15th Triennial World Congress, Barcelona, Spain

# SYSTEM IDENTIFICATION USING FRACTIONAL HAMMERSTEIN MODELS

**Mohamed AOUN[+], Rachid MALTI[*], Olivier COIS[+], Alain OUSTALOUP[+]**

[+] *Laboratoire d'Automatique et de Productique - UMR 5131 CNRS*
*Université Bordeaux 1 - ENSEIRB*
*351 cours de la Libération, 33405 Talence cedex, France*
*Tel: +33 (0)556 842 418 – Fax: +33 (0)556 846 644*
*{cois, aoun, oustaloup}@lap.u-bordeaux.fr*

* *Intelligence dans les Instrumentations et les Systèmes $I^2S$*
*IUT de Sénart-Fontainebleau – Université Paris 12*
*Avenue Pierre Point, 77127 Lieusaint, France*
*Tél: +33 (0)164 134 485 – Fax: +33 (0)164 134 503*
*malti@univ-paris12.fr*

Abstract: Identification of continuous-time non-linear systems characterised by fractional order dynamics is studied. The Riemann-Liouville definition of fractional differentiation is used. A new identification method is proposed through the extension of Hammerstein-type models by allowing their linear part to belong to the class of fractional models. Fractional models are compact and so are used here to model complex dynamics with few parameters. *Copyright © 2002 IFAC*

Keywords: identification, Hammerstein model, non-linear modelling, fractional differentiation.

## 1. INTRODUCTION

The goal of system identification is to establish a mathematical model able to reproduce the dynamic behaviour of a system. For many systems, when there is a wide operating area rather than a unique operating point, a linear model cannot be used. In this case non-linear models such as multi-models, Volterra series, neural networks, fuzzy logic, Hammerstein and Wiener-type models can be used. In this paper, a Hammerstein model, with a static non-linear part connected to a fractional linear part, is considered.

To our knowledge, no identification study is available in the literature on non-linear systems whose dynamics have a fractional differentiation character. Here, Hammerstein models with local fractional order dynamics are considered.

Hammerstein models are used in many fields: chemical engineering, e.g., distillation columns, heat exchangers (Özkan, *et al.*, 2001) pH systems (Zhu, Seborg, 1994); electrical engineering (Haber and Unbehauen, 1990) and automotive engineering (Ralston, *et al.*, 1997).

Section 2 presents the Hammerstein model. Section 3 focuses on the linear part of the model where some properties of fractional calculus are given. In section 4, two identification algorithms are developed. The first algorithm is based on the minimisation of the quadratic equation error with a linear programming method. The second is based on the minimisation of the quadratic output error with a non-linear iterative optimisation algorithm. The last section illustrates, through a numerical simulation, the advantages of using a fractional Hammerstein model compared to a classical integer Hammerstein model.

## 2. HAMMERSTEIN MODEL

The structure of Hammerstein models is shown in figure (1). It consist of a static non-linear part *N* connected to a dynamic linear part *H*.

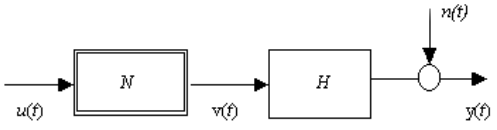


Fig. 1. Hammerstein model

For the non-linear part, we assume that

$$\sup_{u(t)} \left| N(u(t)) \right| < \infty . \quad (1)$$

This condition is not restrictive as long as only stable systems are treated. Let us consider BIBO stability (Bonded Input - Bounded Output). If $H$ is stable, $N(u(t))$ must be bounded for any permitted input data to ensure the stability of the whole system.

When $N$ is unknown, it can be approximated with a polynomial expansion (Ljung, 1987):

$$v(t) = N(u(t)) = \sum_{k=1}^{M} \alpha_k u^k(t) \quad (2)$$

where $M \in \mathbb{N}$ and $\alpha_k \in \mathbb{R}$. As $N$ is static, $u(t)$ and $v(t)$ have the same transition time with different amplitudes. It is assumed, as is usual for Hammerstein models, that $v(t)$ is not accessible for measurement (Billings and Fakhouri, 1982; Haber and Unbehauen, 1990). The virtual signal $v(t)$, resulting from $N(u(t))$, is mapped through the linear part $H$ which could be described with a differential equation, transfer function or state space representation. In this paper, $H$ is extended to the class of fractional linear models (see eq. (3)). **This is the original part of our contribution**.

