

**“MODEL-ON-DEMAND” IDENTIFICATION FOR  
CONTROL: AN EXPERIMENTAL STUDY AND  
FEASIBILITY ANALYSIS FOR MOD-BASED  
PREDICTIVE CONTROL**

**M. W. Braun<sup>\*</sup> B. A. McNamara<sup>\*</sup> D. E. Rivera<sup>\*,1</sup>  
A. Stenman<sup>\*\*</sup>**

*<sup>\*</sup> Department of Chemical, Bio, and Materials Engineering,  
Control Systems Engineering Laboratory, Manufacturing  
Institute, Arizona State University, Tempe, Arizona, USA  
85287-6006*

*<sup>\*\*</sup> Division of Automatic Control, Department of Electrical  
Engineering, Linköping University, SE-581 83 Linköping, Sweden*

Abstract: An experimental study of “Model-on-Demand” (MoD) identification is made on a pilot-scale brine-water mixing tank. MoD estimation is compared against semi-physical modeling techniques using identification data generated from a systematically designed m-level Pseudo Random Sequence (PRS) input. The estimated models are the basis for evaluating the usefulness of MoD-based Model Predictive Control (MPC). For this application, MoD-MPC is shown to provide better performance at high bandwidths compared to a linear MPC controller.

Keywords: nonlinear model identification, local modeling, nonlinear model predictive control, integrated identification and control, experimental application

## 1. INTRODUCTION

“Just-In-Time” (JIT) or “Model-on-Demand” (MoD) modeling is a novel paradigm first proposed by (Cybenko, 1996). The philosophy of MoD modeling is that a model is not estimated until it is really needed. All observations of the process are stored in a database, which is accessed to estimate a local model at the current operating point. The variance/bias tradeoff is optimized locally by adapting the number of data and their relative weighting. The MoD approach enhances local modeling and provides an alternative to nonlinear black-box techniques that optimize model predictions globally (such as neural networks).

Previous work derived guidelines for “plant-friendly” control relevant m-level Pseudo-Random Sequences (m-level PRS) for nonlinear system identification (Braun *et al.*, 1999b). m-level PRS input signals designed for identification of a rapid thermal processing (RTP) wafer reactor simulation were shown to provide a database rich enough to support MoD estimation of the output. For the RTP simulation, additional work showed that MoD estimation using a m-level PRS database performed as well as semi-physical models combined with feed-forward neural networks (Braun *et al.*, 1999a).

The focus of this paper is a pilot-scale brine-water tank experiment present at the ASU unit operations laboratory (Figure 1). It consists of a main mixing tank with agitator, a fresh water feed stream, a salt water feed stream, and an exiting

---

<sup>1</sup> To whom all correspondence should be addressed.  
phone: (480) 965-9476 fax: (480) 965-0037; e-mail:  
daniel.rivera@asu.edu

flow. A range of the dominant time constant of the experimental system is determined from a first principles model. Next, design variables for a m-level PRS are calculated and a signal is generated and introduced into a pilot-scale brine-water mixing tank experiment. The predictive ability of the MoD estimator is compared against semi-physical models on data recorded from the tank. This work lends experimental support to previous results which were tested in simulation only. The semi-physical model is used as the “plant” for a MoD based MPC controller simulation. The simulation serves as a proof of concept for the SISO MoD-MPC formulation and illustrates the advantages of MoD-MPC over linear MPC.

