval Y_j. Then the various hypotheses can be considered such as H_S: divide the interval into subintervals Y_j. For each subinterval Y_j, for $j = 1, 2, \ldots, h$, suppose a state space model with parameters denoted by θ^j is fitted using the CVA method with the Akaike information criterion (AIC) used to select the best model state order.

By successive application of the Markov property, the joint probability density of the observations conditioned on the initial state x_0 is given by

$$\log p(Y_1, \ldots, Y_h | x_0, \Theta^S) = \sum_{j=1}^{h} \log p(Y_j | Y_{j-1}, \theta^j) \qquad (3)$$

where $\Theta^{ST} = (\theta^{1T}, \ldots, \theta^{hT})$ is the parameter vector for the composite model consisting of all of the CVA models fitted over each of the h subintervals. This gives the log likelihood as the sum of the conditional log likelihoods on each subinterval.

To evaluate the model fit of using a set of time intervals $(Y_1, \ldots, Y_h)$ and associated models Θ^S, the AIC is used which is defined for a particular time interval and model as

$$AIC(Y_j, \hat{\theta}_j) = -2 \log p(Y_j | Y_{j-1}, \hat{\theta}_j) + 2M_j \qquad (4)$$

where M_j is the number of estimated parameters in fitting $\hat{\theta}_j$. Thus the AIC corresponding to (3) is given by

$$AIC(Y_1, \ldots, Y_h, \hat{\Theta}^S) = \sum_{j=1}^{h} AIC(Y_j, \hat{\theta}^j) \qquad (5)$$

The optimal selection of data subintervals is chosen as the S minimizing the above AIC.

An efficient procedure for selecting optimal interval lengths can be developed using an interval doubling procedure. Consider the following interval sets for model fitting: S_1 consists of intervals of some shortest length L; S_2 is the set of intervals of length $2L$; ...; S_{2^i} is the set of intervals of length $2^i L$. Now if the AIC's are saved for each of the subintervals, then comparisons can be made for composite models over some total data interval of say $2^m L$. The interval length with the minimum AIC gives the optimum choice data length for identification of the process dynamics.

3 Model Predictive Control

Modern control systems usually utilize a model of the process to be controlled. If a reasonably accurate physical model of the process is not available, an empirical model can be identified from input-output data. The CVA identification technique has

shown considerable promise for estimating a reliable, accurate low-order model of the process from input-output data. The identified process model can then be used as a basis for the implementation of advanced control algorithms.

Model predictive control (MPC) has found widespread acceptance in the chemical industry (Garcia et al, 1989; Richalet, 1993). Several companies have specialized in the development and installation of MPC algorithms for plants and processes, and various industrial applications have been reported in recent years in the open literature. It is the ability of MPC to achieve high performance control on rather complex systems without constant expert intervention which has made it so successful. Using MPC, problems such as inverse response, large dead times, and inequality constraints on inputs and outputs can be handled in a rather straightforward manner.

Common features of MPC algorithms are, among others, the following:

(i) A performance index is employed to calculate the control action which optimizes the process output over some finite future time interval.

(ii) A reference trajectory and a vector of predicted outputs is used.

(iii) A process model is used for system representation and prediction.

The two best known MPC algorithms are Model Algorithmic Control (MAC; see Richalet et al, 1993) and Dynamic Matrix Control (DMC; see Garcia et al, 1989).

