Monograph:

Billings, S.A. and Fadzil, M.B. (1984) The Practical Identification of Systems with Nonlinearities. Research Report. Acse Report 264 . Dept of Automatic Control and System Engineering. University of Sheffield

[redacted]

The Practical Identification of Systems with Nonlinearities

S.A.Billings, B.Eng., Ph.D., C.Eng., MIEE, AFIMA, M.Inst.MC

M.B.Fadzil B.Sc., M.Sc.

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

Research Report No. 264

August 1984

Submitted to 7th IFAC Symposium on
Identification and System Paramater
Estimation, York 1985.

Abstract

The practical aspects of identifying nonlinear systems are discussed by using the identification of a heat exchanger as a case study example. Methods of determining if the system is linear or nonlinear prior to parameter estimation are illustrated. Techniques for estimating the unknown parameters in a nonlinear difference equation model called a NARMAX model are discussed and the application of model validity tests which detect the existence of unmodelled linear or non-linear terms in the residuals are demonstrated. Throughout the emphasis is placed on the practical implementation of the techniques introduced.

1. Introduction

The identification and digital control of linear systems is primarily based on the linear difference equation model relating sampled outputs to sampled inputs. Numerous parameter estimation routines and controller design procedures have been developed based on this description {Goodwin & Payne, 1977; Isermann, 1981; Ljung & Soderstrom, 1984, Harris, Billings, 1981}. If the response is dominated by nonlinear characteristics however it will be necessary to use a nonlinear model rather than an approximate linear description and this immediately raises the problem of what class of models to use {Marmarelis & Marmarelis, 1978; Billings, 1980}. The choice of model is vitally important since this will influence its usefulness in prediction and control, and in view of the success of the linear difference equation model it is natural to search for a nonlinear extension of this description.

In the present study the identification of a nonlinear difference equation model of a heat exchanger is discussed based on a NARMAX or Nonlinear AutoRegressive Moving Average model with eXogeneous inputs. The derivation and conditions for the existence of such models have been discussed elsewhere {Leontaritis & Billings, 1985a}. and the emphasis in the present investigation will be on the practical implementation of identification procedures based on this description. Methods of detecting if the system is linear or nonlinear prior to parameter estimation are illustrated and the selection of input signals are discussed. Techniques for estimating the unknown parameters in the NARMAX model are described and the application of model validity tests which detect the existence of unmodelled linear or nonlinear terms in the residuals are demonstrated. Throughout it is shown that blind application of linear identification techniques to the estimation of nonlinear models usually leads to incorrect results.

2. Identification of a Heat Exchanger

2.1 The System

The heat exchanger which is illustrated schematically in Fig. 1a consists of a radiator through which heated water is passed and a fan which blows air across the radiator. Water is pumped through the radiator around a closed loop which includes a heater tank. The control objective is to control the temperature drop across the radiator together with the air flow rate across it by adjusting the inputs to the heater and the fan. A block diagram of the system is shown in Fig. 1b. Extensive experimentation has shown that the loops G_{22} and G_{12} are linear while G_{11} is nonlinear. Lack of space precludes a complete multivariable analysis here and so only the nonlinear loop will be considered.

2.2 Testing for Nonlinearities

In the early stages of experimentation on a process it is important to determine if the process under test exhibits nonlinear characteristics which will warrant a nonlinear model. The simplest method of achieving this objective is to inject step inputs of varying amplitude into the process and plot the system gain against input amplitude. If the process cannot be taken off normal production however or if the data is pre-recorded or the analysis relates to the system residuals alternative tests must be used {Billings & Voon , 1983}.

It can readily be shown that whenever the input u(t)+b, $\overline{u(t)} = 0$, $b \neq 0$ is applied to a system, the system cannot be linear if $\overline{z_b(t)} \neq \overline{z(t)}$ where $\overline{z_b(t)}$ and $\overline{z(t)}$ are the mean levels of the system output for the inputs b (i.e. u(t) = 0) and u(t)+b respectively. Alternatively, 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 these properties) then the process is linear iff

$$\phi_{z' z'^2}(\nabla) = E\left[z'(t+\nabla)(z'(t))^2\right] = 0 \;\forall\; \nabla \qquad (1)$$

where the dash superscript indicates that the mean has been removed. The test will distinguish between additive noise corruption of the measurements and distortion due to nonlinear effects providing the noise and input are independent.

Application of step tests to loop G_{11} of the heat exchanger indicated significant nonlinear effects and this was confirmed by $\phi_{z'z'^2}(\nabla)$, illustrated in Fig 2, which is clearly well outside the confidence bands. The latter was computed based on a Gaussian white input with mean -0.1177 and bandwidth 0.5Hz. Similar tests on the loops G_{22} and G_{12} indicated that these could be adequately described by a linear model.

2.3 Input Design

Experiment design is probably the most important step in a system identification study because all the results thereafter are dependent upon the quality and information content of the data collected. Whilst normal operating records could be used as a basis for the identification it is preferable whenever possible to inject externally generated inputs which can be tailored to the process. These inputs are often designed, based on preliminary experiments on the process, to be persistently exciting. For a nonlinear system this means that the input should be selected to excite all the modes and amplitudes of interest within the system. Pseudo random sequences are not in general appropriate for nonlinear systems since they exhibit discontinuous probability density functions and will not yield a persistently exciting input over the full amplitude range of any input nonlinearities. The design of inputs for nonlinear system identification is complex but general rules have been derived from information and theoretic arguments {Leontaritis and Billings, 1985b}. These indicate that for a power or amplitude constraint on the input, the input should be an independent sequence. In addition the input should have a gaussian distribution for a power constraint, and a uniform distribution for an amplitude constraint.