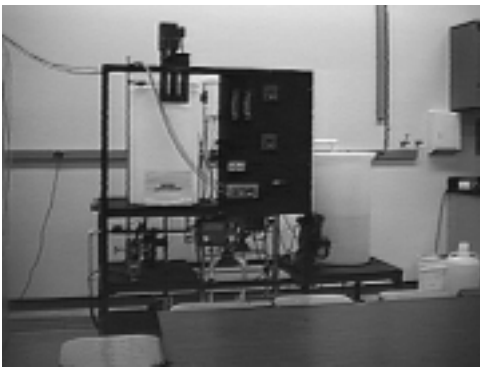


Fig. 1. Experimental Brine Tank Setup

Section 2 discusses m-level PRS database design, MoD estimation, and incorporation of the MoD estimator into an MPC controller. Section 3.1 outlines the features of the brine tank experiment and the identification experiment. Section 3.2 presents the results of the MoD-MPC vs. linear MPC comparison. Finally, conclusions are drawn and advantages of the MoD-MPC controller are noted.

## 2. MODEL-ON-DEMAND IDENTIFICATION AND CONTROL

This section summarizes the theory involved in creating a database for MoD estimation, briefly treats the derivation of the MoD estimator and derives the MoD-MPC formulation to be used in the case study section.

### 2.1 Input Signal Design

The MoD estimator is a data-driven methodology which requires that the input signal introduced to the plant excite the nonlinearities within the expected operating ranges of both the input and output spaces. While many possible signals could

be selected for excitation, we have chosen m-level PRS signals because they are periodic, deterministic, have an autocorrelation function close to that of white noise, and are capable of exciting nonlinear dynamics. The use of minimum crest factor multisines is also considered in this conference (Braun *et al.*, 2000).

In order to design a m-level PRS signal using the guidelines previously derived in (Braun *et al.*, 1999b), the user must have or estimate the following *a priori* information:

- desired closed-loop speed of response
- estimated range of dominant time constant
- order of nonlinearity of the plant and the structure of the model to be fit
- acceptable signal length and amplitude

With these in hand, the guidelines determine low and high frequency limits for the input signal. There is also a lower limit on the number of levels  $m$  to be used in the signal, based on the expected order of nonlinearity of the plant. The next part of the procedure involves selecting design variables to meet these limits as well as making decisions regarding signal length and harmonics to be suppressed. Once the design variables are selected, the signal is scaled to the acceptable amplitude. It may take several iterations of the input signal design procedure in order to meet the user’s criteria. There may also be some compromise required by the user. More details regarding the input signal design procedure can be found in (Braun *et al.*, 1999b; Braun *et al.*, 1999a).

### 2.2 MoD Estimation

The MoD modeling formulation pursued in this paper follows from the approach of (Stenman *et al.*, 1996; Stenman, 1999). Consider a SISO process with nonlinear ARX structure

$$y(k) = m(\varphi(k)) + e(k), \quad k = 1, \dots, M, \quad (1)$$

where  $m(\cdot)$  is an unknown nonlinear mapping and  $e(k)$  is an error term modeled as i.i.d. random variables with zero mean and variance  $\sigma_k^2$ . The MoD predictor attempts to estimate a value  $\hat{y}$  based on a local neighborhood of the regressor space  $\varphi(t)$  (Stenman *et al.*, 1996). The regressor vector is of the form

$$\begin{bmatrix} \varphi(t) \\ u(t - n_k) \dots u(t - n_b - n_k) \end{bmatrix}^T \quad (2)$$

where  $n_a$ ,  $n_b$ , and  $n_k$  denote the number of previous outputs, inputs, and delay in the model.

A local estimate  $\hat{y}$  can be obtained from the solution of the weighted regression problem

$$\hat{\beta} = \arg \min_{\beta} \sum_{k=1}^N \ell(y(k) - m(\varphi(k), \beta)) \cdot W\left(\frac{\|\varphi(k) - \varphi(t)\|_M}{h}\right), \quad (3)$$

where  $\ell(\cdot)$  is a scalar-valued and positive norm function,  $\|u\|_M \triangleq \sqrt{u^T M u}$  is a scaled distance function on the regressor space,  $h$  is a *bandwidth* parameter controlling the size of the local neighborhood, and  $W(\cdot)$  is a window function (usually referred to as the *kernel*) assigning weights to each remote data point according to its distance from  $\varphi(t)$ . The window is typically a bell-shaped function with bounded support. A tricube window function is chosen since it has a continuous derivative and smoothly descends to zero at the boundaries. Assuming a local *linear* model structure,