As suggested by Morari and Lee (1991), the optimal estimation of the open-loop output predictions $\tilde{Y}_t$ (not related to the previous notation Y_j) can be represented in a two-step form:

(i) Model Prediction

In the model prediction step the predicted future (open loop) output vector $\tilde{Y}_t$ is updated, taking into account the input at $t-1$.

$$\tilde{Y}_{t|t-1} = M\tilde{Y}_{t-1} + S\Delta u_{t-1} \qquad (6)$$

M is a $n \times n$ shift matrix which has the form

$$M = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & \vdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \qquad (7)$$

and n is the model horizon. It is chosen in such a way that $n\Delta t$ covers 99% of the open-loop response time of the process to be controlled. Δt is the sampling period of the system. S is the first column of the step response matrix A.

$$A = \begin{pmatrix} a_1 & 0 & \cdots & \cdots & 0 \\ a_2 & a_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ a_n & a_{n-1} & \cdots & a_2 & a_1 \end{pmatrix} \qquad (8)$$

with

$$a_k = H(\sum_{i=1}^{k} \Phi^{i-1})G + D \qquad (9)$$

where a_k is the kth step response coefficient. The matrices H, Φ, G and D are evaluated from input/output data using the CVA method.

The change in input Δu_{t-1} is defined as

$$\Delta u_{t-1} = u_t - u_{t-1} \qquad (10)$$

(ii) Prediction Correction

In the prediction correction step the model prediction of step (i) is corrected for disturbances and model errors, according to

$$\tilde{Y}_t = \tilde{Y}_{t|t-1} + K[y_t - \hat{y}_{t|t-1}] \qquad (11)$$

y_t is the current, measured output and $\hat{y}_{t|t-1}$ is the predicted output, based on measured data up to and including $t-1$ (see (14)). K is a gain vector and has the form

$$K = [ffff...]^T; \quad 0 < f < 1 \qquad (12)$$

As pointed out by Morari and Lee (1991), the value of f lies in (0,1) and is related to the disturbance-to-noise ratio. When only small modeling errors and insignificant noise are expected, f can be chosen close to one, which results in a deadbeat observer. The evaluation of the future open-loop outputs $\tilde{Y}$ using a simple Kalman filter is completely decoupled from the control part of the algorithm. Thus, as this ratio goes to infinity, f goes to 1. As the ratio goes to 0, f tends to 0.

The predictive control problem, as formulated and solved in Morari and Lee (1991), involves minimization of a performance index which includes the (closed-loop) prediction error E_t:

$$E_t = R_t - \hat{Y}_t = \begin{pmatrix} r_t \\ r_{t+1} \\ \vdots \\ r_{t+p} \end{pmatrix} - \begin{pmatrix} \hat{y}_{t|t-1} \\ \hat{y}_{t+1|t-1} \\ \vdots \\ \hat{y}_{t+p|t-1} \end{pmatrix} \qquad (13)$$

R_t is the vector of future set points (reference trajectory); $\hat{Y}_t$ is a vector of predicted future outputs based on data up to $t-1$. Parameter p is referred to as the prediction horizon. The vector of predicted future outputs $\hat{Y}_t$ is evaluated from

$$\hat{Y}_t = A_p^m \Delta U_t + \tilde{Y}_{pt} \qquad (14)$$

where A_p^m is a submatrix of the step response coefficient matrix A, comprised of the first p rows and m columns. The control horizon m is the number of control actions that are calculated in order to affect the predicted outputs over the prediction horizon p. ΔU_t is the vector of future input changes which will be the subject of the optimization:

$$\Delta U_t = \text{col}[\Delta u_t, \Delta u_{t+1}..., \Delta u_{t+m}] \qquad (15)$$

$\tilde{Y}_{pt}$ includes the first p elements of the predicted output vector $\tilde{Y}_t$ from (11).

Now consider the general quadratic measure of performance J to be minimized by appropriate selection of the future incremental control inputs ΔU_t to solve the optimal deterministic predictive control problem,

$$J = (R_t - \tilde{Y}_t - A\Delta U_t)^T W_1 (R_t - \tilde{Y} - A\Delta U_t) + \Delta U_t^T W_2 \Delta U_t \qquad (16)$$

where W_1 and W_2 are weighting matrices for the output and input, respectively. This optimization problem can be solved over p time steps into the future. The minimization of the objective function J is then performed by differentiating J with respect to ΔU_t and setting this derivative to zero. The explicit control law can then be easily evaluated as

$$\Delta U_t = (A^T W_1 A + W_2)^{-1} A^T W_1 (R_t - \tilde{Y}_t) \qquad (17)$$

under the assumption that the weighting matrices W_1 and W_2 are diagonal. These concepts can be readily extended to the multi-input/multi-output (MIMO) case and when the process exhibits integrating behavior, as shown by Morari and Lee (1991).

