Non-linear modelling of a one-degree-of-freedom twin-rotor multi-input multi-output system using radial basis function networks

S M Ahmad[1], **M H Shaheed**[2], **A J Chipperfield**[3] and **M O Tokhi**[3*]
[1]Department of Mechanical and Marine Engineering, The University of Plymouth, UK
[2]Department of Engineering, Queen Mary University of London, UK
[3]Department of Automatic Control and Systems Engineering, The University of Sheffield, UK

Abstract: Modelling of innovative aircraft such as unmanned air vehicles (UAVs), X-wing, tilt body and delta-wing aircraft is not easy. It is argued in this paper that non-linear system identification is suitable for modelling air vehicles of complex configuration. This approach is demonstrated through a laboratory test rig. Extensive time and frequency-domain model-validation tests are employed in order to instil confidence in the estimated model. The estimated model has a good predictive capability and can be utilized for non-linear flight simulation studies. Some aspects of the modelling approach presented may be relevant to flight mechanics modelling of new generations of air vehicle.

Keywords: helicopter, non-linear system identification, radial basis function networks, twin rotor MIMO system

NOTATION

c_i	RBF centres
DOF	degree-of-freedom
$e(t)$	zero mean white noise sequence
E	residual vector
ERR	error-reduction ratio
$f(\cdot)$	non-linear functional form
F1, F2	thrust generated by rotors in the vertical and horizontal planes respectively
$n = 1, \ldots, N$	data length
n_r	number of centres or hidden neurons
n_y, n_u	maximum lags in the output and the input respectively
NARX	non-linear autoregressive model with exogenous inputs
NN	neural network
OLS	orthogonal least squares
P_i	fixed non-linear functions of lagged outputs and inputs (regressors)
P	regression matrix
PSD	power spectral density
RBF	radial basis function
TRMS	twin rotor multi-input multi-output system
$u(t)$	system input
UAVs	unmanned air vehicles
w_i	weights or parameters
w_0	bias or d.c. level
$\tilde{x}_i^n$	normalized variable x_i
$\bar{x}_i$	mean of variable x_i
$\mathrm{x}(t)$	input vector of RBF network
$y(t)$	system output response
$\hat{y}(t)$	model predicted output
β_i	Gaussian spread constant
$\delta(\tau)$	impulse function
$\varepsilon(t)$	prediction error or residual
θ_i	ith parameter
$\boldsymbol{\theta}$	weight or parameter vector
σ	variance
$\hat{\sigma}_{\varepsilon^2}$	variance of residual $\varepsilon(t)$
$\phi_{u\varepsilon}(\tau)$	cross-correlation function between $u(t)$ and $\varepsilon(t)$
$\|\cdot\|$	Euclidean norm

The MS was received on 4 February 2002 and was accepted after revision for publication on 10 September 2002.
* *Corresponding author: Department of Automatic Control and Systems Engineering, The University of Sheffield, Mappin Street, Sheffield S1 3JD, UK.*

1 INTRODUCTION

Recent advances in aircraft technology have led to the development of many new concepts in aircraft design, which are strikingly different from their predecessors. The differences are in both aircraft configuration and control paradigms. This trend can be attributed to the increasing emphasis on the need for aircraft to be agile (i.e. a high angle of attack), low-observable (stealth), multi-purpose, etc., for varied civilian and military operations. These new generation air vehicles have presented a variety of unprecedented challenges and opportunities to aerodynamicists and control engineers. The expectations of the new generation air vehicles to be highly agile and multi-functional demands that they perform over a large flight envelope. Enhanced agility in control terms implies a large excursion from the trim condition. In such a situation the linearized models can no longer describe the aircraft dynamics well enough. Hence, there is a need for high-fidelity non-linear dynamic models. Such models are essential for the design of control systems, validation and piloted simulation.

This paper presents a suitable modelling technique for such air vehicles. In this work, a non-linear system identification technique based on the radial basis function (RBF) is utilized for modelling an experimental test rig representing a *complex* twin rotor multi-input multi-ouput system (TRMS). The TRMS depicted in Fig. 1 is a laboratory set-up designed for control experiments by Feedback Instruments Limited [**1**]. The behaviour of the TRMS in certain aspects resembles that of a helicopter. From the control point of view it exemplifies a high-order non-linear system with significant cross-coupling.

