Process Control Applications of Subspace and Regression-based Identification and Monitoring Methods

(Invited Paper)

Ben C. Juricek
Toyon Research Corporation,
75 Aero Camino, Suite A,
Goleta, CA 93106 USA
Email: bjuricek@toyon.com

Dale E. Seborg
Department of Chemical Engineering,
University of California,
Santa Barbara, CA 93106 USA
Email: seborg@engineering.ucsb.edu

Wallace E. Larimore
Adaptics, Inc,
1717 Briar Ridge Road,
McLean, VA 22101 USA
Email: larimore@adaptics.com

Abstract— This tutorial paper summarizes the application of a variety of identification techniques to simulations of two realistic chemical processes, a continuous stirred-tank reactor (CSTR) and the Tennessee Eastman challenge process. Both subspace identification methods (N4SID and CVA) and regression techniques (PLS and CCR) are considered. Emphasis is placed on the relative performance of the various identification methods, and their strengths and weaknesses. Also, the use of these identification methods in monitoring and fault detection is discussed.

In the CSTR case study, Dynamic ARX and FIR models are identified using two regression techniques, PLS and CCR, and the predictive error method. are compared with state-space models identified using two subspace algorithms, CVA and N4SID. The objective functions for PLS and CCR are shown to be related. A comprehensive simulation study of the CSTR with different characteristics and noise properties is used to compare the identification methods. The results indicate that, if the time delay structure is known or estimated accurately, the identified subspace models tend to be more accurate than the models identified using regression. The state-space models identified using the CVA algorithm are especially accurate.

The Tennessee Eastman challenge process is a realistic simulation of a chemical process that has been widely used in process control studies. In this case study, several identification methods are examined and used to develop MIMO models that contain seven inputs and ten outputs. ARX and finite impulse response models are identified using reduced-rank regression techniques (PLS and CCR) and state-space models identified with prediction error methods and subspace algorithms. For a variety of reasons, the only successful models are the state-space models produced by two popular subspace algorithms, N4SID and canonical variate analysis (CVA). The CVA model is the most accurate. Important issues for identifying the Tennessee Eastman challenge process and comparisons between the subspace algorithms are also discussed.

Index Terms— Subspace system identification, multivariate regression, industrial processes, monitoring, fault detection.

I. INTRODUCTION

A key requirement for many applications for advanced control and monitoring techniques is an accurate process model. Developing dynamic models based on fundamental physico-chemical relationships is often prohibitively difficult for industrial applications due to, e.g., unknown chemical reactions, poorly known or unknown kinetic coefficients, etc. Thus, decades of research have been devoted to developing models from empirical input-output data.

Most industrial processes, and almost all found in the chemical industry, are multivariable (that is, two or more inputs and outputs), nonlinear and are constantly responding to disturbances that are unmeasurable and occurring at unknown times. Although almost all processes are nonlinear, in practice linear models are commonly used for control and monitoring. This practice is not entirely unjustified, since several processes are operated within a localized region, and the nonlinearities for many processes (e.g., refinery processes) are quite mild within this operating region. The important issue for industrial processes is therefore estimating a multivariable, linear model that is appropriate for control and monitoring applications. The word "appropriate" is rather subjective but reflects the ability of the model to describe the system, and not just fit the data – what Ljung [1] refers to as the difference between system identification and curve fitting.

The field of system identification is quite mature. Published in the late-1980's, the textbooks of Box et al. [2], Ljung [3] and Söderström and Stoica [4] contain the fundamental theory for identifying discrete, stochastic, linear models. Both books describe the necessary steps and decisions for identifying empirical models: selecting the model structure, the model order, and parameter estimation method. For multi-input multi-output (MIMO) models, these issues (in particular the model structure) are especially challenging and generally left as a design parameter for an "expert" to specify. Several researchers have since studied strategies that systematically identify large-dimensional models. A systematic methodology is particularly important for, e.g., model predictive control applications, where an expert in MIMO system identification is not always available. This research compares methods that purport to identify accurate linear models for multivariable systems with little or no influence on the part of the user – a true "black-box" model.