In the case of the heat exchanger simple step tests were initially performed to provide estimates of the system bandwidth. The amplitude range of interest was the total allowable variation on the input transducers. Based on this information a gaussian white excitation was designed with a mean of -0.1177 and a bandwidth of 0.5Hz. To ensure that all possible modes of the process were excited inputs either side the designed bandwidth, in this case 0.05Hz, 0.15Hz and 1.5Hz, were also injected into the system. Whilst this is not normally necessary it can often be very worthwhile providing experimentation time on the process is available. The range of input bandwidths increases the probability of capturing the true dynamics of the process and the different pieces of data can always be used in model validation studies.

In all cases 1000 data pairs were digitized at a sampling rate of 0.3 secs.

As a rule of thumb in linear identification the sampling rate is fixed by the bandwidth of the input. This will not necessarily follow when the process is nonlinear because there may be high frequency harmonics in the output which reflect the nonlinear dynamics of the process.

2.4 Linear Identification

Because the structure detection test $\phi_{z'z'^2}(t)$ was nonzero this has already established that the process is nonlinear. It is worthwhile however illustrating what happens if a linear model is fitted to the data.

Many simulation studies have indicated that fitting linear models to data with significant nonlinearities can provide very misleading results. This occurs because the linear parameter estimation routines tend to yield an estimated combination of process and noise model which visually provide a good prediction of the system output. The model is however significantly biased and prediction over a different data set usually reveals this. The application of the traditional linear covariance tests for checking the validity of the model can fail to indicate the inadequacy of the estimated model when the process is nonlinear {Billings and Voon, 1983}. This can mislead the experimenter into believing that the model is adequate when it is not.

The first of these problems can be overcome by splitting the data set into an estimation and a prediction set. Better still, estimate the model using data from one experimental condition (the estimation set) and analyse the prediction over data from a second experimental condition (the prediction set).

A detailed theoretical study {Billings and Voon, 1983} of model validity tests based on correlation analysis has shown that 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 hold iff

$$\begin{aligned} \phi_{\xi\xi}(\tau) &= \delta(\tau) \\ \phi_{u\xi}(\tau) &= 0 \ \forall \tau \\ \phi_{\xi\xi u}(\tau) &= E[\xi(t)\xi(t-1-\tau)u(t-1-\tau)] \\ &= 0 \ \forall \tau \geq 0 \end{aligned} \qquad (2)$$

Notice that for nonlinear systems the traditional linear tests $\phi_{\xi\xi}(\tau)$ and $\phi_{u\xi}(\tau)$ 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

$$\begin{aligned} \phi_{u\xi}(\tau) &= 0 \ \forall \tau \\ \phi_{u^{2'}\xi}(\tau) &= 0 \ \forall \tau \\ \phi_{u^{2'}\xi^2}(\tau) &= 0 \ \forall \tau \end{aligned} \qquad (3)$$

Experience has shown that when using a prediction error algorithm the tests in both equns (2) and (3) 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.

For both the estimation of the linear and the nonlinear models 500 data pairs from the 0.5Hz bandwidth experiment will be used as the estimation set. The prediction set will consist of 500 data pairs from the 0.05Hz bandwidth experiment.

Extensive preliminary analysis of the data involved estimating the coefficients in linear models of varying process and noise model orders and with various time delays computing the loss function and analysing the residuals. Linear model validity tests for a second order process model with a second order noise model are illustrated in Fig. 3. Although $\phi_{u\xi}(\tau)$ is slightly outside the 95% confidence bands $\phi_{\xi\xi}(\tau)$ is well outside indicating a deficiency in the noise model. Increasing the noise model order to four alleviated this difficulty and gave what appeared to be the best linear model

$$z'(t) = 0.851z'(t-1)-0.1571z'(t-2)+0.265u'(t-1)-0.333u'(t-2)$$
$$+\varepsilon(t)-0.08942\varepsilon(t-1)+0.339\varepsilon(t-2)+0.227\varepsilon(t-3)+0.0813\varepsilon(1+4) \quad (4)$$

The coefficients in this model were estimated using a prediction error algorithm. A comparison of the process output and the predicted output of the estimated linear model eqn (4) is illustrated in Fig. 4. Notice that whilst the predicted and the process outputs are virtually coincident over the estimation set (the first 500 points) the deficiency of the linear model is clearly evident over the prediction set (last 500 points). This illustrates the importance of analysing the prediction errors over a different data set. Note the limiting in the output at the beginning of the prediction set Fig. 4.

Computing the residuals and applying the model validity tests of eqns (2) and (3) gave the results illustrated in Fig. 5.