While the Newtonian mechanics or the Lagrange equations of motion can be used to find the non-linear differential equations in a *generic* form, the unknown parameters must still be identified. Such model-based identification is commonly employed with practical systems. There are numerous examples that demonstrate the applicability, feasibility and versatility of the model-based concepts. For instance, neural networks have been employed for estimating the aerodynamic coefficients of unmanned air vehicles (UAVs) [**2**]. More recently, Kim and Calise [**3**] have used RBF networks to capture variations in aircraft Mach number. Here the neural network (NN) is used to perform the dual roles of (a) identifying the input–output model parameters (off-line learning) using the mathematical model of an aircraft and (b) an adaptive network that compensates for imperfect inversion and in-flight changes in the actual aircraft dynamics. An innovative time-domain non-linear mapping-based identification method has been presented by Lyshevski [**4**] for identification of unsteady flight dynamics. Lately, B-splines have been investigated in modelling and identification of non-linear aerodynamic functions of aircraft [**5**]. In all these cases the model structure is *known*. However, in the present study, no model order was assumed *a priori*, which implies that no physical insight is available or used. However, the selected model structure belongs to well-established methods (see Section 3). Such an approach yields input–output models with neither *a priori* defined model order nor specific parameter settings reflecting any physical aspects, i.e. black-box modelling. The approach is thus useful in modelling a class of air vehicles whose dynamics are not well understood or difficult to model from first principles using laws of physics. An initial account of this work was presented at *NAECON 2000* [**6**].

In this study RBF networks are used to demonstrate these concepts by successfully modelling the dynamical

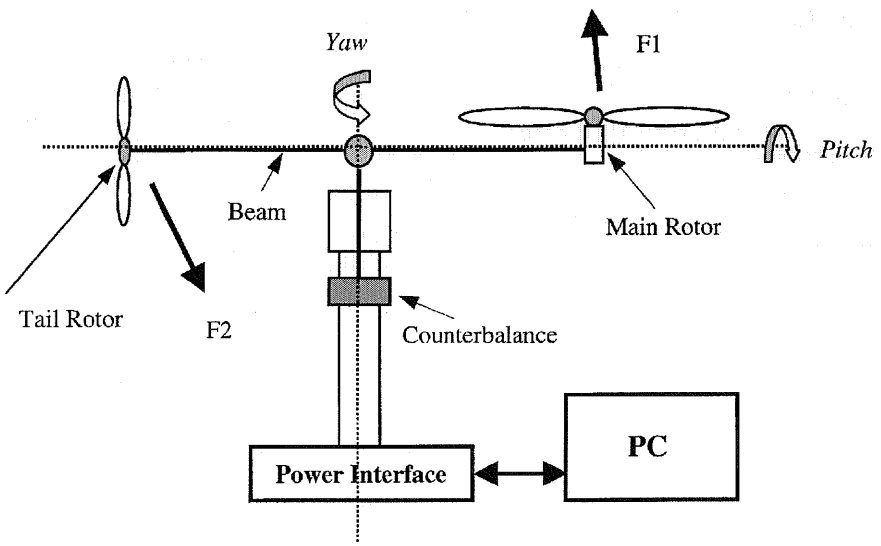


Fig. 1 The twin rotor multi-input multi-output system

behaviour of a 1 degree-of-freedom (DOF) TRMS. Such a high-fidelity non-linear model is often required for the non-linear flight simulation studies. Since there is no reliance on the mathematical model, the estimated RBF model has to be thoroughly verified using rigorous time and frequency domain tests. If the model structure and the estimated parameters are correct then the residuals (the difference between the model and system output) should be unpredictable from all linear and non-linear combinations of past inputs and outputs. This is ensured by carrying out higher-order cross-correlation tests, proposed by Billings and Voon [7].

The paper first describes the TRMS system in Section 2 and the non-linear modelling approach adopted in Section 3. This is followed by a discussion of RBF networks in Section 4. System experimentation, the type of excitation signal and data pre-processing needed to identify the non-linear model are outlined in Section 5. Implementation and results are presented in Section 6. Finally, the main findings of this study are summarized in Section 7.

2 THE TRMS SYSTEM

The TRMS considered in this work is described in Fig. 1. This consists of a beam pivoted on its base in such a way that it can rotate freely in both its horizontal and vertical planes. There are rotors (the main and tail rotors), driven by d.c. motors, at both ends of the beam. A counterbalance arm with a weight at its end is fixed to the beam at the pivot. The state of the beam is described by four process variables: yaw and pitch angles measured by position sensors fitted at the pivot, and two corresponding angular velocities. Two additional state variables are the angular velocities of the rotors, measured by tachogenerators coupled with the driving d.c. motors.

