93_60



SYSID '94

10th IFAC Symposium on System Identification
4 - 6 July 1994

Editors:

Mogens Blanke and Torsten Söderström

Preprints Vol. 1

Published by Danish Automation Society 1994

93-60 C

COMPARISON OF THREE CLASSES OF IDENTIFICATION METHODS

Yucai Zhu*, Peter Van Overschee‡, Bart De Moor¶ and Lennart Ljung§

* SETPOINT IPCOS bv, De Waal 32, NL-5684 PH Best, The Netherlands
Tel.: +31.4998.64412, fax: +31.4998.97820

‡ Research Assistant of the Belgian National Fund for Scientific Research
¶ Research Associate of the Belgian National Fund for Scientific Research
ESAT, Dept. EE, Katholieke Universiteit Leuven
Kardinaal Mercierlaan 94, B-3001 Leuven (Heverlee), Belgium
Tel.: +32.16.220931, fax: +32.16.221855
e-mail: perter-vanoverschee@esat.kuleuven.ac.be

§ Dept. EE, Linköping University
S-581 83 Linköping, Sweden,
Tel. +46.13.281310, fax +46.13.139282
e-mail: ljung@isy.liu.se

Abstract. Three linear discrete-time identification methods are compared in this work. First, comparing criteria are proposed from an application viewpoint. Then the methods, namely, a subspace method, the prediction error method family and an asymptotic method are briefly introduced. All the methods are used to identify two multivariable industrial processes. The comparison is made using the criteria and based on the performances of the methods for the real life data. The purpose of this work is to make some contribution in unifying the field of identification and in bridging the gap between theory and application.

Key words. Identification, MIMO processes, order/structure selection, model predictive control

1 Introduction

System (process) identification is a very diversified discipline, many theories and methods have been developed and studied, numerous applications are reported. Identification methods that can treat multivariable (MIMO) problems hold high promise for industrial process control applications such as model predictive control (MPC). Academic researchers have paid relatively little attention to the problem of identification for MPC. Often rather primitive methods are used by process industries.

Therefore, our intention in this work is to narrow this gap between theory and application. We will propose some criteria for measuring the quality of identification methods for use in MPC (Section 2). Then we put our own methods under test. The methods, namely the family of prediction error methods (Ljung, 1987), the subspace methods (N4SID) (Van Overschee and De Moor, 1994) and the asymptotic method (Zhu, Backx and Eykhoff, 1991) are introduced in Section 3. The three methods have many things in common: they use time domain, discrete-time data; they are based on a stochastic frame- work and can supply both process model and disturbance model; and they can identify multivariable processes. Hence they are good candidates for MPC applications. However, these methods are developed from different viewpoints, they use different model parametrizations, different model order/structure selection rules and different numerical algorithms. In Section 4 the data from two multivariable industrial processes, a glass tube drawing process and a four effect (stage) evaporator, are used to test the methods. In order to make a fair play, each method is used by its author(s) to identify the processes. The methods are compared in Section 5. Section 6 contains the conclusions.

2 How to Evaluate Multivariable Identification Methods

From an application point of view, an identification method should include at least input design, model order/structure selection and an estimation algorithm. The criteria for measuring the quality of an identification method should be determined by the intended model use. Here we fix our goal of identification for model predictive control (MPC). In the following the criteria are proposed and briefly discussed.

1) **Model accuracy.** In general, we can say that a more accurate model will result in better control performance. Model accuracy is defined here as the simulation and prediction accuracy for the validation data set which has not been used in model estimation.
2) **Speed and robustness in model estimation.** Here speed is related to the computing time needed for calculating a model. Of course, a fast algorithm is always an advantage, but minor differences in speed are not that important for a small process with only few inputs and outputs. The speed of an algorithm will become important when the identified process has more than 10 inputs and 10 outputs, which is not uncommon in a refinery. A numerically robust algorithm does not suffer from round-off errors, local minima, convergence problems.
3) **Model order/structure selection.** This is a very important issue and it contains several problems.
— *Complexity.* Does the order/structure selection need exhaustive test for each order/structure, and does each candidate model need to be estimated?
— *Robustness against undermodelling and disturbance.* Does the criterion for model order/structure selection function well in face of disturbances and when the process dynamics are not contained in the model set?