$$m(\varphi(k), \beta) = \beta_0 + \beta_1^T (\varphi(k) - \varphi(t)), \quad (4)$$

a quadratic norm,  $\ell(\varepsilon) = \varepsilon^2$  is used and the model is then linear in the unknown parameters. The estimate can be easily computed using least squares methods. If  $\hat{\beta}_0$  and  $\hat{\beta}_1$  denote the minimizers of (3) using the model from (4), a one-step ahead prediction is given by

$$\hat{y}(t) = m(\varphi(t), \hat{\beta}) = \hat{\beta}_0. \quad (5)$$

Each local regression problem produces a single prediction  $\hat{y}(t)$  corresponding to the current regression vector  $\varphi(t)$ . To obtain predictions at other locations in the regressor space, the weights change and new optimization problems have to be solved. This is in contrast to the global modeling approach where the model is fitted to data only once and then discarded. However, in a neighborhood around  $\varphi(t)$ , the local linear model provides an input-output linearization

$$A(q^{-1})y(t) = B(q^{-1})u(t - n_k) + \alpha, \quad (6)$$

where  $A(q^{-1})$  and  $B(q^{-1})$  are polynomials in the backward time-shift operator  $q^{-1}$  obtained from the components of  $\hat{\beta}$ , and

$$\alpha = \hat{\beta}_0 - \hat{\beta}_1^T \varphi(t) \quad (7)$$

is an offset term.

The bandwidth  $h$  controls the neighborhood size and has a critical impact on the resulting estimate since it governs a trade-off between the bias and variance errors of the estimate. Traditional bandwidth selectors produce a single global bandwidth; in MoD estimation, bandwidth is computed adaptively at each prediction. Various measures are available for this purpose (Stenman, 1999). The method used in this paper is a localized Akaike Information Criterion (AIC).

### 2.3 MoD-MPC

To incorporate the MoD estimator into an MPC controller, we start with the MPC objective function given in (Meadows and Rawlings, 1997). Given the model description in Equation 6 and knowledge of the current system state, we seek a control that minimizes the objective function

$$J = \sum_{k=0}^{N-1} Q_e(k) (r(t+k+1) - \hat{y}(t+k+1))^2 + Q_u(k) u^2(t+k) + Q_{\Delta u}(k) \Delta u^2(t+k) \quad (8)$$

where  $Q_e(k)$ ,  $Q_u(k)$  and  $Q_{\Delta u}(k)$  represent penalty weights on the control error, control signal, and control move size, respectively. Of the  $N_u$  future control actions that minimize  $J$ , only the first one is used. As new measurements become available, a new optimization problem is formulated whose solution provides the next control action. This is referred to as the receding horizon principle. A special feature of the formulation is the presence of the move size,

$$\Delta u(t+k) = u(t+k) - u(t+k-1), \quad (9)$$

in the objective. In some process control applications, the move size is restricted due to actuator constraints. The main advantage of MPC is that hard constraints can be specified by the user on move size, input magnitude, states or output. Since optimization of Equation 8 can be computationally intensive for large prediction horizons  $N$ , often only 4 or 5 control increments are solved, with the implicit assumption that the control action will be held constant for the remaining moves. The practice of using smaller control horizons  $N_u$  has the effect of producing less aggressive controllers and providing stable control for non-minimum phase systems (Meadows and Rawlings, 1997).