In a typical helicopter, the aerodynamic force is controlled by changing the angle of attack of the blades. The laboratory set-up is constructed such that the angle of attack of the blades is fixed. The aerodynamic force is controlled by varying the speed of the motors. Therefore, the control inputs are supply voltages of the d.c. motors. A change in the voltage value results in a change of the rotational speed of the propeller, which results in a change of the corresponding angle (in radians) of the beam. F1 and F2 in Fig. 1 represent the thrust generated by the rotors in the vertical and horizontal planes respectively.

3 NON-LINEAR MODELLING

There are a number of different types of non-linear models that are potentially suited to this problem. Some examples are the output-affine model, the polynomial model and the rational model. In this investigation, a non-linear autoregressive model with exogenous inputs (NARX) [8], which provides a concise representation for a wide class of non-linear systems, is employed. The model is of the form:

$$y(t) = f\Big(y(t-1),\ldots,y(t-n_y),u(t-1),\ldots,u(t-n_u)\Big) + e(t) \quad (1)$$

where $y(t)$ is the output, $u(t)$ is the input and $e(t)$ accounts for uncertainties, possible noise, unmodelled dynamics, etc., n_y, n_u are the maximum lags in the output and the input respectively, $\{e(t)\}$ is assumed to be a zero mean white noise sequence and $f(\cdot)$ is some vector-valued non-linear function of $y(t)$ and $u(t)$ respectively. The NARX model is also referred to in the literature by various other names, such as one-step-ahead predictor or series–parallel model. Because the system noise $e(t)$ is generally unobserved, it can only be replaced by the prediction error or residual $\varepsilon(t)$, and equation (1) can be rewritten as

$$y(t) = f\Big(y(t-1),\ldots,y(t-n_y),u(t-1),\ldots,u(t-n_u)\Big) + \varepsilon(t) \quad (2)$$

where the residual is defined as

$$\varepsilon(t) = y(t) - \hat{y}(t) \quad (3)$$

with $\hat{y}(t)$ representing the model predicted output.

Two considerations are of practical importance for the application of the NARX approach. Firstly, the non-linear functional form $f(\cdot)$ should be capable of describing the non-linear input–output space. Secondly, an efficient identification procedure for selecting a parsimonious model structure is required. The present study employs an RBF network to model the input-output relationship. This is depicted in Fig. 2. The non-linearity within the RBF can be selected from a small set of typical non-linear functions, such as the thin-plate–spline function, the Gaussian function, the multiquadratic and the inverse multiquadratic functions. A generally held opinion is that the choice of the non-linearity is not crucial for performance [9]. The non-linear functional form $f(\cdot)$ in the RBF expansion used in this study is the Gaussian function. Orthogonal least squares (OLS) [9] provides an elegant method for determination of model parameters. If the OLS is employed with the polynomial NARX model, it selects a parsimonious model structure as well as estimates the selected model parameters. However, if the NARX–RBF model structure is adopted then the OLS routine yields optimal model parameters, i.e. weights and centres.

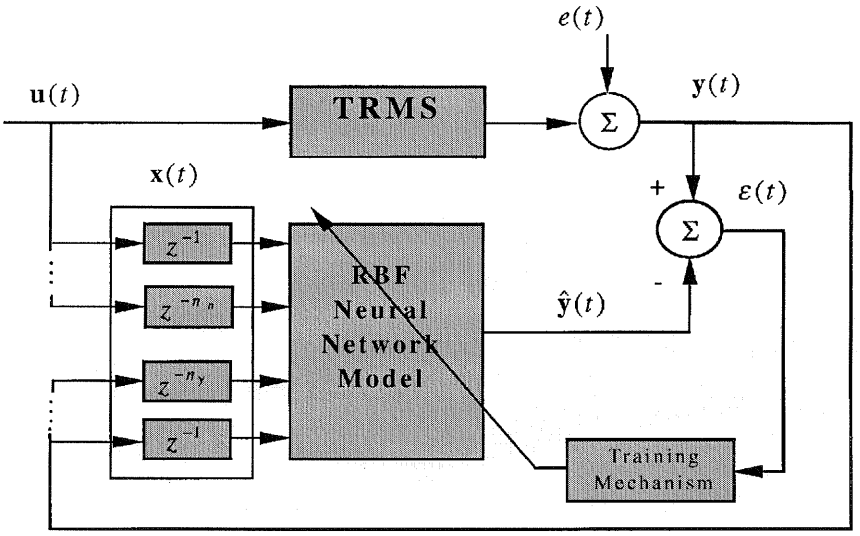


Fig. 2 NARX model identification with RBF networks

4 RADIAL BASIS FUNCTION

An RBF network can be regarded as a special two-layer network which is linear in the parameters provided that all the RBF centres are prefixed. Given fixed centres, i.e. no adjustable parameters, the first layer or the hidden layer performs a fixed non-linear transformation, which maps the input space on to a new space. The output layer then implements a linear combiner on this new space and the only adjustable parameters are the weights of this linear combiner. These parameters can therefore be determined using the linear least squares method, which is an important advantage of this approach.

