
Monograph:
Billings, S.A. (1985) Introduction to Nonlinear Systems Analysis and Identification. Research Report. Acse Report 277 . Dept of Automatic Control and System Engineering. University of Sheffield

---

INTRODUCTION TO NONLINEAR SYSTEMS

ANALYSIS AND IDENTIFICATION

S. A. Billings, BEng,PhD,CEng,MIEE,AFIMA,MInstMC

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

Research Report No. 277

June 1985

Invited contribution to the SERC Vacation School on "Signal Processing for Control", University of Warwick, 15-20 September 1985.

SERC Vacation School

"SIGNAL PROCESSING FOR CONTROL"

University of Warwick

15-20 September 1985

Introduction to Nonlinear Systems Analysis and Identification

Dr. S. A. Billings

Department of Control Engineering, University of Sheffield

1. Introduction

Most systems encountered in practice are nonlinear to some extent due to inherent distortion introduced by the components of the system such as saturation or because they include deliberately introduced nonlinear effects (e.g. bang-bang controllers). Any system for which the superposition principle does not hold is defined to be nonlinear. Nonlinear systems exhibit phenomena like jumps, limit cycles, hysteresis and chaotic motions which are not possible in linear systems. It is these characteristics which often dictate that the study of nonlinear systems is restricted to specific system structures.

The statistical analysis of nonlinear systems is in general an extremely difficult problem and a unified theory applicable to a broad class of systems does not exist. Systems which contain two or more single-valued nonlinear elements, multivalued nonlinearities or nonlinear functions of two or more system variables are particularly difficult to analyse and recourse is often made to either simulation or piece-wise linear analysis.

The present study briefly reviews some of the methods which are available for the statistical analysis of static and dynamic nonlinear systems including linearisation methods, system identification algorithms, and stochastic control.

2. Static Nonlinear Systems

Consider the system illustrated in Fig.1 where $u(t)$ is applied as an input to a single-valued instantaneous nonlinear element $N(\cdot)$ to produce an output $y(t)$.

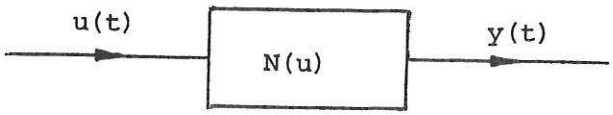


Fig.1

If the input is stationary in the strict sense the k'th order probability density function of y can be obtained from $f_u(u_1,u_2,\ldots u_k;\ t_1,t_2\ldots t_k)$ the density of u [1]. To determine $f_y(y_1..y_k;\ t_1,..t_k)$ solve the system of equations $y_1 = N(u_1),\ldots y_k = N(u_k)$ for $u_1,\ldots u_k$ and assuming a unique solution

$$f_y(y_1,..y_k;\ t_1,..t_k) = \frac{f_u(u_1,\ldots u_k;\ t_1,\ldots t_k)}{|N'(u_1)|\ldots|N'(u_k)|} \qquad \ldots(1)$$

For example consider the evaluation of the density $f_y(y;t)$ when $y(t) = u^2(t)$. When $y>0$, the eqn $y = x^2$ has two solutions $u_1 = \sqrt{y}$ and $u_2 = -\sqrt{y}$. Further since $|dy/du| = 2\sqrt{y}$, then from eqn (1)

$$f_y(y;t) = \frac{f_u(\sqrt{y};t) + f_u(-\sqrt{y};t)}{2\sqrt{y}}$$

If $y<0$, then $f_y(y) = 0$.

The moments of the output y, Fig.1, can however be expressed directly in terms of the probability density function of the input. The auto-correlation function of the output is for example given by

$$R_{yy}(\tau) = \int_{-\infty}^{\infty}\int y_1 y_2\ f(y_1,y_2;\tau)\,dy_1 dy_2 \qquad \ldots(2)$$

or $$R_{yy}(\tau) = \iint N(u_1)N(u_2)f(u_1,u_2;\tau)\,du_1 du_2 \qquad \ldots(3)$$

An analytical expansion of this integral can be obtained when u is stationary and normally distributed

$$f(u_1,u_2;\tau) = \frac{1}{2\pi\sigma^2(1-\rho^2)^{\frac{1}{2}}}\ \mathrm{Exp}\left\{-\left(\frac{u_1^2+u_2^2-2\rho u_1 u_2}{2\sigma^2(1-\rho^2)}\right)\right\} \qquad \ldots(4)$$

where
$$\begin{aligned} &E[u(t)] = 0 \\ &E[u^2(t)] = \sigma^2 \\ &E[u(t)u(t+\tau)] = \sigma^2\rho(\tau) \end{aligned} \qquad \ldots(5)$$

and $\rho(\tau)$ is the normalised covariance function of the input. Using Mercer's formula [2] eqn (4) can be expanded as

$$\frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\left(\frac{a_1^2+a_2^2-2\rho a_1 a_2}{2(1-\rho^2)}\right)\right\}$$

$$= \left\{\exp\left\{-\left(\frac{a_1^2}{2}+\frac{a_2^2}{2}\right)\right\}/2\pi\right\} \sum_{n=0}^{\infty} \rho^n Q_n(a_1) Q_n(a_2) \qquad \ldots(6)$$

where $Q_n(a)$ is the n'th order Hermite polynomial

$$Q_n(a) = \exp\left(\frac{a^2}{2}\right) \cdot \frac{(-1)^n}{\sqrt{n!}} \frac{d^n}{da^n}\left\{\exp\left(\frac{a^2}{2}\right)\right\} \qquad \ldots(7)$$

and $a_i = u_i/\sigma$.

Combining eqn's (3), (4) and (6) yields

$$R_{yy}(\tau) = \sum_{n=0}^{\infty} \rho^n k_i^2 \qquad \ldots(8)$$

where $k_i = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} N(\sigma a) Q_n(a) \exp\left(\frac{a^2}{2}\right) da \qquad \ldots(9)$

The autocorrelation function of the output eqn (8) is given therefore as a power series of the normalised autocorrelation function of the input where the coefficients k_i, eqn (9), depend on the form of nonlinearity.

For symmetrical non-linearities, $N(u) = -N(-u)$, all even coefficients k_{2_j} vanish, and for nonlinearities which can be expressed as truncated power series with zero coefficients for powers greater than j, then $k_i = 0$ for $i>j$. As an example, the output autocorrelation function for the bang-bang nonlinearity

$y = 1 \quad , \quad u>0$

$y = -1 \quad , \quad u<0$

is given by

$$R_{yy}(\tau) = \frac{2}{\pi} \mathrm{Sin}^{-1} \rho(\tau) \qquad \ldots(10)$$

Although the input-output cross-correlation function $R_{uy}(\tau)$ can be determined by following a similar procedure to the above a slightly more

general result can be obtained by introducing separable processes [3], [23].

Let $f(u_1,u_2;\tau)$ be the second order probability density function of the stationary process u(t) in Fig.1 and define

$$g(u_2,\tau) = \int_{-\infty}^{\infty} u_1 f(u_1,u_2;\tau)\,du_1 \qquad \ldots(11)$$

If the g-function separates as

$$g(u_2,\tau) = g_1(u_2)g_2(\tau) \quad \forall\ u,\tau \qquad \ldots(12)$$

then u(t) is said to be a separable process, where

$$g_1(u_2) = \frac{u_2 f(u_2)}{g_2(0)} \qquad \ldots(13)$$

$$g_2(\tau) = \frac{g_2(0)R_{uu}(\tau)}{R_{uu}(0)} \qquad \ldots(14)$$

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

Define the cross-correlation function

$$R_{uy}(\tau) = \iint_{-\infty}^{\infty} u_1 N(u_2) f(u_1,u_2;\tau)\,du_1 du_2 \qquad \ldots(15)$$

substituting from eqn's (12) through (14) yields

$$\begin{aligned} R_{uy}(\tau) &= \int N(u_2) g_1(u_2) g_2(\tau)\,du_2 \\ &= \frac{R_{uu}(\tau)}{R_{uu}(0)} \int N(u_2) u_2 f(u_2)\,du_2 \\ &= C_F R_{uu}(\tau) \end{aligned} \qquad \ldots(16)$$

where C_F is a constant scale factor.

Equation (16) which is known as the invariance property, shows that $R_{uy}(\tau)$ is directly proportional to $R_{uu}(\tau)$ for any static nonlinear characteristic $N(\cdot)$ providing u(t) is separable.