Since the offset term in Equation 6 changes stepwise, we choose to utilize a Control ARIMA model (CARIMA), shown in Equation 10, where  $\Delta = 1 - q^{-1}$ . This forces the controller to have integral action. While nominal performance is the focus of the formulation presented in this paper, current research on MoD-MPC addresses robustness and stability considerations.

$$A(q^{-1})\Delta y(t) = B(q^{-1})\Delta u(t) \quad (10)$$

To express the output prediction at time  $t+k$  as a function of future controls, introduce the following identity (Åström and Wittenmark, 1995)

$$1 = A(q^{-1})F_k(q^{-1})\Delta + q^{-k}G_k(q^{-1}) \quad (11)$$

i.e.,

$$\hat{y}(t+k) = B(q^{-1})F_k(q^{-1})\Delta u(t+k-1) \quad (12)$$

$$+ G_k(q^{-1})y(t).$$

By partitioning  $B(q^{-1})F_k(q^{-1})$  as

$$B(q^{-1})F_k(q^{-1}) = S_k(q^{-1}) + q^{-k}\tilde{S}_k(q^{-1}), \quad (13)$$

where  $\deg S_k(q^{-1}) = k-1$  and  $\deg \tilde{S}_k(q^{-1}) = n_b - 2$ , the output prediction in Equation 12 can be rewritten as

$$\hat{y}(t+k) = S_k(q^{-1})\Delta u(t+k-1) + \bar{y}(t+k). \quad (14)$$

The first term depends on future control actions whereas the remaining terms depend on measured quantities only. By introducing

$$\hat{\mathbf{y}} \triangleq (\hat{y}(t+1) \quad \dots \quad \hat{y}(t+N))^T,$$

$$\Delta \tilde{\mathbf{u}} \triangleq (\Delta u(t) \quad \dots \quad \Delta u(t+N-1))^T,$$

$$\bar{\mathbf{y}} \triangleq (\bar{y}(t+1) \quad \dots \quad \bar{y}(t+N))^T,$$

$$\tilde{\mathbf{S}} \triangleq \begin{pmatrix} s_0 & 0 & \dots & 0 \\ s_1 & s_0 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ s_{N-1} & s_{N-2} & \dots & s_0 \end{pmatrix},$$

where  $s_i$  are the coefficients of  $S_k(q^{-1})$  and

$$\hat{\mathbf{y}} = \bar{\mathbf{y}} + \tilde{\mathbf{S}}\Delta \tilde{\mathbf{u}} \quad (15)$$

Therefore

$$\hat{\mathbf{y}} = \bar{\mathbf{y}} + \mathbf{S}\Delta \mathbf{u} \quad (16)$$

where

$$\Delta \mathbf{u} \triangleq (\Delta u(t) \quad \dots \quad \Delta u(t+N_u-1))^T, \quad (17)$$

and

$$\mathbf{S} \triangleq \tilde{\mathbf{S}}\Lambda, \quad \Lambda \triangleq \begin{pmatrix} I \\ 0 \end{pmatrix}. \quad (18)$$

The control moves can also be expressed in vector form

$$\mathbf{u} = \mathbf{T}\Delta \mathbf{u} + \bar{\mathbf{u}} \quad (19)$$

where  $\mathbf{T}$  is matrix with 1's in the lower triangle and diagonal, and  $\bar{\mathbf{u}}$  is a vector containing only  $u(t-1)$ 's. The objective can then be simplified as

$$J = \| \mathbf{r} - \bar{\mathbf{y}} - \mathbf{S}\Delta \mathbf{u} \|_{Q_c}^2 + \quad (20)$$

$$\| \mathbf{T}\Delta \mathbf{u} + \bar{\mathbf{u}} \|_{Q_u}^2 + \| \Delta \mathbf{u} \|_{Q_{\Delta u}}^2$$

where  $r$  denotes the desired reference trajectory. For the *unconstrained* case, the minimizing control sequence is obtained explicitly by ordinary least squares theory. For the *constrained* case, the constraints can be re-formulated in matrix/vector form and the problem is solved efficiently using standard numerical optimization algorithms.