$$\begin{aligned} a_0 y(t) + a_1 D^{n_{a_1}} y(t) + \cdots + a_L D^{n_{a_L}} y(t) = \\ b_1 D^{n_{b_1}} v(t) + \cdots + b_J D^{n_{b_J}} v(t) + n(t), \end{aligned} \quad (3)$$

$n(t)$ is a white noise, with zero mean and finite variance: $\mathrm{E}(e(t)) = 0$ and $\mathrm{var}(e(t)) < \infty$,

$a_0 .., a_L, b_1 .., b_J, n_{a_1} .., n_{a_L}, n_{b1} .., n_{b_J}$ are real numbers.

This new class of models includes the class of Hammerstein integer linear models.

Studies on real systems such as thermal (Battaglia, *et al*., 1999), (Battaglia, *et al*., 2000) or electrochemical (Oustaloup, 1995) reveal inherent fractional behaviour. Identifying such non-linear systems using fractional non-linear models may be more suitable for the general case.

## 3. FRACTIONAL LINEAR MODELS: MATHEMATICAL BACKGROUND

During the 19th century some mathematicians such as Abel, Liouville, Riemann and Cauchy were interested in the extension of classical integer differentiation to real orders. Some definitions and proprieties of this mathematical tool are now provided.

### *3.1. Fractional integration*

Let $f(t)$ be a continuous real function. The fractional integral of a function $f(t)$ is defined by Samko, *et al*., (1993):

$$\left(I_{a+}^{n} f\right)(t) \stackrel{\Delta}{=} \frac{1}{\Gamma(n)} \int_a^t \frac{f(\tau)}{(t-\tau)^{1-n}} d\tau , \quad (4)$$

where $t > a$ . $n$ is the real integration order. $\Gamma(n)$ is the Gamma Euler function:

$$\Gamma(n) = \int_0^{\infty} \mathrm{e}^{-x} x^{n-1} dx . \quad (5)$$

When $n$ is real, the integral in equation (4) is the area of the surface generated by $f(t)$ weighted with the factor $\frac{1}{\Gamma(n)(t-\tau)^{1-n}}$ .

The Laplace transform of the integral of a function $f(t)$ is :

$$\begin{aligned} \mathrm{L}\left\{I_0^n f(t)\right\} &= \int_0^{\infty} \mathrm{e}^{-pt} \left( \frac{1}{\Gamma(n)} \int_0^t \frac{f(\tau)}{(t-\tau)^{1-n}} d\tau \right) dt \\ &= \frac{1}{s^n} F(s), \end{aligned} \quad (6)$$

where $F(s)$ is the Laplace transform of $f(t)$.

### *3.2. Fractional differentiation*

The Riemann-Liouville fractional derivative of order $n$ of $f(x)$ is defined as (Miller and Ross, 1993):

$$D_{t_0}^{n} f(t) \stackrel{\Delta}{=} \left( \frac{d}{dt} \right)^{n+1} \left( I_{t_0}^{1-n} f(t) \right) \quad (7)$$

Second definition (Grünwald's definition) is:

$$D_{t_0}^{n} f(t) = \lim_{h \to 0} \frac{1}{h^n} \sum_{j}^{(t-t0)/h} (-1)^j \binom{n}{j} f(t - jh) \quad (8)$$

where $\binom{n}{j} = \frac{\Gamma(n+1)}{\Gamma(j+1)\Gamma(n-j+1)}$ and $\binom{n}{0} = 1$

The two definitions (7) and (8) are equivalent when $f(t_0) = D_{t_0}^{1} f(t_0) = ... = D_{t_0}^{\infty} f(t_0) = 0$ .

From equation (8), we note that fractional differentiation is not a local operator. The value of the fractional derivative function at $t$ depends on the whole past of the function. However, in the case where the differentiation order is an integer value, the derivative function depends only on some local points. For example when $n$ = 1:

$$D_0^1 f(t) = \lim_{h \to 0} \frac{f(t) - f(t-h)}{h} . \quad (9)$$

When $f(0) = D_0^1 f(0) = ... = D_0^{\infty} f(0) = 0$ the Laplace transform of $D_0^n f(t)$ (Oldham and Spanier, 1974):

$$\mathcal{L}\left\{D_0^n f(t)\right\} = s^n F(s) . \quad (10)$$

This result is coherent with the classical case where $n$ is an integer. Consequently, it is easy to define a symbolic representation of a dynamic system, such as a transfer function representation.