4   Controller Design

In the general MPC formulation, four design parameters have to be specified: the control horizon m, the prediction horizon p, and the two weighting matrices W_1 (output weighting) and W_2 (input weighting). Ricker (1991) proposes a controller design methodology which allows m, p and W_1 to be specified in a rather straightforward manner. He then suggests using the input weighting matrix W_2 as the main tuning parameter. These

tuning rules have one significant disadvantage: there is no available method to specify W_2 a priori. Trial and error simulations have to be employed to evaluate a suitable W_2.

As pointed out by Maurath et al (1988), using a control horizon $m > 1$ immediately necessitates the specification of a nonzero weighting matrix W_2, with all its related problems. One way to avoid these is to use a control horizon of $m = 1$. In this case the prediction horizon is used as the tuning parameter, and W_2 can be set to zero. Maurath et al (1988) show that this design method produces an effective controller that can be easily tuned, or even automated in a straightforward manner. In their paper, Maurath et al (1988) propose a design method where the prediction horizon p is chosen such that $a_p > 0.5K_p$, where the process gain K_p is assumed for simplicity to be positive. In the following examples, this method is used.

5 CSTR Simulations

A high fidelity simulation of a continuous stirred tank reactor (CSTR) with a single irreversible reaction, $A \to B$, is used to represent a common nonlinear process. This is used to show how the model obtained from CVA and the MPC algorithm interact. Closed-loop simulations are used to evaluate the CVA method under closed-loop conditions. The CSTR is also used to show how the controller design works in a MIMO control configuration.

The CSTR is a rather common process in the chemical industries. Typically the reactor concentration C_A is controlled using the jacket temperature T_j as the manipulated variable. The CSTR can be described using energy and component balances. By assuming i) perfect mixing, ii) negligible ambient heat loss, iii) first-order irreversible reaction, and iv) constant physical properties of the inlet and outlet streams, the model equations are:

$$\frac{dC_A}{dt} = \frac{Q_f}{V}(C_{AF} - C_A) - r \tag{18}$$

$$\frac{dT}{dt} = \frac{Q_f}{V}(T_F - T) + \frac{(-\Delta H)}{c_p \rho}r - \frac{UA}{V c_p \rho}(T - T_j) \tag{19}$$

$$r = C_A k_0 e^{(\frac{-B}{RT})} \tag{20}$$

The notation and nominal conditions are tabulated in Seborg et al (1989). Typically the reactor concentration is measured using a gas chromatograph. Here we assume an analysis time of 4 minutes which determines the sampling period and introduces a time delay of the same magnitude in the measured response.

Figure 1 shows a comparison between the step response of the process and a step response generated from a CVA model. This CVA model was evaluated from input and output data vectors containing 450 data points, where the output data were corrupted by noise ($\sigma = 1 \times 10^{-2}$Kg mol/m^3). The input was a PRBS signal ($\pm 5.56°K$) imposed on the jacket temperature. The CVA model describes the process rather well despite its nonlinearity. Figure 2 shows a comparison of the performance of a well-tuned PI controller ($K_c = -155\frac{°K}{\text{Kgmol/m}^3}, \tau_i = 8.25$ min) and the MPC algorithm $m = 1$, $p = 4$, f=0.5). At $t = 12$ minutes a setpoint change from 0.73 to 0.80 occurs. At $t = 70$ minutes an inlet flow disturbance (+10% change) is introduced. At t=155 minutes another disturbance, this time in the inlet concentration (+10% change) is made. For the setpoint change and the inlet flow disturbance, the MPC algorithm performs markedly better than the PI controller; for the inlet concentration disturbance, they perform equally well. Overall, these simulations show that a CVA-based SISO MPC can perform very well when applied to processes typically encountered in industry.

