Monograph:

Billings, S.A. (1982) Structure Detection and Model Validity Tests in the Identification of Nonlinear Systems. Research Report. ACSE Report 196 . Department of Control Engineering, University of Sheffield, Mappin Street, Sheffield

STRUCTURE DETECTION AND MODEL VALIDITY TESTS

IN THE IDENTIFICATION OF NONLINEAR SYSTEMS

BY

S.A. Billings, B.Eng., Ph.D., C.Eng., M.I.E.E., A.F.I.M.A., M.Inst.M.C.
and W.S.F. Voon, B.Eng.

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

Research Report No. 196

October 1982

ABSTRACT

A structure detection test which distinguishes between linear and nonlinear dynamic effects in the system response, and model validity checks which indicate deficiencies in estimated nonlinear models are derived.

1. INTRODUCTION

Structure detection and model validity tests are a fundamental part of most system identification procedures. Whereas structure detection involves the determination of the form of model which will most appropriately fit the data model validity checks are designed to indicate the inadequacy of the fitted model. Most studies relating to these procedures assume that the system under investigation is linear. Structure detection then reduces to the problem of determining the model order and time delay of the system[1,2,3]. Estimates of the parameters in the model can then be determined and a linear covariance analysis of the residuals can be applied to test the adequacy of the fitted model. This procedure is iterative and a number of models may be fitted and analysed before the final model is selected. All these tests however, depend critically on the assumption of linearity. If the process can be taken off line this assumption can easily be verified by performing a series of step tests over the amplitude range of operation. When the data is prerecorded and additional experimentation is precluded or when the analysis relates to the system residuals, tests for linearity become much more involved.

Very few authors have investigated this problem. West[4] considered nonlinear signal distortion correlation by studying static nonlinear characteristics. By splitting the output from the nonlinear element into two portions one proportional to the input signal and the other a distortion noise West showed that there is no correlation between the input and distortion signals whenever the input belongs to the separable class of random process. Douce[5] proved that the same property occurs for a specific class of nonlinear dynamic systems. The nonlinear distortion can however, be detected by cross-correlating the residual with a test signal obtained by passing the system input through a specified nonlinearity and Douce developed an identification procedure based on this result. Subba Rao and

Gabr[6] investigated the use of bispectral density functions for testing for linearity and Rabjman[7] introduced dispersion functions to measure the degree of nonlinearity of systems.

In the present study higher order correlation functions are introduced as simple to compute measures of nonlinearity. The techniques are shown to avoid the complicated computations involved in determining bispectral densities or dispersion functions and apply to a wider class of inputs than separable processes. The inadequacy of linear covariance techniques for structure detection and model validity tests is demonstrated in the next section which includes a problem statement together with some illustrative examples. Linear and nonlinear detection is investigated in section 4 and a simple correlation test is derived. These results are extended to include tests on model adequacy in Section 5. Confidence intervals are derived in Section 6 and simulated examples are included to illustrate the algorithms.

2. PROBLEM STATEMENT

The problems of structure detection and model validity testing although similar in some respects are quite different. Structure detection in the present context will be defined as a method of detecting nonlinearity and of distinguishing this from linear effects and additive noise distortion. Model validity testing however involves detecting terms in the residuals which if ignored will cause bias in the parameter estimates. There is no need in this latter case to distinguish between linear, nonlinear or correlated noise effects since any one of these can induce bias into the estimates. Unfortunately the traditional linear covariance tests which are now a fundamental part of linear system identification can easily be shown to be inadequate for both of the above problems[8,9]. This is best illustrated by a simple example:-

Assume that in estimating the parameters of a system various terms in the model were inadvertantly omitted and these appear in the residuals $\xi(k)$ as

$$\xi(k) = cu(k-1)e(k-1) + e(k) \tag{1}$$

where e(k) is white Gaussian noise and e(k) and u(k) are independent zero mean. It is easily shown that computing the normalised autocorrelation function of the residuals and the normalised cross-correlation function between the system input u(k) and the residuals yield

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

$$\phi_{u\xi}(\tau) = 0 \ \forall \ \tau \tag{2}$$

According to the linear analysis therefore the residuals contain no further information and appear white. Inspection of equation (1) clearly shows this is false and $\xi(k)$ will undoubtedly introduce severe bias into the parameter estimates[8,9]. This is a very disturbing situation which clearly demonstrates that linear covariance techniques do not in general detect predictable nonlinear effects[10,11].

Returning to the structure detection problem. In any identifaction procedure the first stage of the analysis ought to involve some simple calculations on the input/output data which indicate if the relationship between input and output is linear or nonlinear[12]. In other words will it be worthwhile trying to fit a nonlinear model?

Subba Rao and Gabr[6] suggested a solution to this problem using higher order spectra and defining

$$X_{ij} = \frac{|S(\omega_i, \omega_j)|^2}{S_\omega(\omega_i)S_\omega(\omega_j)S_\omega(\omega i/\omega_j)} \tag{3}$$

where $S_\omega(\omega_i, \omega_j)$ is the bispectral density and $S_\omega(\omega_i)$ the power spectral density of the time series. They proved that $S_\omega(\omega_i, \omega_j) = 0 \ \forall \ \omega_i, \omega_j$ implies that the process is linear or the third order moment of the driving input is zero $\mu_3 = 0$, whereas if X_{ij} is a constant the process is linear and $\mu_3 \neq 0$. There are two disadvantages to this approach. Firstly, the requirement to estimate spectral densities introduces problems of windowing and averaging. Secondly the method can fail when applied to system identification problems. This latter problem arises because in time series analysis there is no measured input, the output is assumed to be generated by passing a ficticious input with defined properties through a model. In system identification there are three signals: input, output and noise, and generally the input and noise are assumed to be independent. It is this latter property that defeats Subba- Rao's test when applied to the identification problem. The sequence

$$z(k) = \alpha_1 u^3(k-1) + e(k)$$

for example, where the input $u(k)$ and noise $e(k)$ are independent zero mean processes with symmetric distributions, remains undetected by the test. It is important to emphasise that these comments are not criticisms of Subba Rao's test which was derived for, and works well for time series. The comments do indicate the differences between the two problems of identification and time series analysis and the need to develop new tests tailored to the system identification problem[8,9,12].