Inspection of Fig. 5 shows that the standard linear covariance tests $\phi_{\xi\xi}(\tau)$, $\phi_{u\xi}(\tau)$ now indicate that the model is adequate. A normal linear analysis would therefore terminate at this point although it would appear that there is room for a significant improvement in the prediction accuracy of the model, Fig. 4. The poor prediction would probably be assumed to be a result of a low s/n ratio. The model validity tests designed to detect unmodelled nonlinear terms in the residuals Fig. 5 however, confirm the results of the $\phi_{z'z'^2}(t)$ test, Fig.2, which indicated that the process is nonlinear. Since $\phi_{\xi\xi u}(t)$ is acceptable this shows that there are no terms of the form $u^q(t-m)\varepsilon(t-n)$ $\forall n,m,$odd$q$ in the residuals. $\phi_{u^{2'}\xi^2}(\tau)$ is just outside the confidence bands and $\phi_{u^{2'}\xi}(\tau)$ is unacceptably large. This combination strongly suggests that even terms (e.g. $u^2(\cdot)$) and/or internal noise terms of the form $u^k(t)\varepsilon^\ell(t)$ $\forall$ t, ℓ,k even should be added to the model. The effects of introducing these nonlinear terms into the model was therefore investigated.

2.5 Nonlinear identification

Before we can discuss the estimation of a nonlinear model of the heat exchanger we need to consider the introduction of a class of non-linear models and develop estimation routines based on these.

2.5.1 Parameter estimation for the NARMAX model

It can be proved {Leontaritis and Billings,1985a} that under fairly mild conditions nonlinear r-input m-output multivariable stochastic systems can be represented by the model

$$\begin{aligned} z_i(t+p) = q_i\big[& z_1(t+n-1), \;\ldots\; z_1(t) \\ & \vdots \\ & z_m(t+n_m-1), \;\ldots\; z_m(t) \\ & u_1(t+p), \;\ldots\; u_1(t) \\ & \vdots \\ & u_r(t+p), \;\ldots\; u_r(t) \\ & \varepsilon_1(t+p-1), \;\ldots\; \varepsilon_1(t) \\ & \vdots \\ & \varepsilon_m(t+p-1), \;\ldots\; \varepsilon_m(t)\big] + \varepsilon_i(t+p) \end{aligned} \qquad (5)$$

where i = 1, 2, ... m, the integeters $n_1,n_2 \ldots n_m$ are the observability indices, q_i are nonlinear functions and $p = \max(n_1,n_2 \ldots n_m)$. This model is called the multi-structural input-output prediction error or innovation model. For single-input single-output systems the model, eqn(5) becomes

$$\begin{aligned} z(t) = q\big[& z(t-1), \;\ldots\; z(t-n_y), \; u(t-d), \;\ldots\; u(t-d-n_u+1), \\ & \varepsilon(t-1) \;\ldots\; \varepsilon(t-n_\varepsilon)\big] + \varepsilon(t) \end{aligned} \qquad (6)$$

where $q[\cdot]$ is some nonlinear function, d is the time delay and $\varepsilon(\cdot)$ represents the prediction errors

$$E[\varepsilon(t) \mid z(t-1),\ z(1-2), \ldots\ u(t)\ \ldots\ldots] = 0 \qquad (7)$$

The nonlinear difference equation model, eqn (6), is referred to as the nonlinear ARMAX or NARMAX model.

The models in eqns. (5) - (6) are valid provided the system is finitely realizable and could be approximated by a linear model if operated in a region close to the equilibrium point. A rigorous derivation of these results and a comparison of the models with other well known nonlinear representations such as the volterra and state-affine models is given elsewhere {Leontaritis & Billings (1985a)}.

Expanding eqn (6) as a polynomial and regrouping terms

$$\begin{aligned} z(t) = {} & G^{zu}[z(t-1). \ \ldots\ z(t-n_y),\ u(t-d)\ \ldots\ u(t-d-n_u+1)] \\ & + G^{zu\varepsilon}[z(t-1),\ \ldots\ z(t-n_y), u(t-d)\ \ldots\ u(t-d-n_u+1), \\ & \varepsilon(t-1),\ \ldots\ \varepsilon(t-n_\varepsilon)] + G^{\varepsilon}[\varepsilon(t-1),\ \ldots\ \varepsilon(t-n_\varepsilon)] \\ & + \varepsilon(t) \end{aligned} \qquad (8)$$

Separating out the unknown parameters gives

$$\begin{aligned} z(t) &= \psi(t)^T\theta(t-1) + \varepsilon(t) \\ &= \begin{bmatrix} \psi_{zu}^T(t) & \psi_{zu\varepsilon}^T(t) & \psi_{\varepsilon}^T(t) \end{bmatrix} \begin{bmatrix} \theta_{zu}(t-1) \\ \theta_{zu\varepsilon}(t-1) \\ \theta_{\varepsilon}(t-1) \end{bmatrix} + \varepsilon(t) \end{aligned} \qquad (9)$$

where $G^{zu}[\cdot] = \psi_{zu}^T(t)\ \theta_{zu}(t-1)$, $G^{zu\varepsilon}[\cdot] = \psi_{zu\varepsilon}^T(t)\theta_{zu\varepsilon}(t-1)$ $G^{\varepsilon}[\cdot] = \psi_{\varepsilon}^T(t)\ \theta_{\varepsilon}(t-1)$ and the definitions of the ψ's and θ's follows.