### *3.3. Modelling of fractional system*

Consider a SISO LTI system $H$, relaxed at $t$=0. $H$ can be described by the differential equation:

$$a_1 D^{n_{a_1}} y(t) + \cdots + a_L D^{n_{a_L}} y(t) = b_1 D^{n_{b_1}} y(t) + \cdots + b_J D^{n_{b_J}} y(t) + n(t), \quad (11)$$

where $a_1 .., a_L, b_1 .., b_J$ are real. $u(t)$ and $y(t)$ are the input and the output signals. they are differentiated to the real orders (integer or non integer) $n_{a_1} .. n_{a_L}$ and $n_{b_1} .. n_{b_J}$. As $H$ is relaxed, the Laplace transform of $D^{n_a} y(t)$ and $D^{n_b} u(t)$ are respectively $s^{n_a} Y(s)$ and $s^{n_b} U(s)$ (see eq. (10)).

Applying the Laplace transform to equation (11) the following is obtained:

$$a_1 s^{n_{a_1}} Y(s) + \cdots + a_L s^{n_{a_L}} Y(s) = b_1 s^{n_{b_1}} U(s) + \cdots + b_J s^{n_{b_J}} U(t) + e(s) \quad (12)$$

from which the generalised transfer function is defined :

$$H(s) = \frac{y(s)}{u(s)} = \frac{b_1 s^{n_{b_1}} + \ldots + b_J s^{n_{a,J}}}{a_1 s^{n_{a_1}} + \ldots + a_L s^{n_{a_L}}} = \frac{B(s)}{A(s)} \quad (13)$$

### *3.4. Stability condition of fractional systems*

As BIBO stability is considered, sufficient stability condition is given by:

$$\int_0^{\infty} \left| \mathcal{L}^{-1}\{H(s)\} \right| dt = K < +\infty , \quad (14)$$

where $H(s)$ is the Laplace transform of the system impulse response. Structural decomposition of non integer systems, allows the definition of a stability condition (Matignon, 1996) :

$$Arg(\lambda_k) > n\frac{\pi}{2} \quad (15)$$

where $\lambda_k$ are the eigenvalues of the system $H$

## 4. OPTIMISATION METHODS

Degree of the polynomial $N(u(t))$, $M$, and the differentiation orders are assumed to be *a priori* known by the user. In the case of many thermal systems, Battaglia *et al.* (2000) show that best differentiation orders are multiples of 0.5. When fractional differentiation orders cannot be fixed due to a lack of information, techniques providing estimation of both coefficients and differentiation orders can be applied as done in the linear case in (Cois, *et al*., 2000) and (Trigeassou, *et al*., 1999).

The optimal values of $\hat{a}_1 ,..., \hat{a}_L \hat{b}_1 ,.., \hat{b}_J , \hat{\alpha}_1 ,..., \hat{\alpha}_M$ are obtained according to a criterion based on the equation error or output error.

### *4.1. Equation error algorithm*

From equations (2) and (3)

$$y(t) = \sum_{j=1}^{J} b_j D^{n_{b_j}} \left( \sum_{k=1}^{M} \alpha_k u^k(t) \right) - \sum_{l=1}^{L} a_l D^{n_{a_l}} y(t) + \varepsilon(t) \quad (16)$$

then

$$y(t) = \sum_{i=1}^{I} b_i \sum_{k=1}^{m} \alpha_k D^{n_{b_j}} u^k(t) - \sum_{l=1}^{L} a_l D^{n_{a_l}} y(t) + \varepsilon(t) \quad (17)$$

where $\varepsilon(t)$ is the equation error.

For same input/output data the Hammerstein model is not unique. There are many equivalent models. Two Hammerstein models $\{N(.),H(s)\}$ and $\{N'(.),H'(s)\}$ are equivalents if

$$\frac{N(.)}{N'(.)} = \frac{H'(s)}{H(s)} \quad (18)$$

We are interested in one of the equation models. Hence, without any lost of generality we suppose that $b_1 = 1$.

