[redacted]

AN OVERVIEW OF NONLINEAR SYSTEMS IDENTIFICATION

by

S. A. Billings B.Eng, Ph.D., C.Eng. MIEE, AFIMA, M.Inst. MC
Department of Control Engineering
University of Sheffield
Sheffield
S1 3JD

Research Report No. 271

January 1985

To be presented at the 7th IFAC Symposium on Identification and System Parameter Estimation York, 1985.

7th IFAC Symposium on Identification and System Parameter Estimation, York, 1985.

AN OVERVIEW OF NONLINEAR SYSTEMS IDENTIFICATION

S. A. Billings

Department of Control Engineering, University of Sheffield,
Mappin Street, Sheffield S1 3JD, England.

Abstract. There are basically three distinct approaches to the identification of nonlinear systems and these are based on functional series expansions, block oriented systems and nonlinear differential or difference equation models. The present study is an attempt to provide a brief introduction and overview of some of the better known algorithms for each of these system descriptions.

Keywords. Identification, nonlinear systems, prediction, signal processing.

INTRODUCTION

There are basically three distinct approaches to the identification of nonlinear systems and these are based on functional series expansions, block oriented systems and parameter estimation methods applied to nonlinear differential or difference equation models. Each of the three approaches has distinct features and characteristics which arise both from the mathematical description which is used to characterise the system and the method of identification that is applied. The present study is an attempt to provide a brief introduction and overview of some of the better known algorithms for each of these system descriptions.

LINEAR OR NONLINEAR

The first stage in the identification procedure should be designed to indicate if the relationship between the input and output data is linear or nonlinear. The simplest method of achieving this objective is to inject step inputs of varying amplitude and plot the system gain against the input amplitude. If this is precluded because the process cannot be taken off normal production or if the data is pre-recorded or the analysis relates to the residuals then alternative tests must be used. Several tests are available and these have been surveyed by Haber (1985) who recommended the time domain test, higher order autocorrelation and nonlinear cross-correlation methods. The time domain test is just a generalisation of the step test whereby multilevel inputs are injected into the process, and the nonlinear correlation test is discussed in the section on block structured systems - see eqn (21). The higher order autocorrelation test which can be applied whenever the third order moments of the input are zero and all even order moments exist (e.g. a sine wave, gaussian input etc., etc.) consists of computing (Billings and Fadzil, 1985).

$$\phi_{z'z'^2}(\tau) = \overline{(z(t+\tau)-\bar{z})(z(\tau)-\bar{z})^2} \qquad (1)$$

where the bar indicates time average. It can readily be proved that $\phi_{z'z'^2}(\tau) = 0 \ \forall \ \tau$ iff the process is linear. The test will distinguish between additive noise corruption of the measurements and distortion due to nonlinear effects providing the noise and input are independent.

SYSTEM IDENTIFICATION

If an analytical model of the process under study has been derived but the coefficients are unknown standard optimisation procedures can be applied to yield estimates of the model parameters. This situation however seldom exists and often only a limited amount of a priori information is available and the process is treated as a black box. The choice of identification routine then becomes dependent upon the expansion which is chosen to represent the nonlinear system. There are several expansions which can be selected (Billings 1980, Mehra 1979), but most of these can be categorised as either a functional series, a block structured system or a differential/difference equation and each induces different identification algorithms which are briefly discussed in the following sections.

FUNCTIONAL SERIES

Early in the twentieth century Frechet showed that a large class of nonlinear time invariant systems can be represented by the Volterra series (Barrett 1979, Schetzen 1980).

$$y(t) = \sum_{n=1}^{\infty} \int_{-\infty}^{\infty} \ldots \int h_n(\tau_1, \ldots, \tau_n) \prod_{i=1}^{n} u(t-\tau_i) d\tau_i \qquad (2)$$

where $h_n(\tau_1, \ldots \tau_n)$ is the n'th order Volterra kernel. Convergance of the Volterra series for both deterministic and stochastic inputs has been studied in the literature. Identification of nonlinear systems based on the Volterra representation requires the measurement of the kernels $h_i(\tau_1, \tau_2, \ldots \tau_i)$. There are several approaches to this problem the most common being the extension of the correlation methods used for linear systems.