Grouping all terms involving $\varepsilon(t)$ and defining

$$\xi(t) = \psi_{zu\varepsilon}^T(t)\ \theta_{zu\varepsilon}(t-1) + \psi_{\varepsilon}^T(t)\theta_{\varepsilon}(t-1) + \varepsilon(t) \qquad (10)$$

gives

$$z(t) = \psi_{zu}^T(t)\ \theta_{zu}(t-1) + \xi(t) \qquad (11)$$

Any noise which enters the nonlinear system as an internal disturbance cannot in general be translated to be additive at the output as in linear systems. This induces the multiplicative terms, represented by $G^{zu\varepsilon}[\cdot]$ in eqn. (8), between the prediction errors and the measured inputs and outputs. This problem arises even when the noise is additive at the output because the NARMAX model maps past inputs and outputs into the present output. Consequently, as eqn. (10) shows $\xi(t)$ will in general be highly correlated with the elements of $\psi_{zu}^{T}(t)$ and the direct application of many of the parameter estimation routines developed for linear systems will yield biased estimates {Billings & Leontaritis 1982}.

Fortunately, the recursive extended least squares algorithm can be readily applied to the NARMAX model eqn.(9) by expanding $\psi_{zu}(t)$, $\psi_{zu\varepsilon}(t)$, $\psi_{\varepsilon}(t)$ as polynomials based on a previous estimate of the prediction errors $\varepsilon(t)$. The development of a recursive maximum likelihood algorithm is slightly more involved and required a major rederivation. It can be shown that instrumental variables will yield unbiased estimates providing the noise terms in the NARMAX model can be represented as a purely linear map. This restriction can be widened slightly by employing a new suboptimal least squares routine based on the model

$$z(t) = q\left[\hat{y}(t-1) \;....\; \hat{y}(t-uy), u(t-1-d) \;...\; u(t-d-n_u)\right] + \varepsilon(t) \qquad (12)$$

where $\hat{y}$ (t) represents the predicted output. The algorithm will yield unbiased estimates whenever the noise is additive at the output. The properties of the noise required by the latter two algorithms are quite restrictive but can be verified {Billings & Voon 1984} by suitable application of the model validity tests in eqns. (2) and (3).

The direct application of a maximum likelihood algorithm is not possible unless the distribution of the prediction errors, which in general will not be gaussian, is known. However, by considering the loss function

$$J(\theta) = \frac{1}{2N} \log_e \det \sum_{t=1}^{N} \varepsilon(t;\theta)\, \varepsilon(t;\theta)^T \qquad (13)$$

it can be shown that the prediction error estimates obtained by minimising eqn (13) have very similar asymptotic properties to the maximum likelihood estimate even when $\varepsilon(t)$ is non-gaussian {Goodwin & Payne 1977:Ljung & Soderstrom, 1984} A prediction error algorithm for the NARMAX model eqn (9) has been developed based on this result.

Before any of these algorithms can be applied to the NARMAX model a method of determining which terms to include in the model must be developed. Direct estimation based on a polynomial expansion of eqn (9) will involve an excessive number of terms. Simply increasing the order of the dynamic terms (n_y, n_u) and the order of the polynomial expansion to achieve the desired prediction accuracy will in general result in an excessively complex model and possibly numerical ill-conditioning. Several procedures have therefore been developed based on the NARMAX model and the algorithms mentioned above to detect the structure of the model. These are based on a new liklihood ratio test, Akaike tests and forward, backward and stepwise regression algorithms.

2.5.2 Estimation of a nonlinear model of the heat exchanger

A prediction error estimation algorithm coupled with a stepwise regression procedure was used to estimate the coefficients in the non-linear model of the heat exchanger. The data set was exactly the same as for the linear model estimation (500 data pairs in the estimation set from the 0.5Hz bandwidth experiment and 500 data pairs in the prediction set from the 0.05Hz bandwidth experiment) and the initial specification for the NARMAX model was:-

order of lagged inputs $n_y = 2$

order of lagged outputs $n_y = 2$

order of lagged prediction errors $n_\varepsilon = 2$

delay of input $d = 1$

degree of nonlinearity of the input and output $= 3$

degree of nonlinearity of the prediction errors $= 2$

With this initial specificiaiton the total number of possible terms in the model was eighty three. Allowing the stepwise regression algorithm linked with a prediciton error routine to sort through all the possible terms using a Fisher F-ratio test operating with 95% confidence bounds produced the follwoing model

$$\begin{aligned} z(t) &= 2.072 + 0.9158z(t-1) + 0.4788u(t-1) \\ &- 0.01572z^2(t-1) - 0.01133u^2(t-1) - 0.002244z^2(t-1)u(t-1) \\ &- 0.002239u^3(t-1) + \varepsilon(t) \end{aligned} \tag{14}$$

A constant term ot dc level is included in the estimated model eqn (14) to accommodate any mean levels that are present in the raw data $u(\cdot)$, $z(\cdot)$. The nonlinear model must be estimated using the raw data. Estimation based on normalised data $z'(\cdot) = z(\cdot) - \bar{z}$, $u'(\cdot) = u(\cdot) - \bar{u}$ as in the linear case eqn (4) will in general result in a model which is input sensitive {Billings and Voon, 1984} . This means that the model parameters become a function of the variance and higher order statistics of the input signal. A model estimated for one particular input would therefore be invalid for prediction based on any other input with different statistics. This problem can be avoided by operating on the raw data. Differencing the data should also be avoided since this will often induce a large number of nonlinear terms into the model.