3. Nonlinear Systems with Dynamics

3.1 Functional Series Methods

A functional representation of nonlinear systems which is a generalisation of the linear convolution integral was first studied by Volterra early in the twentieth century. Volterra investigated analytic functionals and introduced the representation [4]

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

$$= \sum_{n=1}^{\infty} y_n(t) \qquad \ldots(17)$$

which has become known as the Volterra series. The functions $h_i(\tau_1, \ldots \tau_i)$ in eqn (17) are referred to as Volterra kernels. The kernels are bounded and continuous in each τ_j, symmetric functions of their arguments, and for causal systems $h_i(\tau_1, \ldots \tau_i) = 0$ for any $\tau_j < 0$. Systems which contain nonlinear memory elements such as hysteresis or backlash are excluded from the description of eqn (17).

Consider the Volterra series representation of the system illustrated in Fig.2.

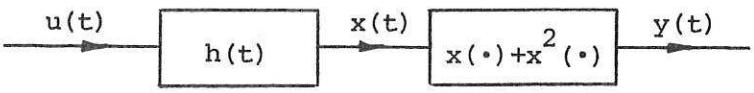


Fig.2

From the convolution integral

$$x(t) = \int h(\tau) u(t-\tau) d\tau$$

But $y(t) = x(t) + x^2(t)$ and therefore

$$y(t) = \int h(\tau_1) u(t-\tau_1) d\tau_1 + \iint h(\tau_2) h(\tau_3) u(t-\tau_2) u(t-\tau_3) d\tau_2 d\tau_3 \qquad \ldots(18)$$

is the Volterra series representation.

Taking the Fourier Transform of the n'th order kernel in eqn (17) yields the n'th order kernel transform [4]

$$H_n(j\omega_1, \ldots j\omega_n) = \int_{-\infty}^{\infty} \cdots \int h_n(\tau_1, \ldots \tau_n) \exp\{-j(\omega_1 \tau_1 + \ldots + \omega_n \tau_n)\} d\tau_1 \ldots d\tau_n$$

$$\ldots(19)$$

Similarly, considering just the n'th order component of the output $y_n(t)$ and taking Fourier transforms relates the multispectral density

$$Y_n(j\omega_1,\ldots j\omega_n) = H_n(j\omega_1,\ldots j\omega_n)U(j\omega_1)\ldots U(j\omega_n) \qquad \ldots(20)$$

to the input spectrum $U(j\omega)$, from which the output spectrum can be evaluated as

$$Y_n(j\omega) = \frac{1}{(2\pi)^{n-1}} \int_{-\infty}^{\infty} \ldots \int Y_n(j\omega - j\mu_1, j\mu_1 - j\mu_2, \ldots j\mu_{n-1}) d\mu_1 \ldots d\mu_{n-1} \qquad \ldots(21)$$

The statistical analysis of systems described by a Volterra series when the kernels are known has been studied extensively particularly with reference to communication systems [5-8] and Gaussian, sine wave plus Gaussian or random pulse train inputs. An excellent review of the use of functionals in the analysis of nonlinear systems is given by Barnett [9].

Consider the evaluation of the autocorrelation of the output, input-output cross-correlation and associated spectral densities for a system with known Volterra kernels assuming the output is strict-sense stationary. Define the correlation functions

$$R_{uy}(\tau) = E[y(t)u(t-\tau)]$$

$$= \sum_{n=1}^{\infty} R_{uy_n}(\tau) \qquad \ldots(22)$$

$$R_{yy}(\tau) = E[y(t)y(t-\tau)]$$

$$= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} R_{y_n y_m}(\tau) \qquad \ldots(23)$$

To evaluate the above expressions it is necessary to determine the partial correlation functions $R_{uy_n}(\tau)$, $R_{y_n y_m}(\tau)$. Rather than evaluating the general expressions which are given in the literature [5-8] consider the system illustrated in Fig.2 to illustrate the procedure.

Thus from eqn (18)

$$R_{uy}(\tau) = \int h(\tau_1) E\{u(t-\tau_1)u(t-\tau)\} d\tau_1$$

$$+ \int\int h(\tau_2)h(\tau_3) E\{u(t-\tau_2)u(t-\tau_3)u(t-\tau)\} d\tau_2 d\tau_3 \qquad \ldots(24)$$

$$R_{yy}(\tau) = \int\int h(\tau_1)h(\tau_2)E\{u(t-\tau_1)u(t-\tau-\tau_2)\}d\tau_1 d\tau_2$$

$$+ \int\int\int h(\tau_1)h(\tau_2)h(\tau_3)E\{u(t-\tau_1)u(t-\tau-\tau_2)u(t-\tau-\tau_3)\}d\tau_1 d\tau_2 d\tau_3$$

$$+ \int\int\int h(\tau_1)h(\tau_2)h(\tau_3)E\{u(t-\tau_1-\tau)u(t-\tau_2)u(t-\tau_3)\}d\tau_1 d\tau_2 d\tau_3$$

$$+ \int\int\int\int h(\tau_1)h(\tau_2)h(\tau_3)h(\tau_4)E\{u(t-\tau_1)u(t-\tau_2)u(t-\tau-\tau_3)u(t-\tau-\tau_4)\}$$
$$d\tau_1 d\tau_2 d\tau_3 d\tau_4 \qquad \ldots(25)$$

In general all the moments of the input process up to order n must be known before eqn's (24) and (25) can be evaluated. When the input is a zero mean white Gaussian process where

$$E\{u(t_1)u(t_2)\ldots u(t_i)\} = 0 \qquad i \text{ odd}$$
$$= \sum_i \prod_{n\neq m} \delta(t_n - t_m) \qquad i \text{ even} \qquad \ldots(26)$$

and the summation is over all ways of dividing i objects into pairs, eqn's (24) and (25) reduce to

$$R_{uy}(\tau) = h(\tau) \qquad \ldots(27)$$

$$R_{yy}(\tau) = \int h(\tau+\tau_2)h(\tau_2)d\tau_2$$
$$+ 2\int\int h(\tau+\tau_3)h(\tau+\tau_4)h(\tau_3)h(\tau_4)d\tau_3 d\tau_4 \qquad \ldots(28)$$

Spectral densities are computed by taking Fourier Transforms of the associated correlation functions and using eqn's (19)-(21). An algebra of nonlinear systems based on the Volterra series has been developed by George [10] and this simplifies the notation considerably in many problems.

3.2 The Fokker-Planck-Kolmogorov Equation

Consider the class of dynamic systems which can be represented by the stochastic vector differential equations

$$d\underline{x}/dt = A(\underline{x},t) + C(\underline{x},t)V(t) \qquad \ldots(29)$$

where $\underline{x} = \{x_i\}$ are the n-state variables, $A(x,t) = \{a_i\}$ and $C = \{c_{ij}\}$ are coefficient matrices and $\underline{V}(t)$ is an m-dimensional Gaussian white noise vector with the properties $E[\underline{V}(t)] = 0$, $E[\underline{V}(t)\underline{V}(s)^T] = Q\delta(t-s)$, $\underline{Q} = \text{diag}\{\sigma_{ii}^2\}$. $\underline{V}(t)$ can be used to represent random external disturbances, modelling discrepancies and random parametric perturbations.