Consider for example the identification of a system which can be described by a Volterra series with just two terms

$$y(t) = \int_0^{\infty} h_1(\tau_1) u(t-\tau_1) d\tau_1 + \int_0^{\infty} \int h_2(\tau_1, \tau_2) u(t-\tau_1) u(t-\tau_2) \, d\tau_1 \, d\tau_2 \qquad (3)$$

Computing the first and second order cross-correlation functions yields

$$\phi_{uy}(\sigma) = \int_o^\infty h_1(\tau_1)\overline{u(t-\tau_1)u(t-\sigma)}\, d\tau_1 + \iint_o^\infty h_2(\tau_1,\tau_2)\overline{u(t-\tau_1)u(t-\tau_2)u(t-\sigma)}\, d\tau_1\, d\tau_2 \quad (4)$$

$$\phi_{uuy}(\sigma_1,\sigma_2) = \overline{y(t)u(t-\sigma_1)u(t-\sigma_2)} = \int_o^\infty h_1(\tau_1)\overline{u(t-\tau_1)\, u(t-\sigma_1)u(t-\sigma_2)}\, d\tau_1 + \iint_o^\infty h_2(\tau_1,\tau_2)\overline{u(t-\tau_1)u(t-\tau_2)u(t-\sigma_1)u(t-\sigma_2)}d\tau_1 d\tau_2 \quad (5)$$

The solution of these equations for a general stochastic input is extremely difficult. If however the input is selected to be white Gaussian the moments of the input are

$$E[u(t_1)u(t_2)\ldots u(t_1)] = 0 \quad i \text{ odd}$$
$$= \sum\prod \delta\,(t_n - t_m) \quad i \text{ even} \quad (6)$$
$$i \quad n\neq m$$

where the summation is over all ways of dividing i objects into pairs and eqn's (4) and (5) reduce to

$$\phi_{uy}(\sigma) = h_1(\sigma) \quad (7)$$

$$\phi_{uuy}(\sigma_1,\sigma_2) = \bar{y}\,\delta\,(\sigma_1-\sigma_2) + 2h_2(\sigma_1,\sigma_2) \quad (8)$$

Removing the mean of the output eliminates the first term in eqn (8) and estimates of the system kernels follow directly. This is of course a very idealised example. In practice the number of terms in the Volterra series would be unknown and the above method would break down whenever significant higher order terms were present.

Fakhouri (1980) has introduced an alternative approach to this problem based on the representation of the kernels in terms of their multidimensional pulse transfer functions. The algorithm is based on non-white Gaussian inputs, correlation and least squares and Fakhouri proves that using the approach each kernel can be identified sequentially. Korenberg (1983) uses a different approach, again based on correlation, but where the unidentified residue is successively approximated by a single basic model. Unlike the algorithm of eqn's (7) and (8) both of these methods will work for systems which are represented by Volterra series with more than two terms.

To facilitate the identification of nonlinear systems Wiener used the Volterra series as a basis and applied a Gram-Schmidt orthoganlisation procedure to construct a new functional series (Schetzen, 1980)

$$y(t) = \sum_{n=o}^{\infty} \{G_n(k_n,u(t)\} \quad (9)$$

where the functionals are orthogonal to one another for a Gaussian white stimulus and can therefore be identified in isolation. The first few terms of the Wiener functional series are

$$G_1\{k_1,u(t)\} = \int_{-\infty}^{\infty} k_1(\tau)u(t-\tau)\, d\tau$$

$$G_2\{k_2,u(t)\} = \iint_{-\infty}^{\infty} k_2(\tau_1,\tau_2)u(t-\tau_1)u(t-\tau_2)d\tau_1 d\tau_2 - P\int_{-\infty}^{\infty} k_2(\tau_1,\tau_1)\, d\tau_1 \quad (10)$$

where P represents the power spectral density of the input. Although the Wiener series is equivalent to the Volterra series, the former spans the function space more efficiently due to the orthoganality of the functionals. In general the Wiener kernels are not equal to the Volterra kernels.