In the next section, the generic identification problem is presented, followed by a discussion of methods from multivariable statistics and subspace algorithms for identifying dynamic models. Then, several comprehensive simulation studies are discussed comparing the techniques and demonstrating a number of significant issues.

II. IDENTIFYING LINEAR MULTIVARIABLE MODELS

For this paper, linear, discrete-time, stochastic models are of interest. The books by Ljung [3] and Söderström and Stoica [4] are excellent references for the classical system identification theory. A summary of the relevant issues is presented within this section.

Identifying an empirical process model requires exciting the process to collect experimental data, and using these data for estimating a mathematical model. After the experimental data have been collected, the user makes three important selections: whether to use a parametric or non-parametric model, the model structure (i.e., the type of model and the model order), and the method for estimating parameters. All three factors have significant effects on the identified model, and in the simulation section the relative importance of each factor will be examined.

A multivariable linear, stochastic model is written as a linear difference equation of the form

$$y(k) = \sum_{i=1}^{N_A} A_i y(k-i) + \sum_{i=0}^{N_B} B_i u(k-i) + \sum_{i=0}^{N_C} C_i e(k-i) \quad (1)$$

where N_A, N_B and N_C are respectively the AR, X and MA orders. The function A_i is the pulse response function from the past outputs to the present output $y(k)$, and B_i and C_i are the pulse responses from the past and present inputs $u(k)$ and noise $e(k)$, respectively, to the present outputs. These are all equivalently expressed as transfer functions in the backshift operator.

Several cases are of particular interest. The finite impulse response (FIR) model is where $A_i = 0$ and $C_i = 0$ and only the model order N_B is specified. Although the coefficient matrices must be estimated, the FIR model is labeled "non-parametric" because the FIR model acts as a function for describing the system response to impulse input [4].

FIR models are frequently used because there are few restrictions about the processes that can be modeled, are estimated using least squares regression, and the FIR structure is intuitive yet meaningful to control engineers. The major disadvantage of FIR models is that the model order can be quite large even for very simple models. This is particularly true for MIMO models of stiff processes that will have wide range of time constants. Additionally, for high-order models with large input and output dimensions, the estimates of B_i can be poor due to a poorly conditioned matrix of regressors.

Parametric models include autoregressive models with exogenous inputs (ARX) where $C_i = 0$, and autoregressive moving average models with exogenous inputs (ARMAX) where all of the coefficients A_i, B_i and C_i are estimated. A special case of the ARMAX model structure is the output error (OE) structure, where $C_i = A_i$.

The ARMAX structure provides a flexible description of the noise process and remains parsimonious. The difficulty with identifying ARMAX models is (1) cannot be directly written as a linear regression problem as before. This paper assumes that the noise model is unknown. If it were known, the noise model could act as a filter on the residuals and linear regression used for estimating model parameters [3]. Several estimation methods have been proposed for estimating ARMAX models including pseudo-linear regression, correlation methods (e.g., instrumental variables), and more recently, subspace methods.

The subspace methods identify a state-space model,

$$x(t+1) = Ax(t) + Bu(t) + Gw(t) \quad (2)$$
$$y(t) = Cx(t) + v(t) \quad (3)$$

where x is a $(n_x \times 1)$ vector of state variables, w is the $(n_x \times 1)$ vector of state noise variables, v is the $(n_y \times 1)$ vector of measurement errors. The state and measurement noise vectors are assumed to be uncorrelated with one another and distributed as normal random variables with covariance matrices Q and R, respectively. As described by [7], [3], [8], a state-space model is parametrically equivalent to the ARMAX model in (1). The state-space model can be identified using nonlinear optimization and a prediction error criterion, i.e., by using a prediction error method (PEM). However, the structure of the state-space matrices can be difficult to specify, and estimating the matrices requires solving a nonconvex optimization problem. Hence, the computational burden is large and there is no guarantee of a global minimum, i.e., the identified model using PEM may not converge to the "true" model even if the correct model order is specified.