Because Gaussian white noise is not mathematically meaningful, rewrite eqn (29) in terms of the incremental Wiener process $d\underline{W}(t) = \underline{V}(t)dt$ to yield the Itô [11,12] stochastic differential equation

$$d\underline{x} = \underline{A}(\underline{x},t)dt + \underline{C}(\underline{x},t)d\underline{W}(t) \qquad \ldots(30)$$

where

$$E\{d\underline{W}\} = 0 \ , \quad E\{d\underline{W}(t)d\underline{W}(s)^T\} = \underline{Q}\delta(t-s)dt$$

Equation (30) generates a Markov process $\underline{x}(t)$ since $\underline{x}(t)$ and $\underline{w}(t)$ are independent with independent increments. Consequently, the solution of eqn (30) is completely characterized by the first order probability density function $f(\underline{x},t)$ and the transitional probability density function $f(\underline{x},t_2/\underline{x},t_1)$ for $t_2>t_1$, both of which can be shown to satisfy the Fokker-Planck-Kolmogorov equation [1,11,12,13]

$$\frac{\partial f(\cdot)}{\partial t} = -\sum_{i=1}^{n} \frac{\partial}{\partial x_i}(a_i(\underline{x},t)f(\cdot)) + \tfrac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j}((\underline{C}\underline{Q}\underline{C}^T)_{ij}f(\cdot)) \qquad \ldots(31)$$

In general it is of more practical value to determine the moments $m_n = E[x_i^n]$ of the system states rather than the probability density function $f(\underline{x},t)$. Using either the method of moments or Itô's fundamental lemma it can be shown, for example, that the first two moments are given by the solution of the ordinary differential equations [13,14]

$$\frac{dE(x_i)}{dt} = E(a_i(\underline{x},t)) \qquad \ldots(32)$$

$$\frac{dE(x_i x_j)}{dt} = E(a_i x_i + a_j x_j) + E((\underline{C}\underline{Q}\underline{C}^T)_{ij}) \qquad \ldots(33)$$

for $(i,j) = 1,2\ldots n$, and given initial conditions $\underline{x}(t_o)$.

To illustrate the procedure consider a linear first order system with transfer function $\frac{1}{1+sT}$ driven by unity variance Gaussian white noise. The system model in state-space form is

$$Tdx = -xdt + dw \qquad \ldots(34)$$

and from eqn (31) its associated Fokker-Planck-Kolmogorov equation is

$$\frac{\partial f}{\partial t} = \frac{1}{T}\frac{\partial(xf)}{\partial x} + \frac{1}{2T}\frac{\partial^2 f}{\partial x^2} \qquad \ldots(35)$$

The first two moments are, from eqn's (32) and (33)

$$\frac{dm_1}{dt} = -\frac{1}{T} m_1 \qquad \ldots(36)$$

$$\frac{dm_2}{dt} = -\frac{2}{T} m_2 + \frac{1}{T^2} \qquad \ldots(37)$$

with solutions

$$m_1(t) = m_1(0)e^{-t/T} \qquad \ldots(38)$$

$$m_2(t) = m_2(0)e^{-2t/T} + \frac{1}{2T}(1-e^{-2t/T}) \qquad \ldots(39)$$

where $m_1(0)$ and $m_2(0)$ are initial conditions.

If the system equations (29) are nonlinear in the states the lower order moment eqn's (32) and (33) are in general functions of higher order moments and closed analytic solutions are not possible. Similarly the probability density function $f(\underline{x},t)$ which satisfies the Fokker-Planck-Kolmogorov equation cannot be found except by linearization or approximation methods. An alternative is to simulate the stochastic difference eqn (29) and evaluate the required moments by averaging over the realisations [13-15].

4. Linearisation Methods

The relative simplicity of the methods of statistical analysis for linear systems compared with the inherent complexity of the analysis outlined above has led to the development of approximation methods based on linearisation techniques [16].

The simplest form of linearisation is based upon the expansion of the nonlinear function in a Taylor series about some operating point and retaining only the linear terms in the analysis. Thus the nonlinear function $g(x_1,\ldots x_n)$ is replaced by the approximate expression

$$g(x_1,\ldots x_n) \approx g(\bar{x}_1,\ldots,\bar{x}_n) + \sum_{i=1} g'_{x_i}(\bar{x}_1,\ldots,\bar{x}_n)(x_i-\bar{x}_i) \qquad \ldots(40)$$

where $\bar{x}_i$ is the mean of x_i and $g'_{x_i} = \partial f/\partial x_i$.

Although eqn (40) is linear with respect to fluctuations it is nonlinear with respect to expectations.

Equation (40) is valid only for continuous functions with continuous first derivatives, and cannot therefore be used to study the characteristics of discontinuous components such as relays or limiters. To linearise such characteristics the method of statistical linearisation was developed.

The earliest method of statistical linearisation was developed by Booton [17] for static nonlinearities and stochastic inputs with zero mean. Booton's method consists of replacing the nonlinearity N(•) by an equivalent gain which is selected so as to minimise the mean square of the difference between the output of the devices.

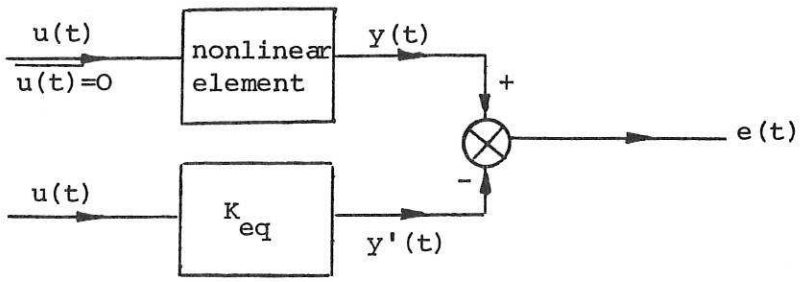


Fig.3

Consider the system illustrated in Fig.3 where

$$e(t) = y(t) - K_{eq}u(t)$$

$$\overline{e^2(t)} = \overline{y^2(t)} - 2K_{eq}\,\overline{u(t)y(t)} + K^2_{eq}\,\overline{u^2(t)} \qquad \ldots(41)$$

Selecting K_{eq} so as to minimise $\overline{e^2(t)}$ yields

$$K_{eq} = \frac{\overline{u(t)y(t)}}{\overline{u^2(t)}} = \frac{\int yuf(u)\,du}{\int u^2 f(u)\,du} \qquad \ldots(42)$$

When the input is Gaussian white $\int u^2 f(u)\,du = \sigma^2$ and hence

$$K_{eq} = \frac{1}{\sigma^2}\int yuf(u)\,du \qquad \ldots(43)$$

It can readily be shown [18] that K_{eq} in eqn (43) is equivalent to the first term in the Wiener series representation eqn (53) of a nonlinear zero memory system.

Somerville and Atherton [19] extended Booton's method to include cases of non-zero mean input signals as illustrated in Fig.4 to yield

$$K_{dc} = \frac{\overline{y(t)}}{\overline{u(t)}}$$

$$K_{eq} = \frac{\overline{u(t)y(t)} - \overline{u(t)}\cdot\overline{y(t)}}{\overline{u^2(t)} - \overline{u(t)}^2} \qquad \ldots(44)$$

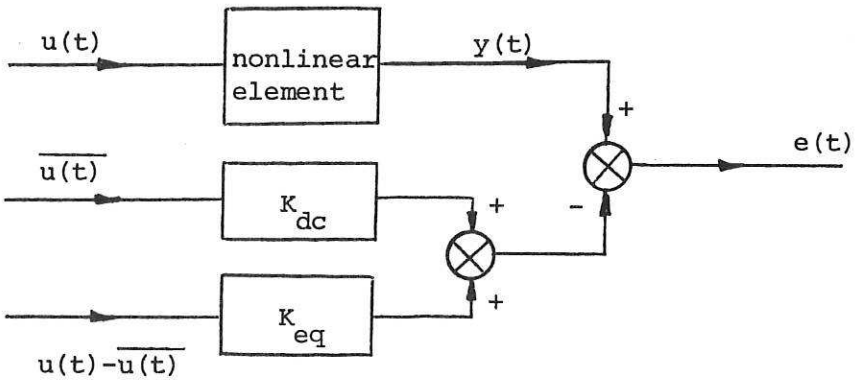


Fig.4

To conserve the spectrum of the output Pupkov [20] proposed replacing the zero memory nonlinear component N(•) by a dynamic stationary linear system

$$N(u(t)) = K_{dc}\overline{u(t)} + \int_{-\infty}^{t} h(t_1)(u(t-t_1) - \overline{u(t-t_1)})dt_1 \qquad \ldots(45)$$

where h(t) is determined to ensure the autocorrelation equivalence of the left and right hand sides of eqn (45)

$$\int_{-\infty}^{\infty}\int h(t_1)h(t_2)R_{uu}(\tau-t_1-t_2)dt_1dt_2 = R_{yy}(\tau) \qquad \ldots(46)$$

and K_{dc} is given by eqn (44).

If h(t) in eqn (45) ($\overline{u(t)} = 0$) is selected to minimise the mean squared error this leads to the Wiener-Hopf equation

$$\int h(t_1)R_{uu}(t_1-\tau)d\tau = R_{uy}(\tau) \qquad \ldots(47)$$

Although this equation is in general difficult to solve for h(t), when the input is a separable process [3] from eqn (16)

$$R_{uy}(\tau) = C_F R_{uu}(\tau)$$

and the optimum linear approximation to the nonlinear element is Booton's equivalent gain $C_F = K_{eq}$.