Numerous methods have developed to identify the kernels in Wiener's series (Marmarelis and Marmarelis 1978) the most popular being a correlation method by Lee and Schetzen (Schetzen 1980). The procedure consists of computing multidimensional correlation functions between the white Gaussian input and the system output to yield

$$\hat{k}_n(\tau_1,\tau_2,\ldots\tau_n) = \frac{1}{n!P}\{y(t) - \sum_{m=o}^{n-1} G_m[k_m,u(t)]\}\, \overline{u(t-\tau_1)\ldots u(t-\tau_n)} \quad (11)$$
$$\forall\ \tau_1, \ldots\tau_n.$$

The second term on the rhs of eqn (11) only has a value on the diagonal and is included to remove impulse functions which would otherwise appear when $\tau_1=\tau_2\ldots\tau_n$. The amount of computation associated with eqn (11) can be excessive with computing time increasing almost exponentially with the order of the kernel to be evaluated. It has recently been shown (Palm and Poggio, 1978) that errors associated with the diagonal kernel estimates in the continuous time formulation can introduce fundamental difficulties in the identification of third and higher order kernels but these problems can easily be avoided by using appropriate discrete stochastic inputs.

Algorithms based on the expansion of eqn (9) are usually referred to as Wiener II methods. In the original formulation, referred to as the Wiener I method (Billings 1980), Wiener represented each functional term by a Fourier-Hermite series and expanded the kernels in a series of Laguerre functions $l_m(\tau)$ as

$$k_n(\tau_1,\ldots\tau_n) = \sum_{m_o=o}^{\infty}\ldots\sum_{m_n=o}^{\infty} C_{m_o\ldots m_n}\, l_{m_o}(\tau_1)\ldots l_{m_n}(\tau_n)$$

$$y(t) = \sum_{n=o}^{\infty} H_n\{k_n,u(t)\} \quad (12)$$

Laguerre functions were selected because they could be represented by a series of phase shifted electrical networks. Identification of the coefficient $C_{m_o\ldots m_n}$ is achieved by correlating the output of the unknown system $y_b(t)$ with the output $y_a(t)$ of a known system for a white Gaussian input to yield

$$C_{m_o\ldots m_n} = \frac{1}{n!P}\overline{y_a(t)y_b(t)} \quad (13)$$

The known system is constructed such that its output for a white Gaussian input is $G_n\{k_n,u(t)\}$ where the kernel in the leading term is synthesised as a product of Laguerre functions. Unfortunately identification of even a simple system containing a second order nonlinearity would require the evaluation of typically 10^{10} coefficients $C_{m_o\ldots m_n}$.

It is relatively easy to show that the coefficients in the series eqn (10) are the coefficients of the Hermite polynomials so that

$$y(t) = \sum_{n=o}^{\infty} G_n\{k_n,u(t)\} = \sum_{n=o}^{\infty} H_n\{k_n,u(t)\} \quad (14)$$

which illustrates the relationship (Billings, 1980) between the two distinct approaches which were developed by Wiener.

Since almost all the identification algorithms based on functional series expansions use a white Gaussian excitation the possibility of replacing this by pseudorandom sequences has been widely investigated in relation to the correlation based algorithms (Barker and Pradisthayon, 1970). Whilst pseudorandom sequences were found to reduce the computation time significantly.. The higher order moments contained anomalies and these induce errors in the kernel estimates. It has recently been shown however that the use of prbs inputs for non-linear systems can lead to disasterous results where even identifiability is lost.

Identification in the frequency domain (Tsoi,1979) has been studied by several authors based on higher order or polyspectra (Priestley, 1981) and FFT algorithms and often results in a considerable reduction in the computational requirement.