The estimated output is:

$$\hat{y}(t) = \sum_{j=1}^{J} \hat{b}_{ji} D^{n_{b_{ji}}} \left( \sum_{k=1}^{M} \hat{\alpha}_k u^k(t) \right) - \sum_{l=1}^{L} \hat{a}_l D^{n_{a_l}} \hat{y}(t) \quad (19)$$

Let

$$Y(t) = [y(t),...,y(0)]^T = \Phi\theta \quad (20)$$

where

$$\theta = [a_1 ,\cdots, a_L, \alpha_1 , \alpha_1 b_2 \cdots, \alpha_1 b_J, \cdots; \alpha_M, \alpha_M b_2, \cdots, \alpha_M b_J]^T \quad (21)$$

$$\Phi = [\Phi_0 \ \Phi_1 \ \cdots \Phi_M ];$$

$$\Phi_0 = \begin{bmatrix} -D^{n_{a_1}} y(t) & \cdots & -D^{n_{a_L}} y(t) \\ \vdots & \vdots & \vdots \\ -D^{n_{a_1}} y(0) & \cdots & -D^{n_{a_L}} y(0) \end{bmatrix};$$

$$\Phi_k = \begin{bmatrix} D^{n_{b_1}} u^k(t) & \cdots & D^{n_{b_J}} u^k(t) \\ \vdots & \vdots & \vdots \\ D^{n_{b_1}} u^k(0) & \cdots & D^{n_{b_J}} u^k(0) \end{bmatrix}; k = 1 \ldots M.$$

The minimisation criterion is based on the quadratic predictive error (for $K$+1 observations of input-output data):

$$J(\theta) = \int_0^K (y(t) - \hat{y}(t|\theta))^2 dt \quad (22)$$

The optimum $\theta$ is given by the standard least squares equation:

$$\hat{\theta} = (\Phi^T \Phi)^{-1} \Phi Y(t) \quad (23)$$

The values of $b_2 ,..,b_J , \alpha_1 ,..., \alpha_M$ are not given in $\theta$.

The drawback of this method is that initial Hammerstein coefficients cannot be computed once the vector $\theta$ is known. The reason is that coefficients $\alpha_i$ and $b_i$ are coupled in (21). Hence, if one tries to solve for $\alpha_i$ and $b_i$, knowing $\theta$, then one will obtain

more equations than unknowns. Often, the solution is impossible. However, a model can be expressed using the coefficient vector $\theta$ and the regression matrix $\Phi$.

### *4.2. Output error identification algorithm*

To optimise parameters of the Hammerstein model $\{N(.),H(s)\}$, length of $\theta$ must be equal to the number of unknown parameters. $\theta$ can be defined as:

$$\theta = [a_1,...,a_J;b_2,...b_J;\alpha_1,\alpha_M]^T .$$

This implies that a non-linear optimisation method must be used. A new criterion $J$ based on the quadratic output error is defined:

$$J(\theta) = \int_0^K (e(t))^2\,dt = \int_0^K (y(t)-\hat{y}(t))^2\,dt \qquad (24)$$

where

$$\hat{y}(t) = \mathcal{L}^{-1}\left(\frac{s^{n_{b_1}} + b_2 s^{n_{b_2}} + ... + b_J s^{n_{b_J}}}{1 + a_1 s^{n_{a1}} + ... + a_L s^{n_{aL}}}\right) * \sum_{k=1}^{M} \alpha_k u^k(t) \qquad (25)$$

The optimal value of $\hat{\boldsymbol{\theta}}$ is then obtained iteratively using a non-linear programming technique, herein the Marquardt algorithm (Marquardt, 1963) was used.

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i - \left\{\left[\mathbf{J}''_{\boldsymbol{\theta\theta}} + \xi\ \mathbf{I}\right]^{-1} \mathbf{J}'_{\boldsymbol{\theta}}\right\}_{\hat{\boldsymbol{\theta}}=\boldsymbol{\theta}_i}, \qquad (26)$$

with

$$\begin{cases} \mathrm{J}'_{\theta} = -2\int_0^t e(t)\mathrm{S}(t,\hat{\theta})dt : \text{gradient} \\ \mathrm{J}''_{\theta\theta} \approx 2\int \mathrm{S}(t,\hat{\theta})\mathrm{S}^{\mathrm{T}}(t,\hat{\theta})dt : \text{hessian} \\ \mathrm{S}(t,\hat{\theta}) = \dfrac{\partial\,\hat{y}(t,\hat{\theta})}{\partial\,\theta} : \text{output sensitivity function} \\ \xi : \text{Marquardt parameter} \end{cases}$$

This algorithm, often used in non-linear optimisation, ensures robust convergence.