Choosing between parametric and nonparametric methods, and selecting the model order and structure (for both the process and the noise model) both depend on the end use of the model. For example, several model predictive control software packages use non-parametric models [9]. Many modern control and monitoring applications are developed for state-space models. For multivariable FIR and ARX models, estimating the model parameters requires methods that are robust to a large degree of correlation among the regressors. In the next section, multivariable statistical methods for the regression problem are described. After discussing the multivariable statistical methods, subspace algorithms for identifying state-space models are described.

III. MULTIVARIABLE REGRESSION

For a linear relation between predicted variables, y, and regressor variables, x, the ith sample is,

$$y(i) = B^T x(i) + e(i) \quad (4)$$

where y is dimension n_y, x is dimension n_x and the noise $e(i)$ is assumed to be uncorrelated in time, and normally distributed with zero mean and constant covariance matrix Σ. Additionally, x and y are both assumed to be mean-centered. Equation (4) can be written for N measurements of y and x as,

$$Y = XB + E \quad (5)$$

where the rows of matrices Y (dimension $N \times n_y$), X (dimension $N \times n_x$) and E contain the individual observations of y^T, x^T, and e^T. The ordinary least squares (OLS) estimate of B is,

$$\hat{B}_{OLS} = (X^T X)^{-1}(X^T Y) \qquad (6)$$

The OLS normally assumes $\Sigma = I$. Generalized least squares (GLS) [10] is appropriate for general $\Sigma \neq I$. The GLS solution is equivalent to scaling both x and y by $\Sigma^{-\frac{1}{2}}$; thus, Σ is known and invertible.

The standard OLS solution can be inaccurate if $X^T X$ is ill-conditioned. If there are collinearities within the matrix X, the inverse of $X^T X$ does not exist. If there are large interactions between X variables, the inverse may be poorly conditioned, resulting in an inaccurate estimate of $\hat{B}_{OLS}$. Additionally, the OLS solution ignores any interactions between the predicted variables, Y, treating each variable independently. A number of methods have been proposed for multivariable regression that are appropriate for high-correlation in both the regressor and response matrices: principal component regression (PCR), ridge regression, curds & whey, canonical correlation regression (CCR) and partial least squares (PLS). There are two different PLS approaches: one that treats outputs independently ("PLS1") and one that incorporates correlation between outputs ("PLS2"). This paper focuses on PLS2 and the term "PLS" refers to the PLS2 algorithm. Several excellent review articles analyze the various properties of these regression methods [11], [12], [13]. The two methods of interest in this paper are PLS and CCR because they are the most commonly used techniques in the control engineering literature.

The chemometric methods such as PCA, PCR, and PLS were proposed to handle the large amounts of data that were stored as part of the "information age". These methods identify the coefficient matrices by forming reduced-rank matrices of the regressors and predictors. That is, correlations and collinearities within the input or output matrices can be used for accurately estimating B, unlike the ordinary least squares (OLS) solution.

Canonical correlation regression is a related multivariate statistical method derived as the maximum likelihood estimate for the rank-deficient regression problem, i.e., when the X and Y matrices may not be full-rank [14] [15]. The use of CCR for identifying dynamic FIR models were examined in [16].

It has been shown [17] that PCA, PCR, PLS, and CCA are all special cases of a generalized singular value decomposition, where the weighting used in the SVD determines the which computational procedure is used. This considerably clarifies the relationship among these methods. For PCA, PCR, and PLS, the weighting is chosen as a fixed value for each. For CCA the weighting is determined by the data itself and results in a maximum likelihood procedure for any specified choice of rank. This plays a central role

in the subspace system identification methods discussed below.

IV. SUBSPACE METHODS

Subspace identification methods are a recent development in the system identification field. The canonical variate analysis (CVA) algorithm was proposed by Larimore [17], [19], [20], [21], and is based on the time series analysis methods developed by Akaike [23], [24]. The N4SID algorithm, developed by van Overschee and De Moor [25], [26], is more closely related to engineering linear systems theory. Both algorithms identify a stochastic state-space model,

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + Ke(k) \qquad (7) \\ y(k) &= Cx(k) + Du(k) + e(k) \end{aligned}$$

where x is the $(n_x \times 1)$ vector of state variables, u is the vector of measured inputs $(n_u \times 1)$, and y is the vector of measured outputs $(n_y \times 1)$. For subspace algorithms, the state vector has a very particular meaning which will be described below. As mentioned earlier, any linear model structure (e.g., ARX, ARMAX, OE) can be written as a state-space model.