Volterra series have been widely applied in the analysis of nonlinear systems (Hung and Stark, 1977; Hung, Stark and Eykhoff, 1982, Barrett 1979) and several practical systems have been identified based on Wiener's series (Marmarelis and Marmarelis 1978, Schetzen 1980).

BLOCK STRUCTURED SYSTEMS

Block structured systems are systems which can be represented by interconnections of linear dynamic models and static nonlinear elements (Tsoi 1979, Haber and Keviczky, 1976). Although this class of models can be treated using the functional series expansions simplified identification techniques can in many cases be derived by exploiting the structural form of the model. The basic philosophy underlying this approach has been to avoid a black-box description by identifying the systems in terms of the individual component subsystems. The ideas can best be illustrated by considering the general model illustrated in Fig.1. which has been analysed by many authors.

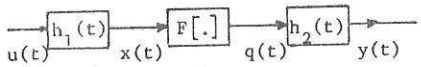

Fig.1. The General Model

Inspection of Fig.1. shows that the major difficulty in identifying this system will arise because of the presence of the static non-linearity $F[.]$. It is convenient therefore to consider this in isolation initially. Assume that the signal $x(t)$ can be measured and let $f(x_1,x_2;\tau)$ be the second order probability density function of the process $x(t)$ and define

$$g(x_2,\sigma) = \int_{-\infty}^{\infty} x_1 f(x_1,x_2;\sigma)\, dx_1 \qquad (15)$$

If this function separates as

$$g(x_2,\sigma) = g_1(x_2)g_2(\sigma) \;\forall\; x,\sigma \qquad (16)$$

then $x(t)$ is said to be a separable process (Billings and Fakhouri, 1982) where

$$g_1(x_2) = \frac{x_2 f(x_2)}{g_2(o)}\,,\quad g_2(\sigma) = \frac{g_2(o)\phi_{xx}(\sigma)}{\phi_{xx}(o)} \qquad (17)$$

The separable class of random processes is fairly wide and includes the Gaussian process, sine wave process, pseudorandom binary sequence etc.

Define the cross-correlation function

$$\phi_{xq}(\sigma) = \int\int_{-\infty}^{\infty} x_1 F[x_2] f(x_1,x_2;\sigma)\,dx_1 dx_2 \qquad (18)$$

Substituting from eqn's (15) and (17) yields

$$\begin{aligned}\phi_{xq}(\sigma) &= \int F[x_2] g_1(x_2) g_2(\sigma)\,dx_2 \\ &= \frac{\phi_{xx}(\sigma)}{\phi_{xx}(o)} \int F[x_2] x_2 f(x_2)\,dx_2 \\ &= C_F \phi_{xx}(\sigma) \end{aligned} \qquad (19)$$

where C_F is a constant. This is known as the invariance property and forms the basis of an identification algorithm for the general model in Fig.1. Realistically the internal signals in Fig. 1. $x(t)$ and $q(t)$ will not be available for measurement.

However by considering separability under linear and nonlinear transformation the result of eqn (19) can be generalised such that correlating between the input and output in Fig.1. for a non-zero mean Gaussian white input yields (Billings and Fakhouri 1982)

$$\begin{aligned}\phi_{uy}(\sigma) &= C_{FG} \int h_1(\tau_1) h_2(\sigma-\tau_1)\, d\tau_1 \\ \phi_{u^2y}(\sigma) &= C_{FFG} \int h_1^2(\sigma-\tau_1) h_2(\tau_1)\, d\tau_1 \end{aligned} \qquad (20)$$

where C_{FG} and C_{FFG} are constants providing $h_1(t)$ is stable bounded-inputs bounded outputs. The results of eqn (20) are, except for the constance C_{FG} and C_{FFG}, completely independent of the nonlinear element and this effectively decouples the identification of the linear and nonlinear component subsystems. The two unknowns $h_1(t)$ and $h_2(t)$ can be determined to within a constant scale factor by decomposing eqn (20) and the nonlinearity can then be determined.