Closed-loop data generated through several setpoint changes were then used to evaluate the CVA (setpoints were changed regularly from 0.8 kg mol/m^3 to 0.6 kg mol/m^3 and back – altogether 7 changes within 500 minutes). The closed-loop CVA model is shown in Figure 3. The overall quality of the identification is about the same as in the open-loop case. Figure 4 shows a comparison between the cases where the open-loop and the closed-loop CVA identified model were used to determine the A matrix and the MPC controller.

The suggested SISO controller design method can be readily extended to the MIMO case. A slightly modified CSTR model ($\Delta t = 1$ min.; no time delay in the outputs) is used to show the controller performance for a MIMO configuration. The two controlled CSTR outputs are now the reactor concentration C_A and the reactor temperature T. The two manipulated inputs are the jacket temperature T_j and the reactor inlet concentration C_{AF}. This control configuration includes a rather challenging feature: a change in the inlet concentration results in an inverse response in the reactor concentration C_A. As can be seen in Figure 5, the CVA identified model describes the process quite well. The MIMO CVA model was consequently used for control purposes. The performance of the MIMO controller with a prediction horizon of 13 is shown in Figure 6. At $t = 10$ min., the setpoint of the reactor concentration was changed from 0.729(Kg mol/m^3) to 0.8(Kg mol/m^3). At $t = 100$ min., the reactor temperature setpoint was changed from 317.11° K to 318° K. At $t = 200$ min., an unmeasured disturbance, namely a 10 % increase in the feed flow rate Q_f was introduced.

Overall the implementation of the suggested controller design method was straightforward; subsequently the controller showed very satisfactory performance.

6 Summary

Implementation of an MPC is quite straightforward using the controller design method outlined here. It works well for both SISO and MIMO systems and requires the specification of relatively few design parameters and no weighting functions. The parameters are chosen using simple ad hoc rules.

CVA models obtained from both open-loop and closed-loop data characterize the CSTR process studied here quite well, even for the cases of inverse dynamics or for a nonlinear underlying process model. The MPC controller obtained using a CVA model worked as well as one obtained from the actual process model itself (no modeling error) and a comparison of MPC controllers — one designed on the basis of a CVA model obtained from open-loop data, the second using a model obtained with closed-loop data — shows few differences.

Thus we conclude that MPC techniques, when applied with CVA models obtained under either open- or closed-loop conditions, provide excellent properties for the controlled system. The advantages obtained from use of the CVA technique to obtain process models under operating conditions (discussed above) are quite significant compared to alternative approaches.

7 Acknowledgement

8 References

Garcia, C.E., D.M. Prett, and M. Morari (1989), "Model Predictive Control: Theory and Practice - a Survey", Automatica, 25, 335-348.

Larimore, W.E. (1994), "The Optimality of Canonical Variate Identification by Example," 10th IFAC Symposium on System Identification, July 4-6, 1994, Copenhagen.

Larimore, W.E. (1992), ADAPTx Automated System Identification Software Users Manual, Adaptics, Inc, 40 Fairchild Drive, Reading, MA 01867.

Larimore, W.E. (1990), "Canonical Variate Analysis for System Identification, Filtering, and Adaptive Control, " Proc. 29th IEEE Conference on Decision and Control, Honolulu, Hawaii, December, Vol. 1, pp. 635-9.

Larimore, W.E. (1986), "Development of Statistical Methods Using Predictive Inference and Entropy," Final Report for Air Force Office of Scientific Research, Scientific Systems, Inc.

Larimore, W.E. (1983). "Predictive Inference, Sufficiency, Entropy, and an Asymptotic Likelihood Principle", Biometri-

ka, Vol 70, pp. 175-81.

