Proceedings of the American Control Conference
San Diego, California • June 1999

A Tutorial on Inferential Control and its Applications

Babu Joseph[*], Department of Chemical Engineering
Washington University, St.Louis, MO 63130
E-mail: joseph@wuche.wustl.edu

Abstract--In this tutorial, we discuss approaches to control of one or more primary variables in a process using secondary measurements. This approach is useful when the primary variable is not easily measured, or has large time delays or lags associated with it. It is also useful when the secondary measurements contain information about disturbances that affect the primary variable. We start by discussing some classical approaches to this problem. Then we present inferential control strategies that use process models to predict the effect of the disturbance variable on the primary output and use this prediction to regulate the process. Next we present a framework for incorporating the inferential control strategy within the framework of the often-used model predictive control (MPC). This framework, termed as model predictive inferential control (MPIC), is general enough to accommodate multiple secondary measurements as well as nonlinear estimators and controllers. The concept is also extended to end product quality control in batch processes using intermediate measurements available during the middle of the batch. The advantages of inferential control are established using the Shell challenge case study problem, which employs linear transfer function models. Problems of collinearity among the secondary measurements (which frequently arises) is addressed using principal component analysis (PCA) during the construction of the dynamic estimator. Numerous applications demonstrate the advantages of the inferential control strategy.

1. Introduction

There exist a number of processes in which the primary variable to be controlled is difficult to measure or is a sampled measurement with a long delay in the sampling and analysis process. Sometimes the quantity to be controlled is a calculated variable. In such cases control of the process is usually accomplished by measuring secondary variables (for which sensors are more reliable, cheaper or more readily available and installed) and setting up a feedback control system using these secondary variables. Such control strategies are referred to as inferential control. Figure 1 shows a few examples from process industry where inferential control has found application.

In distillation the primary variables to be regulated are the product compositions (bottoms and distillate purity) as shown in Figure 1(a). Gas chromatographs typically used to measure these are expensive, difficult to maintain and calibrate and introduce significant measurement delays because of the time needed to purge the sample line and to heat the sample. In this case, control is accomplished using temperature measurements on the intermediate trays. Joseph and Brosilow (1978a, 1978b) developed methods for construction of optimal and sub-optimal estimators and compared the two methods for inferential control of product composition on a simulated multi-component distillation column. Willis et al. (1991) discuss a neural network based estimation procedure for feedback control of the product composition from an industrial distillation tower using measured quantities such as overheads temperatures. Ye et al. (1993) report improved control of both the product flow and compositions with a neural-net based inferential control approach for the Tennessee Eastman industrial test process. Abdel Jabbar and Alatiqi (1997) present an inferential-feedforward control scheme for a petroleum fractionator with undefined blends of hydrocarbons as the feed. Unmeasured feed composition disturbances are estimated from secondary measurements and the manipulated variables are varied to maintain desired product quality.

In polymerization reactors, primary variables of interest are the molecular weight and viscosity of the product, as shown in figure 1(b). Control is accomplished using secondary measurements such as temperature and pressure in the reactor. MacGregor et al. (1991) report how some common reactor operating problems while producing Low Density PolyEthylene (LDPE) can be detected from online measurements such as the temperature profile down the reactor and the solvent flow rate, which are available on a more frequent rate than fundamental polymer properties. Irwin et al. (1995) describes inferential estimation of polymer viscosity in a polymerization reactor. Here the measurement from the viscometer is subject to a significant time delay but the torque from a variable speed drive provides an instantaneous indication of reactor viscosity. McKay et al. (1996) applied nonlinear inferential models to predict polymer viscosity using available measurements for an industrial extrusion process. Figure 1(c) shows a third application of inferential control. This pertains to control of industrial drying processes. Koppel et al. (1995) describe simulation and control of such drying processes with inferential determination of solids moisture, a variable not usually measurable, and using this to manipulate the temperature of drying air until the desired target in solids moisture is obtained.

In addition to the above, inferential control is also applied to control concentration in reactor effluents. Budman et al. (1991) applied an inferential control technique to an experimental fixed bed reactor where two controlled variables (the exit temperature and maximum bed temperature) are inferred from a single temperature measurement. Parrish and Brosilow (1985,1986) discuss nonlinear inferential control of reactor effluent concentration from temperature and flow measurements. They also present rules to tune inferential controllers online to enable them to outperform conventional feedback control systems.

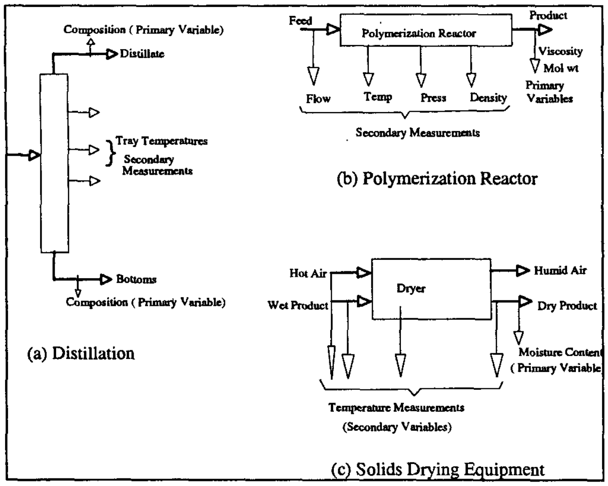

Figure 1. Examples of Processes with Inferential Control Problems

A common problem in inferential control is the estimation from a large number of process variables, which are highly correlated with one another. Kresta et al. (1994) discuss how Partial Least Squares (PLS) techniques can be applied to build good inferential estimators in such cases. Mejdell and Skojestad (1992) present a PCR (Principal Component Regression) estimator for estimating product composition in distillation columns from secondary temperature and flow rate measurements. It was demonstrated that the performance of the estimate is generally improved by adding more temperature measurements once the strong collinearity between the temperature measurements is tackled with PCR.

In this paper, we show how model predictive and inferential control strategies can be combined to cut costs by improve product quality and quantity. In Section 2, we describe the generic problem of inferential control. Limitations of some of the classical control strategies are discussed. In Section 3, we demonstrate how the inferential control framework can be extended to a model-predictive control framework, which is widely used in industry and applicable to multi-variable and nonlinear systems. In Section 4, we discuss some practical issues such as use of multiple secondary measurements, collinearity among the measurements, extension to nonlinear systems and application to end product quality in batch processes. In Section 5, we present a case study of the Shell process control challenge problem to demonstrate the superior disturbance rejection properties of this scheme.

2. Generic problem

Figure 2 shows the schematic of the generic problem tackled by inferential control. The control objective typically is to keep the primary variable on target in presence of unmeasured disturbances. We first look at some classical techniques used, which while being simple to implement can be costly because of poor controller performance.

