Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

System Identification using Slow and Irregular Output Samples

Yucai Zhu* and Herman Telkamp

Abstract— **Identification of systems with very slowly and irregularly sampled output data is considered. While the input is sampled uniformly and frequently enough, the output sampling time is assumed very long, say, much longer than the dominant time constants. The output sampling may be irregular. Dual-rate system with slow output samples is a special case of the problem studied here. Model structures for the identification of the fast rate models will be proposed. Output error method is proposed and motivated for parameter estimation. Consistence of the estimate is established and covariance expression of the parameters is derived. Simulation studies will be used to illustrate the method. The developed identification method can also be used in other fields where measurement delay is a major problem.**

I. Introduction

Inferential models or soft sensors are often used in model predictive control (MPC) controllers to provide real-time prediction of the product compositions (qualities) that are measured at a much lower frequency than the control frequency. The quality sampling time is often more than 100 times longer than the sampling time of inputs and output sampling time may also vary. The problem may also occur in other fields. In process control community, a misconception exists that says: *if the output sample time is very long, say, greater than the system settling time, then it is impossible to identify the fast dynamic system using fast input data and slow output data.* It is based on a wrong intuition and/or a misunderstanding of the Shannon's Sampling Theorem. This may explain why most inferential models in the process control are static models. Ignoring dynamics in inferential modelling is one of the major causes of inaccuracy.

The problem has not been studied extensively, may due to the above-mentioned misconception. Telkamp (1994) studied the problem and proposed an output error method without theoretical analysis. If the output is sampled uniformly, the problem can be addressed in the dual-rate system identification where a popular solution is the so-called lifting technique; see, e.g., Li *et al* (2001, 2002, and 2003) and the references therein. In a listing technique, first a lifted model with higher dimension will be identified and then the original fast rate model is extracted. This technique is not suitable for the given problem because the resulting lifted model will have too high order and because the output sampling may be irregular.

Another approach is to create the fast rate output data by interpolation and then apply a single-rate identification method; see, e.g., Raghavan *et. al.* (2004). When the output data sampling time is much greater than relevant time constants, the interpolation error will be too large and this approach will result in erroneous models.

A third method is to use an auxiliary model to predict the fast rate output and then identify the fast rate model. Based on the idea, Ding and Chen (2004) have developed a recursive least-squares algorithm. The quality of their approach depends heavily on the quality of the auxiliary model.

In this work, we continue the work of Telkamp (1994) and propose an output error method to solve the problem. Consistency will be established and variance expression of parameter estimation will be given. Simulation studies are used to show the relevance of the method.

II. Statement of the Problem

An inferential model has a single output that is often a composition of a product; it has multiple inputs that are often temperatures, flows (or flow ratios) and pressures. The inputs are measured in real time at the sampling time of the MPC controller, which is often 1 minute. The model output, the composition, can only be sampled at a much longer sampling time, much longer than the dominant time constants, ranging from 30 minutes to 24 hours. The output may be sampled irregularly. Denote the fast inputs as $u_1(t)$, ..., $u_m(t)$ and fast output as $y(t)$. Assume that the fast inputs and output data are generated by a fast sampled linear multi-input single-output (MISO) system:

$$\begin{aligned} y(t) &= G_1(q)u_1(t) + ... + G_m(q)u_m(t) + v(t) \\ v(t) &= H(q)e(t) \end{aligned} \qquad (1)$$

where

Manuscript received February 22, 2005.
Yucai Zhu is with Control System Group, Department of Electrical Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands, y.zhu@tue.nl
Herman Telkamp is at Jan van Beverwijckstraat 104, NL-5017 JA Tilburg, The Netherlands, herman_telkamp@hotmail.com.
*Corresponding author. Also at Tai-Ji Control, Grensheuvel 10, 5685 AG Best, The Netherlands, y.zhu@taijicontrol.com.

$$G_i(q) = \frac{B_i(q)}{A_i(q)} = \frac{(b_1^i q^{-1} + ... + b_n^i q^{-n})q^{-d_i}}{1 + a_1^i q^{-1} + ... + a_n^i q^{-n}}$$

is the transfer function from $\underline{u}_i$ to $y(t)$ which is stable, d_i is the delay from ith input to the output, $v(t)$ is the unmeasured output disturbance and

$$H(q) = \frac{c_0 + c_1 q^{-1} + ... + c_n q^{-n}}{1 + d_1 q^{-1} + ... + d_n q^{-n}}$$

is a stable and invertible (minimum-phase) filter, and q^{-1} denotes unit delay operator. The disturbance source $e(t)$ is a white noise signal.