A schematic of the RBF network with n inputs and a scalar output is shown in Fig. 3. Such a network can be represented as

$$\hat{y}(t) = w_0 + \sum_{i=1}^{n_r} w_i f_r\Big(\|\mathrm{x}(t) - c_i\|\Big) \tag{4}$$

where $\hat{y}(t)$ is the network predicted output, $\mathrm{x}(t)$ is the input vector containing all regressors of equation (1), i.e.

$$\mathrm{x}(t) = \Big[(y(t-1,\ldots,y(t-n_y), u(t-1),\ldots,u(t-n_u)\Big]^{\mathrm{T}} \tag{5}$$

where w_i are the weights or parameters, w_0 is the bias or the d.c. level term at the output, c_i are known as RBF centres and n_r represents the number of centres or the hidden neurons. Once the functional form $f(\cdot)$ and the centres c_i are fixed, and the set of input $\mathrm{x}(t)$ and the corresponding desired output vector [$y(t)$ in this study] are provided, the weights w_i can be determined using the linear least squares method. Clearly, $\hat{y}(t)$ is the non-linear model predicted output determined by the past values of the system output vector $y(t)$, and the input vector $u(t)$ with maximum lags n_y and n_u respectively.

The Gaussian form for the RBF, $f_i(x)$, is

$$f_i\Big(x_1(t), x_2(t), \ldots, x_n(t)\Big) = \exp\left\{-\frac{[x_1(t)-c_{1i}]^2 + [x_2(t)-c_{2i}]^2 + \ldots + [x_n(t)-c_{ni}]^2}{\beta_i^2}\right\} \tag{6}$$

where $\mathrm{c}_i = (c_{1i}, c_{2i}, \ldots, c_{ni})$ is a vector that defines the centre of the RBF f_i in neuron i and β_i^2 is the 'shape' of the function or the spread constant. Input patterns x activate the nodes according to their distance $\|\mathrm{x} - c_i\|$ from the node centres c_i. Thus, each hidden neuron responds only to inputs that are in a region (the *receptive field*) around its centres c_i. Other functions can also be used as the activation functions of radial basis nodes, without significantly affecting the performance of the RBF network. The scalar output $\hat{y}(t)$ is the sum of a linear combination of the RBF outputs, $f_i(x)$, with the weights w_i of the connections from the hidden to the

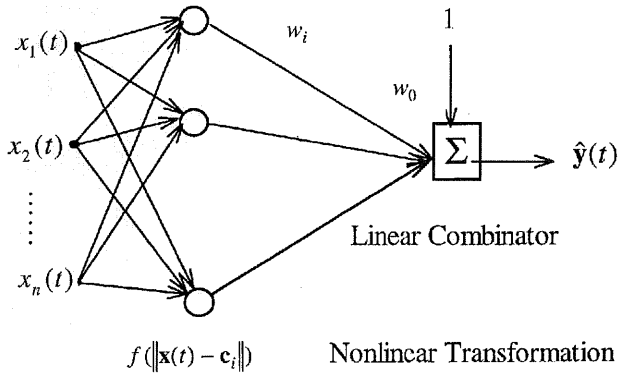


Fig. 3 Radial basis function network

output nodes:

$$\hat{y}(t) = \sum_{i=1}^{n_r} w_i f_i\Big(x_1(t), x_2(t), \ldots, x_n(t)\Big) \tag{7}$$

or

$$\hat{y}(t) = \sum_{i=1}^{n_r} w_i f_i\Big(x(t)\Big) \tag{8}$$

This discussion can be best understood by assuming the RBF network in equation (4) to be a special case of the linear regression model:

$$y(t) = \sum_{i=1}^{M} p_i(t)\theta_i + \varepsilon(t) \tag{9}$$

where $y(t)$ is the desired output and p_i are known regressors, which are non-linear functions of lagged outputs and inputs, i.e.

$$p_i(t) = p_i\Big(\mathrm{x}(t)\Big) \tag{10}$$

with $\mathrm{x}(t)$ as defined in equation (5). A constant term (w_0 in Fig. 3) can be included in equation (9) by setting the corresponding term $p_i(t) = 1$. The residual $\varepsilon(t)$ is assumed to be uncorrelated with the regressors $p_i(t)$. It is clear that a given centre c_i with a given non-linear function $f(\cdot)$ corresponds to $p_i(t)$ in equation (9).