— *Possibility of automation.* Is the selection method easy to use by a non expert user, e.g., can it be done systematically, or, can the selection procedure be automated?
— *MIMO versus MISO model structure.* For a diagonal form matrix fraction description (MFD) of a linear process, each output can be treated separately as a MISO (multi-input single-output) process. The problem of (possible) non minimality can be solved by model reduction. However, some researchers believe that by this approach one cannot capture the multivariable nature of the process. We hope that the case studies in Section 4 will be helpful for clarifying this matter.

4) Model uncertainty description. The errors of identified models are not negligible for industrial MPC applications. Therefore, being able to quantify model errors of obtained models for use in model validation, in robust control analysis and design is a very strong point.

5) Optimal input design for control. Input signals have a important influence on model quality for control. The requirement here is to find input test signals so that the identified model is optimal or suboptimal for MPC applications (therefore, not for minimizing parameter errors).

6) Closed loop identification, unstable processes. Closed loop identification experiments have many practical advantages and sometimes it is a must in industrial process control. Recent research shows that a model from closed loop data is more suitable for closed loop control than that from open loop data. The requirement here for an identification method is to maintain the model quality, e.g., consistency, for closed loop data. An integrator is the most common unstable dynamics encountered in process industries.

7) Recursive estimation. For the purpose of adaptive control and online monitoring, it is desirable to have a recursive version of the estimation algorithm.

3 The Identification Methods

Consider a sampled process with m inputs and p outputs. A general linear time-invariant model for the relationship between inputs and outputs can be written

$$y(t) = G(q)u(t) + v(t) \tag{3.1}$$

where

$$G(q) = \sum_{k=1}^{\infty} G_k q^{-k}$$

is called the transfer function of the process model; q^{-1} is the unit delay operator; $y(t)$ is the p-dimensional column output vector at time t; $u(t)$ is the m-dimensional column input vector at time t; $\{G_k\}$ is the impulse response of the process, which is a sequence of $p \times m$ matrices; $\{v(t)\}$ is a p-dimensional stochastic stationary process with zero mean values.

The frequency response matrix of the model is defined as

$$G(e^{i\omega}) = \sum_{k=1}^{\infty} G_k e^{-i\omega k} \qquad -\pi \le \omega \le \pi \tag{3.2}$$

The identification problem is to estimate an approximate model of the process and of the disturbances from observed input/output data. Denote the data sequence by Z^N:

$$Z^N := y(1), u(1), \ldots \ldots, y(N), u(N) \tag{3.3}$$

where N is the number of samples of the data sequence.

3.1 The Family of Prediction Error Methods (PEM)

The family of prediction error methods is the most mature part of identification theory that covers most of the identification methods developed from 1960 to 1980. A unified treatment of the methods can be found in Ljung (1987).

If we have a parametrized model:

$$y(t) = G(q,\theta)u(t) + H(q,\theta)e(t) \tag{3.4}$$

where θ is a d-dimensional parameter vector, then a common way to determine the parameters is to minimize the squared sum of the prediction errors

$$\hat{\theta} = \arg \min_{\theta \in D_n} \frac{1}{N} \sum_{t=1}^{N} \varepsilon^T(t,\theta)\varepsilon(t,\theta) \tag{3.5}$$

where $\varepsilon(t,\theta)$ is the prediction error vector:

$$\varepsilon(t,\theta) = H^{-1}(q,\theta)[y(t) - G(q,\theta)u(t)] \tag{3.6}$$

and $D_n \subset R^d$ is the parameter space.

The general MIMO black-box model is defined as (Ljung, 1987):

$$A(q)y(t) = F^{-1}(q)B(q)u(t) + D^{-1}(q)C(q)e(t) \tag{3.7}$$

where $A(q)$, $B(q)$, $C(q)$, $D(q)$ and $F(q)$ are polynomial matrices with dimension $p \times p$, $p \times m$, $p \times p$, $p \times p$ and $p \times p$ respectively. The parameter vector θ is formed from the coefficients of the polynomial matrices.

Specific methods can be obtained by taking a specific model parametrization. Examples are:
— Least-squares (LS) method (equation error method or ARX model):

$$A(q)y(t) = B(q)u(t) + e(t)$$

— Generalized least-squares method:

$$A(q)y(t) = B(q)u(t) + D^{-1}(q)e(t)$$

— Maximum likelihood method (or ARMAX model):

$$A(q)y(t) = B(q)u(t) + C(q)e(t)$$

The prediction error method family can also in a straightforward way be applied to arbitrarily parametrized state space models. The whole class of prediction error methods, applied to arbitrarily linear model structures is well known and they are implemented in Matlab's Identification Toolbox (Ljung, 1991). A nonlinear minimization algorithm is needed for calculating all of the model structures, except for the ARX model (LS method). An iterative Gauss-Newton algorithm is used. The initial estimate is obtained by a special four stage LS and instrumental-variable (IV) method. It can work with multivariable canonical forms. Various methods methods can be used for model order/structure selection, such as cross validation, Akaike's Final Prediction Error (FPE) criterion.