Denote θ as the parameter vector of process model $G_1(q),...,G_m(q)$

$$\theta = [a_1^1,...,a_n^1, b_0^1,...,b_n^1,......, a_1^m,...a_n^m, b_0^m,...,b_n^m]^T \qquad (2)$$

Then the system model transfer function can be written as $G_1(q,\theta),...,G_m(q,\theta)$ with

$$G_i(q,\theta) = \frac{B_i(q,\theta)}{A_i(q,\theta)}$$

The parameters of the disturbance model $H(q)$ of fast sampling cannot be identified due to slow and irregular sampling of the output.

The input sampling time T_u is the same as the system sampling time; the output sampling time is much greater than T_u ($T_y >> T_u$) and T_y is time-variant meaning that the output is irregularly sampled. The time stamp of the output is known. Here t_y denotes a sample of output. Then the system of slowly sampled output can be denoted as

$$\begin{aligned} y(t_y) &= \left[G_1(q)u_1(t) + ... + G_m(q)u_m(t) + v(t)\right]_{t = t_y} \\ v(t_y) &= \left[H(q)e(t)\right]_{t = t_y} \end{aligned} \qquad (3)$$

Note the output disturbance is not measurable. Denote the number of input samples as N_u and that of output as N_y. Note that $N_u >> N_y$. Denote the data sets as

$$Z = \{u_1(t), ..., u_m(t), t = 1, 2, ..., N_u; y(t_y), t_y = 1, 2, ..., N_y\} \qquad (4)$$

The system studied here is more general than dual rate systems with slowly sampled output. When assuming uniform output sampling, we will obtain a dual rate system.

System and Data Assumptions

- A1 The inputs $u_1(t)$, ..., $u_m(t)$ of process (1) are stationary stochastic processes sampled uniformly at the system sampling time T_u. Inputs are persistent exciting with any finite order and they are not linearly dependent.
- A2. The output $y(t_y)$ is sampled at a much longer sampling time T_y. T_y is a multiple of T_u. It may be obtained at irregular sampling times. Time stamps of the output samples are known. The output sampling time T_y is greater than the length of each process impulse response plus the length of the corresponding input autocorrelation function and T_y is also greater than the length impulse length of the disturbance filter.
- A3. The system operates in open loop, which implies that the disturbance $v(t)$ and inputs $u_1(t)$, ..., $u_m(t)$ are uncorrelated.
- A4. Model order n and time delays d_1, d_2,…, d_m, are correctly determined.
- A5. The models $G_1(q,\theta),...,G_m(q,\theta)$ are smooth functions of parameter vector θ.

The identification of model (1) is the determination of its order, delays and parameters $G_1(q,\theta),...,G_m(q,\theta)$ using data in (4). This work will focus on the determination of model parameters, or, parameter estimation. Note that the noise filter $H(q)$ cannot be obtained due to slow sampling.

III. OUTPUT ERROR METHOD

For a given model order n and given delays d_1, d_2,…, d_m, the problem of parameter estimation is to determine the value of model parameters using measured input-output data in (4) by minimizing a suitable error criterion or loss function.

In the spirit of prediction error method (Ljung, 1987 and Söderström and Stoica, 1989), the best choice of error criteria is to use the error at the point where a white noise disturbance enters the system. Because of the slow sampling as stated in assumption A2, the output disturbance $v(t_y)$ becomes a white noise sequence. Therefore, an output error criterion is a natural choice for parameter estimation.

The output error is defined as

$$\varepsilon_{oe}(t_y) = y(t_y) - \left[\frac{B_1(q,\theta)}{A_1(q,\theta)}u_1(t) + ... + \frac{B_m(q,\theta)}{A_m(q,\theta)}u_m(t)\right]_{t = t_y} \qquad (5)$$

And the output error method estimates parameters by minimizing the loss function

$$V_{OE}^{N_y} = \frac{1}{N_y}\sum_{t_y=1}^{N_y}\varepsilon_{oe}(t_y)^2 \qquad (6)$$

Because the output error $\varepsilon_{oe}(t_y)$ is *nonlinear* in the parameters of denominator polynomials, there exists no analytical solution to this minimization problem. Therefore, a numerical search algorithm is needed to find a minimum. Problems such as local minima and non-convergence can occur.