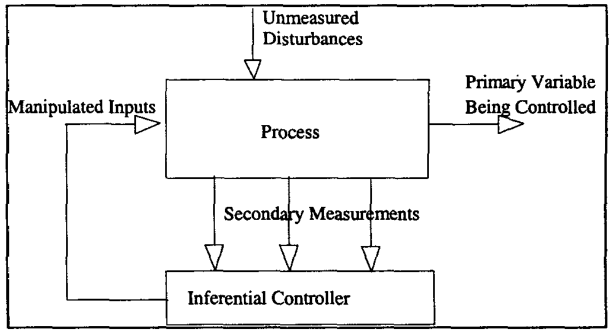

Figure 2. Generic inferential control problem

For the simple case of a linear system with a single disturbance, single primary output and single secondary measurement as in Figure 3, the process can be modeled as:

$$y(s) = g_{uy}(s)u(s) + g_{dy}(s)d(s) \quad \text{: primary output}$$

$$t(s) = g_{ut}(s)u(s) + g_{dt}(s)d(s) \quad \text{: secondary measurements} \qquad (1)$$

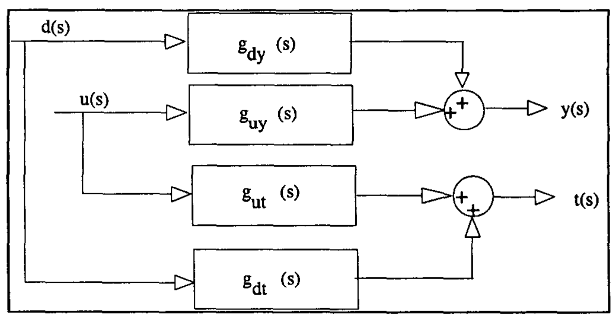

Figure 3. Block diagram of a process with one primary and one secondary measurement

2.1 Some Classical Control Strategies

Consider direct feedback control of the secondary measurement t using the manipulated variable u as shown in Figure 4(a). This strategy can be used if the primary variable and secondary variable are very closely related. For example in distillation, it is well known that temperature is a very good indicator of product composition. Hence, by maintaining one of the tray temperatures constant, we can often maintain good control of the product quality.

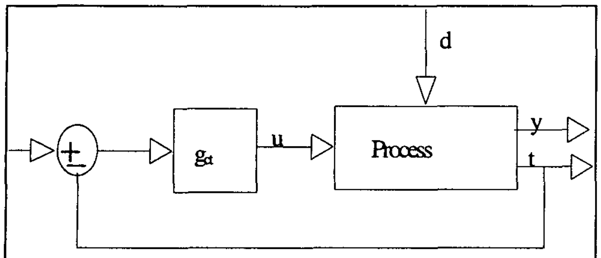

Fig. 4a. A classical approach. Control secondary variable

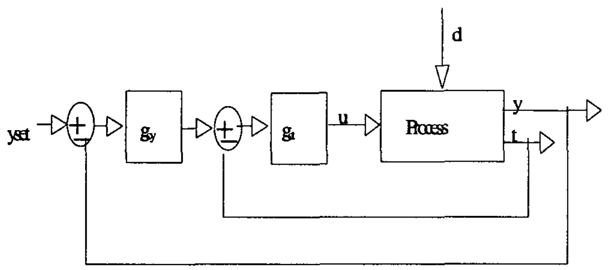

Fig. 4b. Structure of the cascade control system

Taking setpoint of t to be zero without loss of generality, the transfer function for disturbance rejection under perfect control of t can be derived as:

$$y(s) = \left(-\frac{g_{uy}(s)g_{dt}(s)}{g_{ut}(s)} + g_{dy}(s)\right)d(s) \quad (2)$$

Note that y(s) ≈0 if $g_{dt} \approx g_{dy}$ and $g_{uy} \approx g_{ut}$ i.e. when the disturbance d and manipulated variable u affect both t and y in a similar manner. If this is not so (which is often the case) this strategy may result in poor control.

In the cascade control strategy, shown in Figure 4(b), the inner loop tries to maintain the secondary variable at a set point which is adjusted by the outer loop to bring y back to its set point. This strategy is usually employed when there are significant delays and lags associated with measurement of y. To implement this strategy, we must have a measurement of the primary variable. The disturbance rejection transfer function is given by:

$$\frac{y}{d} = (g_{dy} + g_{uy}g_1) \quad (3)$$

where

$$g_1 = -\frac{g_{ct}g_{cy}g_{dy} + g_{ct}g_{dt}}{1 + g_{ct}g_{cy}g_{uy} + g_{ct}g_{ut}} \quad (4)$$

This strategy has the advantage that steady state error in control of y will be reduced to zero if we use integral action in the outer loop controller. But whether this strategy will work well in a process depends on a number of factors. The inner loop should be able to react fast enough to follow frequent set-point changes. If there are significant lags in the inner loop then the system will not have enough time to settle down, and control system performance will be poor. If the disturbances come in at a low frequency such that the outer loop has enough time to correct it, this structure might be acceptable. However if the disturbances come at a frequency that keeps the system from settling down then the controller on the outer loop will not have time to settle down. Because of the large delays involved in the measurement this will usually mean poor performance of the control system.

More importantly, both of the above strategies have no easy extension to the case of multiple secondary measurements. Usually multiple secondary measurements contain more information about the state of the system. Thus methods that use multiple measurements have an advantage over these multi-loop strategies.

2.2 Estimator Based Strategies:

2.2.1 Feedback control using a state estimator for y

In this strategy, an estimator for the unmeasured output y is built first which is then used in a feedback control mode. Figure 5 shows the structure and block diagram of this control strategy.

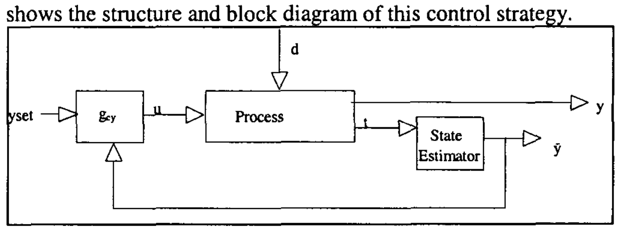

Figure 5. Structure of the controller using state estimator

The disturbance rejection transfer function in this case is given by (assuming perfect state estimation is possible)

$$\frac{y(s)}{d(s)} = \frac{g_{dy}(s)}{1 + g_{uy}(s)g_{cy}(s)} \quad (5)$$

Note that this is the same as what we get if we had direct feedback control on the primary measurement itself. The advantage with this approach is that an estimate of the unmeasured state y is available through the secondary measurements, which is useful for the operator. If $g_{cy}(s)$ and $g_{dy}(s)$ have large lags associated with them, then this can result in poor performance (since the optimum performance achievable using direct feedback control is limited by the timelags and time delays present in the feedback loop). This controller may be worse than direct feedback control on t in some cases if t responds faster to the disturbance and manipulated variable. In addition, state estimators are never perfect and introduce additional errors in the feedback control loop that necessitate detuning the controller to some extent. The construction of state estimators is not trivial.

2.2.2 Inferential Control

In the classical 2-degree of freedom IMC structure an estimate of the disturbance effect on y (Garcia and Morari, 1982; Morari and Zafiriou, 1989) is fed back as shown in Figure 6. The question inferential control tackles is how to generate $d_y = g_{dy}(s)d(s)$ if no measurement of y is available but only t(s) is available.

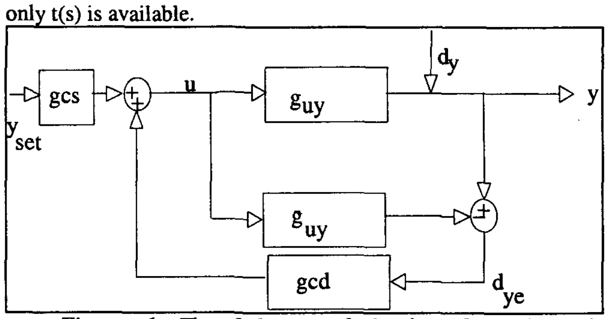

Figure 6. The 2-degree of freedom Internal Model Controller

In this case, we can first compute

$$\begin{aligned} d_t &= g_{dt}(s)d(s) \\ &= t(s) - g_{ut}(s)u(s) \end{aligned} \quad (6)$$

And then obtain an estimate of d(s) (denoted using d_e) as:

$$d_e(s) = g_{dt}^{-1}(s)[t(s) - g_{ut}(s)u(s)] \quad (7)$$

d_y can then be estimated as follows:

$$\begin{aligned} d_{ye}(s) &= g_{dy}(s)d_e(s) \\ &= g_{dy}(s)g_{dt}^{-1}(s)[t(s) - g_{ut}(s)u(s)] \end{aligned} \quad (8)$$

Using this equation we get the structure of the inferential control system as shown in Figure 7 (Tong and Brosilow, 1978).