An alternative method of testing for linearity was developed by Rabjman[7] using dispersion functions. Defining the cross-dispersion function

$$\Theta_{zu}(t_1,t_2) = E_{u(t_2)}\left[\{E_{z(t_1)}\left[z(t_1)\middle|u(t_2)\right] - E_{z(t_1)}\left[z(t_1)\right]\}^2\right] \tag{4}$$

which can be computed from

$$\Theta_{zu}(t_1,t_2) = \int_{-\infty}^{\infty}\{\int_{-\infty}^{\infty} z(t_1)p(z(t_1)|u(t_2))dz(t_1) - \int_{-\infty}^{\infty} z(t_1)p(z(t_1))d\,z(t_1)\}^2 p(u(t_2))d\,u(t_2) \tag{5}$$

if the necessary densities are known. The definition of the auto-dispersion function $\Theta_{uu}(t_1 t_2)$ follows in an obvious manner. Based on these definitions Rabjman introduced the degree of nonlinearity $\nu_{zu}(t_1,t_2)$ as the least mean square deflection of the regression curve $E[z(t_1)|\ u(t_2)]$ from a certain straight line

$$\nu_{zu}(t_1,t_2) = \sqrt{\left[\min_{a,b} \ E_{u(t_2)}\left[\{E_{z(t_1)}[z(t_1)|u(t_2)] - (a + b\ u(t_2))\}^2\right]\right] / \sigma_z^2(t_1)} \qquad (6)$$

which reduces to

$$\nu^2_{zu}(t_1,t_2) = \eta^2_{zu}(t_1,t_2) - \phi^2_{zu}(t_1,t_2)$$

where $$\eta_{zu}(t_1,t_2) = \sqrt{\frac{\Theta_{zu}(t_1,t_2)}{\sigma_z^2(t_1)}}$$

is the normalised cross-dispersion function and $\phi_{zu}(t_1,t_2)$ is the normalised cross correlation function. This approach is even more complicated than Subba- Rao's method since in all practical cases it would be necessary to estimate the conditional density functions in order to compute the dispersion functions.

The present study is an attempt to develop simple measures of nonlinearity using correlation methods[10,12]. It will be assumed that the system under investigation is analytic and can be represented by a Volterra series[12]

$$z(t) = \sum_{k=1}^{\infty} \int_{-\infty}^{\infty} \ldots \int h_n(\tau_1,\tau_2,\ldots\tau_n) \prod_{i=1}^{n} u(t-\tau_i)d\tau_i + e(t) \qquad (7)$$

Using the operator calculus developed by Brilliant and George[12] equation (7) can be expressed as

$$z(t) = \sum_{n=1}^{\infty} H_n[u(t)] + e(t) = \underline{H}[u(t)] + e(t)$$

$$= \sum_{n=1}^{\infty} H_n(u^n(t)) + e(t) \qquad (8)$$

where the square brackets indicate that H operates on u(t) and the parentheses depicts the actual relationship. It is important to emphasise that the continuous time Volterra series model is chosen as a convenient representation for a wide class of nonlinear systems. The fact that all the results are derived for this model does not constrain the applicability of the results to Volterra models only. The final results can be applied to all analytic nonlinear systems whatever form of model is used to characterise the input/output map.

It will be assumed throughout that all random signals are ergodic, so that ensenble averages may be replaced by time averages over one sample function.

3. INPUT SENSITIVITY

To illustrate this phenomena consider a second order Volterra model with an input $u(t) + b$, $\overline{u(t)} = 0$

$$z(t) = H_1 [u(t)+b] + H_2 [u(t)+b] + e(t)$$

$$= H_1(u(t)+b) + H_2(u^2(t) + 2bu(t) + b^2) + e(t)$$

Removing the output mean yields

$$z'(t) = H_1(u(t)) + H_2(u^2(t) + 2bu(t) - \overline{u^2(t)}) + e'(t) \tag{9}$$

where the superscript ' is used throughout to indicate a zero mean process. Inspection of equation (9) shows that any model relating $z'(t)$ and $u(t)$ will be input sensitive. That is it will be dependent on $\sigma_u^2 = \overline{u^2(t)}$ and b and will only yield the correct predicted output for inputs with exactly these statistics. The degree of input sensitivity depends on the model used to represent the system and the identification algorithm. This problem does not arise in linear systems analysis where mean levels of input and output are almost always removed to improve the signal resolution.

Input variance sensitivity can be avoided by measuring the average output $z_b(t)$ of the system with zero input $u(t) = 0$

$\overline{z_b(t)} = H_1[b] + H_2[b] + \ldots + H_n[b] + e(t)$ to yield the input/output description

$$z_b'(t) = z(t) - \overline{z_b(t)} = H_1(u(t)) + H_2(u_2(t) + 2bu(t)) + \ldots + e'(t) \qquad (10)$$

The dependence of the fitted model on the variance of the input has therefore been removed but the model will only be valid for inputs around the operating point b.

Inspection of equations (9) and (10) shows that $\overline{z_b(t)} = \overline{z(t)}$ iff the system is linear and this can be considered as a very simple test for nonlinearity.

4. LINEAR AND NONLINEAR DETECTION

The problem of linear and nonlinear detection can be simply defined as:- is it possible to determine if it is worth trying to fit a nonlinear model to the data?

The solution to this involves the measurement of higher order correlation functions of $z(t)$[10,12].

Assume that the input $u(t)$ and noise $e(t)$ are independent zero mean processes with symmetric probability density functions such that all odd order moments are zero. All even order movements of $u(t)$ are assumed to exist .