3.2 State Space Subspace Methods (N4SID)

These subspace methods were recently developed by Van Overschee and De Moor (1994). N4SID stands for "Numerical algorithms for for Subspace State Space System Identification", and should be read as a Californian license plate: "enforced it". The method directly calculates state space models of the process and disturbances. N4SID requires no *a priori* structure indexes. The order of the state space model is determined through the inspection of the dominant singular values of a matrix that is calculated during the identification. The state space matrices are not calculated in their canonical forms (with a minimum number of parameters), but as full state space matrices in a certain, almost optimally conditioned basis. This implies that the observability indices do not have to be known in advance. Other major advantages of N4SID is that the algorithms are one shot (no iterations) and numerically stable (only QR and SVD decompositions).

A Preliminaries

Consider the combined deterministic-stochastic model as follows.

$$x(t+1) = Ax(t) + Bu(t) + w(t) \tag{3.8}$$

$$y(t) = Cx(t) + Du(t) + e(t) \tag{3.9}$$

with

$$E\left(\begin{bmatrix} w(k) \\ e(k) \end{bmatrix}\begin{bmatrix} w^T(l) & e^T(l) \end{bmatrix}\right) = \begin{bmatrix} Q & S \\ (S)^T & R \end{bmatrix}\delta_{kl} \geq 0 \tag{3.10}$$

$A, Q \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$, $D \in \mathbb{R}^{p\times m}$, $S \in \mathbb{R}^{n\times p}$ and $R \in \mathbb{R}^{p\times p}$, where E denotes the expected value operator and δ_{kp} the Kronecker index. The input vector $u(t)$ (m dimensional) and output vector $y(t)$ (p dimensional) are measured. $e(t)$ (p dimensional) and $w(t)$ (n dimensional) are unmeasurable, Gaussian white noise vectors with zero means. $\{A, C\}$ is assumed to be observable, while $\{A, [B\ (Q)^{1/2}]\}$ is assumed to be controllable. The problem solved is the following: Given the input and output measurements

$$Z^N := y(1), u(1), \ldots, y(N), u(N)\ (N \to \infty)$$

and the fact that these data are generated by an unknown deterministic-stochastic model (3.8)-(3.10), find A, B, C, D, Q, S and R (up to within a similarity transformation.)

Typically, N4SID algorithms make extensive use of block Hankel matrices. Input and output block Hankel matrices are defined as:

$$U_{0|i-1} \stackrel{def}{=} \begin{bmatrix} u(0) & u(1) & \cdots & u(j-1) \\ u(1) & u(2) & \cdots & u(j) \\ \cdots & \cdots & \cdots & \cdots \\ u(i-1) & u(i) & \cdots & u(i+j-2) \end{bmatrix},$$

$$Y_{0|i-1} \stackrel{def}{=} \begin{bmatrix} y(0) & y(1) & \cdots & y(j-1) \\ y(1) & y(2) & \cdots & y(j) \\ \cdots & \cdots & \cdots & \cdots \\ y(i-1) & y(i) & \cdots & y(i+j-2) \end{bmatrix}$$

The subscripts of U and Y denote the subscript of the first and the last element of the first column.

B Algorithm

1. Determine the orthogonal projections:

$$Z_i = Y_{i|2i-1} \Big/ \begin{bmatrix} U_{0|i-1} \\ U_{i|2i-1} \\ Y_{0|i-1} \end{bmatrix}$$

$$= \left[\begin{array}{c|c|c} L_i^1 & L_i^2 & L_i^3 \\ p i \times mi & pi \times mi & p i \times pi \end{array}\right] \begin{bmatrix} U_{0|i-1} \\ U_{i|2i-1} \\ Y_{0|i-1} \end{bmatrix}$$

$$Z_{i+1} = Y_{i+1|2i-1} \Big/ \begin{bmatrix} U_{0|i-1} \\ U_{i+1|2i-1} \\ Y_{0|i} \end{bmatrix}$$

$$= \left[\begin{array}{c|c|c} L_{i+1}^1 & L_{i+1}^2 & L_{i+1}^3 \\ p(i-1)\times m(i+1) & p(i-1)\times m(i-1) & p(i-1)\, i \times p(i+1) \end{array}\right] \begin{bmatrix} U_{0|i} \\ U_{i+1|2i-2} \\ Y_{0|i} \end{bmatrix}$$

where $A/B = AB^T(BB^T)^{-1}B$.