A Minimization Algorithm

There exist a large variety of numerical methods for optimization like Newton-Raphson, Marquardt, and others. Here we will derive the Gauss-Newton method. Denote $\hat{\theta}^k$ as the estimate at iteration *k*. Assume that $\hat{\theta}^k$ is close to a local minimum. For parameter vectors close to $\hat{\theta}^k$ we can approximate the output error residual by truncating the higher order terms of its Taylor expansion so that

$$\begin{aligned}\varepsilon_{\text{oe}}(t_y,\theta) &\approx \varepsilon(t_y,\hat{\theta}^k) + \frac{\partial \varepsilon(t_y,\theta)}{\partial \theta^T}\Big|_{\theta=\hat{\theta}^k}(\theta-\hat{\theta}^k) \\ &= \varepsilon(t_y,\hat{\theta}^k) + \varphi^k(t_y)(\theta-\hat{\theta}^k) \\ &= -[\varphi^k(t_y)\hat{\theta}^k - \varepsilon(t_y,\hat{\theta}^k)] + \varphi^k(t_y)\theta\end{aligned} \tag{7}$$

Here $\varphi^k(t_y)$ is the gradient of the output error residual (5)

$$\varphi^k(t_y) = \frac{\partial \varepsilon(t_y,\theta)}{\partial \theta^T}\Big|_{\theta=\hat{\theta}^k} \tag{8}$$

Under this approximation, the error is linear in the parameters. Hence the least-squares method can be used to find the refined estimate. Then the new estimate is given by

$$\begin{aligned}\hat{\theta}^{k+1} &= \left[\sum_{t_y=1}^{N_y}[\varphi^k(t_y)]^T\varphi^k(t_y)\right]^{-1}\sum_{t_y=1}^{N_y}[\varphi^k(t_y)]^T[\varphi^k(t_y)\hat{\theta}^k - \varepsilon(t_y,\hat{\theta}^k)] \\ &= \hat{\theta}^k - \left[\sum_{t_y=1}^{N_y}[\varphi^k(t_y)]^T\varphi^k(t_y)\right]^{-1}\varepsilon(t_y,\hat{\theta}^k)\end{aligned} \tag{9}$$

Some conditions should be satisfied in order to make the algorithm work. First the models from all iterations must be stable, because otherwise the elements of $\varphi^k(t_y)$ will be unbounded. If an unstable intermediate model is obtained, we can approximate it by a stable one. Secondly $\hat{A}^k(q)$ and $\hat{B}^k(q)$ must have no common factors, which is true when the model order is not higher than the true order. This will ensure that matrix $\left[\sum_{t_y=1}^{N_y}[\varphi^k(t_y)]^T\varphi^k(t_y)\right]$ is invertible when inputs are persistent exciting. These conditions are not restrictive for practical applications.

The Gauss-Newton method is simpler than other optimization methods mentioned above and it is numerically more reliable (Söderström and Stoica, 1989, Chapter 7.)

Initialisation

In order to obtain accurate estimate using the optimization scheme, a good initial estimate should be provided. Two methods are proposed for obtaining an initial model:

1) Estimate a steady gain plus delay model using slowly sampled inputs and the output. Then determine one or two time constants for each model using process knowledge. This method is suitable for situations where the output sampling time is much greater than the input sampling time.
2) Interpolate the output using some technique to create fast sampled output samples. Then estimate a fast sampled model using a single rate identification method, such as output error method, ARMAX method or Box-Jenkins method. This method is suitable for situations where the output sampling time is not much greater than the dominant time constant.

It remains an interesting topic to search for other methods for initial estimate.

Theorem 1. *Given the system (1) - (4), with assumptions A1 – A5 hold. Denote*

$$\hat{\theta} = \arg\min_{\theta} V_{OE}^{N_y}$$

as the parameter estimate that minimizes the output error criterion and $\hat{G}_1(q,\hat{\theta}),\dots,\hat{G}_m(q,\hat{\theta})$ *as the corresponding transfer function estimates. Denote* $G_1^o(q,\theta),\dots,G_m^o(q,\theta)$ *as the true transfer functions of the system and* θ^o *as the true parameter vector. Then*

1) The output error model is consistent meaning that

$$\hat{G}_1(q,\hat{\theta}) \to G_1^o(q,\theta),\dots,\hat{G}_m(q,\hat{\theta}) \to G_m^o(q,\theta) \text{ as } N_y \to \infty \tag{10}$$