Another advantage of this particular algorithm is that it provides information regarding the system structure. If the system under investigation is linear, then

$$\phi_{u^2y}(\sigma) = 0 \;\forall\; \sigma \qquad (21)$$

in eqn (20). Thus the first degree correlation function $\phi_{uy}(\sigma)$ yields an estimate of the linear system impulse response and the second degree correlation function $\phi_{u^2y}(\sigma)$ provides a convenient test for linearity which was referred to as the nonlinear correlation test in section 2. If the first and second degree correlation functions eqn (20) are equal except for a constant of proportionality then the system must have the structure of what is referred to as a Hammerstein model ($h_1(t) = \delta(t)$). However if the second degree correlation function is the square of the first degree correlation function, except for a constant of proportionality, the system has the structure of a Wiener model ($h_2(t) = \delta(t)$).

Similar results which provide estimates of the individual component subsystems are available for feedback, feedforward and multiplicative block oriented systems (Billings, 1980; Billings and Fakhouri, 1982; Korenberg 1973, Tsoi,1979).

DIFFERENCE EQUATION MODELS

Recent results in approximation and realisation

theory have produced nonlinear difference equation models that are suitable as a basis for identification (Fliess and Normand-Cyrot, 1982). The relevant results are briefly:-

(i) Any continuous causal functional can be be arbitrarily well approximated by a bilinear system (Fliess and Normand-Cyrot, 1982).

(ii) Any discrete time nonlinear system can subject to two mild assumptions be represented by the NARMAX (Nonlinear ARMAX) model (Leontaritis and Billings,1985)

$$y(k) = F[y(k-1),\ldots y(k-n_z),u(k-1)\ldots u(k-n_u)] \quad (22)$$

in a region around the equilibrium point where $F[.]$ is some nonlinear function. The model has been derived for multivariable stochastic nonlinear systems although only the SISO case is given here.

(iii) Subject to mild assumptions any discrete-time system can be arbitrarily well approximated by the state-affine model (Sontag, 1979)

$$x(k+1)=(A_o+\sum_{i=1}^{r} P_i(u_1(k),\ldots u_m(k))A_i)x(k)$$

$$y(k)=c(x(k)) \quad (23)$$

where the $P_i(u(k))$ are monomal in the inputs and A_o, A_1,..A_r are square matrices.

It is important to note that with discrete-time bilinear systems it is impossible to approximate the quadratic system $y(k)=u^2(k-1)$ (Fliess and Normand-Cyrot, 1982). Only results (ii) and (iii) above are therefore of interest since in almost all instances the data will be analysed in sampled data form.

Parameter estimation algorithms for the state-affine model have been developed (Dang Van Mien and Normand-Cyrot, 1984) by identifying a series of linear models around different operating points and then patching these together to form the final nonlinear description. This approach, which will work whenever the operating point changes slowly and smoothly, provides a great deal of insight into the operation of the process and has been applied to several industrial processes.

Identification based on the NARMAX model (Billings and Fadzil 1985, Leontaritis and Billings 1985) has concentrated on the stochastic form of eqn(22)

$$z(k+n) = F[z(k+n-1),\ldots z(k),u(k+n)\ldots,u(k),$$

$$\varepsilon(k+n-1),\ldots\varepsilon(k)] + \varepsilon(k+n) \quad (24)$$

where $\varepsilon(k) = E\{\varepsilon(k)|\ z(k-1),\ldots,u(k-1)\ldots\}$ represent the prediction errors and $F[.]$ has been selected as a polynomial. This leads to nonlinear difference equations which include powers and cross-products of delayed inputs, outputs and prediction error terms. The multiplicative terms between $\varepsilon(k-i), i = 1,\ldots n$ and past inputs and outputs occur because internal noise cannot in general be translated to be additive at the output whenever the system is nonlinear and because they are induced by the choice of model. Thus one disadvantage of including past outputs in the model expansion eqn(24) is that even additive output measurement noise induces multiplicative noise terms. As a consequence of this most of the parameter estimation techniques developed for linear systems cannot be applied directly and new algorithms have to be developed (Billings and Fadzil, 1985).