2. Determine the Singular Value Decomposition:

$$\left[\, L_i^1 \middle| L_i^3 \,\right] \begin{bmatrix} U_{0|i-1} \\ Y_{0|i-1} \end{bmatrix} = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} V^T$$

The order of the model is determined from the dominant singular values.

Denote

$$\Gamma_i = U_1 \Sigma_1^{1/2} \text{ and } \Gamma_{i-1} = \underline{U_1} \Sigma_1^{1/2}$$

where $\underline{U_1}$ denotes U_1 without the last p rows.

3. Determine the states $\hat{X}_i$ and $\hat{X}_{i+1}$:

$$\hat{X}_i = \Gamma_i^{\dagger} \left[L_i^1 \middle| L_i^3 \right] \begin{bmatrix} U_{0|i-1} \\ Y_{0|i-1} \end{bmatrix}$$

$$\hat{X}_{i+1} = \Gamma_{i-1}^{\dagger} \left[L_{i+1}^1 \middle| L_{i+1}^3 \right] \begin{bmatrix} U_{0|i} \\ Y_{0|i} \end{bmatrix}$$

4. Determine the least-squares solution:

$$\begin{matrix} & j & & n & m & & j & & j \end{matrix}$$

$$n\begin{bmatrix} \hat{X}_{i+1} \\ \hline Y_{i|i} \end{bmatrix} = n\left[\begin{array}{c|c} \tilde{\mathfrak{x}}_{11} & \tilde{\mathfrak{x}}_{11} \\ \hline \tilde{\mathfrak{x}}_{11} & \tilde{\mathfrak{x}}_{11} \end{array}\right] \cdot \begin{matrix} n \\ m \end{matrix}\left[\begin{array}{c} \hat{X}_i \\ \hline U_{i|i} \end{array}\right] + \begin{matrix} n \\ p \end{matrix}\left[\begin{array}{c} \rho_1 \\ \hline \rho_2 \end{array}\right]$$

5. The system matrices are determined as follows:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \leftarrow \begin{bmatrix} \tilde{\mathfrak{x}}_{11} & \tilde{\mathfrak{x}}_{12} \\ \tilde{\mathfrak{x}}_{21} & \tilde{\mathfrak{x}}_{22} \end{bmatrix} \tag{3.11}$$

$$\left[\begin{array}{c|c} Q & S \\ \hline (S)^T & R \end{array}\right] \leftarrow \frac{1}{J}\left[\begin{array}{c|c} \rho_1\rho_1^T & \rho_1\rho_2^T \\ \hline \rho_2\rho_1^T & \rho_2\rho_2^T \end{array}\right] \tag{3.12}$$

A numerically efficient way to do these computations is described in Van Overschee and De Moor (1994). This algorithm has been implemented in the Interactive System Identification Toolbox of Xmath; see Aling et. al. (1993). It is also included in version 4 of the System Identification Toolbox in Matlab, Ljung (1991).

3.3 Asymptotic Method (ASYM)

This method is proposed by Zhu et al (1991) based on the asymptotic theory of Ljung (1985). The method calculates time domain models using a frequency domain criterion. The method was called *two step method* in Zhu and Backx (1993). We think that *asymptotic method* is a better name due to the extensive use of Ljung's asymptotic theory. The open-loop version of the method is given in the following. For the simplicity we also assume that the inputs are mutually independent. It is possible to develop an algorithm for closed loop data, because the asymptotic theory is also valid for closed loop experiment.

Several steps are involved in the method.

Step 1 *Estimate a high order ARX model and an upper error bound matrix.* The matrix $A(q)$ has a diagonal form:

$$A(q) = \text{diag}\,[\, A_{11}(q) \cdots A_{pp}(q)].$$

Then we have

$$\hat{G}^n(q) = \hat{A}^{-1}(q)\hat{B}(q) \quad \hat{H}^n(q) = \hat{A}^{-1}(q) \tag{3.13}$$

and the spectrum estimates of the disturbances are

$$\hat{\Phi}_v(\omega) = \hat{A}^{-1}(e^{i\omega})\hat{R}\,\hat{A}^{-T}(e^{-i\omega}) \qquad (3.14)$$

where $\hat{R}$ is the estimated covariance of the equation error residuals.