The derivation of subspace algorithms is rather complicated compared with the traditional prediction error methods. Furthermore, the different approaches used in the derivations make comparing the subspace algorithms a challenging task. The derivation of CVA is cast in a mathematical statistics framework. The derivation of N4SID uses geometric arguments and a system theoretic approach, similar to realization theory.

For subspace algorithms, the state vector, $x(k)$, is defined to be a linear combination of past inputs and outputs,

$$x(k) = J p(k) \qquad (8)$$

where,

$$p(k) = [u(k-1), \ldots, u(k-N), y(k-1), \ldots, y(k-N)]^T \quad (9)$$

and $p(k)$ is referred to as the "past" at sample k. The dimension of the past is the number of lags, N. The state vector, $x(k)$, is computed from data, and is not specified a priori. After J has been determined (as will be described below) the state vector can be estimated by (8). The state-space model matrices can then be estimated via linear least squares regression,

$$\begin{aligned} \begin{bmatrix} \hat{A} & \hat{B} \\ \hat{C} & \hat{D} \end{bmatrix} &= \operatorname{Cov}\left(\begin{bmatrix} x(k+1) \\ y(k) \end{bmatrix}, \begin{bmatrix} x(k) \\ u(k) \end{bmatrix} \right) (10) \\ &\times \operatorname{Cov}^{-1}\left(\begin{bmatrix} x(k) \\ u(k) \end{bmatrix}, \begin{bmatrix} x(k) \\ u(k) \end{bmatrix} \right) \end{aligned}$$

The parameter estimation step of subspace algorithms can vary, but all algorithms proceed in the same general fashion: estimate the state vector from the "past", and then estimate the state-space matrices using "current" values for state, input and output vectors.

The calculation of the matrix J distinguishes the various subspace algorithms from one another. In the CVA approach, J is derived from the canonical loadings [18] between the future outputs conditional on the future inputs (i.e., removing the effect of future inputs) and the past [17]. In the N4SID approach, J results from a series of geometric arguments (or linear algebraic arguments) based on a set of matrix equations for the evolution of a linear system. Asymptotically the key step for calculating J can be written as a weighted singular value decomposition (SVD) [27],

$$\text{svd}(W_1 \mathcal{O} W_2) = [U_1 \ U_2] \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \quad (11)$$

where W_1 and W_2 are weighting matrices, and $\mathcal{O}$ is an oblique projection of the future outputs along future inputs on the past inputs and outputs. J is calculated from the lower dimensional subspace defined by U_1, S_1, and V_1 (the details are explained fully in [25]). In the N4SID algorithm, $W_1 = W_2 = I$. For the CVA algorithm, W_1 and W_2 are defined so that J is the matrix of canonical variates for the past inputs and outputs.

The scaling for CVA is particularly noteworthy because it results from the maximum likelihood (ML) solution of a constrained rank regression problem for multivariate regression [17], exactly like CCR. When applied to time series data, an ML procedure incorporates the correct delay structure of the regression coefficients (i.e., the precise structure of the ARMAX polynomials). In CVA, this unknown shift structure is not explicitly imposed, but several simulation studies demonstrate the accuracy of CVA to be essentially equal to the ML solution [5]. A formal proof for this phenomena was recently given by Larimore [22]. ML accuracy permits the computation of confidence bands on model accuracy and precise tests of hypotheses concerning the model structure: e.g., detecting the presence of bias or trends, feedback, and delays. Near ML accuracy in quite small sample sizes has also been demonstrated in a number of complex systems by Larimore [5].