2) The parameter errors follow an asymptotic Gaussian distribution

$$\sqrt{N_y}(\theta^o - \hat{\theta}) \underset{N\to\infty}{\overset{dist}{\rightarrow}} N(0,P_{oe}) \tag{11}$$

with the covariance matrix given by

$$P_{oe} = \sigma^2[E\varphi^T(t_y)\varphi(t_y)]^{-1} \tag{12}$$

where σ^2 *is the variance of* $v(t_y)$ and

$$\varphi(t_y) = \frac{\partial \varepsilon(t_y,\theta)}{\partial \theta^T}\Big|_{\theta=\theta^o} \tag{13}$$

3) The output error model is a minimum variance estimate, meaning that if θ^* *is another estimate, then the following always holds*

$$\operatorname{cov}\theta^* \ge \operatorname{cov}\hat{\theta} \tag{14}$$

A Simplified Proof. For notational simplicity, we will only prove the case of single-input single-output (SISO) system. The proof for MISO systems only needs notation changes.

Denote the model error as

$$\Delta G(q,\hat{\theta}) = G^o(q,\theta) - \hat{G}(q,\hat{\theta})$$

Then

$$\varepsilon_{oe}(t_y,\hat{\theta}) = [\Delta G(q,\hat{\theta})u(t)]_{t=t_y} + v(t_y) \tag{15}$$

Because the output sampling time T_y is very large (see A2), both terms in the right hand side of (15) are white noises and hence stationary processes. Therefore, the output error $\varepsilon_{\text{oe}}(t_y,\hat{\theta})$ is stationary. Then when $N \to \infty$, its sample variances converge to the corresponding expected values:

$$\begin{aligned} V_{\text{oe}}^{N_y} \to EV_{\text{oe}} &= E\left[[\Delta G(q)u(t)]_{t=t_y} + v(t_y)\right]^2 \\ &= E\left[[\Delta G(q)u(t)]_{t=t_y}\right]^2 + E[v(t_y)]^2 \end{aligned} \tag{16}$$

The last equality in (16) is due to an open loop experiment. If the minimization finds the global minimum, we have

$$\Delta G(q) \to 0 \text{ as } N \to \infty$$

which implies that

$$\hat{G}(q,\hat{\theta}) \to G^o(q,\theta) \text{ as } N \to \infty$$

Hence the output error model is consistent.

To prove the asymptotic normal distribution and covariance matrix, note that the output error method is a special case of the predictions error method studied in Ljung, (1987, Chapter 7, 8 and 9) and same conditions are met for the output samples t_y. Hence we can conclude that the vector $\sqrt{N_y}(\theta^o - \hat{\theta})$ follows asymptotically a Gaussian distribution with covariance matrix given by (12).

Minimum variance is a property of the prediction error method.

Remark 1: The consistence can also be proved when the output sampling time T_y ($> T_u$) is not greater than the length of impulse responses and the length of autocorrelations mentioned in A2. The proof is the same if the output sampling is uniform. The proof for irregular output sampling is technically more involved and will not be shown here.

Remark 2: The minimum variance property implies optimality of the output error method. This nice the consequence of A2 which makes the slowly sampled disturbance $v(t_y)$ a white noise sequence. Even when T_y is not greater than the length of impulse responses and the length of autocorrelations mentioned in A2, $v(t_y)$ will still be almost white and the output error estimate will be close to a minimum variance estimate.

Remark 3: For finite output samples N_y, the covariance of the parameter vector can be approximated as

$$\text{cov}\,\hat{\theta} \approx \frac{1}{N_y}\sigma^2[E\varphi^T(t_y)\varphi(t_y)]^{-1} \tag{17}$$

which is proportional to $1/N_y$ and not to $1/N_u$. Therefore, in order to reduce the variance error, one needs to increase the number of output samples N_y.

Remark 4: It is fortunate that the output error has nice properties for identifying systems with slowly sampled output. Other prediction error models used for uniform sampled systems, such as ARX, ARMAX, and Box-Jenkins models, cannot be used here due to slow and irregular output samples.

IV. Simulation Studies

Noise-free data

We will firstly use noise-free data to show that the proposed model is able to fit the dynamic process almost perfectly and to show the errors of using a static model.

The process is noise free and is given as

$$y(t) = \frac{0.2q^{-1}}{1-0.94q^{-1}}u_1(t) + \frac{q^{-1}+0.5q^{-2}}{1-1.5q^{-1}+0.7q^{-2}}u_2(t) \tag{18}$$

The inputs $u_1(t)$ and $u_2(t)$ are generated using low-pass filtered white noises

$$u_1(t) = \frac{0.1}{1-0.98q^{-1}}e_1(t),\ \ u_2 = \frac{0.1}{1-0.978q^{-1}}e_2(t) \tag{19}$$

where $e_1(t)$ and $e_2(t)$ are two white noises independent with each other. The process is simulated for 120000 minutes, or 83.3 days. The input sampling time is 1 minute and the output sampling time is 4 hours (240 minutes). Thus, 500 output samples are available. The first 400 output samples were used in model identification, the later 100 output samples are used for simulation and validation.