Consider the computation of $\phi_{z'z'^2}(\tau)$ where $z'(t)$ is the system response (with mean level removed) to an input $u(t) + b$.

By definition

$$\phi_{z'z'^2}(\tau) = E[z'(t+\tau)(z'(t))^2] \qquad (11)$$

$$z'(t+\tau) = \int h_1(\tau_1)(u(t - \tau_1 + \tau)+b)d\tau_1$$

$$+ \iint h_2(\tau_1,\tau_2)(u(t - \tau_1 + \tau)+b)(u(t-\tau_2+\tau)+b)\, d\tau_1\, d\tau_2$$

$$+ \ldots \quad \ldots + e(t+\tau) - \overline{z}(t+\tau) \qquad (12)$$

interchanging variables in (12)

$$z'(t+\tau) = \int h_1(t-\tau_1+\tau)(u(\tau_1))d\tau_1$$

$$+ \iint h_2(t-\tau_1+\tau,\ t-\tau_2+\tau)\ (u(\tau_1)u(\tau_2)$$

$$+ bu(\tau_1) + bu(\tau_2)\ d\tau_1\ d\tau_2$$

$$- \iint h_2(t-\tau_1 + \tau, t-\tau_2+\tau)\overline{u(\tau_1)\ u(\tau_2)}\,d\tau_1\ d\tau_2$$

$$+ \ldots\ldots + e'(t + \tau) \tag{13}$$

which applying the notation of equation (8) with obvious extensions can be expressed as

$$z'(t+\tau) = H_1^\tau(u(t)) + H_2^\tau(u^2(t)) + 2b\ H_2^\tau(u(t)) - H_2^\tau\ \overline{(u(t)^2)t} +$$

$$+ \ldots\ldots + e'(t + \tau) \tag{14}$$

Consequently, equation 11 becomes

$$\phi_{z'z'^2}(\tau) = E\{[H_1(u) + H_2(u^2 + 2bu - \overline{u^2})$$

$$+ H_3(u^3 + 3bu^2 + 3b^2u - \overline{3bu^2}) + \ldots\ e'(t)]^2$$

$$\cdot\ [H_1^\tau(u) + H_2^\tau(u^2 + 2bu - \overline{u^2}) + H_3^\tau(u^3 + 3bu^2$$

$$+ 3b^2u - \overline{3bu^2}) + \ldots\ldots + e'(t+\tau)]\} \tag{15}$$

$$= E\{[(H_1H_1)(u^2) + 2(H_1H_2)(u^3 + 2bu^2 - u\ \overline{u^2})$$

$$+ (2H_1H_3)(u^4 + 3bu^3 + 3b^2u^2 - \overline{3bu^2}\ \ u) + \ldots + e'^2(t)]$$

$$[H_1^\tau(u) + H_2^\tau(u^2 + 2bu - \overline{u^2}) + H_3^\tau(u^3 + 3bu^2 + 3b^2u - \overline{3bu^2}) + \ldots + e'(t+\tau)]\}$$

(16)

The evaluation of equation (16) is detailed in Appendix I where it is shown that

$$\phi_{z'z'^2}(\tau) = 0 \;\forall\; \tau \tag{17}$$

iff the process is linear (i.e. $H_2, H_3 \ldots H_n = 0$). Whenever $\phi_{z'z'^2}(\tau) \neq 0$ therefore this indicates that the system under test is nonlinear.

Note that the test distinguishes between linear additive noise corruption of the measurements and distortion due to nonlinear efforts.

The test is dependent on the assumption that the third order moment of the input is zero. This can normally be assured in system identification by selecting an input which has a symmetric probability density function such as a sine wave, gaussian signal, uniformly distributed process, ternary pseudorandom sequence etc.etc. The assumption can be readily verified by computing $E\left[(u(t))^3\right]$.

The effect of adding a mean level or d.c. shift b to the input ensures that all terms which reflect the nonlinearity of the system contribute to $\phi_{z'z'^2}(\tau)$. If b were set to zero for example the third term in the expansion (Appendix I) would be zero and $\phi_{z'z'^2}(\tau)$ would not detect odd order nonlinearities.

If $z'_b(t)$ can be measured an analogous result to equation (17) can be derived for $\phi_{z'_b z'^2_b}(\tau)$.

Notice that $\phi_{z'z'^2}(\tau)$ cannot detect the residual sequences of equation (1) whenever $u(k)$ and $e(k)$ are independent and $p(e)$ is symmetric. This does not undermine the test since in structure detection we assume that the noise terms can only be nonlinear if the process itself is nonlinear. In other words if terms such as $u(k-1)e(k-1) + e(n)$ exist in the system output there must also be nonlinear terms in $u(\cdot)$ and $y(\cdot)$ which will be detected by $\phi_{z'z'^2}(\tau)$. The test does therefore yield the correct information.

The situation is however, quite different when considering model validation. If a model has been fitted to the input and output the linear and nonlinear terms in the input and output may have all been included in the model leaving only the sequence $\xi(k)$, equation 1, as residual. Tests which detect this type of distortion must therefore be developed for model validation and these are considered in the next section.

5. MODEL VALIDITY TESTS

Model validity tests are usually defined to detect information in the residuals which if neglected will introduce bias in the parameter estimates. It is not necessary in this context to distinguish between correlated noise, linear or nonlinear dynamics as in the previous section, since if any one of these is present in the residuals biased estimates will result.

If the system under test is linear the residuals should be unpredictable from all past inputs and outputs. When the system is nonlinear the residuals should be unpredictable from all linear and nonlinear combinations of past inputs and outputs. This latter requirement could well involve an enormous amount of data processing and the objective in the present study is to reduce this to just a few easy to implement tests which convey all the required information.