Whatever model structure is assumed the direct application of a maximum likelihood algorithm is not possible unless the distribution of the prediction errors, which for nonlinear systems will usually be non-gaussian, is known. However, by considering the loss function

$$J(\theta) = {}^{1}/_{2N} \log_e \det \sum_{k=1}^{N} \varepsilon(k,\theta)\varepsilon(k-\theta)^T \quad (25)$$

a prediction error estimate can be obtained (Goodwin and Payne 1977). For Gaussian innovations the prediction error estimate will be identical to the maximum likelihood estimate, but for non-gaussian innovations the asymptotic covariance of the prediction error estimates will be larger than the Cramer-Rao bound. The prediction error estimates and the Hessian associated with them will however be asymptotically normally distributed and this allows us to apply many of the classical statistical tests of significance and model order.

It is vitally important in nonlinear identification that the structure of the model is selected carefully prior to complete estimation. The maxim of continually increasing the model order or degree of polynomial expansion to improve the model fit is in many cases inappropriate for nonlinear systems since the latter action may increase the number of parameters in the model dramatically and could induce numerical problems. These difficulties can be avoided by using techniques such as stepwise regression and hypothesis testing routines that indicate which terms linear or nonlinear provide a significant contribution to the output (Billings and Fadzil 1985).

When the system is nonlinear the residuals $\xi(k)$ should be unpredictable from all linear and nonlinear combinations of past inputs and outputs and this condition will be satisfied iff

$$\phi_{\xi\xi}(\tau) = \delta(\tau)$$

$$\phi_{u\xi}(\tau) = 0 \ \forall\ \tau$$

$$\phi_{\xi\xi u}(\tau) = \overline{\xi(t)\xi(t-1-\tau)u(t-1-\tau)} = 0 \forall \tau \geq 0 \quad (26)$$

Experience has shown that the use of the tests in eqn (26) together with

$$\phi_{u^2\xi'}(\tau) = \overline{(u^2(t)-\overline{u^2})\xi(t+\tau)} = 0 \ \forall\ \tau$$

$$\phi_{u^2\xi^2}(\tau) = \overline{(u^2(t)-\overline{u^2})(\xi^2(t+\tau))} = 0 \ \forall\ \tau \quad (27)$$

often gives the experimenter a great deal of information regarding the deficiencies in the fitted model and can indicate which terms should be included in the model to improve the fit (Billings and Fadzil 1985). Notice that the traditional linear tests $\phi_{\xi\xi}(\tau)$ and $\phi_{u\xi}(\tau)$ are not sufficient.

CONCLUSIONS

The identification of nonlinear systems depends on both the construction of model representations which can approximate a large class of nonlinear systems and the development of identification or parameter estimation algorithms based on these descriptions. Although the choice between the various algorithms is wide it will often be dictated by the process, the choice of input, the amount of a priori information and the purpose of the identification.

The functional series methods can be used to represent a very broad class of nonlinear systems but this is achieved at the expense of a large parameter set. Identification of just the first two kernels in a Volterra series would for example require the evaluation of typically 500 kernel values. If it was necessary to identify the higher order kernels the computations could become horrendous. This problem which applies equally to almost all the functional series methods arises because the output is expanded in terms of past input values only. The restriction to a white Gaussian input signal is a further disadvantage of this approach. Nevertheless, numerous applications have been reported, methods of interpreting the kernels have been devised and the use of the expansions in the analysis of nonlinear systems is now well established.

Block structured systems represent a much smaller class of systems than functional series. They are however much easier to identify and this can often be achieved in a manner which preserves the original system structure and provides valuable physical insight and information for control. Even systems with hard nonlinearities can be identified using this approach which requires the estimation of typically 40-100 parameters. The restriction to special inputs such as Gaussian white noise is often required and of course the methods can only be applied if the process has the structure of or can be approximated by a block structured model.