To apply (25) output sensitivities must be computed which is done by differentiating $\hat{y}$ w.r.t. respective parameters:

$$\frac{\partial\hat{y}(t)}{\partial\alpha_k} = \frac{b_1 s^{n_{b_1}} + ... + b_J s^{n_{b_J}}}{a_1 s^{n_{a1}} + ... + a_L s^{n_{aL}}} u^k(t)$$

$$\frac{\partial\hat{y}(t)}{\partial a_l} = \sum_{k=l}^{M} \alpha_k \mathcal{L}^{-1}\left(\frac{-\left(b_1 s^{n_{b_1}} + ... + b_J s^{n_{b_J}}\right) s^{n_{al}}}{\left(a_1 s^{n_{a1}} + ... + a_L s^{n_{aL}}\right)^2}\right) * u^k(t)$$

$$\frac{\partial\hat{y}(t)}{\partial b_j} = \sum_{k=l}^{M} \alpha_k \mathcal{L}^{-1}\left(\frac{s^{n_{b_j}}}{a_1 s^{n_{a1}} + ... + a_L s^{n_{aL}}}\right) * u^k(t)$$

## 5. Example

To illustrate advantages of using fractional Hammerstein models, a system is identified with both fractional and integer Hammerstein models. The estimation data are generated with a fractional Hammerstein model. First, the optimum model is calculated in the class of integer Hammerstein models. Then, it is calculated in the class of fractional Hammerstein models. For both models, a quadratic output error criterion is minimised.

The simulated system is described by the model:

$$\begin{cases} N(u(t)) = u(t) + 2u^2(t) \\ H(s) = \dfrac{1}{1 + 2s^{0.5}} \end{cases} \qquad (27)$$

Output data are corrupted by a stationary zero mean white noise of amplitude:

$$10\log_{10}\left(\frac{\text{Energie of signal}}{\text{Energie of noise}}\right) = 13\text{dB}$$

The degree of the non-linearity and the differentiation orders will be assumed known. All other parameters are identified with the proposed algorithm.

The first model we are looking for belongs to the class of integer Hammerstein models. The non-linearity structure is assumed as known (but not its parameters). The posed problem is to find the optimum order of the linear dynamic part having integer order derivatives and to estimate all the parameters of the system. The optimisation process is then computed for many integer orders.

It was found that the optimal integer Hammerstein model is the following:

$$\begin{cases} N(u(t)) = 0.9177u(t) + 1.6438u^2(t) \\ H(s) = \dfrac{0.0004s^4 + 0.0374s^3 + 2.7177s^2 + 6.3072s + 1}{-0.0118s^4 + 1.0121s^3 + 11.7888s^2 + 11.2286s + 1} \end{cases}$$

The optimum fractional Hammerstein model minimising the quadratic output error is identified as being:

$$\begin{cases} N(u(t) = 0.9587u(t) + 2.0028u^2(t) \\ H(s) = \dfrac{1}{1 + 1.9992s^{0.5}} \end{cases}$$

The number of optimised parameters with an integer Hammerstein model is 11. However, only 3 parameters are optimised with the fractional Hammerstein model. The small number of optimised parameter makes the identification process more efficient.

Figure 2 shows the variation of the output model for the identification data. Figure 4 shows the validation output data. Note that the output of the fractional

Hammerstein model is the closer to the true output especially in validation data fig. 4, as expected.