Equation (9) for $t = 1, \ldots, N$ can be written in matrix form as

$$\mathrm{y} = \mathbf{P}\boldsymbol{\Theta} + \mathrm{E} \tag{11}$$

and the parameter vector $\boldsymbol{\Theta}$ satisfying this equation is given by the well-known least squares (LS) method, provided the centres are fixed.

4.1 RBF–NN learning algorithms

The task of a learning algorithm, or an optimization routine, in an RBF network is to select the centres and to find a set of weights that makes the network perform the desired mapping. In essence, the objective is to minimize the variance or the sum squared of the residual:

$$\hat{\sigma}_{\varepsilon^2} = \sum_{t=1}^{N} \varepsilon^2(t) \tag{12}$$

A number of algorithms are frequently utilized for this purpose [**10**], for instance:

(a) random centre selection and a least squares algorithm,
(b) clustering and a least squares algorithm,
(c) non-linear optimization of all the parameters, i.e. centres, output weights and other free parameters,
(d) the orthogonal least squares (OLS) algorithm.

Among these the OLS algorithm is widely used. The OLS method proposed by Chen *et al.* [**9**] yields both the number of centres c_i, i.e. significant regressors, as well as the corresponding parameter vector $\boldsymbol{\Theta}$. The underlying idea of the algorithm is to transfer the regression equation into an equivalent orthogonal form. Then the RBF centres can be selected and the weights optimized in a simple procedure according to a criterion referred to as the 'error-reduction ratio' (ERR), due to the orthogonality property. Details of the OLS algorithm can be found in Chen *et al.* [**9**].

5 EXPERIMENTATION

The objective of the identification experiments is to estimate a suitable model of the 1 DOF TRMS in hover without any *prior* system knowledge pertaining to the exact mathematical model order. The extracted model is to be utilized for low-frequency vibration control and design of a suitable feedback control law for disturbance rejection and reference tracking. Hence, accurate identification of the rigid-body dynamics is imperative. This would also facilitate understanding of the dominant modes of the TRMS. Since no mathematical model is available, a level of confidence has to be established in the identified model through rigorous frequency and time-domain analyses and cross-validation tests.

It is intuitively assumed that the body resonance modes of the TRMS lie in a low-frequency range of 0–1 Hz, while the main rotor dynamics are at significantly higher frequencies. The rig configuration is such that it permits open-loop system identification, unlike a helicopter, which is open-loop unstable in hover mode. Trim configuration was in a steady state horizontal position of the beam of the TRMS. The system is interfaced through a PC and it is possible to send and record signals through a Matlab/Simulink environment. The duration of the test signal generally should be slightly higher than the system settling time [**11**]. A duration of 60 s was deemed fit for this study, which is higher than the settling time of the TRMS, which is about 15–20 s.

5.1 Excitation signal

In non-linear system identification, the type of input signal to be used plays a crucial role and has a direct bearing on the fidelity of the resulting identified model. The excitation signal should have two important characteristics [**12**]:

1. It should be able to excite all the dynamic modes of interest; i.e. the spectral content of the input signal should be rich in frequency corresponding to system

bandwidth. Such a signal is referred to as persistently exciting.
2. It should be rich in amplitude level, i.e. have different levels of input amplitudes over the whole range of operation.

These two requirements can generally be fulfilled by selecting an input such as a sine wave, Gaussian signal, independent uniformly distributed process or ternary pseudo-random sequence [7].

In order to excite the system modes of interest, i.e. up to 1 Hz, two different signals, (a) an independent uniformly distributed signal (noise) and (b) a pseudo-random binary sequence (PRBS) of 2 and 5 Hz band limit respectively, are employed in this study. Figure 4 shows these two signals along with their amplitude distribution.