Harmonic linearisation, which yields the describing function [21] consists in the simplest case of replacing the nonlinear element by a linear one whose transfer function is equal to the complex ratio of the fundamental component of the output to the sinusoidal input. The method has been used extensively to obtain mathematical descriptions of a large number of nonlinear systems and several modifications have been developed including describing functions which relate all the output harmonics to the input fundamentals.

5. Identification

Identification algorithms for nonlinear systems can be categorised as functional series methods, algorithms for block oriented systems and parameter estimation techniques [22,23].

5.1 Functional Series and Block Structured Algorithms

Identification using the Volterra series representation eqn (17) involves the measurement of the Volterra kernels. To illustrate the approach consider the identification of a system which can be described by just the first two Volterra kernels

$$y(t) = \int_0^\infty h_1(\tau_1)u(t-\tau_1)d\tau_1 + \int_0^\infty\int h_2(\tau_1,\tau_2)u(t-\tau_1)u(t-\tau_2)d\tau_1 d\tau_2 \quad \ldots(48)$$

Defining the mean squared error as $E\{(z(t)-y(t))^2\}$ where z(t) is the measured output and applying calculus of variations yields

$$E\{z(t)\} = \int_0^\infty h_1(\tau_1)E\{u(t-\tau_1)\}d\tau_1 + \int_0^\infty\int h_2(\tau_1,\tau_2)E\{u(t-\tau_1)u(t-\tau_2)\}d\tau_1 d\tau_2 \quad \ldots(49)$$

$$E\{z(t)u(t-\sigma)\} = \int_0^\infty h_1(\tau_1)E\{u(t-\tau_1)u(t-\sigma)\}d\tau_1 + \int_0^\infty\int h_2(\tau_1,\tau_2)E\{u(t-\tau_1)u(t-\tau_2)u(t-\sigma)\}d\tau_1 d\tau_2 \quad \ldots(50)$$

$$E\{z(t)u(t-\sigma_1)u(t-\sigma_2)\} = \int_0^\infty h_1(\tau_1)E\{u(t-\tau_1)u(t-\sigma_1)u(t-\sigma_2)\}d\tau_1$$

$$+ \int_0^\infty\int h_2(\tau_1,\tau_2)E\{u(t-\tau_1)u(t-\tau_2)u(t-\sigma_1)u(t-\sigma_2)\}d\tau_1 d\tau_2 \quad \ldots(51)$$

The solution of this set of equations for a general stochastic input is extremely difficult. However, if the system input is white Gaussian substituting eqn (26) in eqn's (49)-(51) yields

$$\overline{z(t)} = \int_0^\infty h_2(\tau,\tau)d\tau$$

$$R_{uz}(\sigma_1) = h_1(\sigma_1) \quad \ldots(52)$$

$$R_{uuz}(\sigma_1,\sigma_2) = \bar{z}\delta(\sigma_1-\sigma_2) + 2h_2(\sigma_1,\sigma_2)$$

and the solution for $h_1(t)$ and $h_2(t_1,t_2)$ is direct providing the mean level $\bar{z}$ is removed. Identification of systems which contain higher than second order kernels is very difficult using this approach. Alternative schemes involve approximating the kernels by an expansion of orthogonal functions and estimating the coefficients [22,23].

Wiener used a Gram-Schmidt orthogonalisation procedure to construct a new functional series where the functionals $\{G_n\}$ are orthogonal for a Gaussian white stimulus. The first two terms in the Wiener series are

$$G_1[k_1,u(t)] = \int_{-\infty}^\infty k_1(\tau)u(t-\tau)dt \quad \ldots(53)$$

$$G_2[k_2,u(t)] = \int_{-\infty}^\infty\int k_2(\tau_1,\tau_2)u(t-\tau_1)u(t-\tau_2)d\tau_1 d\tau_2$$

$$- P\int_{-\infty}^\infty k_2(\tau_1,\tau_1)d\tau_1 \quad \ldots(54)$$

where P is the power spectral density of the white noise input. In general the Wiener kernels are not equal to the Volterra kernels. Numerous methods have been developed to identify the kernels in Wiener's series [22,23] the most popular being a correlation method by Lee and Schetzen [4]. The procedure consists of computing multidimensional correlation functions between the white Gaussian input and the system output to yield

$$k_n(\tau_1,\ldots\tau_n) = \frac{1}{n!P}\overline{\{y(t) - \sum_{m=o}^{n-1} G_m[k_m,u(t)]\}u(t-\tau_1)\ldots u(t-\tau_n)}^{\tau_1,\tau_2\ldots\tau_n} \quad \ldots(55)$$

In an attempt to reduce the computational burden associated with the functional-series methods various authors have considered the identification of block oriented systems [23,24] which can be represented by interconnections of linear dynamic systems and static nonlinear elements. Consider the system illustrated in Fig.5 to illustrate the approach.

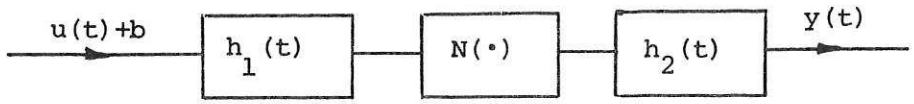


Fig.5

By extending the theory of separable processes [22-24] and using the result of eqn (16) it can readily be shown that for a Gaussian white input with mean level b

$$R_{uy'}(\sigma) = C_{FG}\int h_1(\tau_1)h_2(\sigma-\tau_1)d\tau_1 \quad \ldots(56)$$

$$R_{u^2y'}(\sigma) = C_{FFG}\int h_1^2(\sigma-\tau_1)h_2(\tau_1)d\tau_1 \quad \ldots(57)$$

where providing $h_1(t)$ is stable bounded-inputs bounded outputs C_{FG} and C_{FFG} are constants and the superscript ' indicates that the mean level has been removed from the signal. Estimates of the individual linear subsystems $h_1(t)$ and $h_2(t)$ can be obtained by decomposing eqn's (56) and (57) [24] and the nonlinearity can then be determined. The results of eqn's (56), (57) inherently provide information regarding the structure of the nonlinear system and this can be used to determine the position of the nonlinear element prior to complete identification. Similar results, which provide estimates of the individual component subsystems, are available for feedback, feedforward and multiplicative block oriented systems [24].

5.2 Parameter Estimation Algorithms

Parameter estimation methods for nonlinear systems where the structural form of the describing differential equations are known are now well established [22,23]. When little a priori information is available and the process is treated as a black-box, the usual approach is to expand the input/output using a suitable model representation. Two particular choices of model expansion, the NARMAX model and piecewise linear models will be briefly considered in the present analysis.

5.2.1 The NARMAX Model

If a system is linear then it is finitely realizable and can be represented by the linear difference equation model

$$y(k) = \sum_{i=1}^{n_y} (a_i y(k-i)) + \sum_{i=1}^{n_u} (b_i u(k-i)) \qquad \ldots(58)$$

if the Hankel matrix of the system has finite rank. When the system is nonlinear a similar representation can be derived by utilizing concepts from Nerode realization, multistructural forms and results from differential geometry to yield the nonlinear difference equation model [25]

$$y(k) = F^*[y(k-1),\ldots y(k-n_y),u(k-1),\ldots u(k-n_u)] \qquad \ldots(59)$$

where $F^*[\cdot]$ is some nonlinear function of $u(\cdot)$ and $y(\cdot)$. The model of eqn (59) can be shown [25] to exist whenever

(i) the state-space of the Nerode realization does not have infinite dimensions (i.e. we exclude distributed parameter systems), and

(ii) the linearized system around the origin has a Hankel matrix of maximum rank (i.e. a linearized model would exist if the system were operated close to an equilibrium point).

Equation (59) represents the single-input single-output case but the results have been extended to include multivariable systems. The Hammerstein, Wiener, bilinear, Volterra and other well known nonlinear models can be shown to be special cases of eqn (59).