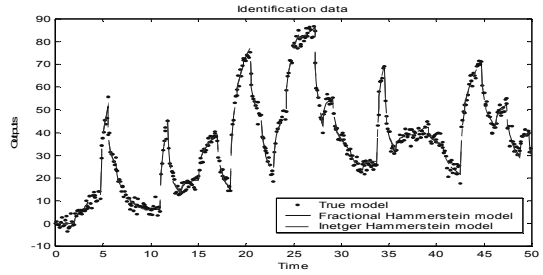


Fig. 2. Model output for identification data

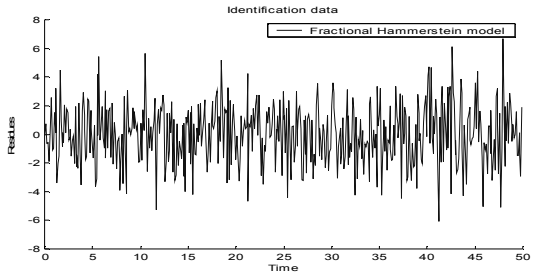


Fig. 3. Output error with fractional Hammerstein model for identification data

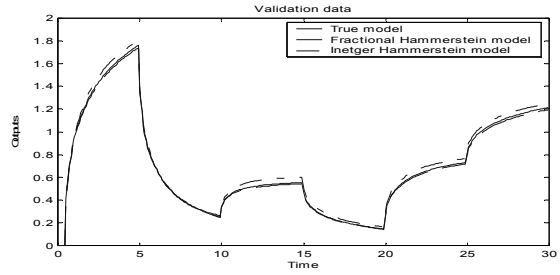


Fig. 4. Output model for validation data.

## 6. CONCLUSION

A new approach for identifying non-linear systems whose dynamics have a fractional order is proposed. It is based on Hammerstein models linear part of which is extended to the class of fractional systems. As shown through the example, this kind of models is more compact. Hence, when the system belongs to the class of non-linear systems whose dynamics have a fractional order; the number of parameters to optimise is reduced considerably. Hence, This approach is especially useful when the *a priori* information about system's dynamics reveals fractional behaviour.

An interesting perspective would be to extend this approach to Winner-type models.

## 7. REFERENCES


Battaglia, J.L, O. Cois, L. Puigsegur and A. Oustaloup. Solving an inverse heat conduction problem using a non-integer identified model, *International Journal of Heat and Mass Transfer*, 2001, n°14, vol. 44.

Battaglia, J.L., L. Le Lay, J.-C. Batsale, A. Oustaloup and O. Cois. Heat flow estimation through inverted non integer identification models, *International Journal of Thermal Science*, 1999, n°39, vol. 3.

Billings, S.A. and S.Y. Fakhouri. Identification of systems containing linear dynamic and static non-linear elements, *Automatica* 18 (1) (1982) 15-26.

Cois, O., A. Oustaloup, E. Battaglia, J.-L. Battaglia, *Non integer model from modal decomposition for time domain system identification*. 5th Symposium on system Identification.2000. Santa Barbara.

Haber, R. and H. Unbehauen. Structure identification of non linear dynamic systems : a survey on input/output approachs, *Automatica* 26 (4) (1990) 651-677.

Ljung, L. (1987). *System identification theory for the user*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey 07632.

Marquardt, D.W. (1963), *An algorithm for least squares estimation of non-linear parameters*, J. Soc. Industr. Appl. Math, 11(2), 431-441.

Miller, K.S. and B. Ross. (1993), *An introduction to the fractional calculus and fractional differential equations*. A Wiley-Interscience Publication.

Oustaloup, A. (1995). *La dérivation non entière: théorie, synthèse et applications*, Hermès, Paris.

Oldham K. B. and J. Spanier (1974). *The fractional calculus*. Academic Press, New York and London.

Özkan, G., S. Özen, S. Erdogan, H. Hapoglu and M. Alpbaz. Nonlinear control of polymerization reactor, *Computers & Chemical Engineering*, Volume 25, Issues 4-6, 1 May 2001, Pages 757-763.

Ralston, J.C., Zoubir, A.M.and Boashash, B., Identification of a class of nonlinear systems under stationary non-Gaussian excitation, *IEEE Transactions on Signal Processing*, Volume 45, Issue 3, 1997, Pages 719-735.

Samko, S.G., A.A. Kilbas and O.I. Marichev (1993), *Fractional integrals and derivatives: theory and applications*, Gordon and Breach Science Publishers, Amsterdam.

Trigeassou J.-C., T. Poinot, J. Lin, A. Oustaloup, F. Levron. *Modeling and identification of a non integer order system*. Proc ECC'99, European Control Conference. 1999. Karlsruhe, Germany.

Zhu, X. and D.E. Seborg. Nonlinear predictive control based on Hammerstein models, *Control Theory & Applications*, Volume 11, Issue 5,1994, Pages 564-575.