Maurath, P., D.A. Mellichamp and D.E. Seborg (1988), "Predictive Controller design for Single-Input/Single-Output (SISO) Systems," *Ind.Eng.Chem.Res.* Vol. 27, p. 956-963.

Marchetti, J.L., D.A. Mellichamp and D.E. Seborg (1983), "Predictive Control Based on Discrete Convolution Models," *Ind. Eng. Chem. Process Des. Dev.* Vol. 22, p. 488-495.

Morari, M. and J.H. Lee (1991) "Model Predictive Control: The Good, the Bad and the Ugly" *Proc. Chemical Process Control Conference IV* South Padre Island, TX, p. 419-444, Eds. Y. Arkun and W.H. Ray, AIChE, New York, NY.

Richalet, J. (1993), "Industrial Applications of Model Based Predictive Control", Automatica, 29, 1251-74.

Ricker, N.L. (1991) "Model Predictive Control: State of the Art," *Proc. Chemical Process Control Conference 1991 IV* South Padre Island, TX, Eds. Y. Arkun and W.H. Ray, AIChE, New York, NY, p. 271-296.

Schaper, C.D., W.E. Larimore, D.E. Seborg, D.A. Mellichamp (1994), "Identification of Chemical Processes using Canonical Variate Analysis," *Computers & Chemical Engineering*, Vol. 18, pp. 55-69.

Seborg D.E., T.F. Edgar and D.A. Mellichamp, *Process Dynamics and Control*. Wiley, NY (1989)
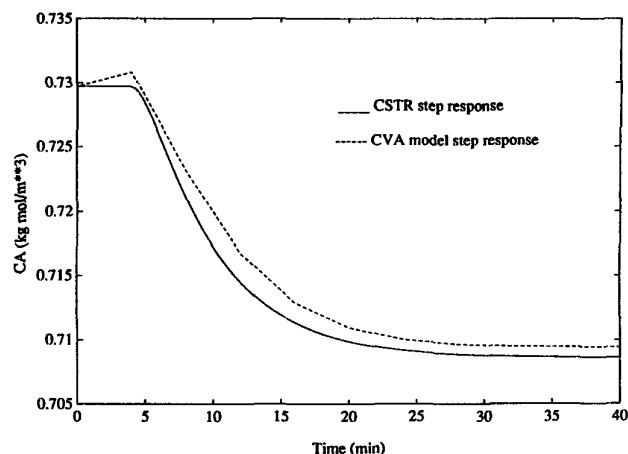
Figure 1: Comparison of actual and estimated CSTR models.
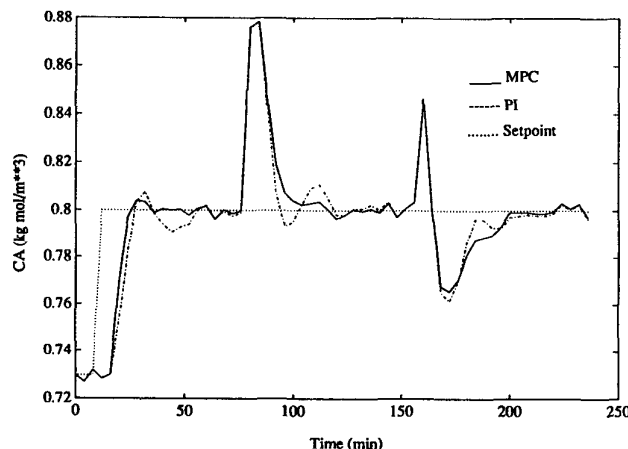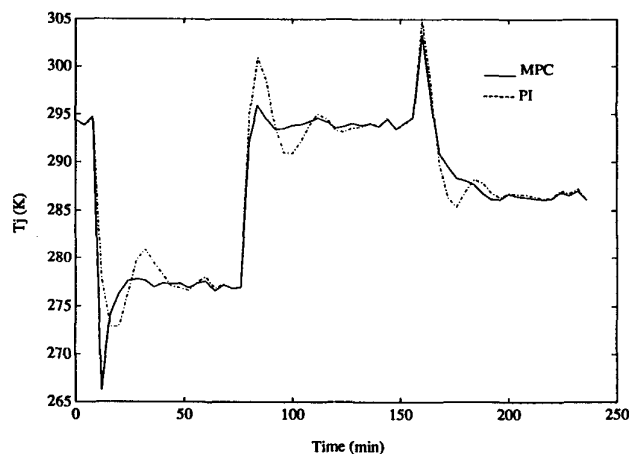
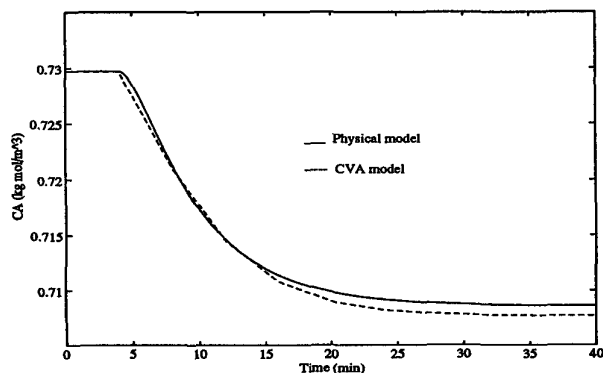Figure 2: Comparison of PI and MPC performance for setpoint and load changes ($\Delta t = 4$ min.).