The assumptions stated in section 4 will be carried over to the present analysis. During model validation testing we have very little control over the signals, the input and the residuals, these are prespecified either by the experiment or the model. Under such circumstances we must derive tests which work under the worst possible combinations of signal properties and will assume therefore that $u(\cdot)$ and $e(\cdot)$ are independent zero mean processes, all odd order moments are zero, $e(\cdot)$ is white and $u(\cdot)$ maybe white.

Consider the situation when the system under test is known to be linear[13]. This information will be available from the structure detection test $\phi_{z'z'^2}(\tau)$ described in the previous section. Suppose that the true model of the system is known to be

$$z(t) = G(u(t)) + L(e(t)) \tag{18}$$

where G and L are linear operators. A model is fitted to the input/output data to yield an estimate

$$z(t) = \hat{G}(u(t)) + \hat{L}(\hat{e}(t)) \tag{19}$$

where the residuals are defined by

$$\hat{e}(t) = \hat{L}^{-1}\{(G-\hat{G})(u(t)) + L(e(t))\} \tag{20}$$

If the process model is correct $\hat{G} = G$ but the noise model is incorrect $\hat{L} \neq L$

$$\hat{e}(t) = \hat{L}^{-1}\{L(e(t))\} \tag{21}$$

and although the residuals $\hat{e}(t)$ will be autocorrelated

$$\phi_{\hat{e}'\hat{e}'}(\tau) \neq \delta(\tau)$$

they will not be correlated with the input

$$\phi_{u\hat{e}'}(\tau) = 0 \ \forall \ \tau$$

Alternatively, if the noise model is correct $\hat{L} = L$ and the process model is biased $\hat{G} \neq G$

$$\hat{e}(t) = \hat{L}^{-1}\{(G-\hat{G})(u(t))\} + e(t) \tag{22}$$

and the residuals are both autocorrelated $\phi_{\hat{e}'\hat{e}'}(\tau) \neq \delta(\tau)$ and correlated with the input $\phi_{u\hat{e}}(\tau) \neq 0$. If both process and noise model estimates are correct then $\phi_{\hat{e}'\hat{e}'}(\tau) = \alpha\delta(\tau)$, $\phi_{u\hat{e}'}(\tau) = 0 \ \forall \ \tau$.

Consequently, if the system is linear $\phi_{z'z'^2}(\tau) = 0 \ \forall \ \tau$ and if $\phi_{u\hat{e}}(\tau) \neq 0$ this indicates that the estimate of the process model is deficient. Once the process model has been improved to yield $\phi_{u\hat{e}}(\tau) = 0$ any correlation in the residuals $\phi_{\hat{e}'\hat{e}'}(\tau) \neq \delta(\tau)$ indicates that the noise model is incorrect. It is possible using these simple correlations therefore to distinguish between deficiencies in the process and noise models.

for example, which is very common in parameter estimation for nonlinear systems and occurs whenever noise enters internally or for certain parameterisations when the noise is purely additive at the output,[8,9,11] is extremely difficult to detect.

It appears to be impractical therefore to develop a simple procedure which distinguishes between discrepancies in the process and noise models. In fact three tests are required to detect all the terms in equation (24) as detailed below.

The autocorrelation of the residuals $\phi_{\xi'\xi'}(\tau)$ is the primary test. An expression for $\phi_{\xi'\xi'}(\tau)$ is derived in Appendix 2 where it is shown that the test fails under two conditions:-

(i) The test incorrectly indicates that the residuals are unpredictable from past inputs and outputs if either

$$\xi(k) = u(k-n) + e(k) \text{ or} \tag{25}$$

$$\xi(k) = u(k-m)e^2(k-n) + e(k) \quad \forall\, n,m \tag{26}$$

whenever $u(\cdot)$ and $e(\cdot)$ are independent white noise sequences.

(ii) Similarly the test incorrectly fails to detect all terms of the form

$$\xi(k) = u^q(k-m)e(k-n) + e(k) \tag{27}$$

$\forall$ n,m; odd q and arbitrary input u(k).

These deficiencies in $\phi_{\xi\xi'}(\tau)$ can be corrected by computing two additional correlations. The terms in (i) are detected by computing $\phi_{u\xi'}(\tau)$ which for the residuals in equations (25),(26) is easily shown to yield

$$\phi_{u\xi'}(\tau) = \phi_{uu}(\tau+n) \neq 0$$

for equation (25), and

$$\phi_{u\xi'}(\tau) = \phi_{ee}(0)\phi_{uu}(\tau+m) \neq 0$$

for equation (26). One test which correctly detects the cross product term in equation (27) is

$$\phi_{\xi'\xi'u}(\tau) = E[\xi'(k)\xi'(k-1-\tau)u(k-1-\tau)] \quad \forall \ \tau \geq 0 \tag{28}$$

Notice that $\phi_{\xi'\xi'u}(\tau)$ only detects the cross-product terms in equation (27) and should ideally be zero.

There may well be higher order terms with the same properties as equation (26) and (27). For example the residuals in equation (27) would have the same property if $u^q(t-m)$ were replaced by $u(t-m_1).u(t-m_2)\ldots\ldots u(t-m_q)$, for all q odd etc.

To summarise in the general case when it is known that the system under test is nonlinear ($\phi_{z'z'^2}(\tau) \neq 0$ with input $u(t) + b$ and $p(u)$ symmetric) the residuals will be unpredictable from all past inputs and outputs iff

$$\phi_{\xi'\xi'}(\tau) = \delta(\tau) \tag{29}$$

$$\phi_{u\xi'}(\tau) = 0 \ \forall \ \tau \tag{30}$$

$$\phi_{\xi'\xi'u}(\tau) = 0 \ \forall \ \tau \tag{31}$$

When the system is nonlinear it is very difficult to distinguish between bias in the process model or bias in the noise model as in the linear case equation (18). The cross-correlation $\phi_{u\xi'}(\tau)$ for example detects only odd terms in $G^u[u(t)]$ (all even terms make no contribution) and can have a value for certain cross-product terms in the noise model $G^{ue}[u(t),e(t)]$. If, therefore any of the conditions equation (29) through (31) are violated all that can be said is that the model is deficient in some way.