5.2 Data pre-processing

Processing of the raw input–output data obtained from the experiments is recommended for system identification. Pre-processing could involve removal of outliers, stray data points and normalization. In the case of identifying a system model using NN, it is advantageous to apply pre-processing transformations to the input data before presenting it to the network. Reducing the difference of magnitude of input variables used to train the network leads to faster convergence. One of the common methods of pre-processing is linear rescaling of the input variables. The normalized data are obtained by carrying out the following data manipulation:

$$\tilde{x}_i^n = \frac{x_i^n - \bar{x}_i}{\sigma_i^2} \tag{13}$$

where $\bar{x}_i$ is the mean and σ_i^2 is the variance of each variable of the training set, defined as

$$\bar{x}_i = \frac{1}{N}\sum_{n=1}^{N} x_i^n \tag{14}$$

$$\sigma_i^2 = \frac{1}{N-1}\sum_{n=1}^{N} (x_i^n - \bar{x}_i)^2 \tag{15}$$

where $n = 1, \ldots, N$ represents the number of data points or the data length. The rescaled variables defined by $\tilde{x}_i^n$ have zero mean and unit standard deviation. The target values are also subjected to similar linear rescaling.

6 IMPLEMENTATION AND RESULTS

In this section the results of modelling the TRMS with neural networks are described. Modelling with the NN was carried out with the TRMS pitch response to a uniformly distributed noise signal, as described in the previous section. The rationale of using the noise signal is that the two-level PRBS signal may not be good enough to capture non-linearities, if present, in the system. For the sake of comparison the two-level PRBS input shown in Fig. 4b is also utilized for modelling the 1 DOF TRMS. Results obtained with the main rotor input and the pitch output are described below.

Selection of the model structure, *estimation* of parameters and *verification* are the fundamental issues in the system identification exercise. Since the RBF is chosen as the model structure, the remaining two issues of estimation and verification are addressed in this section.

6.1 Mode determination

In order to detect the dominant system modes, spectral plots of the TRMS output and model output are analysed. The solid line curve in Fig. 5 shows the power spectral density (PSD) plot of the actual pitch response of the TRMS to the independent uniformly distributed input signal of 2 Hz bandwidth. As noted, this shows closely spaced modes between 0 and 1 Hz as expected, with a main resonant mode at 0.34 Hz, which can be attributed to the main body dynamics. A model order of 2, 4 or 6 corresponding to prominent normal modes at 0.25, 0.34 and 0.46 Hz is thus anticipated.

The next step is to capture or model the plant dynamics using an RBF network. The Matlab neural network toolbox [13] is utilized to carry out the *parameter estimation*, which uses an OLS learning algorithm. An iterative procedure can be devised to identify the NARX model using the RBF expansion by linking the OLS routine and the model validity tests. The non-linear function in the RBF expansion is the Gaussian function. The main steps in the identification can be summarized as follows:

1. Choose n_y and n_u. Initially the set of candidates centres are all

$$\mathrm{x}(t) = \left[(y(t-1), \ldots, y(t-n_y), u(t-1), \ldots, u(t-n_u)\right]^{\mathrm{T}}$$

2. Select the Gaussian spread constant β_i and define the error goal.
3. An iterative loop is then entered to update the model based on the 'error-reduction ratio' (ERR) criteria [9].
4. Different time- and frequency-domain validity tests are performed to assess the model. If the model is good enough the procedure is terminated. Otherwise go to step (1).

The OLS learning method selects a suitable set of centres c_i (regressors) from a large set of candidates as well as estimates of the linear parameters w_i, or the weights. The iterative procedure described above is used