In principle different subspace algorithms could identify equivalent state-space models if scaled appropriately. However, other important differences exist among the subspace algorithms: the numerical procedure for calculating the state-space matrices, determining the dimension of the past (N in (8)) and selecting the model order. Larimore has shown that the CVA weighting can identify models that are nearly equivalent to prediction error methods (PEM), which require that the exact model structure be known a priori.

Ideally, the model order corresponds to the number of singular values in (11) that are greater than zero, or some very small value, ϵ. For N4SID, the order is selected by the user, usually by looking for a "knee" in the plot of singular values vs. model order, or where the singular values fall below a specified critical value. Frequently there is more than one knee, or no knee at all. For CVA, the model order is selected via Akaike's Information Criteria (AIC),

TABLE I

CSTR Results for Closed-loop Identification

Method	$R^2(y_1)$ (%)	$R^2(y_2)$ (%)
PLS-FIR	72.7	84.6
PLS-ARX	86.4	89.6
CCR-FIR	84.0	88.0
CCR-ARX	86.3	89.6
N4SID	81.4	85.7
CVA	87.1	90.3
ARX	81.9	82.2
PEM	87.3	90.4

a statistical method for hypothesis testing. The singular values (or equivalently, the canonical correlations) can also be used for order selection and hypothesis testing.

Subspace algorithms provide simple and fast methods for identifying a MIMO state-space model. The calculations are non-iterative and computationally efficient. Thus a rich model structure can be identified without the need of optimization or an expert in system identification.

V. PROCESS IDENTIFICATION COMPARISONS

In Juricek et al. [31], two simulation examples were used extensively to compare regression and subspace methods. The first example was a continuous stirred-tank reactor (CSTR) with a first order reaction rate $A \rightarrow B$ [33]:

$$\frac{dC_A}{dt} = \frac{F}{V}(C_0 - C_A) - k_0 \exp\left(-\frac{E_A}{RT}\right) C_A \quad (12)$$

$$\frac{dT}{dt} = \frac{F}{V}(T_0 - T) + \frac{-\Delta H}{\rho C_p} k_0 \exp\left(-\frac{E_A}{RT}\right) \quad (13)$$

$$+ \frac{UA}{\rho C_p V}(T_J - T)$$

The concentration of A, C_A, and the temperature, T are measured variables. The manipulated variables are the jacket temperature, T_J, and flowrate, F.

In [31], extensive simulations were done for a number of simulation models using various identification methods. A typical result is shown in Table I for the case of a closed-loop simulation model. The R^2 measure of model fit for each of the outputs y_1 and y_2 is shown, and it is seen that the best are PEM and CVA with N4SID, ARX and PLS-FIR the worst. Extensive comparisons were also made looking at empirical frequency response functions that are much more descriptive of the accuracy of the various parametric modeling methods.

A second example in [31] was the "quality control" example from Dayal and MacGregor [16], given by:

$$x(t) = K(q^{-1})\left(\begin{bmatrix} 1 & 0.7 \\ 0.7 & -1 \end{bmatrix} u(t) + \begin{bmatrix} 1 & 0.2 \\ 0.2 & 1 \end{bmatrix} w(t)\right)$$

$$y(t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0.2 \\ 0.2 & 1 \end{bmatrix} x(t) + v(t) \quad (14)$$

where $K(q^{-1}) = 0.2q^{-5}/(1 - 0.8q^{-1})$. Because of the collinearity between outputs, a reduced-rank method is necessary. Two simulations were performed:

1) $w(t) = 0$, or an output error process.
2) $w(t) \in N(0, Q)$, or an ARMAX process.

The measurement noise $v(t)$ was uncorrelated with $w(t)$ and was normally distributed, $v(t) \in N(0, R)$ with covariance matrix R chosen to provide a signal-to-noise ratio of approximately three for each output.

The results of the simulation studies in [31] provide more evidence that the subspace methods, particularly CVA, are very adept at producing parsimonious models that are quite robust to different noise processes. However the quality control example demonstrates that subspace methods are particularly sensitive to time delays, and that FIR and ARX models could be more accurate when time delays are not specified. Although the accuracy of any identified model will improve if time delays are known, time delay estimation methods, such as that provided by the ADAPT$_X$ software [36], are especially important for subspace methods.