6. COMPUTATION ASPECTS

All the tests derived above are based on correlation functions which for sampled input and output signals are calculated according to the formulae

$$\hat{\phi}_{xy}(k) = \frac{\frac{1}{N}\sum_{t=1}^{N-k}(x(t) - \bar{x})(y(t+k) - \bar{y})}{\sqrt{\phi_{xx}(o)\ \phi_{yy}(o)}} \tag{32}$$

$$-1 \leq \hat{\phi}_{xy}(k) \leq 1$$

In reality confidence intervals plotted on the graphs indicate if the correlation between variables is significant or not. If N is large the standard deviation of the correlation estimate is $1/\sqrt{N_1}$ the 95% confidence limits are therefore approximately $\pm 1.96/\sqrt{N}$

7. SIMULATION RESULTS

The algorithms described above have been tested by simulating various linear and nonlinear systems. A sine wave input has been used throughout for the structure detection tests. This satisfies the conditions of symmetric density functions, is much easier to generate than white noise etc., and yields good results. Alternative inputs can be chosen providing they satisfy the conditions specified in section 4. A first order linear system with pulse transfer function

$$y(k) = \frac{0.4z^{-1}}{1-0.8z^{-1}} u(k)$$

$$z(k) = y(k) + \eta(k) \tag{33}$$

was simulated where $u(k) = u'(k) + b$, $b = 0.2$, and $\eta(k)$ was a Gaussian white sequence N(0,0.1). The mean levels $\bar{z} = 0.40$, $\bar{z}_b = 0.41$ and the structure detection test $\phi_{z'z'^2}(\tau)$ illustrated in Fig. 1 clearly indicate that the system is linear. The estimated model

$$z(k) = 0.7871\, z(k-1) + 0.4123\, u(k-1) \tag{34}$$

was fitted using a recursive least squares algorithm and inspection of $\phi_{u\xi'}(\tau)$ and $\phi_{\xi'\xi'}(\tau)$ illustrated in Fig. 2 where $\xi(k)$ are the residuals indicate that the estimates are biased. Since $\phi_{z'z'^2}(\tau) = 0 \;\forall\; \tau$ this bias must be induced by additive linear noise. Applying an extended least squares algorithm yields the estimated model

$$z(k) = 0.8020z(k-1) + 0.4008\, u(k-1) - 0.8319\, \eta(k-1) + \eta(k) \tag{35}$$

and as expected the correlation functions $\phi_{u\xi'}(\tau)$ and $\phi_{\xi'\xi'}(\tau)$ were reduced to be within the 95% confidence intervals.

An implicit nonlinear system defined by a first order NARMAX model[8,9]

$$y(k) = 0.5y(k-1) + 0.3y(k-1)u(k-1) + 0.2u(k-1) + 0.6u^2(k-1) + 0.05y^2(k-1) \quad (36)$$

$$z(k) = y(k) + \eta(k) \quad (37)$$

was simulated with b = 0.2, and $\eta(k)$ a discrete white noise sequence distributed as N(0,0.1). The class of systems which can be represented by nonlinear difference equations or NARMAX models has been studied by considering the observability of nonlinear systems[8,9]. Substituting (37) into (36) yields

$$\begin{aligned} z(k) &= 0.5z(k-1) + 0.2u(k-1) + 0.3z(k-1)u(k-1) + 0.6u^2(k-1) \\ &\quad + 0.05z^2(k-1) \\ &\quad + \{\eta(k) - 0.5\eta(k-1) - 0.3\eta(k-1)u(k-1) - 0.1z(k-1)\eta(k-1) \\ &\qquad + 0.05\eta^2(k-1)\} \end{aligned} \quad (38)$$

which clearly shows that the noise enters the model multiplicatively. Standard parameter estimation algorithms derived for linear systems will therefore yield biased estimates and modified procedures have to be derived[8,9].

The mean levels for equation (38), $\bar{z}_b = 0.147$ and $\bar{z} = 1.235$ indicate that the process is nonlinear and this is confirmed by inspection of $\phi_{z'z'^2}(\tau)$ and $\phi_{z_b'z_b'^2}(\tau)$ illustrated in Fig. 3. It has been found that $\phi_{z_b'z_b'^2}(\tau)$ usually provides a much clearer indication of nonlinear effects if they exist compared to $\phi_{z'z'^2}(\tau)$ and we would therefore recommend that $\phi_{z_b'z_b'^2}(\tau)$ is implemented whenever z_b' is available.

To demonstrate the effectiveness of the model validity tests derived in section 5 assume that the term 0.3 $u(k-1)\eta(k-1)$ in equation (38) is not included in the estimated model. Parameter estimation using a modified extended least squares algorithm[8,9] then yields

$$z(k) = 0.5266z(k-1) + 0.1961\ u(k-1) + 0.3116z(k-1)\ u(k-1)$$
$$+ 0.5702\ u^2(k-1) + 0.04032\ z^2(k-1)$$
$$+ \{\xi(k) - 0.593\ \xi(k-1) - 0.07403\ z(k-1)\ \xi(k-1)$$
$$+ 0.3407\ \xi^2(k-1)\} \qquad (39)$$

As expected $\phi_{\xi'\xi'u}(\tau)$ illustrated in Fig. 4 indicates that the model (39) is biased because a cross-product term has been omitted. Notice that the traditional linear covariance tests $\phi_{\xi'\xi'}(\tau)$ and $\phi_{u\xi'}(\tau)$ both fail to detect that a term is missing from the model.