### 3. BRINE-WATER MIXING TANK CASE STUDY

#### 3.1 Identification

The experiment consists of an agitated mixing tank into which salt water and fresh water are fed. The inlet flow rates of the two streams are controlled via regulatory control points FIC100 and FIC101. The outlet flow rate is used to control the level in the tank via LIC100. The inlet salt water concentration and outlet salt concentration are measured via conductivity meters. In this experiment, inlet salt concentration ( $q_c$ ) is held constant, fresh water flow rate ( $q_w$ ) is the manipulated variable and outlet salt concentration ( $c$ ) is the response variable. A schematic of the tank is included in Figure 2.

03 Oct 95 11:08:23 1

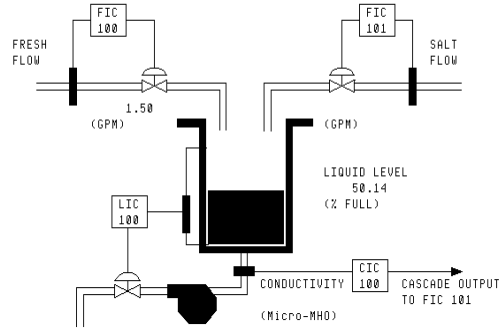


Fig. 2. Experimental Brine Tank Setup

$$\frac{dc}{dt} = \frac{q_c}{V}(c_c - \frac{\rho_c}{\rho}c) - \frac{q_w c}{V}, \quad V = hA \quad (21)$$

Initially, a first principles model shown in Equation 21 was derived. Cross-sectional area of the tank ( $A$ ), inlet/outlet stream densities ( $\rho_c$  and  $\rho$ ) and  $c_c$ , inlet brine concentration, are constant-valued parameters in the model. The volume of the material in the tank ( $V$ ) and therefore the level in the tank ( $h$ ) was assumed constant. This model was used in dynamic simulation to estimate a dominant time constant range of  $3.5 \text{ min} \leq \tau \leq 18.13 \text{ min}$  for the real system. With this and other information as summarized in Table 1, a 7-level PRS signal was designed for identification of the true system. The final design variables for the signal are shown in Table 1.

The input signal was programmed into a custom point implemented via a Honeywell TDC3000<sup>TM</sup> plant information and control system interfaced to the tank. Outlet concentration, inlet salt water concentration, inlet salt water flow rate, and fresh water flow rate were recorded at a sampling time

Table 1. m-level PRS design parameters

Desc.:	Symbol:	Value:
Closed-loop parm.	$\alpha$	3
Low Freq. parm.	$\beta$	6
Levels	$m$	7
Galois Field	$q$	7
Shift Registers	$n_r$	2
Switching Time	$T_{sw}$	3 min.
Sup'd Harmonics	$h$	2

of 10 seconds. Figure 3 shows the input/output data. A validation data sequence, designed to amplify the effects of nonlinearity in the plant and test the interpolative ability of the MoD estimator, was injected following the 7-level PRS.

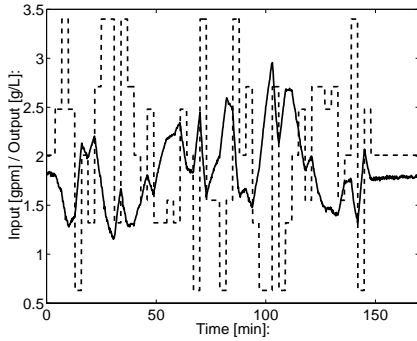


Fig. 3. Brine Tank Estimation Data; Solid: Output; Dashed: Input

In system identification, the plant under study may exhibit significant nonlinear behavior, resulting in poor performance of linear models. The model requirements may not necessarily require the time and effort for construction of a first principles model. The user may decide to include regressor terms in the model based on physical insight of the plant. This is the basic concept of semi-physical modeling (Lindskog, 1996). By including nonlinear regressor terms which are linear in the estimated parameters, it may be possible to improve model performance, while still using linear least-squares parameter estimation. This leads to a predictor of the form

$$\hat{y}_{sp}(t) = \theta_{sp}^T \varphi(t) \quad (22)$$

where

$$\theta_{sp} = [\theta_1 \theta_2 \dots \theta_n]^T. \quad (23)$$

For the mixing tank, a variety of different semi-physical models were feasible. Out of a family of possible semi-physical models, Equation 24 provided the lowest RMS and Maximum errors in comparison with the validation data. These errors were 0.0414g/L and 0.121g/L, respectively.