It was shown that the estimated regressor matrices and the objective functions for the CCR and PLS methods are asymptotically equivalent if scaled appropriately. For dynamic models, the scaling issue is further complicated by the presence of the noise process. For FIR or ARX model identification, these simulation results and those of previous studies [16], [31] suggest that the choice of CCR or PLS could depend on the input excitation and the frequency range of interest, because CCR remains sensitive to variation at high frequencies.

VI. TENNESSEE EASTMAN CHALLENGE PROCESS

A detailed study of CVA identification of the Tennessee Eastman process is given in Juricek et al. [30]. The Tennessee Eastman Challenge Process (TE) was published by the Tennessee Eastman Company [34] as a process simulation for academic research. By academic standards, the problem is quite large: it contains 41 measured variables and 12 manipulated variables. Based on a real chemical process, the TE produces two products (labeled G and H) from four reactants (labeled A, C, D and E). An explicit mathematical representation of the process is not given; instead the simulation is distributed as purposely convoluted FORTRAN code [35]. In addition to the process description, the problem statement defines process constraints, 20 types of process disturbances, and six operating modes corresponding to different production. In order to define a process model that was large, but manageable, seven inputs and ten outputs were selected for the simulation study.

The Tennessee Eastman Challenge Process was used to compare dynamic models identified using the CVA, N4SID and ARX methods. Although the TE is a nonlinear system, linear models for the base operating mode were reasonably accurate for most of the seven inputs and ten outputs that were included in the model. Comparisons of actual and predicted responses are shown in Fig. 1 from u_6 to y_3 for prediction horizons of 1, 15, and ∞, and in Fig. 2 at a prediction horizon of 15 for two other input-output pairs, u_5 to y_4 and u_7 to y_1. The models identified by the CVA algorithm were particularly accurate, and should be well-suited for model-based control and monitoring applications. In general, the set of MISO ARX models was less accurate than the CVA model, but better than the N4SID state-space model, as indicated by the R^2 values for two validation data sets. Compared with a set of MISO ARX models, the single state-space model identified by CVA was more accurate and simpler to interpret. For example, the statistical properties of the noise and the modes of the system are easier to analyze using the single state-space model than seven ARX models. Although the subspace algorithms are related by the generalized singular value decomposition, using the N4SID algorithm with the CVA weighting did not produce the same state-space model as the CVA model that was identified using the ADAPT$_X$ software [36].

VII. PROCESS MONITORING AND FAULT DETECTION

A very brief summary of the use of subspace methods in fault detection is given in this section. A model identified using CVA can use several methods for fault detection. Three methods are investigated in Juricek et al. [32], the first based on Kalman filter residuals for the CVA model, the second based on canonical variable residuals. In addition, a third method is proposed that uses the local statistics approach [29] for detecting changes in the canonical variable coefficients. The proposed methods use multivariable versions of the classical Shewhart and CUSUM control charts. Thus while the statistics are non-traditional, they still use standard SPC charts and "SPC thinking"; i.e., alarm limits and limit violations have their usual meanings.

The major motivation for developing the proposed method is sensitivity. That is, slow-developing process changes can go undetected using standard, residual-based detection methods. The "local statistics" approach allows detecting changes in parameters, and the proposed method makes use of the parametric changes for detecting process changes.