Denote $\hat{G}^n_{ij}(e^{i\omega})$ and $G^o_{ij}(e^{i\omega})$ as the (i,j) element of high order model $\hat{G}^n(e^{i\omega})$ and the true model $G^o(e^{i\omega})$ respectively. Then according to the asymptotic theory of Ljung (1985) (see Zhu, 1989 for MIMO extension), we can define the 3σ bound for $|G^o_{ij}(e^{i\omega}) - \hat{G}^n_{ij}(e^{i\omega})|$ as

$$|G^o_{ij}(e^{i\omega}) - \hat{G}^n_{ij}(e^{i\omega})| \le 3\sqrt{\frac{n}{N}\Phi^{-1}_{u_j}(\omega)\Phi_{v_i}(\omega)} \quad \text{w.p. } 99.99\% \qquad (3.15)$$

where $\Phi_{u_j}(\omega)$ is the spectrum of $u_j(t)$, and $\Phi_{v_i}(\omega)$ is the spectrum of $v_i(t)$.

Step 2 *Model reduction.* Perform model reductions on $\hat{G}^n_N(q)$ and on $\hat{H}^n_N(q)$ to obtain reduced process and disturbance models.

We can view the high order ARX model as the noisy observations of the true process (Wahlberg, 1989). Because we know the asymptotic distribution of the high order estimates is jointly normal, it is natural to apply the *maximum likelihood* method to these observations to find the low order model. Denote $G^l(q)$ and $H^l(q)$ as the reduced order process model and disturbance model respectively, then it can be shown (Wahlberg, 1989) that the *asymptotic maximum likelihood* loss function of the process model is:

$$V_i = \sum_{i=1}^{p}\sum_{j=1}^{m}\frac{1}{2\pi}\int_{-\pi}^{\pi}|\hat{G}^n_{ij}(e^{i\omega}) - \hat{G}^l_{ij}(e^{i\omega})|^2\frac{\Phi_{u_j}(\omega)}{|H^o_{ii}(e^{i\omega})|^2}\,d\omega \qquad (3.16)$$

The minimization of these loss functions involves a nonlinear optimization algorithm. The implemented optimization algorithm is an output error identification method. Again a diagonal form MFD model is obtained.

The order of the i-th MISO model is determined such that the following relation holds in the frequency band which is important for control purpose:

$$\sum_{j=1}^{m}|\hat{G}^n_{ij}(e^{i\omega}) - \hat{G}^l_{ij}(e^{i\omega})|^2 \le \sum_{j=1}^{m}\frac{n}{N}\hat{\Phi}^{-1}_{u_j}(\omega)\hat{\Phi}_{v_i}(\omega) \qquad (3.17)$$

It can be shown (Zhu, 1993) that the reduced model obtained using this order selection rule will have best accuracy in the frequency domain and from a stochastic viewpoint. The same model reduction order section can be performed for the disturbance model using white noises as its inputs.

Step 3. *Obtain a minimal state space model.* Transform the diagonal form MFD model into a state space realization and perform a balanced model reduction. This is to guarantee that the final model that will be used for controller design is controllable and observable.

This algorithm has been implemented in the CACSD Package of SETPOINT IPCOS bv.

4 Industrial Case Studies

4.1 The Two Processes

A Glass tube drawing process

The outline of the process is shown in Fig. 4.1. By indirect electric heating the glass is melted. It then flows down through a ring shaped hole along a mandril. Shaping of the tube takes place at, and just below, the end of the mandril. The glass tube is pulled down due to gravity and a drawing machine. Two measures of dimension, wall thickness and diameter of the tube, are the most important quantities to be regulated; hence they are the outputs of the process. The mandril gas pressure and the drawing speed can affect the wall thickness and diameter most directly and easily; so they are chosen as control inputs. Other variables, such as the power supplied to melt the glass, the pressure in the melting vessel and the room temperature, will be considered as disturbances. Therefore the process can be modelled as a 2-input 2-output process with disturbances.