$$c(t) = c(t-1) + \theta_1 q_c(t-1)c_c(t-1) + \theta_2 q_c(t-1)c(t-1) + \theta_3 q_w(t-1)c(t-1) + \theta_4 \quad (24)$$

The MoD estimator with a first order local polynomial was applied to the input-output data. An initial fit of linear ARX models to the data provides a good estimate of the structures to use for the MoD estimation. For MoD estimation within a set of ARX structures ( $1 \leq n_a \leq 2$ ,  $1 \leq n_b \leq 2$ ,  $n_k = 1$ ), an ARX211 structure provided the lowest RMS and maximum errors of 0.0520 g/L and 0.238 g/L, respectively. For a MoD estimator with a second order local polynomial, an ARX111 structure provided RMS and maximum errors of 0.0500 g/L and 0.118 g/L, respectively. The best fitting semi-physical and the first and second order MoD estimators are shown in Figure 4.

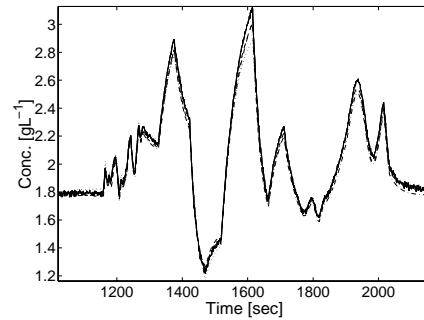


Fig. 4. Model Validation; Solid: Measured Data; Dashed: Semi-physical; Dotted: MoD First Order; Dash-Dot: MoD Second Order

### 3.2 Control

Previous work (Stenman, 1999) has tested the performance of MoDMPC controllers in tracking a setpoint trajectory (*servo control*). In this paper focus is on Type I and Type II disturbance rejection (*regulatory control*). A ramp disturbance is active from  $t = 10 \text{ min.}$  to  $t = 40 \text{ min.}$  Two step disturbances are active from  $60 \text{ min.} \leq t \leq 75 \text{ min.}$  and  $75 \text{ min.} \leq t \leq 90 \text{ min.}$ , respectively. The disturbances were added to the manipulated variable of the process.

The best fit semi-physical model was used as the “truth model” to compare performance of MoD-MPC against that of linear MPC. The same m-level PRS design used to identify the experimental system was used to generate input-output data from this model. The data was used for the subsequent MoD estimation and determination of the linear model. The MoD estimator with the first order polynomial and ARX211 database was implemented in the MoD-MPC formulation discussed in Section 2. A linear model (ARX221) was constructed using the MATLAB System Identification Toolbox and selected based on AIC. The

model was converted to step response format and used with the Model Predictive Control Toolbox.

High closed loop speed of response was desired for control of the outlet concentration, hence a setpoint tracking weight  $Q_y$  of 1 and a move size weight  $Q_{\Delta u}$  of 0.000001 were selected based on simulation results. The prediction horizon  $N_p$  covers the 95% settling time of the open-loop plant. Manipulated variable constraints were chosen within the range of the input during identification. The MPC tuning parameters used for both the linear MPC controller and MoD-MPC controller are presented in Table 2.