An equivalent representation for nonlinear stochastic systems can be derived by considering input-output maps based on conditional probability

density functions to yield the model

$$z(k) = F\left[z(k-1),\ldots z(k-n_z),u(k-1),\ldots u(k-n_u),\varepsilon(k-1),\ldots \varepsilon(k-n_\varepsilon)\right] + \varepsilon(k) \qquad \ldots(60)$$

where $\varepsilon(k)$ is the prediction error. This model is referred to as the Nonlinear AutoRegressive Moving Average model with eXogenous inputs or NARMAX model [25].

A NARMAX model with first order dynamics expanded as a second order polynomial nonlinearity would for example be represented as

$$\begin{aligned} y(k) &= F_2\left[y(k-1),u(k-1)\right] \\ &= C_1 y(k-1) + C_2 u(k-1) + C_{11} y^2(k-1) + C_{12} y(k-1)u(k-1) + C_{22} u^2(k-1) \end{aligned} \qquad \ldots(61)$$

Assuming that the output measurements are corrupted by additive noise

$$z(k) = y(k) + e(k)$$

gives the input-output model

$$\begin{aligned} z(k) = {} & C_1 z(k-1) + C_2 u(k-1) + C_{11} z^2(k-1) + C_{12} z(k-1)u(k-1) \\ & + C_{22} u^2(k-1) + e(k) - C_1 e(k-1) - 2C_{11} z(k-1)e(k-1) \\ & + C_{11} e^2(k-1) - C_{12} e(k-1)u(k-1) \end{aligned} \qquad \ldots(62)$$

Because the NARMAX model maps the past input and output into the present output multiplicative noise terms are induced in the model even though the noise was additive at the output. In general the noise may enter the system internally and because the system is nonlinear it will not always be possible to translate it to be additive at the output. This situation will again result in multiplicative noise terms in the NARMAX model with the added complication that the noise source and the prediction error will not in general be equal. Since most of the parameter estimation techniques derived for linear systems assume that the noise is independent of the input, biased estimates result when they are applied to nonlinear systems eqn (60).

The recursive extended least squares (RELS) algorithm can however be readily adapted to the NARMAX model, by defining the following vectors

$$Q(k) = [z(k-1), u(k-1), z^2(k-1), z(k-1)u(k-1), u^2(k-1), \varepsilon(k-1),$$
$$\varepsilon(k-1)z(k-1), u(k-1)\varepsilon(k-1), \varepsilon^2(k-1)]^T$$

$$\hat{\theta} = [\hat{c}_1, \hat{c}_2 \ldots \hat{c}_s]^T$$

$$\varepsilon(k+1) = z(k+1) - Q(k+1)^T \hat{\theta}(k) \qquad \ldots(63)$$

for the model of eqn (62) for example. With these definitions the standard RELS algorithm can be applied to yield unbiased parameter estimates. The development of recursive maximum likelihood and instrumental variable algorithms for the NARMAX model is not quite so straightforward [26].

The direct application of an offline maximum likelihood algorithm is not possible because in general the prediction errors will not have a Gaussian distribution. However, by considering the loss function

$$J(\theta) = \frac{1}{2N} \log_e \det \sum_{k=1}^{N} \varepsilon(k;\theta)\varepsilon^T(k;\theta) \qquad \ldots(64)$$

it can be shown that the prediction error estimates obtained by minimising eqn (64) have very similar asymptotic properties to the maximum likelihood estimates even when $\varepsilon(k)$ is non-gaussian. A prediction error algorithm has been developed for the NARMAX model based on this result. This together with least squares derived algorithms [26] have been augmented with a stepwise regression algorithm, a likelihood ratio test and Akaike tests to detect the model structure or significant terms in the model prior to final estimation [28].

Whichever model formulation or identification algorithm is implemented it is important to test that the identified model does adequately describe the data set. When the system is nonlinear the residuals $\zeta(k)$ should be unpredictable from all linear and nonlinear combinations of past inputs and outputs and this condition will hold iff [27]

$$\begin{aligned} R_{\zeta\zeta}(\tau) &= \delta(\tau) \\ R_{u\zeta}(\tau) &= 0 \;\forall\; \tau \\ R_{\zeta\zeta u}(\tau) &= E[\zeta(k)\zeta(k-1-\tau)u(k-1-\tau)] = 0 \;\forall\; \tau \geq 0 \end{aligned} \qquad \ldots(65)$$

Notice that for nonlinear systems the traditional linear tests $R_{\zeta\zeta}(\tau)$ and $R_{u\zeta}(\tau)$ are not sufficient. If instrumental variables or suboptimal least squares are used the residuals may be coloured. It can be shown

that in this case the process model is unbiased iff

$$R_{u\zeta}(\tau) = 0 \;\forall\, \tau$$

$$R_{u^{2'}\zeta}(\tau) = E\left[(u^2(k)-\overline{u^2})\zeta(k+\tau)\right] = 0 \;\forall\, \tau \qquad \ldots(66)$$

$$R_{u^{2'}\zeta^2}(\tau) = E\left[(u^2(k)-\overline{u^2})\zeta^2(k+\tau)\right] = 0 \;\forall\, \tau$$

Experience has shown that when using a prediction error algorithm the tests in both eqn's (65) and (66) often give 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 [28].

5.2.1.1 An Example

To illustrate some of the ideas associated with parameter estimation based on the NARMAX model consider the identification of a model relating the input volume flow rate u(t) and the level of liquid z(t) in the interconnected tanks illustrated in Fig.6.

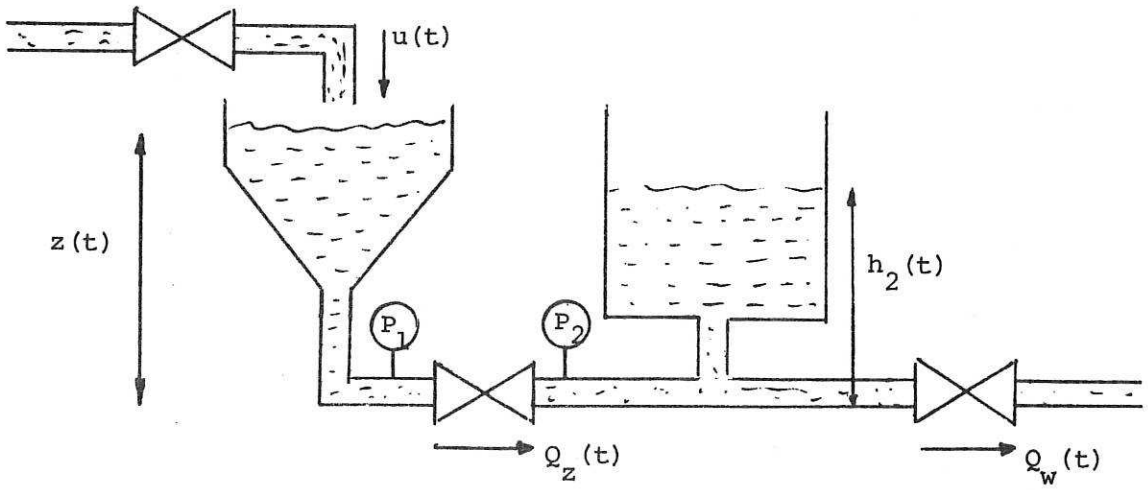


Fig.6. Interconnected Tanks

A zero mean Gaussian signal was used to perturb the input u(t) and 1000 data pairs were recorded by sampling the input and output at 9.6 secs.

In the early stages of any identification procedure it is important to establish if the process under test exhibits nonlinear characteristics which will warrant a nonlinear model. This can readily be achieved using a simple correlation test [23,27]. If the third order moments of

the input are zero and all even order moments exist (a sine wave, gaussian or ternary sequence would for example satisfy the properties) then the process is linear iff

$$R_{z'z'^2}(\tau) = E\left[(z(k)-\bar{z})(z(k+\tau)-\bar{z})^2\right] = 0 \;\forall\, \tau \qquad \ldots(67)$$

$R_{z'z'^2}(\tau)$ for the liquid level system is illustrated in Fig.7 and clearly shows that, as expected, the liquid level system is highly nonlinear.