$u_1(t)$: input 1, gas pressure

$u_2(t)$: input 2, drawing speed

$y_1(t)$: output 1, wall thickness

$y_2(t)$: output 2, diameter

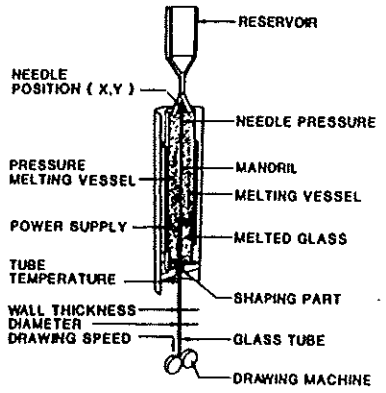


Fig. 4.1 The glass tube drawing process

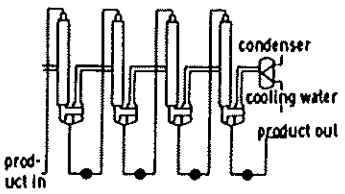
Two independent PRBS signals have been used as test signals. Two data sets are obtained, one for model estimation (600 samples) and one for model validation (669 samples).

B Four-effect (stage) evaporator

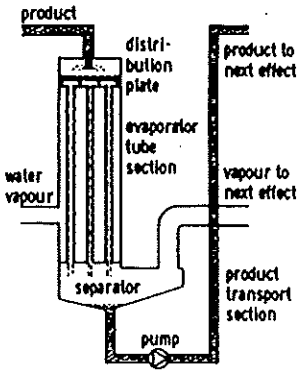
Evaporators are used to reduce the water content of a product, for example, milk. Fig. 4.2 shows a four-effect (stage) evaporator which is used in milk powder production as a concentration process before the drying process. A multiple-effect evaporator is composed of similar effects. The product enters the effect and first reaches the distribution plate. This plate distributes the product to all evaporating tubes equally. The product flows as a film along the inside wall of the evaporation tubes while it is heated by condensing water vapour on the outside of the tubes. A two-phase flow of concentrate and vapour reaches the separator where product and vapour are separated. The product is pumped to the distribution plate of the next effect and the vapour is directed to the next effect as well to condense on the outside of the evaporator tubes. The vapour from the last effect is condensed in a condenser and the product from this effect is directed to the drying stage.

This process is modelled as a 3-input 3-output process with disturbances:

- input 1: flow of feed;
- input 2: vapour flow to first evaporator stage;
- input 3: cooling water flow to condenser;
- output 1: dry matter content of the product;
- output 2: flow of product; and
- output 3: temperature of the product.



A) The complete process



B) One effect of the evaporator

Fig. 4.2 The four-effect evaporator (by courtesy of Netherlands Institute for Diary Research (NIZO))

Variables which also influence the outputs are dry matter content of feed, temperature of the feed, temperature of cooling water entering the condenser, and surrounding temperature. These are disturbances.

PRBS test signals are used in the identification experiments. The input/output data for model estimation and model validation are collected from two different days, each consists of about 3000 samples.

Remark—It is really a luxury that we could collect both estimation data set and validation data set with sufficient data points from the two processes. This may not be the case for many slow and/or capital intensive processes.

4.2 The Results of Identification

Two quality measures of methods are used, model accuracy and speed computation. All models are used for the simulation and prediction using the validation data sets. The square root of the mean square errors (MSE) of the simulation and prediction at each output are used to measure the model accuracy. They are denoted as:

- MSE-s: MSE's of simulation at all outputs
- MSE-p1: MSE's of one-step-ahead prediction at all outputs
- MSE-p5: MSE's of five-step-ahead prediction at all outputs

The speed of a method is measured by the required *flops* (floating point operations) for the model computation in Matlab.

The results of identification for the glass tube process are given in Table 4.1, where the two lines correspond to the two outputs. The results for the evaporator are contained in Table 4.2. In the tables, PEM denotes a prediction error method, N4SID the subspace method and ASYM the asymptotic method; order denotes the McMillan degree of models.

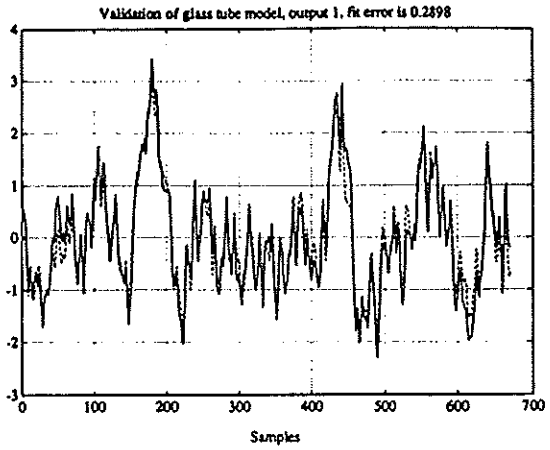
In order to give an impression of model accuracy, the measured and simulated outputs of the glass tube process are plotted in Fig. 4.3. The ASYM model of order 14 is used for simulation. The plots for other models look very much the same.