Similarly if the term 0.3 $z(k-1)\ u(k-1)$ in equation (38) is omitted from the model, modified extended least squares yields

$$z(k) = 0.3269\ z(k-1) + 0.3628\ u(k-1) + 0.9874\ u^2(k-1)$$
$$+ 0.1318z^2(k-1)$$
$$+ \{\xi(k) - 0.1179\xi(k-1) + 0.2985\ u(k-1)\ \xi(k-1)$$
$$- 0.08305\ z(k-1)\xi(k-1) - 0.2793\xi^2(k-1)\} \qquad (40)$$

which is biased as indicated by the model validity tests illustrated in Fig. 5 . Including all the terms in the model yields the final model

$$z(k) = 0.5377\ z(k-1) + 0.192\ u(k-1) + 0.3258z(k-1)\ u(k-1)$$
$$+ 0.5519z^2(k-1) + 0.03541z^2(k-1)$$
$$+ \{\xi(k) - 0.6145\xi(k-1) - 0.4403\ u(k-1)\ \xi(k-1)$$
$$+ 0.02091\ z(k-1)\ \xi(k-1) + 0.2506\ \xi^2(k-1)\} \qquad (41)$$

8. CONCLUSIONS

Structure detection and model validation methods have been investigated for a broad class of nonlinear systems. It has been demonstrated that if it is possible to inject a non zero mean input which has a symmetric probability density function the correlation test $\phi_{z'z'^2}(\tau)$ indicates prior to parameter estimation if it is worth fitting a nonlinear model to the data.

The use of traditional linear covariance techniques for model validation have been shown to be inappropriate when the system under test is nonlinear and additional methods which detect all terms in the residuals which are predictable from all linear and nonlinear combinations of past inputs and outputs have been developed.

Although there may be alternative structure detection and model validity tests which could be derived hopefully the techniques developed in the present paper are amongst the simplest to compute and interpret.

Acknowledgements

The authors gratefully acknowledge that this work is supported by SERC under grant GR/B 31163.

Appendix 1

Evaluation of $\phi_{z'z'^2}(\tau)$

Consider the evaluation of each term in the expression for $\phi_{z'z'^2}(\tau)$ equation (16) assuming that the probability density function of the input signal u(t) is symmetric such that all odd moments are zero and all even moments exist:-

$$\text{(i)}\quad E\left[(H_1H_1)(u^2)H_1^\tau(u)\right] = E\left[(H_1H_1H_1^\tau)(u^3)\right] = 0 \tag{42}$$

$$\text{(ii)}\quad E\left[(H_1H_1)(u^2)H_2^\tau(u^2 + 2bu - \overline{u^2})\right]$$

$$= E\left[(H_1H_1H_2^\tau)(u^4 + 2bu^3 - \overline{u^2}\,u^2)\right] \neq 0 \tag{43}$$

$$\text{(iii)}\quad E\left[(H_1H_1)(u^2)H_3^\tau(u^3 + 3bu^2 + 3b^2u - 3b\overline{u^2})\right]$$

$$= E\left[(H_1H_1H_3^\tau)(u^5 + 3bu^4 + 3b^2u^3 - 3b\overline{u^2}\,u^2)\right] \neq 0 \tag{44}$$

$$\text{(iv)}\quad E\left[(2H_1H_2)(u^3 + 2bu^2 - u\,\overline{u^2})H_1^\tau(u)\right]$$

$$= E\left[(2H_1H_2H_1^\tau)(u^4 + 2bu^3 - \overline{u^2}\,u^2)\right] \neq 0 \tag{45}$$

Similarly it can be shown that all the remaining terms are non zero and contribute to the final expression $\phi_{z'z'^2}(\tau)$ except the terms involving the noise process $e(t)$ which all tend to zero

$$E\left[(H_1H_1)(u^2)e'(t+\tau)\right] = 0$$

$$E\left[(2H_1H_2)(u^3 + 2bu^2 - \overline{u^2}u)\,e'(t+\tau)\right] = 0 \tag{46}$$

$$\vdots$$

$$E\left[e'^2)\,e'(t+\tau)\right] = 0$$

Appendix 2

Derivation of $\phi_{\xi'\xi'}(\tau)$

An expression for $\phi_{\xi'\xi'}(\tau)$, the autocorrelation of the residuals defined in equation (24).

$$\xi(t) = G^{u}[u(t)] + G^{ue}[u(t),e(t)] + G^{e}[e(t)] \tag{24}$$

is computed below assuming all mean levels are zero and utilizing the assumptions defined in section 5.

By definition

$$\phi_{\xi'\xi'}(\tau) = E\{(G^{u}[u(t)] + G^{ue}[u(t),e(t)] + G^{e}[e(t)])$$

$$(G^{u}[u(t+\tau)] + G^{ue}[u(t+\tau),e(t+\tau)]$$
$$+ G^{e}[e(t+\tau)])\} \tag{47}$$

$$= \psi_{uu}(\tau) + \psi_{u(ue)}(\tau) + \psi_{ue}(\tau)$$

$$+ \psi_{(ue)u}(\tau) + \psi_{(ue)(ue)}(\tau) + \psi_{(ue)e}(\tau)$$

$$+ \psi_{eu}(\tau) + \psi_{e(ue)} + \psi_{ee}(\tau) \tag{48}$$

where the ψ's are polynomial correlation functions with $\psi_{uu}(\tau)$ defined as

$$\psi_{uu}(\tau) = E\,\{G^{u}[u(t)]G^{u}[u(t+\tau)]\}$$

and the other terms defined in an analogous manner. Since it is important to consider the worst case and detect every isolated term in equation (24) which could possibly exist all terms in equation (48) which are not polynomial autocorrelation functions will be neglected in the analysis given below. Equation (48) therefore reduces to