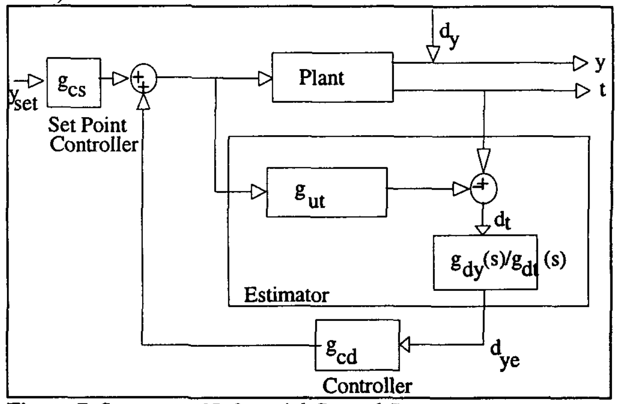

Figure 7. Structure of Inferential Control System

The design of the disturbance controller, $g_{cd}(s)$ is as in IMC. The output response is given by

$$y = g_{uy}u + d_y \quad (9)$$

To keep y close to the set point (= 0) we choose the controller so that

$$u = -g_{uy}^{-1}d_y \quad (10)$$

Using the estimate of d_y in place of d_y we get:

$$u = -g_{uy}^{-1}g_{dy}g_{dt}^{-1}\left(t - g_{ut}u\right) \quad (11)$$

The controller transfer function derived above may not be realizable since we have to invert the process transfer function. If the process transfer functions contain RHP zeros or time delays then we must add a filter f(s), designed as in the IMC controller, to make the controller realizable:

$$u = -f\,g_{uy}^{-1}g_{dy}g_{dt}^{-1}\left(t - g_{ut}u\right) \quad (12)$$

The disturbance rejection transfer function for the control scheme is given by:

$$\frac{y(s)}{d(s)} = g_{dy}(s)(1 - f(s)) \quad (13)$$

This transfer function is very similar to a feedforward controller. If the filter transfer function used is close to 1 then we have nearly perfect rejection of disturbances as in feed forward control. In general the filter must have lag terms to compensate for modeling errors and make the system robust.

It is important to differentiate the above structure from state estimation techniques like Kalman filter that uses t to predict y and then control y. Estimating y first and then designing a controller for it throws away the feedforward nature of the inferential control structure. If the secondary measurements respond faster to the disturbances then one can take faster corrective action and hence get better control system performance by using t in an inferential structure as illustrated in the example below. If the output y responds to disturbances slower than t, then the state estimator will add additional lag in the feedback loop and hence will adversely affect the performance of the control system. Usually it is possible to identify secondary measurements closer to the disturbance and hence in general the inferential control strategy is preferred.

Both the inferential control scheme and the state-estimation schemes outlined above can be extended to nonlinear and multivariable situations. In the following section, we extend the inferential control structure first to the model predictive framework, which is widely used in industry. This structure is then easily generalizable to nonlinear and multivariable cases.

3. Development of Model Predictive Inferential Control (MPIC) Scheme

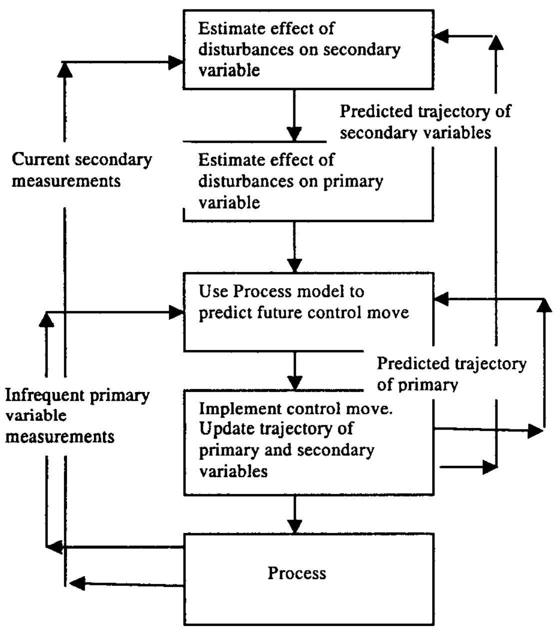

Figure. 8. Structure of MPIC

Model Predictive Control is used extensively in industry to control constrained, multivariable systems (Froisy, 1994). As pointed out by Garcia and Morari (1982), this can be put in the framework of IMC for unconstrained systems. It turns out that the above inferential control strategy is more easily extended to multivariable and nonlinear systems if implemented using the MPC framework. In this section, we develop this framework followed by its extension to multivariable and nonlinear systems.

3.1 Basic MPIC Algorithm

The MPIC algorithm can be summarized as follows. Given

- A predicted trajectory vector of the primary output based on past control actions $\mathbf{y^P} = \left[y_k^P, y_{k+1}^P, \ldots, y_{k+p}^P\right]^T$, and a predicted trajectory vector of the secondary variable based on past control actions $\mathbf{t^P} = \left[t_k^P, t_{k+1}^P, \ldots, t_{k+p}^P\right]^T$ where k is the current time and p is the length of output prediction horizon
- A desired primary output trajectory vector $\mathbf{y^r} = \left[y_k^r, y_{k+1}^r, \ldots, y_{k+p}^r\right]^T$
- Current measurement of primary output y_k^m and secondary variable of t_k^m
- A dynamic matrix $\underline{\underline{A}}$ relating the primary output and the input and dynamic matrix $\underline{\underline{C}}$ relating secondary variable and the input:

$$\mathbf{y} = \mathbf{y}^P + \underline{\underline{A}}\Delta\mathbf{u} + \mathbf{D_y}; \qquad \mathbf{t} = \mathbf{t}^P + \underline{\underline{C}}\Delta\mathbf{u} + \mathbf{D_t} \tag{14}$$

The controller problem is converted into an on-line optimization problem by defining the control objective as:

$$\underset{\Delta u}{Min} \sum_{i=1}^{p} \left\| y^r - y^p - \underline{\underline{A}}\Delta u - D_y \right\|^2 \tag{15}$$

Constraints on input and output variables and move sizes can be imposed while solving the quadratic minimization problem. The first step in the MPC algorithm is to estimate the effect of disturbance on the output variable (given as $\mathbf{D_y}$ in equation 14). In practice, the disturbance estimate is obtained from the difference between the prediction of the current value of y and the current measured value of y. This also takes care of any modeling errors as well. Then the current disturbance and model error effect at the output is computed using:

$$d_k = y_k^m - y_k^P \tag{16}$$

Estimating the future values of disturbances requires some assumptions. The simplest assumption which works well in most cases is that these are constant:

$$d_{y,k+1} = d_{y,k+2} = \ldots.. = d_{y,k} \tag{17}$$

Now suppose y is not measured directly, but only the secondary measurement, t, is available. Then we need a way to estimate $d_{y,k+1}$, $d_{y,k+2}$, ... $d_{y,k+p}$. This is precisely what the estimator derived in the previous section is designed to do. We might view the estimator as a dynamic system which is driven by input d_t and produces estimate d_y (see Figure 9).