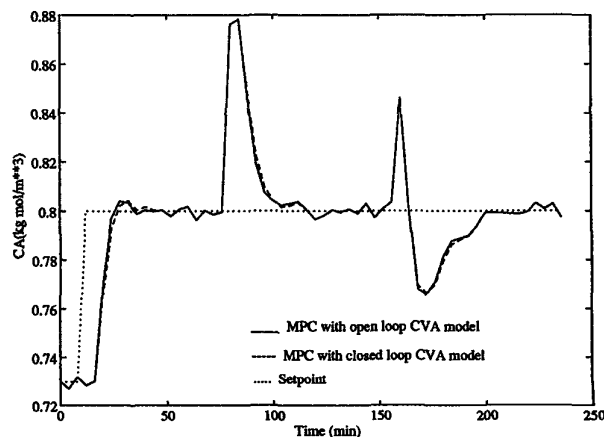Figure 3: Comparison of true and estimated model of the CSTR from closed loop data.

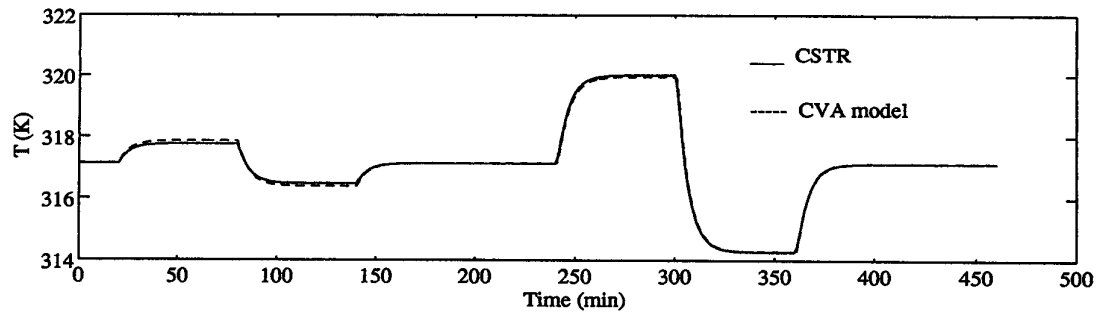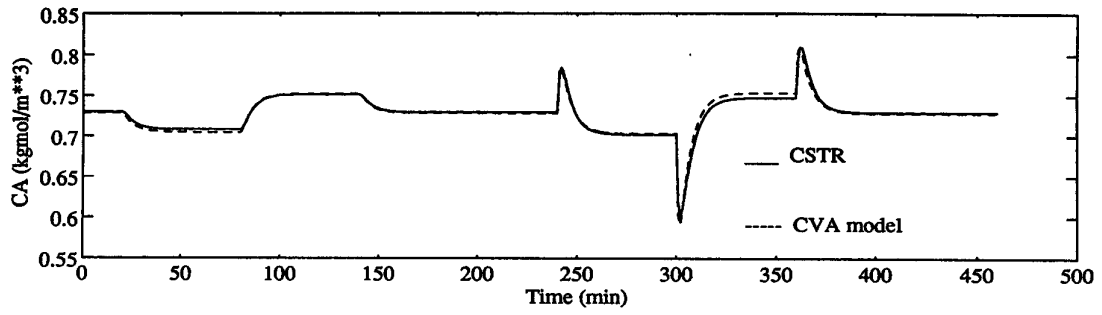Figure 4: Comparison of MPC performance for setpoint and load changes using open-loop end closed-loop CVA midels.