$$\begin{aligned}\phi_{\xi'\xi'}(\tau) = {} & \{\psi_{uu}(\tau) + E[(G^{u}[u(t)]e(t+\tau) + G^{u}[u(t+\tau)]e(t) \\ & + e(t)e(t+\tau)]\} \\ & + \{\psi_{(ue)(ue)} + E[(G^{ue}[u(t),e(t)]e(t+\tau) \\ & + G^{ue}[u(t+\tau),e(t+\tau)]e(t) + e(t)e(t+\tau)]\} + \psi_{ee}(\tau)\end{aligned} \tag{49}$$

$$= \{\Psi_{uu}(\tau) + \phi_{ee}(\tau)\}$$

$$+ \{\psi_{(ue)(ue)} + E\left[(G^{ue}\left[u(t),e(t)\right] e(t+\tau)\right.$$

$$\left. + G^{ue}\left[u(t+\tau),e(t+\tau)\right] e(t)\right] + \phi_{ee}(\tau)\} + \psi_{ee}(\tau) \tag{50}$$

since u(t) and e(t) are independent. Notice that the residuals by definition always include an additive prediction error term e(k) and this induces the additional terms in (49) and (50).

Consider each Ψ term and its associated prediction error component in equation (50) in turn. It follows that

$$\Psi_{uu}(\tau) + \phi_{ee}(\tau)$$

$$= E\{G^{u}\left[u(t)\right] G^{u}\left[u(t+\tau)\right]\} + \phi_{ee}(\tau)$$

$$= E\{G_1^u G_1^{u\tau}(u^2) + G_1^u G_3^{u\tau}(u^4) + \ldots\ldots + G_2^u G_2^{u\tau}(u^4) + \ldots\} + \phi_{ee}(\tau) \tag{51}$$

and assuming that all even moments of u exist all terms G_1^u, G_2^u, G_3^u contribute to $\psi_{uu}(\tau)$ and will therefore be detected except for the case when $G_2, G_3, G_4 \ldots G_n = 0$ and u(.) is white. In this latter case

$$\phi_{\xi'\xi'}(\tau) = \alpha_1 \phi_{uu}(\tau) + \alpha_2\phi_{ee}(\tau) = \beta\delta(\tau) \tag{52}$$

and the test fails. As noted in section 5 this failure is detected by $\phi_{u\xi'}(\tau)$.

Consider the next term in equation (50)

$$\Psi_{(ue)(ue)} + E\left[(G^{ue}\left[u(t),e(t)\right]e(t+\tau)\right.$$

$$\left. + G^{ue}\left[u(t+\tau), e(t+\tau)\right] e(t)\right] + \phi_{ee}(\tau)$$

Expanding this out exactly as in equation (51) shows that $\phi_{\xi'\xi'}(\tau)$ detects all terms except

$$\xi(t) = u(t-m)e^2(t-n) + e(t) \tag{53}$$

whenever u(t) is white and

$$\xi(t) = u^q(t-m)e(t-n) + e(t) \tag{54}$$

$\forall$ n,m odd q and arbitrary input u(.). The autocorrelation of the residuals $\phi_{\xi'\xi'}(\tau)$ with $\xi(t)$ defined as in equations (53) and (54) incorrectly indicates that the residuals are unpredictable from past inputs and outputs. This problem is easily rectified using the tests $\phi_{u\xi'}(\tau)$ and $\phi_{\xi'\xi'u}(\tau)$ outlined in Section 5.

Finally expanding $\Psi_{ee}(\tau)$ in equation (50) shows that all terms are detected.

9. REFERENCES

(1) Goring B., Unbehauen H.: Application of different statistical tests for the determination of the most accurate order of the model in parameter estimation, 3rd IFAC Symp., Ident. & Syst. Par. Est., 1973, pp. 917-928.

(2) Boom van den A.J.W., Enden van den A.W.M.: The determination of the order of process and noise dynamics, 3rd IFAC Symp., Ident. & Sys. Par. Est. 1973, pp. 929-938.

(3) Gustavsson I.,: Survey of applications of identification in chemical and physical processes., 3rd IFAC Symp., Ident. & Sys. Par. Est., 1973, pp. 67-85.

(4) West J.C.: Nonlinear signal distorition correlation., Int. J. Control, 1965, 2, pp. 529-538.

(5) Douce J.L., Identification of a class of nonlinear systems', 4th IFAC Symp., Ident. Syst. Par. Est., 1976, pp. 1-14.

(6) Subba Rao T., Gabr, M.M.: A test for linearity of stationary time series., Res. Rp. UMIST.

(7) Rabjman N.S.: The application of identification methods in the USSR - a survey: Automatica, 1976, 12, pp. 73-95.

(8) Billings S.A., Leontaritis I.J.,: Identification of nonlinear systems using parameter estimation techniques, IEE Conf. Control and its Applications, Warwick, 1981, pp. 183-187.

(9) Billings S.A., Leontarities I.J., : Parameter estimation techniques for nonlinear systems; 6th IFAC Symp., Ident. and Syst. Par. Est., Washington, 1982, pp. 427-432.

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

(11) Granger .W.J., Anderson A.P.,: An introduction to bilinear time series models, Vandenhoech and Ruprecht, Gottingen, 1978.

(12) Billings S.A.: Identification of nonlinear systems - a survey, Proc. I.E.E., Pt. D., 1980, 127, pp. 272-285.

(13) Box G.E.P., Jenkins G.M.: Time series analysis, Holden Day, San Francisco, 1976.