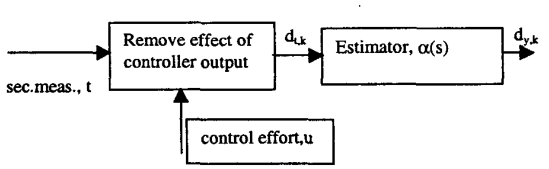

Figure.9. The estimator as a dynamic system

Let $b_1, b_2, b_3 \ldots b_p$ represent the finite step response model coefficients of the estimator α(s). Then $d_{y,k} \ldots d_{y,k+p}$ are estimated as follows:

$$
\begin{aligned}
d_{y,k+1} &= b_1 d_{t,k} \\
d_{y,k+2} &= b_2 d_{t,k} \\
&\ldots. \\
d_{y,k+p} &= b_p d_{t,k}
\end{aligned}
\qquad (18)
$$

Because of the dynamic nature of the estimator, these disturbance estimates are no longer constant. Recall that d_t is the effect of the disturbance on the secondary measurement t:

$$d_t(s) = t(s) - g_{ut}(s)u(s) \qquad (19)$$

$g_{ut}(s)$ can also be represented in discrete form using the step response coefficients:

$$
\begin{aligned}
\Delta t_{k+1} &= c_1 \Delta u_k \\
\Delta t_{k+2} &= c_2 \Delta u_k \\
&.. \\
\Delta t_{k+p} &= c_p \Delta u_k
\end{aligned}
\qquad (20)
$$

Then the disturbance estimate on t can be calculated as

$$d_{t,k} = t_k^m - t_k^P \qquad (21)$$

If there are no constraints, $\Delta \mathbf{u}$ is computed as least square solution to equation (15) and is given by

$$\Delta \mathbf{u} = (\underline{\underline{A}}^T \underline{\underline{A}})^{-1} \underline{\underline{A}}^T (\mathbf{y}^\mathbf{r} - \mathbf{y}^\mathbf{P} - \mathbf{D_y}) \qquad (22)$$

After this, it is necessary to update the prediction trajectory of t taking into account the current control action Δu_k and current disturbance effect on t.

$$
\begin{aligned}
t_{k+1}^P &= t_{k+1}^P + c_1 \Delta u_k + d_{t,k} \\
t_{k+2}^P &= t_{k+2}^P + c_2 \Delta u_k + d_{t,k} \\
&\ldots \\
t_{k+p}^P &= t_{k+p}^P + c_p \Delta u_k + d_{t,k}
\end{aligned}
\qquad (23)
$$

Before going to the next sample time, k+1, the $\mathbf{t^P}$ vector is updated to include the prediction at k+p+1 and discard the prediction at k (since we have a moving horizon and when we get to the next sample time we have to move the trajectory by one sample time forward into the future). This is accomplished by moving

$$t_{k+i}^P(new) \leftarrow t_{k+i+1}^P(old), i = 0,1,\ldots, p-1$$

and replacing $t_{k+p}^P(new)$ with $t_{k+p}^P(old)$ (the last predicted value of t). The entire MPIC algorithm can be summarized as shown in Figure 8.

3.2 Building of inferential estimator

In this section, we address the problem of constructing the estimator for the effect of disturbances using secondary measurements. This is essentially an identification problem. The problem can be stated as follows:

Given a set of past measurements $t_1, t_2, \ldots t_n$ and $y_1, y_2, \ldots, y_n$, construct an estimator for computing d_y from d_t.

We want to emphasize that the fact that we are not estimating y directly but d_y. First to estimate d_t, we need to subtract the effect of the manipulated variable on t. This will require a model relating the two variables. This is accomplished by conducting identification experiments on the plant to relate the manipulated variables to the primary and secondary variables. We could collect data on the plant operational characteristics over a long period of time and then try to relate the two quantities d_t and d_y through regression on the data. This problem may be stated as follows:

Given a set of operating data in the form of a table that tabulates y, t and u, find a linear relationship between the sampled values of d_y and d_t where

$$
\begin{aligned}
d_t &= t - g_{tu} u \\
d_y &= y - g_{yu} u
\end{aligned}
\qquad (24)
$$

Once a time-series of d_y and d_t are obtained as above, an impulse response model relating the two can be written as:

$$d_{y,n} = b_1 d_{t,n} + b_2 d_{t,n} + \ldots.. + b_{nb} d_{t,n-nb+1} \qquad (25)$$

where nb is the number of coefficients used in the FIR model. Put in matrix notation, the identification problem is to obtain a least-square solution to the following equation:

$$\underline{\underline{\mathbf{E}}}\mathbf{X} = \mathbf{D_y} \qquad (26)$$

where

$$\underline{\underline{E}} = \begin{bmatrix} d_{t,nb} & d_{t,nb-1} & . & . & d_{t,1} \\ d_{t,nb+1} & d_{t,nb} & . & . & d_{t,2} \\ . & . & . & . & \\ . & . & . & . & \\ d_{t,N} & d_{t,N-1} & . & . & d_{t,N-nb+1} \end{bmatrix},$$

$$\mathbf{X} = \begin{bmatrix} b_1 \\ b_2 \\ \\ \\ b_{nb} \end{bmatrix}, \quad \mathbf{D}_y = \begin{bmatrix} d_{y,nb} \\ d_{y,nb+1} \\ . \\ . \\ d_{y,n} \end{bmatrix}$$

The advantage with this representation is that prior knowledge about the time delay and/or the order of the system is not needed. One problem here is the fact that y_1, y_2, ... y_n may not be available if y cannot be measured at all. However, if y can be sampled and evaluated later (say in a quality control laboratory) or if sufficiently accurate first principles model is available, then such data can be generated. In a worst case scenario, one may have only steady-state measurements available and in such an instance, a steady-state estimator may be used as an approximation.

4. Practical Issues in Implementation of Inferential Control.

In this section we discuss some practical considerations in the implementation of inferential control. Issues addressed include use of multiple secondary measurements, effect of collinearity among the measurements, application to nonlinear systems and quality control in batch processes.

4.1 Extension to multiple secondary measurements

In most practical situations, more than one secondary measurement is available. A similar strategy may be derived for the case where a single primary variable is to be inferred from multiple secondary measurements. The only difference is that the matrix $\underline{\underline{E}}$ in this case will be augmented to contain the effects of the disturbance on each of h secondary measurements. The vector **X** to be identified will contain the impulse response coefficients between each of the secondary variables and the primary variable.

$$\mathbf{X}^T = [b_{11} \quad b_{12} \quad . \quad b_{1nb} \quad . \quad b_{h1} \quad b_{h2} \quad . \quad b_{h,nb}]$$

and

$$\mathbf{D}_y{}^T = [d_{y,nb} \quad d_{y,nb+1} \qquad\qquad d_{y,N}]$$

4.2 Effect of collinearity among the measurements

Using multiple measurements as inputs for the disturbance estimator has some advantages and some disadvantages. On the one hand, we have more information about the disturbances affecting the process and hence can take more accurate control action. However, this comes at the expense of increased complexity of the estimator and increased susceptibility to measurement errors and sensor failures. Also, the measurements do not respond to input disturbances in an independent way and are correlated. Principal component analysis (PCA) provides a way to compute estimators when the measurements are collinear, i.e. very similar to each other. PCA is used to eliminate nearly singular values and to generate more accurate estimators. The advantage with this approach is that all measurements are included in the estimator, thus maximizing the information content. In this case, PCA of equation 26 needs to be first performed, and the vector **X** is to be identified by including only the first few latent variables, which carry sufficient variance in the data. In the second case study below, such an approach was needed.