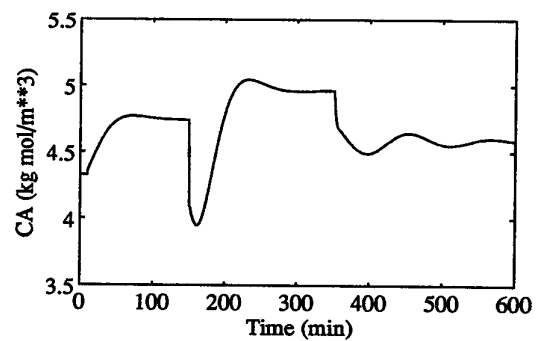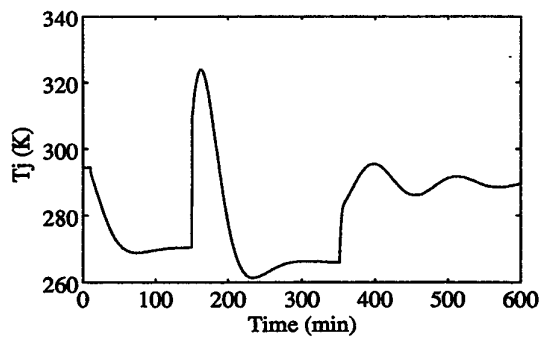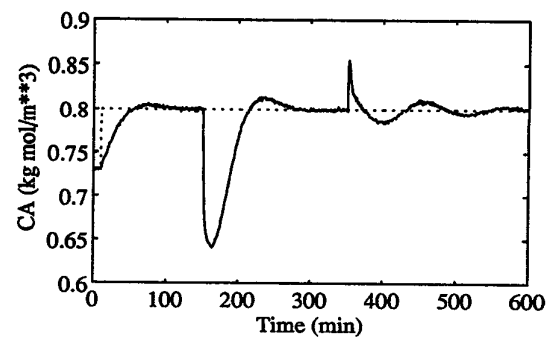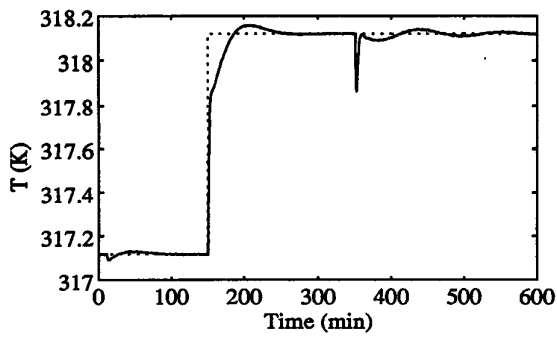Figure 5: Comparison of true and estimated MIMO CSTR models.

Figure 6: MPC performance with set point and load responses for the MIMO CSTR model.