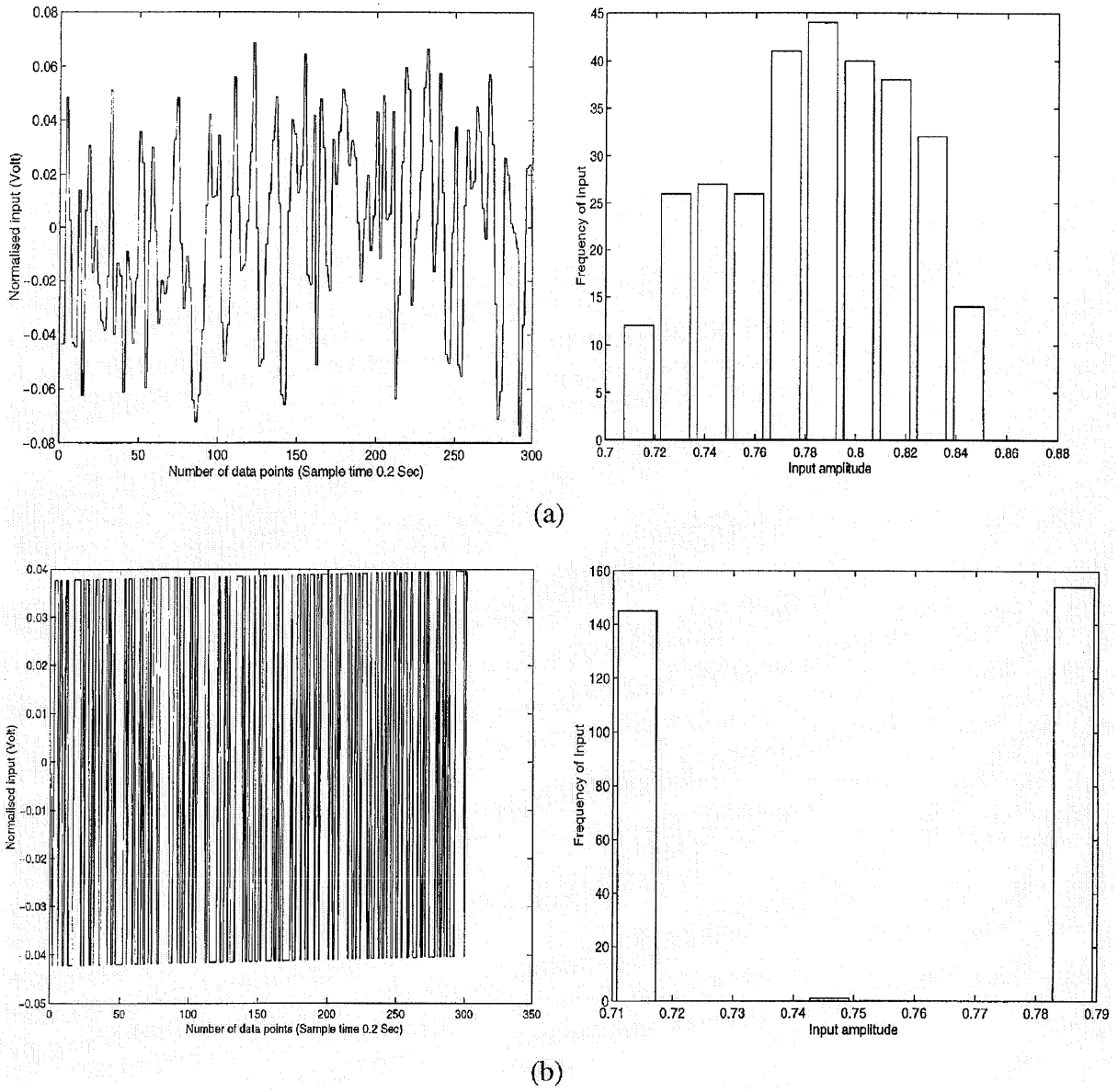


Fig. 4 Excitation signals: (a) noise, (b) PRBS

to identify the RBF model. The RBF model was trained with 300 data points and different combinations of input–output lags were tried. Using a generate-and-test method, an eighth-order NARX model was found to give a better representation of the system dynamics in the frequency domain (see Fig. 5) than the sixth-order model as envisaged. Experience has shown that among all the tunable parameters, a judicious choice of n_y and n_u is central in identifying a reasonable model. This model reached a sum-squared error level of 0.002 after 13 training passes. The identified model included a constant term and 13 centres or neurons. The PSD obtained from the RBF model and the experimental data are superimposed in Fig. 5. It is observed that the dominant modes of the model and the plant coincide quite well, implying good model predicting capability of the important system dynamics. Thus, it is assumed that the identified model is fairly accurate and suitable for system analyses.

6.2 Correlation tests

In the previous section the frequency domain test was employed to detect the system modes. In order to ensure

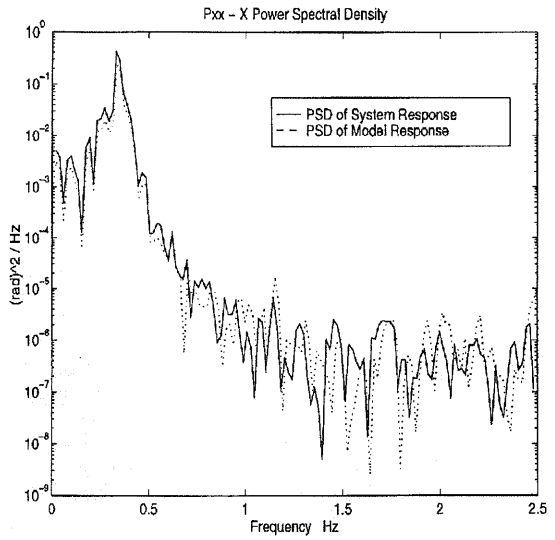


Fig. 5 Power spectral density of output

further confidence in the identified model, time-domain correlation tests are employed next.