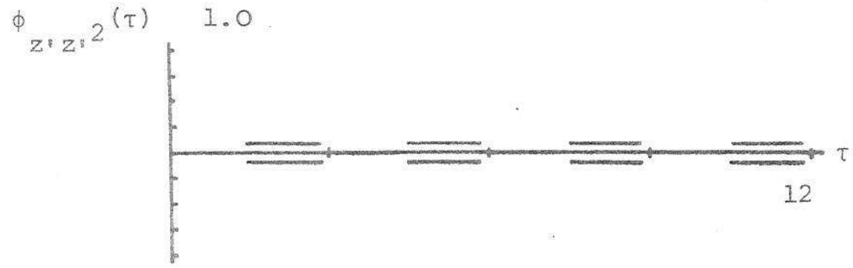

Fig. 1 Detecting nonlinear effects

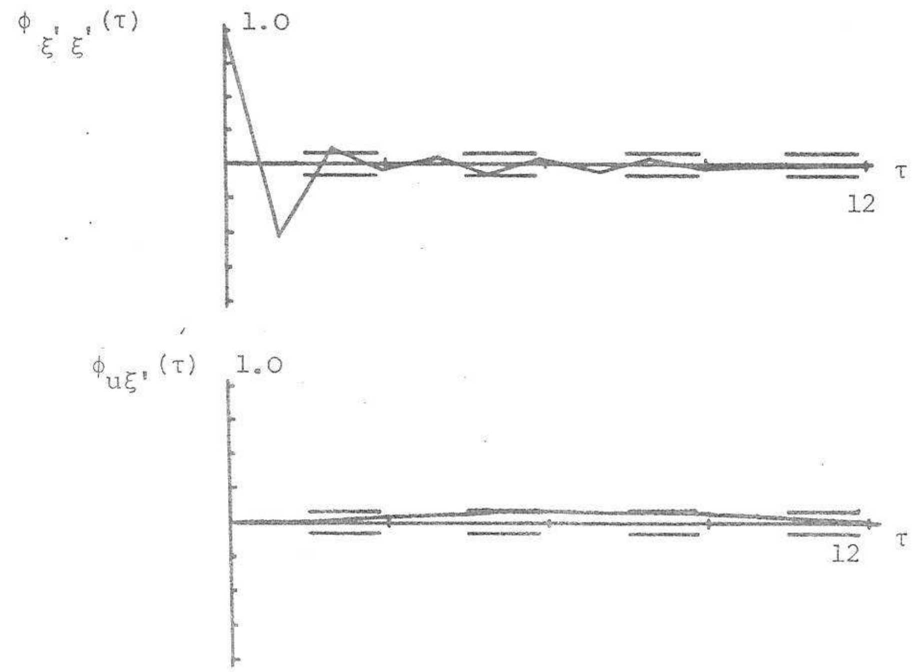

Fig. 2 Model validity tests

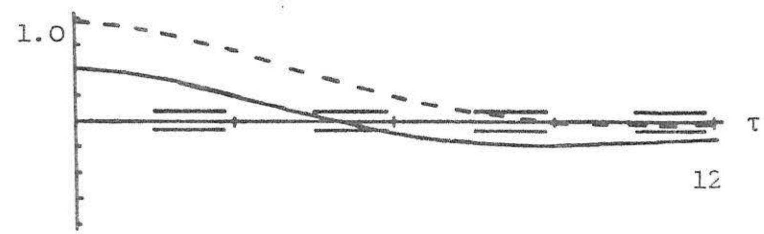

Fig. 3 Testing for nonlinearity

—— $\phi_{z'z'^2}(\tau)$

- - - $\phi_{z'_b z'^2_b}(\tau)$

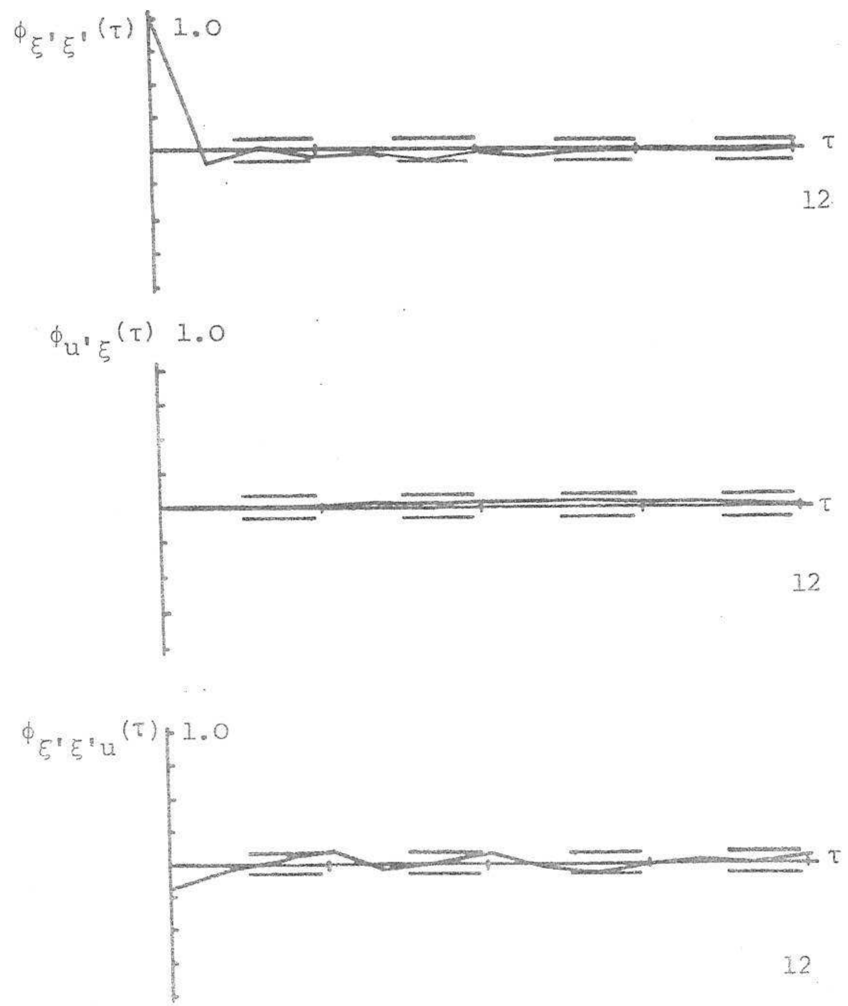

Fig. 4 Model Validity Tests