The nonlinear model eqn(14) includes first order lags in the input and output together with four nonlinear terms including even order terms as previously suggested. From all the possible linear/nonlinear combination of terms within the original specification therefore only those in eqn (14) were found to be significant using the specified confidence levels. In practice we have found that these results provide an excellent basis for experimentation on the structure of the model. Experimentation is performed by analysing the prediction accuracy and validity tests and adding or deleting terms from the model.

The prediction accuracy of the estimated model eqn (14) was good but the model validity tests illustrated in Fig 6 show that the model is deficient in some way. Note that $\phi_{u^{2'}\xi}(\tau)$ which was very large for the linear model, eqn (4) Fig 5, is now acceptable but $\phi_{u^{2'}\xi^2}(\tau)$ is still just outside the confidence bands for lags 11 and 12. This probably indicates that the nonlinearities are correctly modelled since the deviations in $\phi_{u^{2'}\xi^2}(\tau)$ could be caused by unmodelled linear terms. Note that $\phi_{u^{2'}\xi}(\tau)$ would not detect linear terms in the residuals. This hypothesis is supported by $\phi_{\xi\xi}(\tau)$ and $\phi_{u\xi}(\tau)$ which are now unacceptable. A comparison of Fig 6 with Fig 3 shows that $\phi_{\xi\xi}(\tau)$, $\phi_{u\xi}(\tau)$ are very similar in both diagrams and suggests that the addition of a higher order noise model to eqn (14) may improve the situation.

Because a fourth order noise model was found to alleviate the problem in Fig 3 this was added to the model of eqn (14) to yield

$$\begin{aligned}
z(t) &= 2.381 + 0.8345z(t-1) + 0.4828u(t-1) - 0.0106z^2(t-1) \\
&- 0.00867u^2(t-1) - 0.002235z^2(t-1)u(t-1) \\
&- 0.002293u^3(t-1) + \varepsilon(t) - 0.05194\varepsilon(t-1) \\
&+ 0.008461\varepsilon(t-2) + 0.2159\varepsilon(t-3) + 0.08243\varepsilon(t-4) \qquad (15)
\end{aligned}$$

Computation of the model validity tests for this model revealed that $\phi_{\xi\xi}(\tau)$ and all other tests were now acceptable except $\phi_{u\xi}(\tau)$ which was still just outside the confidence limits. This will almost certainly be caused by the omission of a linear term from the model. The conclusion follows from the knowledge that $\phi_{u\xi}(\tau)$ detects deficiencies in linear and/or odd nonlinear terms in the model. If there were odd nonlinear terms missing from the model $\phi_{u^{2'}\xi^2}(\tau)$ should also be unacceptable, and it is not. These conjectures are further supported by the knowledge that a second order model was found to be appropriate in the linear analysis.

In addition to the fourth order noise model therefore a term in u(k-2) was added to the model in eqn (14) nad the coefficients re-estimated to yield

$$z(t) = 2.301 + 0.9173z(t-1) + 0.449u(t-1) + 0.04577u(t-2)$$
$$- 0.01889z^2(t-1) - 0.00999u^2(t-1)$$
$$- 0.002099z^2(t-1)u(t-1) - 0.002434u^3(t-1)$$
$$+ \varepsilon(t) - 0.004\varepsilon(t-1) + 0.0380\varepsilon(t-2)$$
$$+ 0.2745\varepsilon(t-3) + 0.1037\varepsilon(t-4) \qquad (16)$$

The model validity tests for this model which are illustrated in Fig 7 are now all within the confidence bands. Adding a z(t-2) term to the model was investigated but was found to be un-necessary. The predicted output of the final model eqn (16) is illustrated in Fig 8. A comparison with the linear model predictions Fig 4 shows the significant improvement achieved when using the nonlinear model. Notice that this is only evident over the prediction set emphasising the need to split the data to provide a rigorous test of its goodness-of-fit.

3. Conclusions

Several practical and theoretical problems associated with the identification of nonlinear systems have been discussed. The identification of the heat exchanger as a case study has illustrated the significant improvement in prediction accuracy which can be achieved by the addition of just a few nonlinear terms to the model. The simplicity of the final estimated nonlinear model can however only be achieved by combining several realization, structure detection, estimation and model validity tests and using these in an interactive feedback manner. Almost all of these algorithms have been designed to be simple to implement and interpret and their practical application has been demonstrated on the heat exchanger. Other processes including a liquid level system and a 6996 bhp industrial diesel engine have also been successfully identified using this approach. It is important to emphasise however, that direct application of linear methods of data analysis can severely mislead the experimenter to the extent that the erroneous results remain undetected by linear validity tests.

Even if the process to be studied is controlled to operate in a linear region it will often be advantageous to identify a nonlinear model which is valid over the total allowable operating regime. Such a model would provide valuable information to the designer regarding the errors associated with a linear analysis and would allow him to simulate and design for the effects of significant deviations from the chosen operating point. In some instances it will be appropriate to design a nonlinear controller based on the NARMAX description of the process. Providing the amount of work involved in identifying the nonlinear model is not excessively more than would be the case for a linear model, as demonstrated in the present study, then this becomes an attractive proposition.

Acknowledgements.

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

References

Billings S.A. (1980) : Identification of nonlinear systems - a survey; Proc. IEE, Part D, 130, pp 193-199.

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

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

Billings S.A., Voon W.S.F. (1984) : Least squares parameter estimation algorithms for nonlinear systems; Int. J. Syst. Sci, 15, pp 601-615.