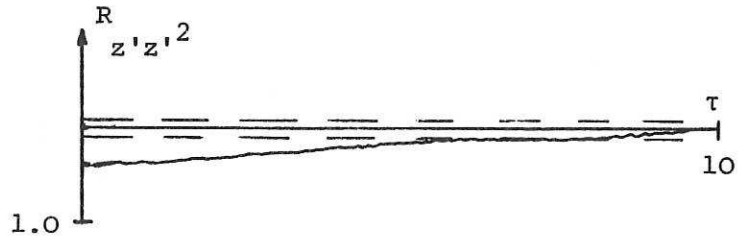


Fig.7. Nonlinear detection test

Initially a linear model was fitted to the data using a maximum likelihood algorithm to give the representation

$$\begin{aligned} z(k) = {} & 0.746z(k-1)+0.340z(k-2)-0.122z(k-3) \\ & +0.471u(k-1)-0.174u(k-2)-0.040u(k-3) \\ & +\varepsilon(k)+0.423\varepsilon(k-1)+0.038\varepsilon(k-2) \end{aligned} \qquad \ldots(68)$$

A comparison of the process and linear model predicted output is illustrated in Fig.8. The model validity tests eqn's (65), (66) for this model are illustrated in Fig.9. Notice that although $R_{\zeta\zeta}(\tau)$ and $R_{u\zeta}(\tau)$ indicate linear adequacy for the model eqn (68), $R_{u^2,\zeta}(\tau)$ and $R_{u^2,\zeta^2}(\tau)$ are well outside the 95% confidence bands indicating that nonlinear terms should be included in the model description.

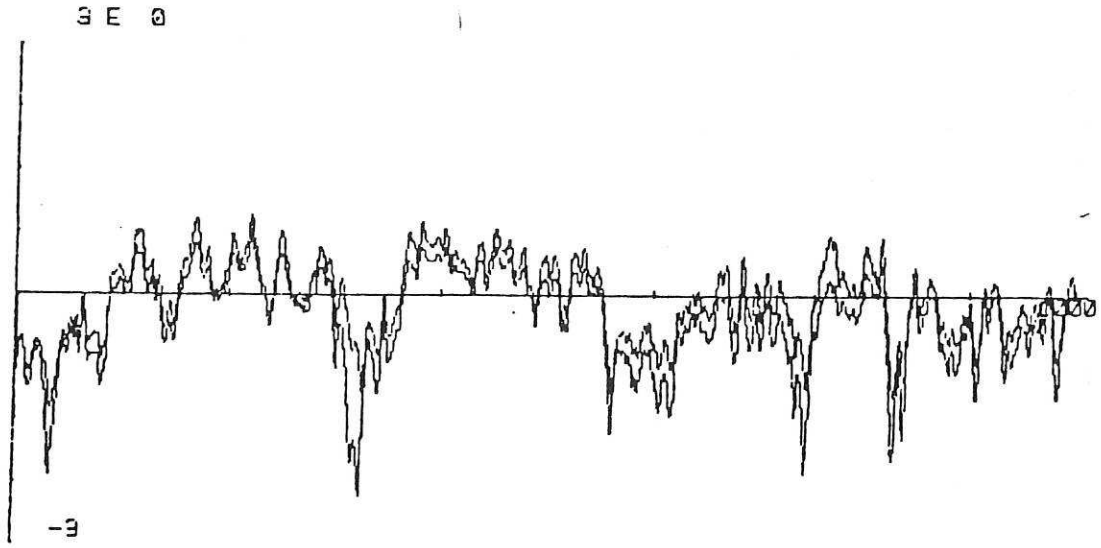


Fig.8. Process and Predicted output-best linear model

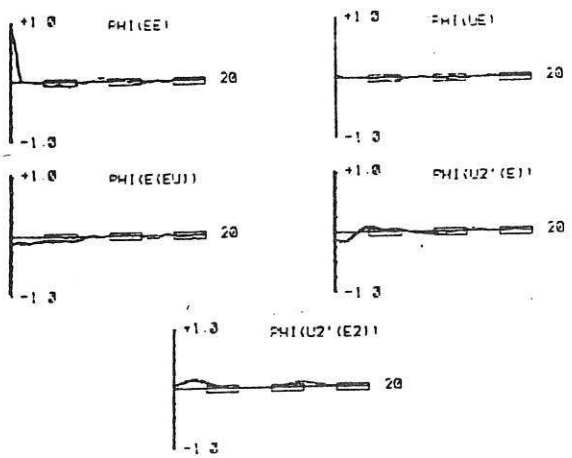


Fig.9. Model Validation - best linear model

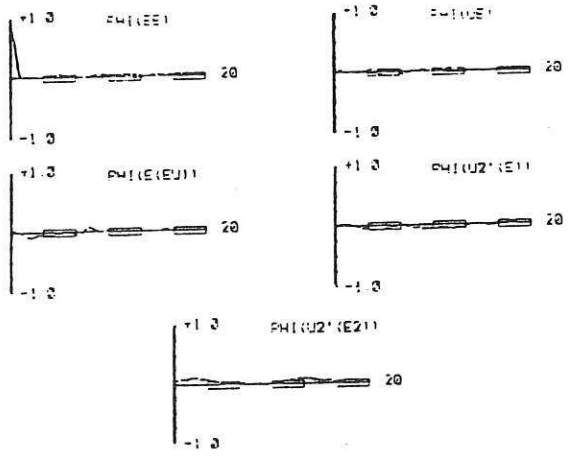


Fig.11. Model validation - best nonlinear model

The effect of introducing nonlinear terms into the model was therefore investigated and a prediction error algorithm yielded the NARMAX model representation

$$\begin{aligned} z(k) = {} & 0.436z(k-1)+0.681z(k-2)-0.149z(k-3) \\ & +0.396u(k-1)+0.014u(k-2)-0.071u(k-3) \\ & -0.351z(k-1)u(k-1)-0.034z^2(k-2) \\ & -0.135z(k-2)u(k-2)-0.027z^3(k-2)-0.108z^2(k-2)u(k-2) \\ & -0.099u^3(k-2)+\varepsilon(k)+0.344\varepsilon(k-1)-0.201\varepsilon(k-2) \end{aligned} \qquad \ldots(69)$$

The model validity tests for the model of eqn (69) are illustrated in Fig.11 and these together with the comparison of the NARMAX model predicted and process output Fig.10 show the considerable improvement in the prediction capabilities of the estimated NARMAX model eqn (69) compared with the best linear model eqn (68).

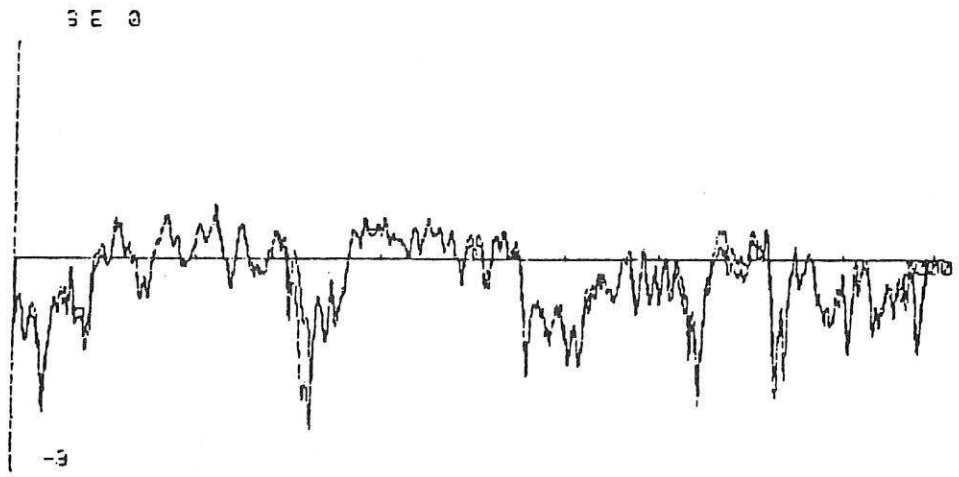


Fig.10. Process and Predicted output for the estimated NARMAX model

5.2.2 Piecewise Linear Modelling

There are several possible ways in which nonlinear systems can be approximated by locally linear models. These involve either the expansions of the NARMAX model eqn (60) using spline functions, spatial piecewise

linear models or linear models with signal dependent parameters [29]. The last of these three representations can be fitted by performing a series of linear experiments and repeatedly utilizing linear parameter estimation routines to build up a series of linearized models of the process. This can offer several advantages compared with fitting a global nonlinear model and may be appropriate in situations where this type of approximate system description is adequate. To illustrate the ideas involved we will just consider signal dependent linear modelling.