4.3 Extension to Nonlinear Systems

The above algorithm and control structure and control structure can also be used if the estimator and process model are both nonlinear. The estimator can be represented abstractly as:

$$d_{y,k} = g_{y,k}(t_k, u_k) \tag{27}$$

where g_y is a nonlinear function of the secondary measurements and the control efforts. Such a nonlinear estimator can be constructed from past measurements using nonlinear regressors such as recurrent artificial neural networks (RANN).

Similarly the process model may be constructed in the form

$$y_k = f_y(u_k, d_y, k) \tag{28}$$

Where f_y is a nonlinear dynamic system driven by current control effort and the current disturbance estimate $d_{y,k}$. Again, if no structure is known, RANN models may be constructed from past measurements.

The MPIC for such systems can be stated as follows:

1. Using current measurements, predict a trajectory for d_y into the future using equation 27.
2. Use the predicted disturbances along with the model to compute future control moves u_k, u_{k+1}, ...etc which will minimize the quadratic $\left\| y^P - y^r \right\|^2$,where y^P is the predicted trajectory and y^r is the desired trajectory. This will involve solution of a non-linear programming problem.
3. Implement control effort. Update trajectories.
4. At the next sample time, take another set of secondary measurements and return to step 1.

Figure 10 shows the schematic of the algorithm. For a more detailed description of how to apply inferential control to nonlinear processes see the paper by Voorakaranam and Joseph (1998).

4.4 Application to Quality Control in Batch Processes

Inferential control concepts can be extended to control of end product quality in batch processes. Note this in this instance the primary variable cannot be measured at all and control must be performed using intermediate measurements during the batch processing. In this case we build estimators to predict end product quality using the available set of intermediate measurements. Corrective control actions consist of changes to the batch recipe to bring the quality back to normal. Because of the shrinking nature of the control horizon, the model predictive control strategy must be modified also. For more details and a case study application of this shrinking horizon model predictive control (SHMPC) see the article by Thomas et al (1997).

4.5 Effect of Modeling Errors

Modeling errors can cause the performance of the inferential controller to deteriorate. Any error in estimation of the primary variable can result in a steady state offset in the control system just as in a feedforward control system. To overcome this problem if possible inferential control should be combined with a feedback control of the primary variable whenever a measurement of the primary variable is available. One simple way to incorporate feedback control is to make a linear correction to the estimate whenever the primary variable is measured. The second case study described below uses this strategy to overcome steady state error in control.
The usual rules about tuning a controller in presence of modeling errors applies to MPIC as well. If the model error is small the controller can be tuned tightly. However if the modeling or linearization error can be large then the controller should be detuned by increasing the move suppression factors.

5. A Case Study: Inferential Control of the Shell Distillation Column

Prett and Morari (1986) have provided a distillation column case study for evaluating control strategies. We use a subset of the problem to illustrate and evaluate the ideas presented above. The process model is:

$$y(s) = \frac{4.05e^{-27s}}{50s+1}u(s) + \frac{1.44e^{-27s}}{40s+1}d(s);$$
$$t(s) = \frac{3.66e^{-2s}}{9s+1}u(s) + \frac{1.27}{6s+1}d(s); \tag{29}$$

Note that this secondary variable responds much faster to both the manipulated variable and the disturbance variable than the primary variable. For direct feedback control of y, the tuning constants for a PID controller are obtained by applying IMC tuning rules to the transfer function between y and u (with $\tau_f = \tau/5 = 10$). Likewise, for feedback control of t, the transfer function between t and u is used to get the controller tuning constants (with $\tau_f = \tau/5 = 2$). Two controllers must be designed for the cascade control scheme. For the inner loop, the same controller designed above for the case of direct feedback control of t is employed. For the outer loop controller, a new step response obtained with the inner loop closed is used to design a PID using IMC tuning rules (with $\tau_f = 10$). The controller tuning constants for all these schemes are shown in Table 1.
Using equation (8), the estimate of d_y is given by

$$d_{ye} = \frac{1.44e^{-27s}}{40s+1}\frac{(6s+1)}{1.27}d_t \tag{30}$$

Control Scheme	Tuning constants
Direct Feedback on y	$K_c = 2.12, \tau_I = 50, \tau_D = 13.5$
Direct Feedback on t	$K_c = 0.81, \tau_I = 9, \tau_D = 1$
Cascade control (outer loop)	$K_c = 0.20, \tau_I = 50, \tau_D = 10$
Cascade control (inner loop)	$K_c = 0.81, \tau_I = 9, \tau_D = 0$

Table 1. Tuning Constants used in the classical control schemes

The controller is

$$g_{cd} = -g_{uy}^{-1} d_{ye}$$
$$= -.279\left(\frac{(50s+1)(6s+1)}{40s+1}\right) \tag{31}$$

To make it realizable we need to add a filter

$$g_{cd} = -.279\frac{(50s+1)}{(40s+1)}\frac{(6s+1)}{(\tau_f s+1)} \tag{32}$$

We can choose the filter time constant to be 1/5 of the numerator time constant or 1.2 following IMC tuning guidelines. The resulting control system is shown in Figure 11.

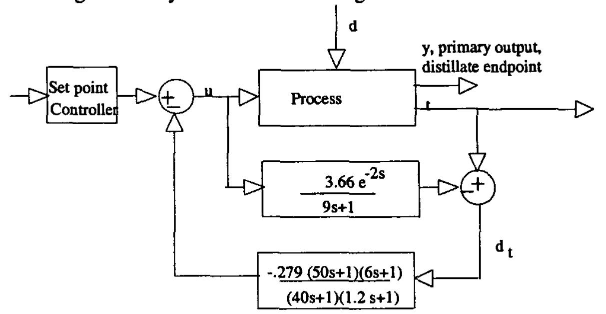

Figure 11. Inferential controller for the shell distillation column

Figures 12 and 13 shows the frequency and step response of controllers using the four different control schemes designed above. The excellent disturbance rejection capabilities of the inferential controller are seen clearly in these graphs. Note that because of the longer time constant and large time delay in the outer loop of the cascade scheme, this controller cannot be expected to respond quickly to input disturbances. Feedback control using t is actually better than direct feedback control of y.