Goodwin G.C., Payne R.L. (1977) : Dynamic system identification; Academic Press.

Harris C.J., Billings S.A. (Eds) (1981) : Self-tuning and adaptive control: Peter Peregrinus.

Isermann R : Digital control systems; Springer Verlag.

Leontaritis I.J., Billings S.A. (1985a): Input-output parametric models for nonlinear systems. Part I - deterministic nonlinear systems. Part II - stochastic nonlinear systems, Int. J. Control, (to appear)

Leontaritis I.J., Billings S.A. (1985b) : Experiment design for nonlinear systems; (in preparation).

Ljung L., Soderstrom T. (1984) : Theory and practice of recursive identification; MIT Press.

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

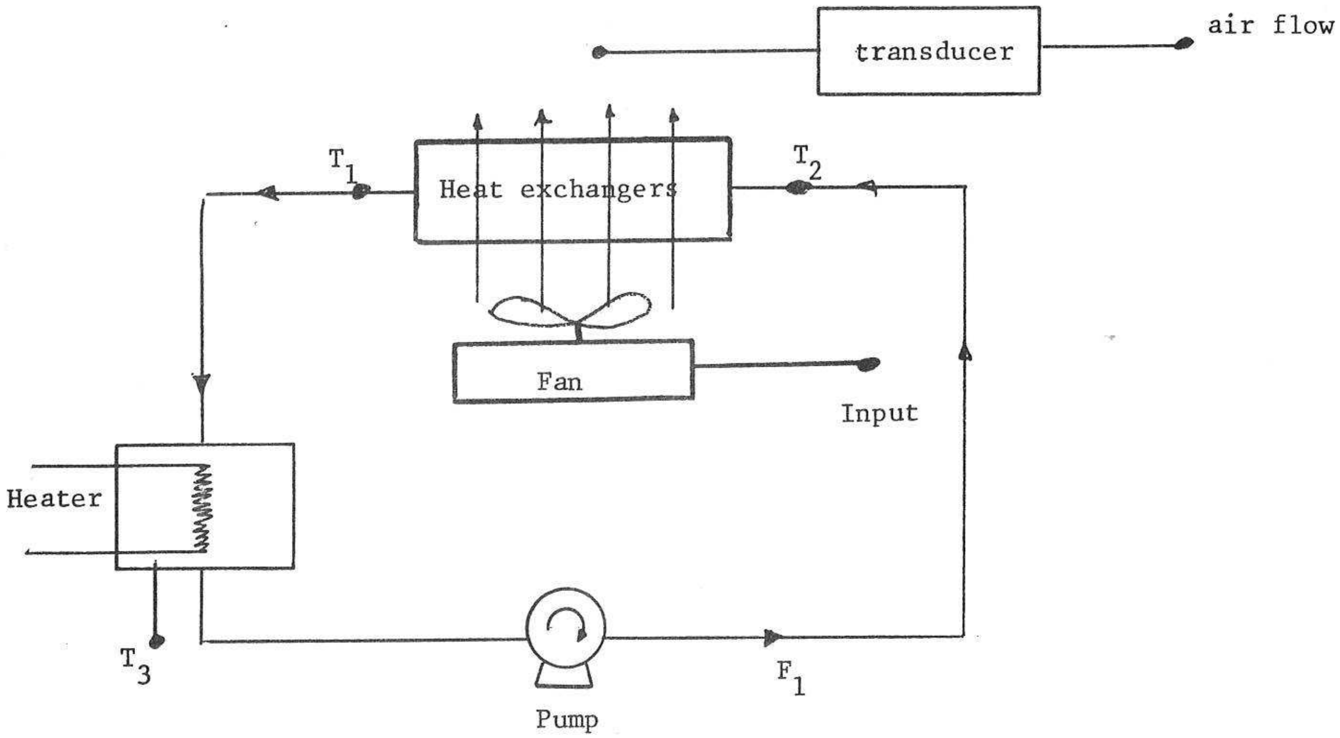

Fig. 1a. The Heat Exchanger

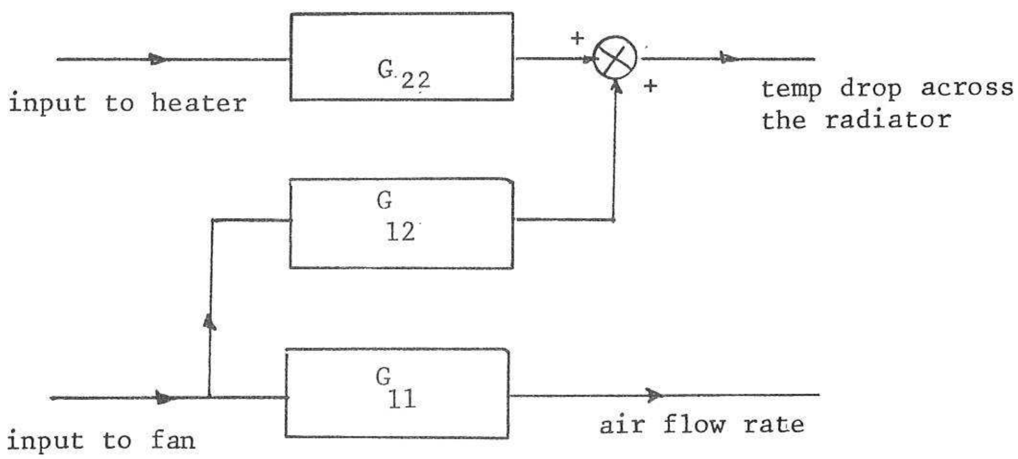

Fig. 1b. Schematic Diagram

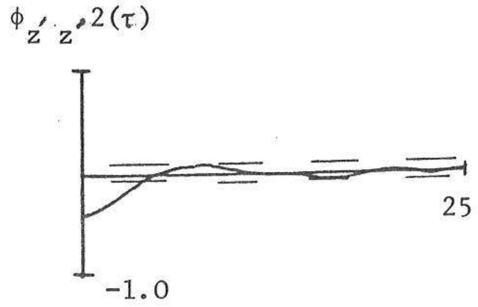

Fig. 2. Nonlinear detection test

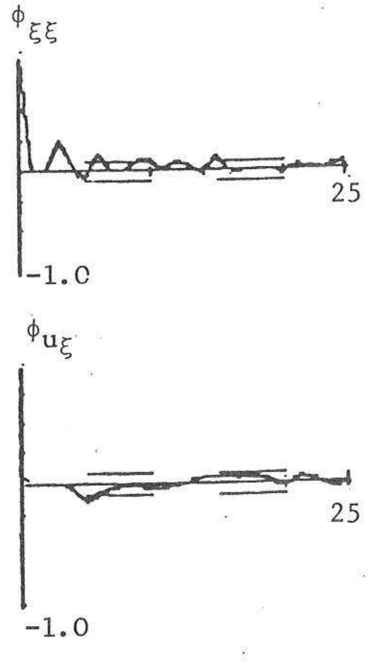

Fig. 3. Model validity tests

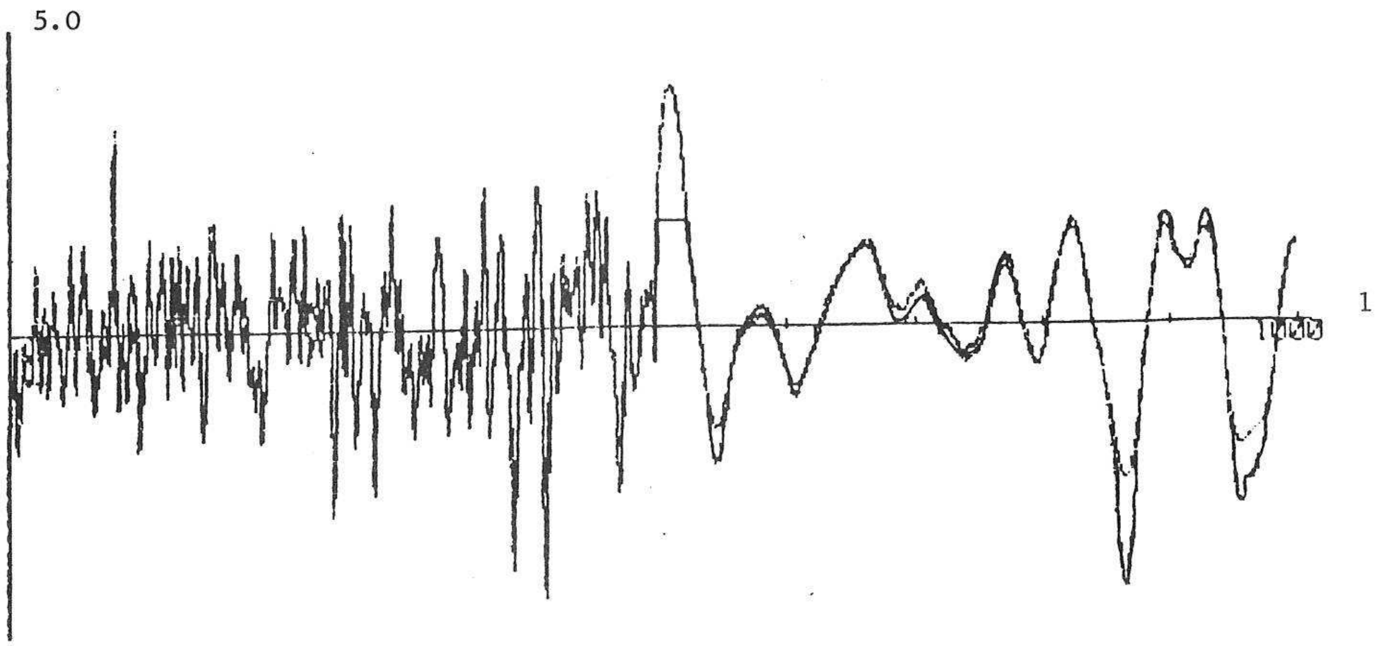

Fig. 4. Precicted output of the linear model.