Linearizing the NARMAX model eqn (60) at a selected operating point $\Delta_k = [z_1,\ldots z_{n_z}, u_1,\ldots u_{n_u}]$, $\Delta_k \varepsilon R$ gives

$$dz(k) = \sum_{i=1}^{n_z} \frac{\partial F[\cdot]}{\partial z(k-i)}\bigg|_{\Delta_k} dz(k-i) + \sum_{i=1}^{n_u} \frac{\partial F[\cdot]}{\partial u(k-i)}\bigg|_{\Delta_k} du(k-i) + d\varepsilon(k) \qquad \ldots(70)$$

where for simplicity the noise is assumed to be represented by a single uncorrelated prediction error term $\varepsilon(k)$. Substituting $dz(k) = z(k)-z(k)|_{\Delta_k}$, $du(k) = u(k)-u(k)|_{\Delta_k}$, $d\varepsilon(k) = \varepsilon(k)-\varepsilon(k)|_{\Delta_k}$ into eqn (70) and manipulating gives

$$z(k) = \theta_o|_{\Delta_k} + \sum_{i=1}^{n_z} \theta_i|_{\Delta_k} z(k-i) + \sum_{i=1}^{n_u} \theta_{nz+i}|_{\Delta_k} u(k-i) + \varepsilon(k) \qquad \ldots(71)$$

where

$$\theta_o|_{\Delta_k} = z(k)|_{\Delta_k} - \varepsilon(k)|_{\Delta_k} - \sum_{i=1}^{n_z} \frac{\partial F[\cdot]}{\partial z(k-i)}\bigg|_{\Delta_k} z(k-i)|_{\Delta_k} - \sum_{i=1}^{n_u} \frac{\partial F[\cdot]}{\partial u(k-i)}\bigg|_{\Delta_k} u(k-i)|_{\Delta_k}$$

$$\theta_i|_{\Delta_k} = \frac{\partial F[\cdot]}{\partial z(k-i)}$$

$$\theta_{n_z+i}|_{\Delta_k} = \frac{\partial F[\cdot]}{\partial u(k-i)} \qquad \ldots(72)$$

If the process is such that the parameters in the linearized model eqn (71) depend on a signal $\omega(k)$ then eqn (71) can be written as

$$z(k) = \theta_1(\omega(k)) + \sum_{i=1}^{n_z} \theta_i(\omega(k))z(k-i) + \sum_{i=1}^{n_u} \theta_{n_z+i}(\omega(k)).u(k-i) \quad \ldots(73)$$

Equation (73) is a signal dependent linear model where $\omega(k)$ is the coefficient dependent signal which may depend on the input, the output or some external variable associated with the operation of the process. In the modelling of a power station for example it may be appropriate to select $\omega(k)$ as the megawatt output. In the liquid level system it would be the level of liquid in the first tank.

Parameter estimation for signal dependent models is usually performed in two stages. Initially the coefficients in a series of linear models as $\omega(k)$ is varied are estimated. These locally linear models are then patched together to form an approximate global nonlinear description of the system under investigation [29]. For example eqn (73) can be expressed as

$$z(k) = \psi^T(k)\underline{\theta}(\omega(k)) + \varepsilon(k) \quad \ldots(74)$$

where the definitions of $\psi(\cdot)$ and $\underline{\theta}(\cdot)$ follow directly. Assuming that the signal dependent parameter vector $\underline{\theta}(\omega(k))$ can be approximated by a finite degree polynomial we can write

$$\theta_i(\omega(k)) = \beta_i^T W(k) \quad , \quad i = 0,\ldots n_z+n_u \quad \ldots(75)$$

where

$$W(k) = \left[1, \omega(k), \omega^2(k), \ldots \omega^{\ell}(k)\right]^T$$

$$\beta_i = \left[\beta_{i0}, \beta_{i1}, \beta_{i2} \cdots \beta_{i\ell}\right]^T$$

Substituting eqn (75) in eqn (74) gives the global nonlinear model

$$z(k) = \psi^T(k)\left[\beta^T W(k)\right] \quad \ldots(76)$$

The advantage of this approach is that it is relatively easy to estimate the model parameters and the results can be readily interpretted using all the well known linear theory. The disadvantage is that the final model will only provide an adequate representation for the system for the particular trajectory of input signal used in the identification unless the process is only mildly nonlinear or the operation of the process moves slowly and smoothly from one operating point to another [29].

5.2.2.1 An Example

The implicit nonlinear model

$$z(k) = 0.5z(k-1)+0.3u(k-1)+0.3z(k-1)u(k-1)+0.5u^2(k-1)+\varepsilon(k) \quad \ldots(77)$$

was simulated over the global input range of ±1.0. Eleven first order linearized models with input range ±0.1 were estimated at the different operating levels where $\omega(k) = u(k)$. The final global nonlinear model eqn (76) was estimated as

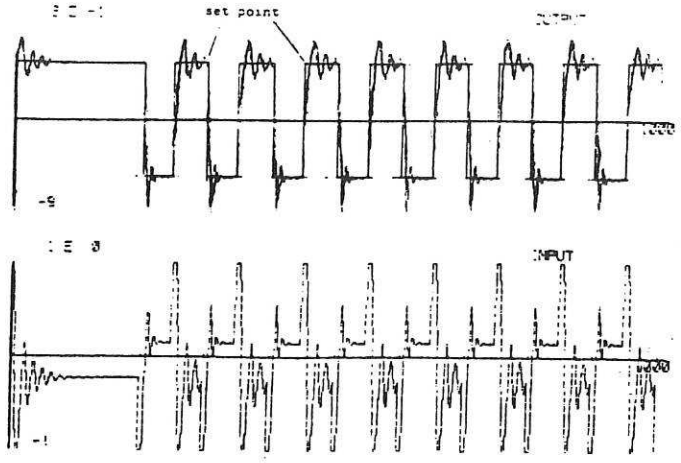
$$z(k) = \left[1, z(k-1), u(k-1)\right] \begin{pmatrix} -0.0019 & 0.0525 & -0.6104 & -0.6156 & -0.3919 \\ 0.4548 & 0.4334 & 0.0580 & -0.1953 & -0.0036 \\ 0.3034 & 1.1151 & 0.3770 & 0.3723 & 0.2348 \end{pmatrix} \times \begin{pmatrix} 1 \\ u(k-1) \\ u^2(k-1) \\ u^3(k-1) \\ u^4(k-1) \end{pmatrix} \quad \ldots(78)$$

The final model eqn (78) provided an adequate representation for the system eqn (77) when perturbed by the input used to initially excite the process but significant deficiencies in the model were visible for other rapidly varying inputs.

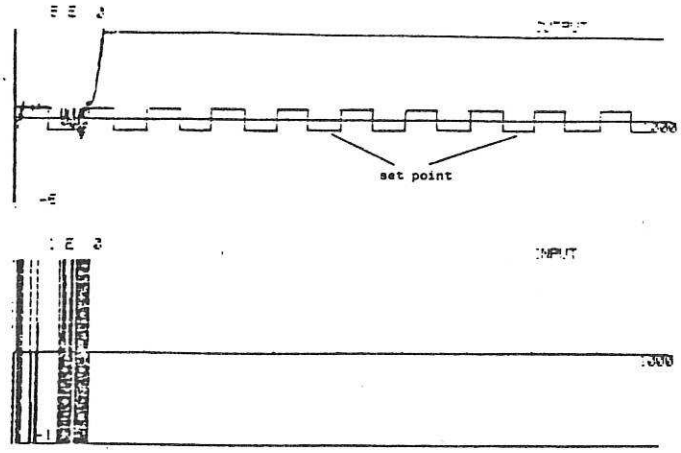
5.3 Control of Nonlinear Sampled-Data Systems

The choice of model representation for nonlinear systems is vitally important since this will influence its usefulness for both prediction and controller design. In view of the success of the linear difference equation model as a basis for linear controller design procedures it is natural to extend these ideas to the NARMAX model. There are many possibilities here that can be investigated [23,30] and only the simple one-step-ahead adaptive controller based on the NARMAX model will be considered here. This is best illustrated by designing controllers for the liquid level system discussed in section 5.2.1.1.