The detection methods are evaluated using three simulation examples; the examples consider the effects of feedback control, process nonlinearities, and multivariable, serially correlated data. The simulations consider several types of common, process faults including sensor faults, load disturbances, and process changes. The simulation results indicate that the local approach provides a very sensitive method for detecting process changes, which are difficult to detect using either the Kalman filter or canonical variable residuals.
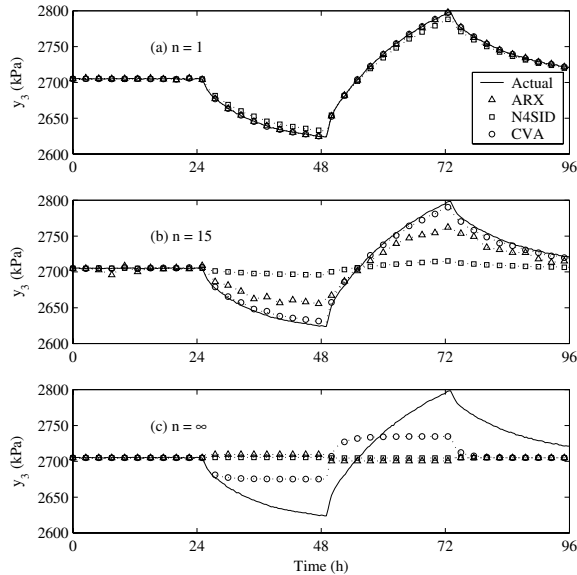
Fig. 1. Predicted and actual responses at different horizons n.
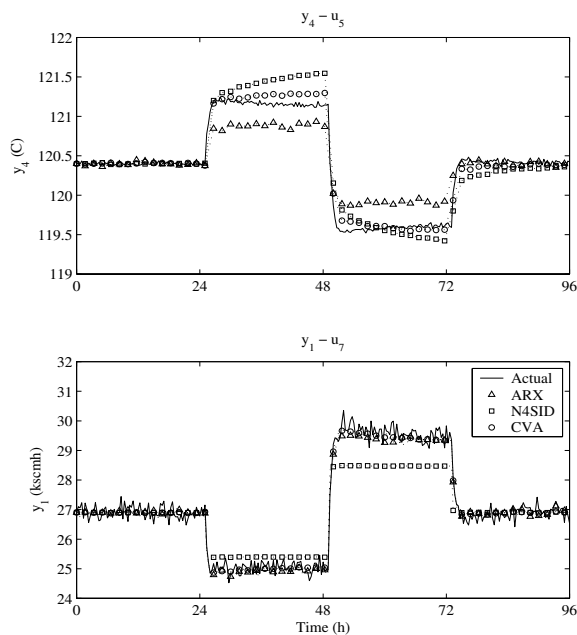

Fig. 2. Predicted and actual responses for two inputs/outputs pairs.

References

[1] Ljung, L. Issues in System Identification. IEEE Control Systems Magazine, 1991 11(1), 25–29.

[2] Box, G.; Jenkins, G.; Reinsel, G. Time Series Analysis: Forecasting and Control; Prentice-Hall: Englewood Cliffs, NJ, 1994.

[3] Ljung, L. System Identification; 2nd edition; Prentice-Hall: Upper Saddle River, NJ, 1999.

[4] Söderström, T.; Stoica, P. System Identification. Prentice-Hall: Hertfordshire, UK, 1989.

[5] Larimore, W. Statistical Optimality and Canonical Variate Analysis System Identification. Signal Processing, 1996 52, 131–144.

[6] Wahlberg, B. Model Reductions of High-Order Estimated Models: the Asymptotic ML Approach. Int. J. Control, 1989 49, 169–192.

[7] Kailath, T. Linear Systems; Prentice-Hall: Englewood Cliffs, NJ, 1980.

[8] Hannan, E.; Deistler, M. The Statistical Theory of Linear Systems; John Wiley: New York, 1988.

[9] Qin, S.; Badgwell, T. An Overview of Industrial Model Predictive Control Technology. AIChE Symposium Series 316: CPC-V. Lake Tahoe, CA, 1996 232–256.

[10] Mardia, K.; Kent, J.; Bibby, J. Multivariate Analysis; Academic Press, New York, 1979.

[11] Burnham, A.; Viveros, R.; MacGregor, J. Frameworks for Latent Variable Multivariate Regression. J. Chemometrics, 1996 10, 31–45.

[12] Breiman, L.; Freidman, J. Predicting Multiple Reponses in Multiple Linear Regression. J. Royal Stat. Soc. B, 1997 59, 3–54.

[13] Frank, I.; Freidman, J. A Statistical View of Some Chemometrics Regression Tools. Technometrics, 1993 35, 109–148.