The model identification computes the parameters of two low order plus delay models using the 96000 samples of 1 minute input data and 400 samples of 4 hour output data. The identified step responses and output simulation are shown in Figure 1. We see that for noise-free data, the model and simulated output fit the process almost perfectly. This shows the feasibility of identifying dynamic models using slowly sampled output data such as lab data.

The same noise-free data were used to identify a static model. The results are shown in Figure 2. We can see that the static model cannot describe the dynamic behaviour of the process. The error variance of the simulation is about 30% of the output!

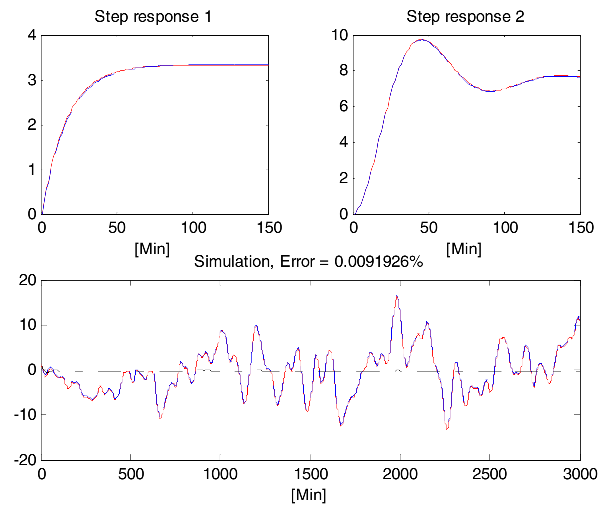

Figure 1. Dynamic model identified using noise-free data. Upper part: Step responses of the process (red/solid) and of the model (blue/dashed). Lower part: Process output (red/solid), model simulated (blue/dashed) and simulation error (black/dashdot)

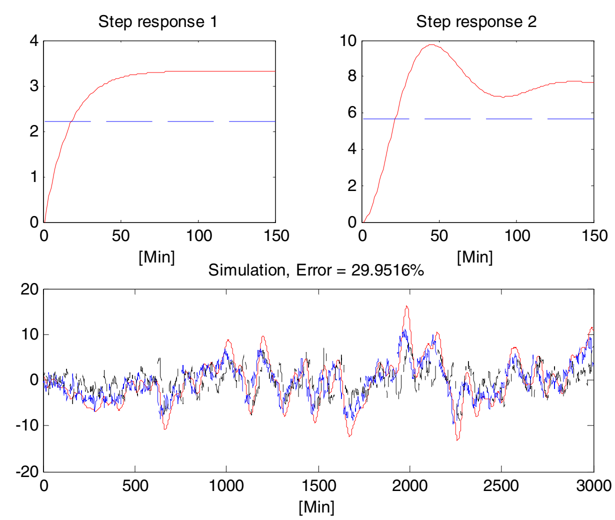

Figure 2. Static model identified using noise-free data. Upper part: Step responses of the process (red/solid) and of the model (blue/dashed). Lower part: Process output (red/solid), model simulated (blue/dashed) and simulation error (black/dashdot)

Noisy data

The process and inputs are the same as given in (18) and (19), but with unmeasured disturbance added at the output. The disturbance $v(t)$ is generated by low-pass filtering a white Gaussian noise $e(t)$. The filter is given as

$$v(t) = \frac{\alpha}{1 - 0.92q^{-1}} e(t) \tag{20}$$

The constant α is adjusted so that the noise variance at the output is 8.7% of the measured output.

The results of dynamic model identification are shown in Figure 3 and that of static model are shown in Figure 4. We can see that dynamic model outperforms static model considerably for the noisy data.

Note that we have compared the fast sampled output with the model simulation. This is only possible in simulation studies. In practice, the comparison can only be done for the slowly samples lab data. The simulation errors for the slow (lab) data have also been checked and they are similar to the simulation errors of the fast sampled data.