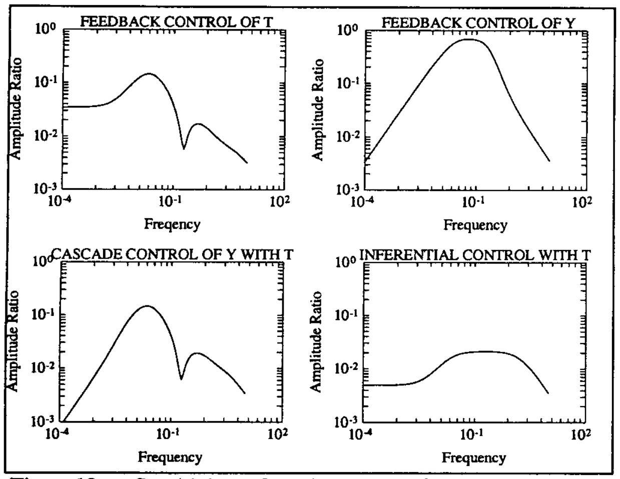

Figure12. Sensitivity of various control systems to input disturbance d. Amplitude ratio of y(s)/d(s) are plotted

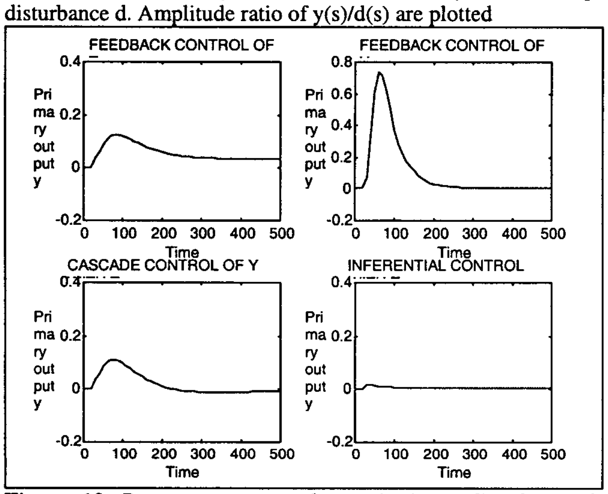

Figure 13. Response to step change in input disturbance for various controllers

6. Case Study: Application of MPIC to an Injection Pultrusion Process

In this section, we demonstrate the application of inferential model predictive control to a simulated Injection Pultrusion (IP) process. We first describe the process and present the control problem. The control strategy is then discussed with details about selection of inferential measurements, generation of dynamic matrices and building of the inferential estimator. We present closed loop results to validate the control scheme and discuss how infrequent measurements of primary variable can be used for better control. We also show how the predictive capabilities of the formulation can help the plant operator to better visualize the process behavior into the future.

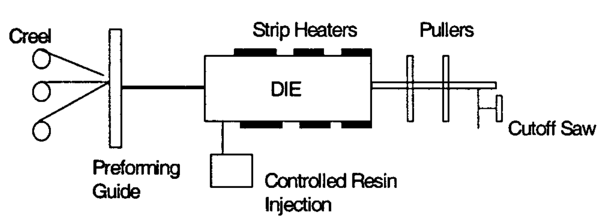

Figure 14. Injection pultrusion process

In IP, the open resin bath of traditional pultrusion is replaced by a machined cavity attached to the curing die, as shown schematically in Figure 14. Preformed dry fibers are continuously pulled into the heated die and resin is injected at high pressures through the cavity into the fiber bundle. The fully impregnated preform is then pulled into a multi-zone heating region of the die where the resin cures around the fibers forming a solid composite. The process variables of interest from a control system point of view are summarized in Table 2.

Variable	Description	Variable Type
T_1	Temperature in Zone 1	Measured secondary variable
T_2	Temperature in Zone 2	Measured secondary variable
T_3	Temperature in Zone 3	Measured secondary variable
q_1, q_2, q_3	Heater inputs to the die	Manipulated variables to control temperatures
V	Pull speed	Optimized manipulated variable
F	Pull Force	Measured secondary variable
P_{inj}	Injection Pressure	Manipulated input
a	Void Content	Infrequent quality measurement
P_B	Backflow pressure	Visually observed processing indicator
S	Mechanical Strength	Infrequent quality measurement
y	Exit degree of cure	Infrequently measured primary control variable

Table 2. Description of Process Variables

6.1 Control Problem

The control objective is to maximize production rates while maintaining 'consistently good quality' of the final part and meeting process constraints shown below. The quality variables of interest in a composite are properties such as degree of cure and void content which determine the part strength. The manipulated variables, and measurements associated with IP are shown in Table 2. The major control objectives are:

1. Control the cure at the exit of the die to the gelpoint of the resin. If the gelpoint occurs within the die, surface imperfections due to friction might appear. In a worse case, material might freeze up in the die and shut down production. Cure at the die exit shouldn't be below the gelpoint in order to avoid part breakage and reduction of strength of the composite. The gelpoint of the resin system is taken as 0.5.
Preventing backflow of resin from the front end of die is a major concern. This can happen if the injection pressure is too high for the pull velocity being used.
2. Avoid void formation which causes reduction of strength in the part. This can happen if the injection pressure is too low for the pull velocity being used.

3. Avoid high temperatures in the composite to prevent resin degradation.

The key disturbances affecting the process are the variations in reaction kinetics due to changes in the resin formulation.

6.2 Model-Predictive Inferential Control

An inferential control strategy which can detect the disturbance and take corrective action before the cure at the die exit itself starts falling steeply is preferable in order to minimize the amount of off-spec product formed. In the worst case scenario, such a strategy can serve as a useful tool to the plant operator to prevent catastrophic scenarios such as resin freezing up within the interior of the die.

In the MPIC approach, we have a steady-state linear program (LP) layer which uses a steady state model of the process to determine optimal operating conditions and an MPC layer as shown in Figure 15 which controls cure in presence of disturbances by manipulating the pull speed. As mentioned previously, this linear program tries to maximize the pull speed, and would drive the process to the upper bounds on the temperature setpoints and the lower bound on the injection pressure. For the MPC layer, the objective is to control the exit cure by manipulating the pull speed and injection pressure. Previous studies (Voorakaranam et al) show that the injection pressure has to follow a linear relationship with the pull speed primarily to prevent back flow and void formation. The injection pressure has negligible effect on the exit cure. Hence the control problem in the MPC layer simplifies to a single input single output problem.

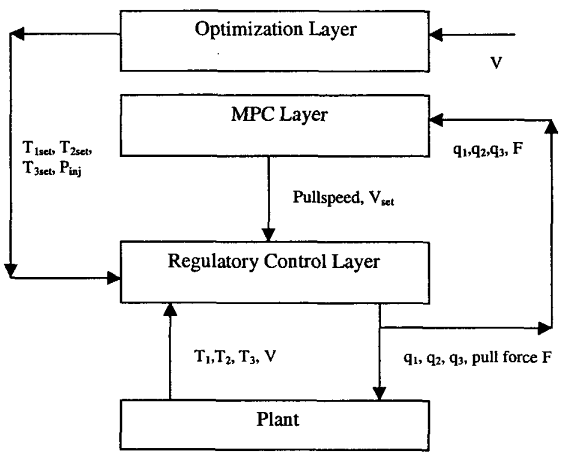

Figure 15. LP-MPC framework for IP Process

The MPC equations relating cure and pullspeed over a projection horizon P can be put in matrix notation as

$$\mathbf{y} = \mathbf{y}^{\mathbf{P}} + \underline{\underline{A}}\ \mathbf{u} + \mathbf{D_y} \quad (34)$$

where the effect of the disturbance on the cure, D_y can be estimated from the effect of the disturbance on the secondary flux measurements q_1, q_2 and q_3.

$$\mathbf{D_y} = \underline{\underline{G}}_{yq_1}\mathbf{D_{q_1}} + \underline{\underline{G}}_{yq_2}\mathbf{D_{q_2}} + \underline{\underline{G}}_{yq_3}\mathbf{D_{q_3}} \quad (35)$$

Similar to the equations for variation of primary variable, the equations for variation of secondary flux variable can be written as:

$$\mathbf{q_1} = \mathbf{q_1}^{\mathbf{P}} + \underline{\underline{G}}_{q_1 u}\Delta\mathbf{u} + \mathbf{D_{q_1}}$$
$$\mathbf{q_2} = \mathbf{q_2}^{\mathbf{P}} + \underline{\underline{G}}_{q_2 u}\Delta\mathbf{u} + \mathbf{D_{q_2}} \quad (36)$$
$$\mathbf{q_3} = \mathbf{q_3}^{\mathbf{P}} + \underline{\underline{G}}_{q_3 u}\Delta\mathbf{u} + \mathbf{D_{q_3}}$$