Method	PEM(OE)	PEM	N4SID	N4SID	ASYM (MISO)	ASYM (Reduced MISO)
Order	8	8	5	9	26	14
Flops	290677022	190776525	33527792	34845534	139715135	140408945
MSE-s	0.3478 0.3145	0.3592 0.3281	0.3964 0.3805	0.3368 0.3368	0.2841 0.3072	0.2898 0.3079
MSE-p1	0.3478 0.3145	0.0941 0.0740	0.2572 0.1723	0.1395 0.1165	0.0880 0.0735	0.1229 0.0928
MSE-p5	0.3478 0.3145	0.3346 0.2209	0.3905 0.3277	0.2814 0.2235	0.2259 0.1994	0.2413 0.2046

Table 4.1

Method	PEM	N4SID	N4SID	ASYM (MISO)	ASYM (Reduced MISO)
Order	13	6	13	22	12
Flops	683132359	22040292	224324099	863582432	864045962
MSE-s	0.3464 0.4033 0.2939	0.4377 0.4178 0.3713	0.3987 0.4048 0.3003	0.3700 0.4115 0.3581	0.3718 0.4102 0.3582
MSE-p1	0.1757 0.3752 0.0743	0.1695 0.3882 0.0727	0.1663 0.3741 0.0741	0.1808 0.3761 0.0745	0.1667 0.4102 0.0751
MSE-p5	0.2453 0.3948 0.1463	0.2550 0.4074 0.1465	0.2575 0.3940 0.1445	0.2728 0.3978 0.1469	0.2651 0.4102 0.1452

Table 4.2



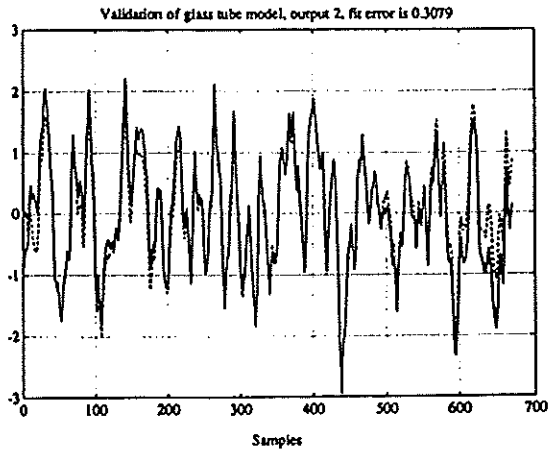


Fig. 4.3 Measured (solid) and model simulation (dashed) using validation data of the glass tube process, the 14th order ASYM model is used.

5 Comparative Discussions

1) Model accuracy. Recent work (Zhu, 1993) shows that PEM and ASYM are both asymptotic frequency domain maximum likelihood estimators. Hence if no numerical problems are encountered, the two methods would deliver models with the same accuracy. A theoretical relation between N4SID and PEM need to be studied.

For the glass tube process, ASYM model is most accurate for both simulation and prediction. For the evaporator, the ASYM model is worse than the other two for simulation. To our experience, the differences in model accuracy is too small to influence control performance.

2) Speed and robustness in model estimation. In the two case studies, the N4SID is about 4 times faster than PEM and ASYM.

The algorithms of N4SID are all one shot ones, therefore, the method does not suffer from the problems of local minima and nonconvergence. Also N4SID does not suffer from non-zero process initial conditions.

PEM, which contains iterative numerical search for the minimum of the criterion may have problems with local minima, and also in case of many parameters nvergence problems may occur. These problems are is pronounced though, with a thoughtful initialization process.

The numerical robustness of ASYM lies in between N4SID and PEM. Numerical iterations are only needed for the model reduction step, where noise-free data from the high order model are used. This greatly reduces the chance of local minima and nonconvergence; and the method does not suffer from non-zero initial conditions. Moreover, a bad local minimum can be detected during order/structure selection.

3) Model order/structure selection. In principle, all the order/structure selection criteria/methods can be combined with any of the model estimation methods. It is outside the scope of this paper to compare different methods of order/structure selection. Here only some notes on MIMO and MISO model structures are given. In principle, all the three methods can use MIMO model structure. At present, the ASYM is implemented only for MISO parametrizations and the final MIMO model is obtained using a model reduction; the PEM uses both MISO and MIMO parametrizations; and N4SID uses MIMO parametrizations. It is interesting to note that in the two case studies the use of MISO models by ASYM does not cause any problems. The final model reduction reduces the order by more than 40%.