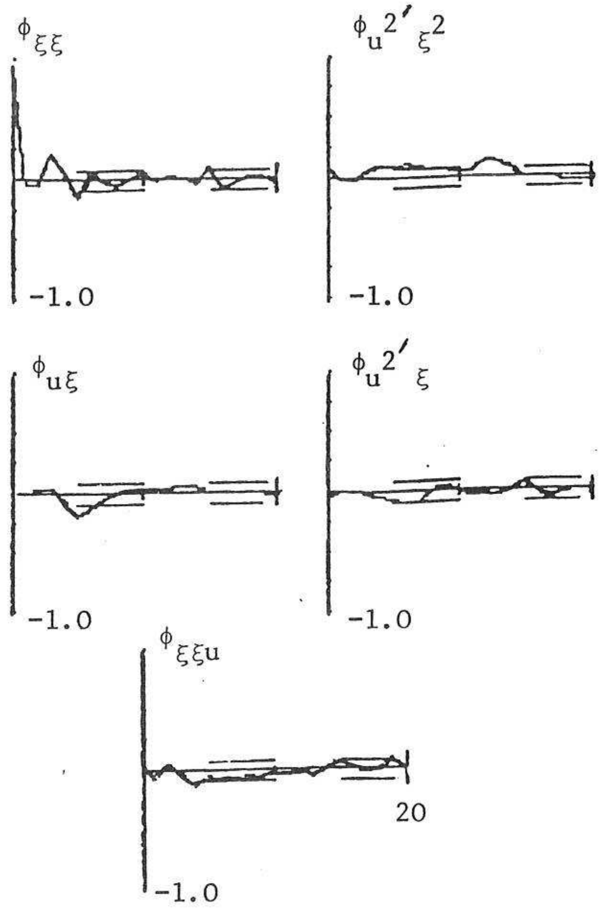

Fig. 6. Validity tests - nonlinear model

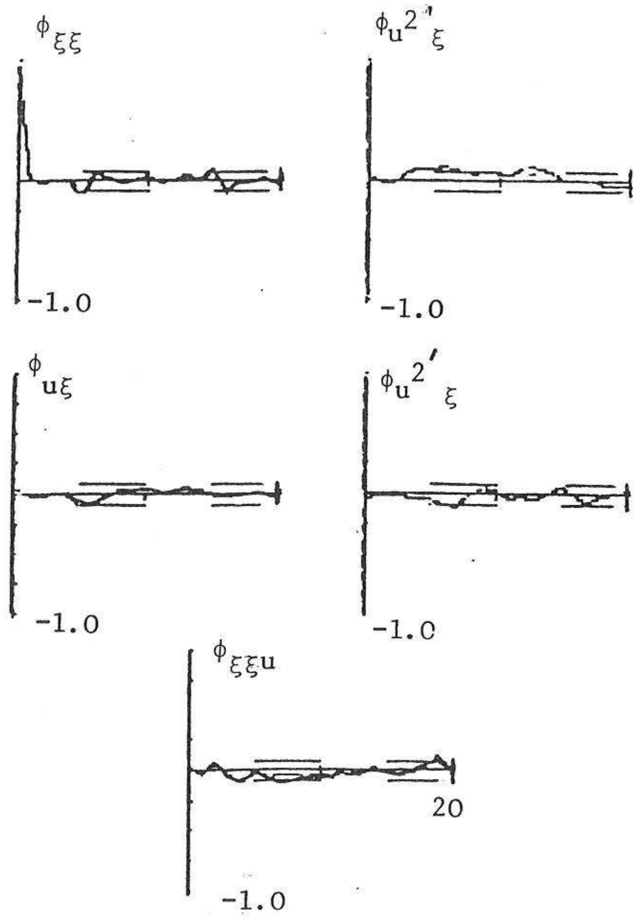

Fig. 7. Validity tests - best nonlinear model

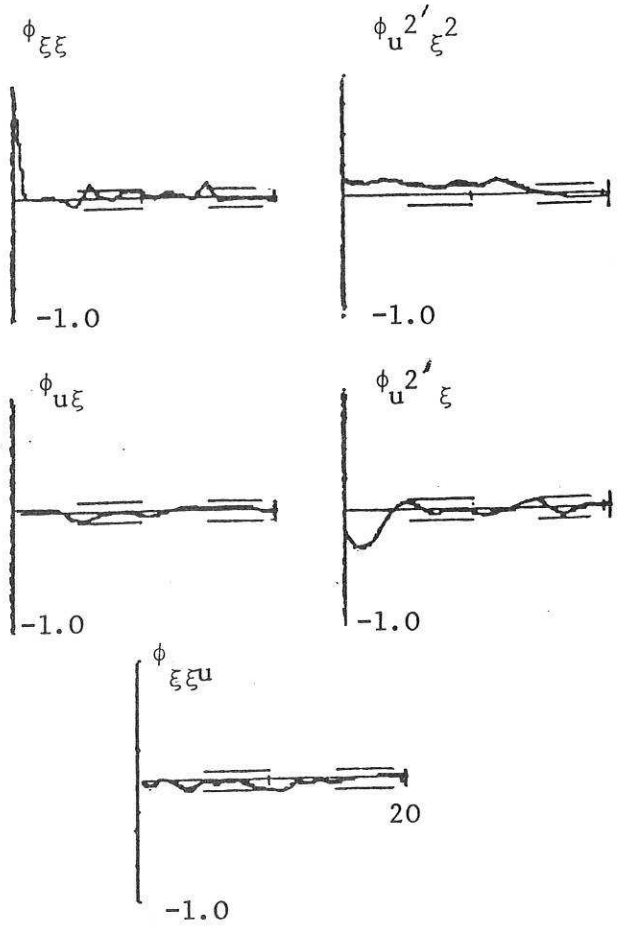

Fig. 5. Validity tests - best linear model