The effect of disturbances on the secondary variables can be obtained as:

$$d_{q_1,k} = q_{1meas,k} - q_{1,k}^{P}$$
$$d_{q_2,k} = q_{2meas,k} - q_{2,k}^{P}$$
$$d_{q_3,k} = q_{3meas,k} - q_{3,k}^{P} \quad (37)$$

All values of the disturbance effect after time k are assumed to stay constant after time k

$$d_{q_1,k+1} = d_{q_1,k+2} = \dots\dots\dots = d_{q_1,k+p}$$
$$d_{q_2,k+1} = d_{q_2,k+2} = \dots\dots\dots = d_{q_2,k+p}$$
$$d_{q_3,k+1} = d_{q_3,k+2} = \dots\dots\dots = d_{q_3,k+p} \quad (38)$$

The disturbances $d_{q_1,k}, d_{q_1,k+1}, \dots d_{q_2,k}, d_{q_2,k}, \dots\dots d_{q_3,k}, d_{q_3,k+1}, \dots$ are used to generate the vectors $\mathbf{D_{q_1}}, \mathbf{D_{q_2}}$ and $\mathbf{D_{q_3}}$:

$$\mathbf{D}_{q_1} = \begin{bmatrix} d_{q_1,k} \\ d_{q_1,k+1} \\ \vdots \\ d_{q_1,k+p} \end{bmatrix} \quad \mathbf{D}_{q_2} = \begin{bmatrix} d_{q_2,k} \\ d_{q_2,k+1} \\ \vdots \\ d_{q_2,k+p} \end{bmatrix}$$

$$\mathbf{D}_{q_3} = \begin{bmatrix} d_{q_3,k} \\ d_{q_3,k+1} \\ \vdots \\ d_{q_3,k+p} \end{bmatrix}$$

Step changes in the pullspeed are simulated keeping the value of disturbances constant in the model. The primary variables and secondary variables are recorded as function of time. The dynamic matrix $\underline{\underline{A}}$ relating the primary variable with the manipulated variable, and the dynamic matrices $\underline{\underline{G}}_{q_1u}, \underline{\underline{G}}_{q_2u}$ and $\underline{\underline{G}}_{q_3u}$ relating the three secondary variables with the manipulated variable can be estimated using standard identification procedures.

6.3 Building of Inferential Estimator

A step change is introduced in the activation energy (the disturbance variable) while keeping the pullspeed (manipulated variable) constant. The resulting profiles of the primary and secondary variables recorded as deviations from their base values. A dynamic model between the primary and secondary variables can then formulated in the form of equation (26). Based on the response time of the exit cure variable, fifty impulse response coefficients are identified for each secondary variable- primary variable relationship. The first *nb*=50 coefficients of solution vector X give the impulse response between the effect of the disturbance on the first secondary variable and the effect of disturbance on the primary variable. The next *nb* coefficients correspond to the second secondary variable. The last *nb* coefficients to the third secondary variable. The impulse response matrices can then be integrated to get the step response coefficient dynamic matrices $\underline{\underline{G}}_{yq_1}, \underline{\underline{G}}_{yq_2}$ and $\underline{\underline{G}}_{yq_3}$ respectively. A predictive horizon of 50, a control horizon of 1 and a sample time of 5 seconds is employed.

As mentioned earlier, a principal component analysis (PCA) needs to be performed first if the multiple measurements used as inputs to the estimator are correlated. A PCA is performed between the matrices $\underline{\underline{E}}$ and D given in equation 26. The first 5 principal components are retained as they provide sufficient estimation accuracy in terms of capturing the variance in the dependent and independent variables.

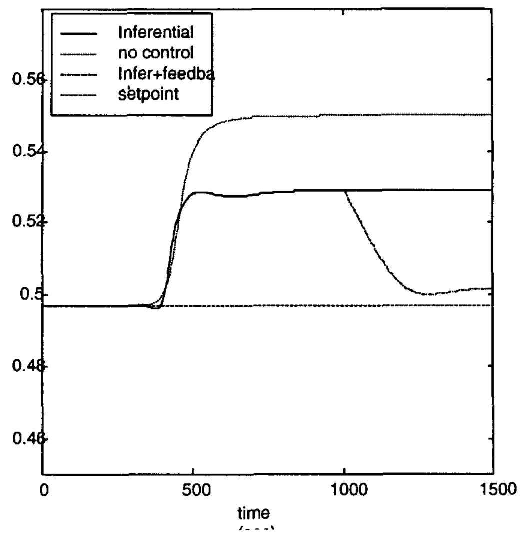

Figure 16. Performance of IFC with infrequent feedback

6.4 Results

Figure 16 shows the performance of the control system for a 25% increase in a second kinetic rate constant. This disturbance is qualitatively of a different nature than any of the disturbances considered above. It pertains to the reaction kinetics for a degree of cure greater than 0.3, and hence it's effect will only be felt towards the end of zone 2 when the degree of cure crosses this value. In contrast, the disturbances based on which our estimator had been built pertain to cure kinetics below a degree of cure of 0.3 and whose effect is felt starting right at the resin injection port. The MPIC still does a reasonable job in recovering from the disturbance but an offset is present. This is to be expected since estimator is not perfect. To get rid of the offset, it is necessary to introduce feedback correction based on cure. However this need not be available at the same sampling rate as our inferential control implementation rate. Even occasional measurements of cure will be able to substantially cancel out any steady-state errors. Such a situation is also demonstrated in the same figure, where a single measurement of the degree of cure available at time t=1000s is used to provide feedback correction for effective steady-state offset elimination. This example demonstrates that the importance of even infrequent quality measurements.

The reason for the good disturbance rejection is due to the fast response of heater inputs to the disturbance. The estimator is able to sense the effect of the disturbance on q and take anticipatory control action. The model predictive inferential control scheme works very well for a range of different types of disturbances which affect the primary variable in a similar way (for instance most disturbances entering through the resin feed).

A further use of the inferential MPC formulation is that at each instant of time we have a projection of how the process is going to behave in the future. Therefore, even if the operator chooses to disable automatic control and prefers manual control, he would still be able to get the future projected trajectories of the process which would enable him to make better decisions. Figure 17 shows such an instance of the snapshots of the exit degree of cure projected into a future time horizon of 400s at different instants of time (without any feedback control action manipulating V. The feedforward nature of the estimator clearly becomes apparent in the Figure 34 as we start getting very reasonable estimates of the process behavior into the future before the effect of the disturbance is felt at the exit of the die.