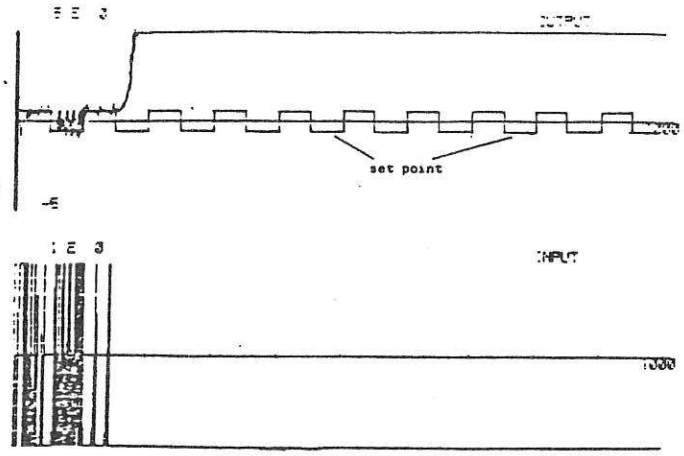
PI, linear adaptive one-step-ahead and self-tuning regulators were designed for the liquid level system described in section 5.2.1.1. The performance of these controllers is illustrated in Fig.12.



(a) PI Controlled System



(b) One-step ahead controlled system



(c) Minimum variance controlled system

Fig.12. Linear controllers applied to the nonlinear liquid level system

Inspection of Fig.12 clearly shows the poor performance of all these linear controllers when applied to the nonlinear liquid level system. The adaptive linear regulators only performed with any degree of satisfaction when the set point signal was slowed right down so that the parameter estimation routines had time to re-adapt to the new operating point.

The adaptive one-step-ahead nonlinear controller was defined, from eqn (69), by the feedback law

$$\begin{aligned} u(k) = \{&y^*-0.436z(k)-0.681z(k-1)+0.149z(k-2) \\ &-0.94u(k-1)+0.071u(k-2)+0.034z^2(k-1) \\ &+0.135z(k-1)u(k-1)+0.027z^3(k-1)+0.108z^2(k-1)u(k-1) \\ &+0.099u^3(k-1)\}/\{0.396-0.351z(k)\} \end{aligned} \qquad \ldots(79)$$

for $\left[0.396-0.351z(k)\right] \geq \gamma > 0$

where γ was set to 0.01 and y^* represents the set point. The performance of the controller eqn (79) when applied to the liquid level system is illustrated in Fig.13.

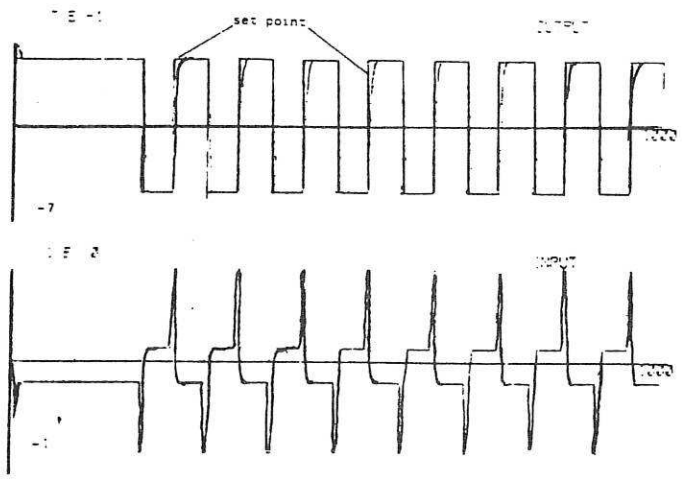


Fig.13. Nonlinear control of the liquid level system

A comparison of the performance of the linear based designs Fig.12 with the nonlinear design Fig.13 clearly shows the excellent response of the system when the nonlinear controller is utilised.

6. Conclusions

The statistical analysis of nonlinear systems is in general a difficult task. Whilst some of the techniques currently available have been briefly described above details of other alternative approaches are readily available in the literature.

Acknowledgements

The author gratefully acknowledges financial support for part of the work presented above from SERC grants GR/B/31163 and GR/D/30587.

References

1. Papoulis, A. 'Probability theory, random variables and stochastic processes', McGraw-Hill, N.Y., 1965.

2. Barrett, J.F., Coales, J.F. 'An introduction to the analysis of nonlinear control systems with random inputs', Proc.IEE, 103C, pp.190-199, 1955.

3. Nuttall, A.H. 'Theory and application of the separable class of random processes', MIT Res. Lab. Elect. Technical Rp 343, 1948.

4. Schetzen, M. 'The Volterra and Wiener theories of nonlinear systems', Wiley, N.Y., 1980.

5. Bedrosian, E., Rice, S.O. 'The output properties of Volterra systems driven by harmonic and Gaussian inputs', Proc.IEEE, 59, pp.1688-1707, 1971.

6. Bussgang, J.J., Ehrman, L., Graham, J.W. 'Analysis of nonlinear systems with multiple inputs', Proc.IEEE, 62, pp.1088-1119, 1974.

7. Rudko, M., Wiener, D.D. 'Volterra systems with random inputs; a formalized approach', IEEE Trans. on Commn., COM-26, pp.217-227, 1978.

8. Barrett, J.F. 'Formula for output autocorrelation and spectrum of a Volterra system with stationary Gaussian input', Proc.IEE, 127, pp.286-289, 1980.

9. Barrett, J.F. 'The use of functionals in the analysis of nonlinear physical systems', J. Elect. & Control, 15, pp.567-615, 1963.

10. George, D.A. 'Continuous nonlinear systems', MIT Res. Lab. Elect., Tech. Report 355, 1959.

11. Jazwinski, A.H. 'Stochastic processes and filtering theory', Academic Press, N.Y., 1970.

12. Doob, J.L. 'Stochastic processes', Wiley, N.Y., 1953.

13. Harris, C.J. 'Stochastic process models and control', SRC Vacation School on Stochastic Processes in Control Systems', Warwick University, April 1978.

14. Harris, C.J. 'Simulation of stochastic processes', ibid.

15. Harris, C.J. 'Modelling, simulation and control of stochastic systems with applications in wastewater treatment', Int. J. Systems Sci., 8, pp.393-411, 1977.

16. Pervozvanskii, A.A. 'Random processes in nonlinear control systems', Academic Press, N.Y., 1965.

17. Booton, R.C. 'The analysis of nonlinear control systems with random inputs', Proc. Symp. on N.L. Ct. Analysis, Brooklyn Poly, 1953.

18. Lawrence, P.J. 'A comparison of nonlinear system approximations', Tech. Note DAG 86, Dynamic Analysis Group, UWIST, 1976.

19. Somerville, M.J., Atherton, D.P. 'Multigain representation for single-valued nonlinearity with several inputs', Proc.IEE, 105C, pp.537-549, 1958.

20. Pupkov, K.A. 'Method of investigating the accuracy of essentially nonlinear automatic control systems by means of equivalent transfer functions', Automn & Remote Contr., 21, pp.126-140, 1960.

21. Atherton, D.P. 'Nonlinear control engineering', Van Nostrand Reinhold, 1975.

22. Billings, S.A. 'Identification of nonlinear systems - a survey', Proc.IEE, 127, pp.272-285, 1980.

23. Billings, S.A., Gray, J.O., Owens, D.H. (Eds). 'Nonlinear System Design', P. Peregrinus, 1984.

24. Billings, S.A., Fakhouri, S.Y. 'Identification of systems composed of linear dynamic and static nonlinear elements', Automatica, 18, pp.15-26, 1982.

25. Leontaritis, I.J., Billings, S.A. 'Input-output parametric models for nonlinear systems, Part I - Deterministic nonlinear systems, Part II - Stochastic nonlinear systems', Int. J. Control, 41, pp.303-344, 1985.

26. Billings, S.A., Voon, W.S.F. 'Least-squares parameter estimation algorithms for nonlinear systems', Int. J. Systems Sci., 15, pp.601-615, 1984.

27. Billings, S.A., Voon, W.S.F. 'Structure detection and model validity tests in the identification of nonlinear systems', Proc.IEE, Part D, 130, pp.193-199, 1983.

28. Billings, S.A., Fadzil, M.B. 'The practical identification of nonlinear systems', 7th IFAC Symp. Ident. Syst. Par. Est., York, 1985.

29. Billings, S.A., Voon, W.S.F. 'Piecewise linear identification of nonlinear systems', (in preparation).

30. Billings, S.A., Tsang, K.M. "Predictive controller design for nonlinear systems', (in preparation).