Table 2. Controller Tunings

Desc.:	Symbol:	Value:
Pred. Horizon	$N_p$	360
Control Horizon	$N_u$	3
Sampling Time	$T_s$	1/6 min
Stpt Tracking Weight	$Q_y$	1
Manip. Var. Weight	$Q_u$	0
Move Size Weight	$Q_{\Delta u}$	0.000001
Hi Manip. Var. Constr.	$u_{max}$	3.226 gpm
Lo Manip. Var. Constr.	$u_{min}$	0.633 gpm

For these tuning variables, manipulated variable and control variable responses from both control simulations are shown in Figure 5a and 5b. Note that the response of the linear MPC controller at these parameters is underdamped and oscillatory in both the controlled and manipulated variables. For the Type II disturbance, the linear MPC controller is not able to maintain setpoint.

#### 4. CONCLUSIONS

Considering the results of Section 3.1 and noting that the physics of the system under study are well understood, it is not surprising that the MoD estimator may not perform as well as the best fit semi-physical model. However, the second order polynomial MoD estimator does provide a lower maximum error compared to all semi-physical models examined. Note that with less engineering effort, the MoD provided performance rivaling that of the semi-physical modeling method. In Section 3.2 we have shown, with minimal additional effort, that a nonlinear MoDMPC controller can be commissioned on a process and provide performance well beyond that of a linear MPC controller for both Type I and Type II input disturbances.

#### 5. REFERENCES

Åström, K.J. and B. Wittenmark (1995). *Adaptive Control*. 2nd ed.. Addison-Wesley.

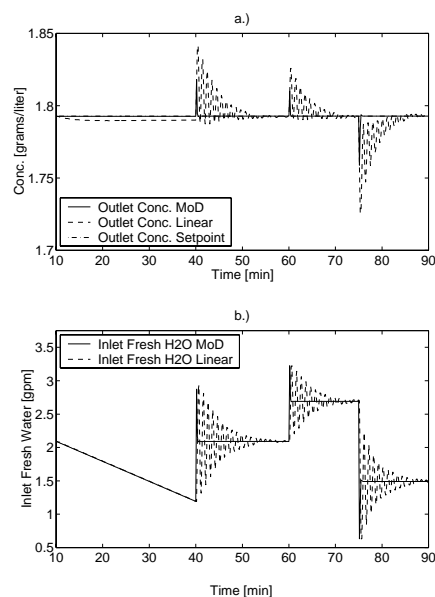


Fig. 5. a.) Controlled variable responses for Linear MPC and MoD-MPC b.) Manipulated responses for Linear MPC and MoD-MPC

- Braun, M. W., R. Ortiz-Mojica and D. E. Rivera (2000). Design of minimum crest factor multi-sinusoidal signals for “plant-friendly” identification of nonlinear process systems. In: *Proc. of IFAC SYSID2000*. Santa Barbara, CA.
- Braun, M.W., D.E. Rivera, A. Stenman and W. Foslien (1999a). Comparison of global nonlinear models and “model-on-demand” estimation applied to identification of a rtp wafer reactor. In: *Proc. IEEE '99 CDC*.
- Braun, M.W., D.E. Rivera, A. Stenman, W. Foslien and C. Hrenya (1999b). Multi-level pseudo-random signal design and “model-on-demand” estimation applied to nonlinear identification of a rtp wafer reactor. In: *Proc. 1999 ACC*.
- Cybenko, G. (1996). Just-in-time learning and estimation. In: *Identification, Adaptation, Learning* (S. Bittani and G. Picci, Eds.), pp. 423–434. NATO ASI Series. Springer.
- Lindskog, P. (1996). Methods, Algorithms and Tools for System Identification Based on Prior Knowledge. PhD thesis. Linköping University.
- Meadows, E. S. and J.B. Rawlings (1997). *Nonlinear Process Control*. Chap. Model Predictive Control. Prentice Hall PTR. Upper Saddle River.
- Stenman, A. (1999). Model on Demand: Algorithms, Analysis and Applications. PhD thesis. Linköping University.
- Stenman, A., F. Gustafsson and L. Ljung (1996). Just-in-time models for dynamical systems. In: *Proc. IEEE '96 CDC*.