[14] Tso, M. Reduced-rank Regression and Canonical Analysis. J. Royal Stat. Soc. B, 1981 43, 183 – 189.

[15] Stoica, P.; Viberg, M. Maximum Likelihood Parameter and Rank Estimation in Reduced-Rank Multivariate Linear Regressions. IEEE Trans. Signal Processing, 1996 44, 3069–3078.

[16] Dayal, B.; MacGregor, J. Multi-output Process Identification. J. Process Control, 1997 7, 269 – 282.

[17] Larimore, W. Optimal Reduced Rank Modeling, Prediction, Monitoring, and Control using Canonical Variate Analysis. IFAC 1997 Int. Symp. on Advanced Control of Chemical Processes. Banff, Canada, 1997 61 – 66.

[18] Anderson, T. An Introduction to Multivariate Statistical Analysis; John Wiley: New York, NY, 1984.

[19] Larimore, W. Canonical Variate Analysis in Control and Signal Processing. Katayama, T.; Sugimoto, S., Ed., Statistical Methods in Control and Signal Processing. Marcel Dekker, New York, 1997.

[20] Larimore, W. Automated Multivariable System Identification and Industrial Applications Proc. American Control Conference, San Diego, CA, 1999 1148–1162.

[21] Larimore, W. Automated and Optimal System Identification by Canonical Variables. Proc. American Control Conference. Baltimore, MD, 1994 1640–1644.

[22] Larimore, W. Large Sample Efficiency for ADAPTx Subspace System Identification with Unknown Feedback Proc. IFAC Conference on Dynamics and Control of Process Systems, Boston, MA. 2004.

[23] Akaike, H. Markovian Representation of Stochastic Processes by Canonical Variables. SIAM J. Control, 1975 13, 162–173.

[24] Akaike, H. Canonical correlation and information. Mehra, R.; Lainiotis, D., Ed. , System Identification: Advances and Case Studies. Academic Press, 1976.

[25] van Overschee, P.; De Moor, B. Subspace Identification for Linear Systems; Kluwer Academic Publishers: Boston, MA, 1996.

[26] van Overschee, P.; De Moor, B. N4SID: Subspace Algorithms for the Identification of Combined Deterministic-Stochastic Systems. Automatica, 1994 30, 75–93.

[27] van Overschee, P.; De Moor, B. A Unifying Theorem for Three Subspace System Identification Algorithms. Automatica, 1995 31, 1853–1864.

[28] Cho, Y.M.; Xu, G.; Kailath, T. Fast Identification of State-Space Models via Exploitation of Displacement Structure. IEEE Trans. Automatic Control, 1994 39, 2004–2017.

[29] Basseville, M.; Nikiforov, I. Detection of Abrupt Changes; Prentice-Hall: Englewood Cliffs, NJ, 1995.

[30] Juricek, B.; Seborg, D.; Larimore, W. Identification of the Tennessee Eastman Challenge Process Using Subspace Methods. Control Engineering Practice, 2001 9, 1337–1351.

[31] Juricek, B.; Seborg, D.; Larimore, W. Reduced Rank ARX and Subspace System Identification for Process Control, Ind. & Eng. Chem. Research, 2002 41, 2185–2203.

[32] Juricek, B.; Seborg, D.; Larimore, W. Fault Detection Using Canonical Variate Analysis Ind. & Eng. Chem. Res., 2004 43, 458–474.

[33] Seborg, D.; Edgar, T.; Mellichamp, D. Process Dynamics and Control, 2nd ed.; John Wiley: New York, 2004.

[34] Downs, J.J.; Vogel, E. A plant-wide industrial process control problem, Comp. & Chem. Eng., 1993 17, 245–255.

[35] Ricker, N.L. Tennessee Eastman challenge archive, bf 2001 Available at: http://depts.washington.edu/control/LARRY/TE/download.html.

[36] Larimore, W. ADAPT$_X$ Automated System Identification Software Users Manual. Adaptics, Inc., McLean, VA, 1996.