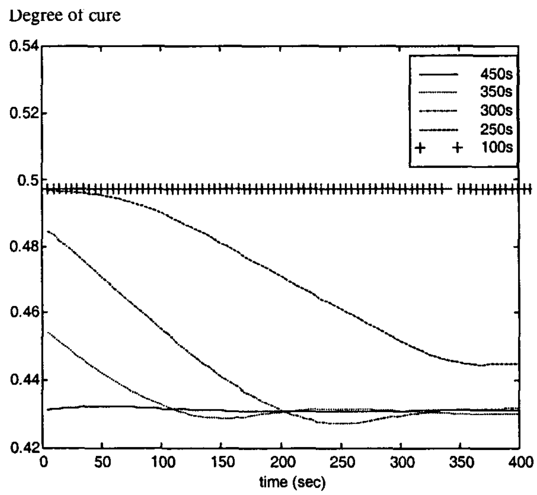

Figure 17. Snapshots of predicted future behavior of exit cure at different times

7. CONCLUSION

In this paper, we presented a model predictive inferential control strategy to tackle the problem of controlling process upsets due to unmeasured disturbances. A case study of the Shell challenge problem was used to highlight the advantages of the control scheme compared to conventional control strategies. We showed that inferential control is a powerful tool that can be applied to a wide variety of systems. When combined with the powerful framework of Model Predictive Control strategy, Inferential Control becomes highly versatile and can overcome the difficulties associated with the inability to measure the primary variable. Another often overlooked quality of inferential control is its feedforward nature, which makes it attractive to use even when the primary variable is measured.

NOMENCLATURE

$\underline{\underline{A}}$ = plant dynamic matrix between input and primary variable
$b_1, b_2, \ldots b_n$ = estimator coefficients
$\underline{\underline{C}}$ = dynamic matrix between input and secondary variable.
d = disturbance
$\mathbf{D}_y$ = vector of effect of disturbances on primary variable
$\mathbf{D}_t$ = vector of effect of disturbances on secondary variable
d_t =effect of disturbance on secondary variable t
$d_{q_1}, d_{q_2}, d_{q_3}$ = effect of disturbance on the heat inputs q_1, q_2 and q_3
d_y = effect of disturbance on primary variable y
d_{ye} = estimate of the effect of disturbance on y
$\underline{\underline{E}}$ = matrix of disturbance effects on secondary variable
f = filter transfer function
f_y = nonlinear plant model
g_{ct} = transfer function for secondary variable controller
g_{cy} = transfer function for primary variable controller
g_{dt} = transfer function between disturbance and secondary variable
g_{dy} = transfer function between disturbance and primary variable
g_{ut} = transfer function between input and secondary variable
g_{uy} = transfer function between input and primary variable
g_y = nonlinear estimator model
g_2 = linear model between maximum injection pressure and velocity
h = number of secondary measurements
k = current instant of time
nb = no of estimator coefficients
p = prediction horizon
t = secondary variable
$\mathbf{t}$ = vector of secondary variables
t^m = measured value of secondary variable
t^p = predicted value of secondary variable
$\mathbf{t^p}$ = vector of predicted values of secondary variable
u = manipulated variable (input)
Δu = step change in input
$\Delta \mathbf{u}$ = vector of step changes in input over control horizon
$\mathbf{X}$ = vector of estimator coefficients
y = primary variable (exit degree of cure for the IP process)
Δy = deviation in primary variable from steady-state.
$\bar{y}$ = estimate of primary variable y
y^P = predicted value of y
$\mathbf{y^P}$ = vector of predicted values of primary variable
y^r = setpoint on y
$\mathbf{y^r}$ = vector of reference setpoint trajectory of primary variable
y_{set} = setpoint on primary variable

Greek symbols

α =estimator
τ_f = filter time constant

REFERENCES

Abdel Jabbar, M.; Alatiqi I.M. Inferential-feedforward control of petroleum fractionators: a PNA approach. *Computers & Chemical Engineering*. **1997**,21(3), 255-62.
Beckman,J.J.; Lehman R.N.; Gauchel J.V.; Werner R.I.; Kusibab Z.; Kufner J.L.; Werner B.D. Injection die meets ladder quality criteria. *Modern Plastics*.**1995**, July, 59-62
Holcomb, T.R.; Morari, M. PLS/neural networks. Computers & Chemical Engineering. **1992**, 16(4), 393-411
Budman, H.M.; Holcomb T.; Morari M. PLS-based robust inferential control for a packed-bed reactor *Proceedings of the 1991 American Control Conference*., Boston, **1991**, 1, 256-61.
Froisy, J.B. Model Predictive Control: Past, Present and Future. *ISA Trans*., **1994**, 33, 235-243
Garcia C.E.; Morari M. Internal Model control. 1. A Unifying Review and Some New Results. *Ind. Eng. Chem. Process Des. Dev*., **1982**, 21(2), 308-323
Irwin, G.W.; Lightbody G.; O'Reilly P. Experience of neural networks for inferential estimation in industrial process control.

IEE Colloquium on 'Intelligent Measuring Systems for Control Applications, London, **1995**, p. 34, 1/1-6,

Joseph, B.; Brosilow, C.B. Inferential control of processes. I. Steady-State analysis and Design. *AIChE J.* **1978a**, 24(3), 485-492

Joseph, B.; Brosilow C.B. Inferential control of processes. III. Construction of optimal and suboptimal dynamic estimators. *AIChE J.* **1978b**, 24(3), 500-509

Koppel, V.; Segal M.; Remer M.; Lopano M.; Tonneguzzo N. Inferential Determination of moisture in solids in drying processes. *Proceedings of symposium on Low cost Automation.* Buenos Aires. **1995**, September, 81-6

Kresta, J.V.; T.E.Marlin T.E.; MacGregor J.F. Development of inferential process models using PLS. *Computers & Chemical Engineering.* **1994**, 18(7), 597-611

MacGregor, J.F.; Skagerberg B.; Kiparissides C. Multivariate statistical process control and property inference applied to low density polyethylene reactors. *IFAC Symposium on Advanced Control of Chemical Processes ADCHEM '91.* Toulouse, **1991**,October, 155-9, Toulouse

Mejdell, T.; and Skogestad S. Estimators for ill-conditioned plants: high-purity distillation. *IFAC Symposium on Advanced Control of Chemical Processes ADCHEM '91.* Toulouse. **1991**, 227-32

Morari, M.; Zafiriou, E. *Robust Process control*, Englewood cliffs, NJ: Prentice Hall. **1989**

Ogden-Swift. "Distributed controls 'the most under used resource' ". *Control & Instrumentation.* **1995**, 27 (5), 33-6

Parrish, J.; Brosilow C.B. Nonlinear inferential control of reactor effluent concentration from temperature and flow measurements. *Proceedings of the 1986 American Control Conference.* Seattle. **1986**, 2, 1027-33

Parrish, J.R.; and Brosilow C.B. Inferential control applications. *Automatica.* 1985, 21(5), 527-38

Prett, D.M.; M.Morari M. *The Shell Process Control Workshop*, Butterworths, Stoneham, MA, **1986**, 355-360

Thomas, M.M, B. Joseph and J.L.Kardos, " Batch Process Quality Control Applied to Curing of Composite Materials", AIChE Journal, 43, 10, pp. 2535-2545, 1997.

Tong, M.; Brosilow C.B. Inferential Control of Processes. II. The Structure and Dynamics of Inferential Control Systems", *AICHE.J.* **1978**, 24(3), 492-500

Voorakaranam, S. and B. Joseph, " Model Predictive Inferential Control with Application to Composites Manufacturing Process.", Paper submitted to Ind. And Eng. Chem. Res., July 1998.

Willis, M.J.; Di Massimo, C.; Montague, G.A.; Tham, M.T.; Morris, A.J. Inferential measurement via artificial neural networks. *IFAC symposium on Intelligent Tuning and Adaptive Control.* Singapore. **1991**,85-90

Ye, N.; McAvoy T.J.; Kosanovich K.A.; Piovoso M.J. Plant-wide control using an inferential approach. *Proceedings of the 1993 American Control Conference.* **1993**, 2, 1900-4

Acknowldgement

This work was supported in part by NSF grants DMI-96-22482 and CTS-95-29578.