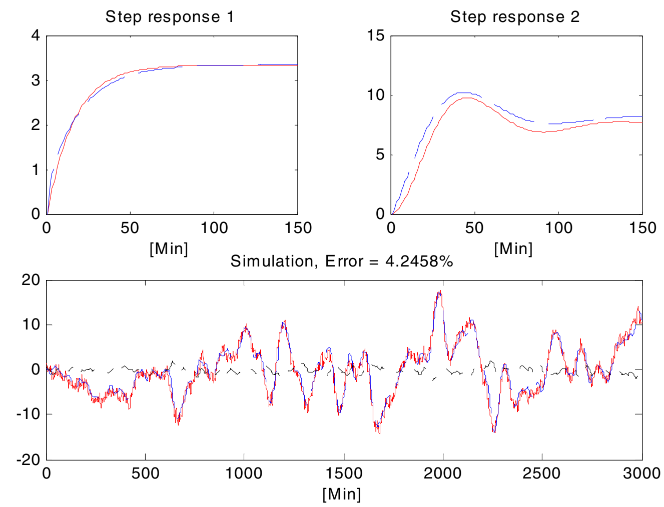

Figure 3. Dynamic model identified using noisy data. Upper part: Step responses of the process (red/solid) and of the model (blue/dashed). Lower part: Process output (red/solid), model simulated (blue/dashed) and simulation error (black/dashdot)

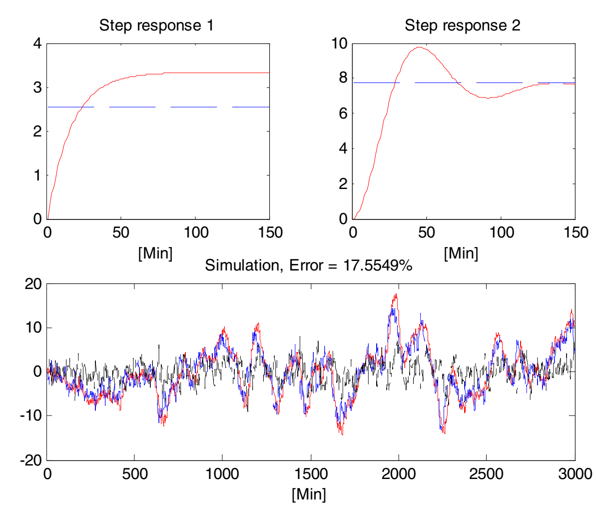

Figure 4. Static model with delays using noisy data. Upper part: Step responses of the process (red/solid) and of the model (blue/dashed). Lower part: Process output (red/solid), model simulated (blue/dashed) and simulation error (black/dashdot)

V. Conclusion and Discussion

System identification for slowly and possibly irregularly sampled output has been studied. The problem often arises in inferential modelling or soft sensor where measurement of certain system output variables contain significant delays. It has been shown that the proposed output error method is a natural approach that will produce most accurate model estimate; namely, the obtained model is consistent and minimum variance. The proposed method can be used for both uniformly sampled output (dual-rate system) and irregularly sampled output. The effectiveness of the method has been demonstrated using simulation examples.

The concept of soft sensor and the proposed method can also be used in other field (rather than process control) for predicting variables where measurement delay is a major problem, for example, in economic forecast and in biomedical engineering.

References

[1] Li, D., S.L. Shah and T. Chen (2002). Analysis of dual-rate inferential control systems. *Automatica*, Vol. 38, pp 1053-1059.

[2] Li, D., S.L. Shah and T. Chen (2001). Identification of Fast-rate Models from Multirate Data. *International J. of Control*, Vol. 74, Num. 7, pp 680-689.

[3] Li, D., S. L. Shah, T. Chen, and K. Z. Qi (2003). Application of Dual-Rate Modeling to CCR Octane Quality Inferential Control, *IEEE TCST*, Volume 11, Number 1, pp 43-51.

[4] Ljung, L. (1987). *System Identification: Theory for the User*. Prentice-Hall, Englewood Cliffs, N.J.

[5] Raghavan, H., R.B. Gopaluni, S.L. Shah, J. Pakpahan, R.S. Patwardhan and C. Robson (2004). Gray-box identification of dynamic models for the Bleaching operation in a BCTMP mill. *Journal of Process Control*, accepted.

[6] Sheng, J., T. Chen, and S.L. Shah (2002). Generalized predictive control for non-uniformly sampled systems. *Journal of Process Control*, vol.12 pp. 875-885.

[7] Söderström, T. and P. Stoica (1989). *System Identification*. Prentice-Hall, New York.

[8] Telkamp, H. (1994). Identification of processes with incomplete output data. Personal correspondence.