Model uncertainty description. Under the assumption that the process is contained in the model set, PEM can supply stochastic error bounds for both parameters and for frequency responses. ASYM supplies stochastic upper error bounds for frequency responses with reported applications;. see Zhu and Backx (1993).

5) Optimal input design for control. There are results for optimal input design for control for the high order model of ASYM; see Yuan and Ljung (1985) for open loop case, Gevers and Ljung (1986) for closed loop case, and Zhu and Backx (1993) for MIMO extension and simplification.

6) Closed loop identification, unstable processes. It is well known that a PEM model from closed loop data is consistent. This is also true for ASYM, because the asymptotic theory holds for closed loop data. The closed loop version of ASYM is, however, not yet implemented. An N4SID model from closed loop data will loose its consistency. N4SID is able to identify an unstable process. ASYM is also able to do so, because the asymptotic theory also holds for unstable processes; see Ljung and Wahlberg (1989). Finally, also PEM handles unstable processes without problems.

7) Recursive estimation. All the three methods can be implemented in a recursive manner.

6 Conclusions

We have looked into how three different approaches to multivariable system identification handle data from two industrial processes. Generally speaking, all three methods come up with very good models. The models are able to reproduce validation data by simulation (as well as prediction) with an accuracy that is quite good, average relative errors of less than 10%. Comparison plots look very good. All the methods thus produce models that are well suitable, for example, for MPC applications in process industries.

We have in the tables given the results for the different methods. One should be careful not to read too much into the decimals of these figures. It would be wrong to consider these comparisons as a "competition" between methods. It rather shows how methods may complement each other, and how it is good to have an arsenal of various tools in one's toolbox. It could be stressed that if we spend days or weeks to collect and secure high quality process data it might be worth while to spend some extra minutes or hours at the computer to try out another model structure and/or another estimation method. Evaluation on the validation data, if available, will show if a better model is obtained in that way.

Finally we would like to point out that the objective of all the identification methods is that they are to be applied on real process data. It is therefore of interest to investigate and show how modern estimation methods are able to cope with data from processes of some complexity. We hope that this will help industrial users to make better use of modern identification methods.

References

Aling H., U. Miletti, R.L. Kosut, M.P. Mesaros, P. Van Overschee and B. De Moor (1993). An interactive system identification module for Xmath. *Proc. American Control Conference*, June 2-4, San Francisco, USA, pp. 3071-3075.

Gevers M. and L. Ljung (1986). Optimal experiment design with respect to the intended model application. *Automatica*, Vol. 22, pp. 543-554.

Ljung, L. (1985). Asymptotic variance expressions for identified black-box transfer function models. *IEEE Trans. Autom. Control*, Vol. AC-30, pp. 834-844.

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

Ljung, L. (1991). *System Identification Toolbox User's Guide*. The MathWorks Inc.

Ljung L. and B. Wahlberg (1989). Asymptotic properties of the least squares method for estimating transfer functions and disturbance spectra. *Report LiTH-ISY-I-1008*, Dept. EE, Linköping University, Sweden.

Van Overschee P. and B. De Moor (1994). N4SID two subspace algorithms for identification of combined deterministic-stochastic systems. *Automatica*, Vol. 30, No. 1, pp. 75-93.

Wahlberg, B. (1989). Model reduction of high-order estimated models: the asymptotic ML approach. *Int. J. Control*, Vol. 49, No. 1, pp. 169-192.

Yuan, Z.D. and L. Ljung (1985). Unprejudiced optimal open loop input design for identification of thansfer functions. *Automatica*, Vol. 21, No. 6 pp. 697-708.

Zhu, Y.C. (1989). Black-box identification of MIMO transfer functions: asym- ptotic properties of prediction error models. *Int. J. Adaptive Control and Signal Processing*, Vol.3, pp. 357-373.

Zhu Y.C., A.C.P.M. Backx and P. Eykhoff (1991). Multivariable process identification based on frequency domain measures. *Proceedings of 30th IEEE CDC*, Brighton, December, 1991, pp. 303-308.

Zhu Y.C. and A.C.M.P. Backx (1993). *Identification of Multivariable Industrial Processes: for Simulation, Diagnosis and Control*. Springer-Verlag, London.

Zhu Y.C. (1993). A frequency domain criterion for MIMO model structure selection and its applications. Submitted to *Automatica*.