Recent results in approximation and realization theory have produced nonlinear difference and state-space equation models that provide concise representations (usually with less than ten terms) of nonlinear sampled data systems which have been used as a basis for identification. The advantage of this approach is the reduction in the number of data points required (typically less than 1000), the small parameter set and the wide choice of input excitation that can be used. The parameter set will be greatly reduced compared to the functional series methods because the information on both past inputs and outputs is used to predict the current output value. Notice however that the number of parameters to be estimated can only be reduced to a small number if structure detection and validity tests are combined with the parameter estimation routines.

The identification of nonlinear systems is a difficult problem and just a few of the better known algorithms and some recent results have been described above. Much work remains to be done to unify the methods, to devise new algorithms and to apply the results to practical systems.

REFERENCES

Because of lack of space it has been impossible to include a comprehensive list of references. As a compromise I have included several previous survey papers together with references to some of the more recent algorithms.

Barker H.A., Pradisthayon T: (1970): High order autocorrelation functions of pseudorandom signals based on m-sequences; ProcIEE, 117, pp.1957-1263.

Barratt J.F. (1979); Bibliography on Volterra series, Hermite function expansions and related subjects; T.H. Report 77-E-71, Eindhoven Univ.

Billings S.A. (1980); Identification of nonlinear systems-a survey: ProcIEE, 127, pp. 272-285.

Billings S.A., Fadzil M.B, (1985); The practical identification of systems with nonlinearities; 7th IFAC Symp.Ident. and Syst. Par.Est, York.

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

Dang Van Mien H., Normand-Cyrot D (1984); Nonlinear state affine identification methods; applications to electrical power plants; Automatica, 20, pp. 175-188.

Fakhouri S.Y. (1980); Identification of the Volterra kernels of nonlinear systems; ProcIEE, Part D, 127, pp. 246-304.

Fliess M., Normand-Cyrot D. (1982); On the approximation of nonlinear systems by some simple state -space models; 6th IFAC Symp. Ident. & Syst. Par. Est, Washington DC, pp.433-436.

Goodwin G.C., Payne R.L. (1977); Dynamic system identification experiment design and data analysis; Academic Press.

Haber R. (1985); Nonlinearity tests for dynamic processes; 7th IFAC Symp. Ident. & Syst, Par. Est. York.

Haber R., Keviczky L.(1976): Identification of nonlinear dynamic systems: 4th IFAC Symp.Ident. & Syst. Par.Est, Tblissi.

Hung G., Stark L (1977): Introductory review: the kernel identification method (1910-1977) Math Biosci, 37, pp. 135-170.

Hung G, Stark L., Eykhoff P. (1982): On the interpretation of kernels; 6th IFAC Symp. Ident. & Syst. Par. Est, Washington DC, pp.417-420.

Korenberg M (1973) Cross-correlation analysis of neural cascades; 10th Rocky Mountain bio.eng. symp. pp. 47-51.

Korenberg M (1983): Statistical kernel identification from non-gaussian inputs; Research Note, Dept. of Electrical Eng. Queen's University, Ontario, Canada.

Leontaritis I.J., Billings S.A. (1985): Input-Output parametric models for nonlinear systems; Part I - Deterministic Nonlinear Systems; Part II - Stochastic Nonlinear Systems; Int. J. Control, 41,303-359.

Marmarelis P.Z., Marmarelis V.Z (1978); Analysis of physiological systems the white noise approach, Plenum Press.

Mehra R.K (1979); Nonlinear system identification selected survey and recent trends: 5th IFAC Symp. on Ident. and Syst. Par. Est, Darmstadt pp. 77-83.

Palm G., Poggio T (1978): Stochastic identification methods for nonlinear systems - an extension of the Wiener theory; SIAM J Appl. Maths, 34, pp. 524-534.

Priestley M.B (1981) Spectral Analysis and Time Series, Vol. 2, Academic Press.

Schetzen M (1980): The Volterra and Wiener theories of nonlinear systems; Wiley.

Sontag E.D. (1979): Polynomial response maps: Springer-Verlag.

Tsoi A.C (1979): Survey of recent advances in nonlinear system identification; Conf. on Control Engineering, Melbourne, 5-7 June, pp.86-92.