A more convincing method of model validation is to use correlation tests. If the model of a system is adequate then the residuals or prediction errors $\varepsilon(t)$ should be unpredictable from (uncorrelated with) all linear and non-linear combinations of past inputs and outputs. This can be tested by means of the following correlation functions [7]:

$$\begin{aligned}
&\phi_{\varepsilon\varepsilon}(\tau) = E[\varepsilon(t-\tau)\varepsilon(t)] = \delta(\tau) \\
&\phi_{u\varepsilon}(\tau) = E[u(t-\tau)\varepsilon(t)] = 0 \qquad \forall\tau \\
&\phi_{u^2\varepsilon}(\tau) = E\left[\left(u^2(t-\tau) - \overline{u}^2(t)\right)\varepsilon(t)\right] = 0 \qquad \forall\tau \\
&\phi_{u^2\varepsilon^2}(\tau) = E\left[\left(u^2(t-\tau) - \overline{u}^2(t)\right)\varepsilon^2(t)\right] = 0 \qquad \forall\tau \\
&\phi_{\varepsilon(\varepsilon u)}(\tau) = E\left[\varepsilon(t)\varepsilon(-1-\tau t)u(t-1-\tau)\right] = 0 \qquad \tau \geqslant 0
\end{aligned} \tag{16}$$

where $\phi_{u\varepsilon}(\tau)$ indicates the cross-correlation function between $u(t)$ and $\varepsilon(t)$, $\varepsilon u(t) = \varepsilon(t+1)u(t+1)$, and $\delta(\tau)$ is an impulse function.

The first two linear correlation tests in equation (16) alone are not sufficient to validate non-linear models. Hence, higher-order correlation tests are included in this study. All five tests defined by equation (16) should be satisfied if the $u(\cdot)$ and $y(\cdot)$ values are used as network input nodes. In practice, normalized correlations are computed. In general, if the correlation functions in equation (16) are within the 95 per cent confidence intervals, $\pm 1.96/\sqrt{N}$, where N is the total number of data points, the model is regarded as satisfactory.

Figure 6 shows the correlation tests described by equation (16). It is important to note that only the first few lags are significant. The lags in the x axis of Fig. 6 are equivalent to the sampling period; i.e. each lag (τ) is equivalent to 0.2 s. The y axis of each plot in Fig. 6 is given by the corresponding correlation function of equation (16).

All the results of the correlation tests, as shown in Fig. 6, are within the 95 per cent confidence interval, indicating a high level of approximation of the actual data set. The model validity tests thus corroborate the fact that the estimated model is adequate. Having accomplished the first two tasks of structure determination and parameter estimation, the final important step is model verification.

6.3 Verification

Verification in the time domain is a key final step. In this process, the predictive quality of the identified model is assessed with data that were not used for modelling. The uniformly distributed noise signal was used for estimation whereas multi-step (3211) and doublet input were used for validation. These are shown as Figs 7 and 8. In non-linear system identification using neural networks, generally one-step ahead (OSA) prediction and model predicted output (MPO) are employed for cross-validating the estimated model. Here, the results of MPO are presented, which is a more robust test and often more difficult to achieve than the OSA prediction. This is expressed as

$$\hat{y}_{\mathrm{d}}(t) = f\Big(u(t), u(t-1), \ldots, u(t-n_u),\ \hat{y}_{\mathrm{d}}(t-1), \ldots, \hat{y}_{\mathrm{d}}(t-n_y)\Big) \tag{17}$$

In Figs 9 and 10 the simulated non-linear MPO and the experimental outputs are compared for the 3211 and the doublet excitation respectively. It is observed that the model and the system response match closely. A slight mismatch between the MPO and the actual system response is expected in any modelling process as the models are only a close 'approximation' of the real system. However, overall the predictive capability of the model is quite good, especially considering the very sensitive nature of the TRMS to ambient disturbances. This has been a major problem in consistently reproducing the same response to an input. An analogous procedure was repeated with the PRBS signal and the corresponding result is shown in Fig. 11.

Comparing the MPO due to noise and PRBS inputs, it is clearly noted that the model obtained with the noise signal has captured the dynamics better than with the PRBS. This is primarily due to the excitation of system dynamics across the input range, unlike the PRBS where only two levels of amplitude are present in the input, thereby making it unable to excite the non-linear dynamics associated with the other input amplitudes. Although PRBS is routinely

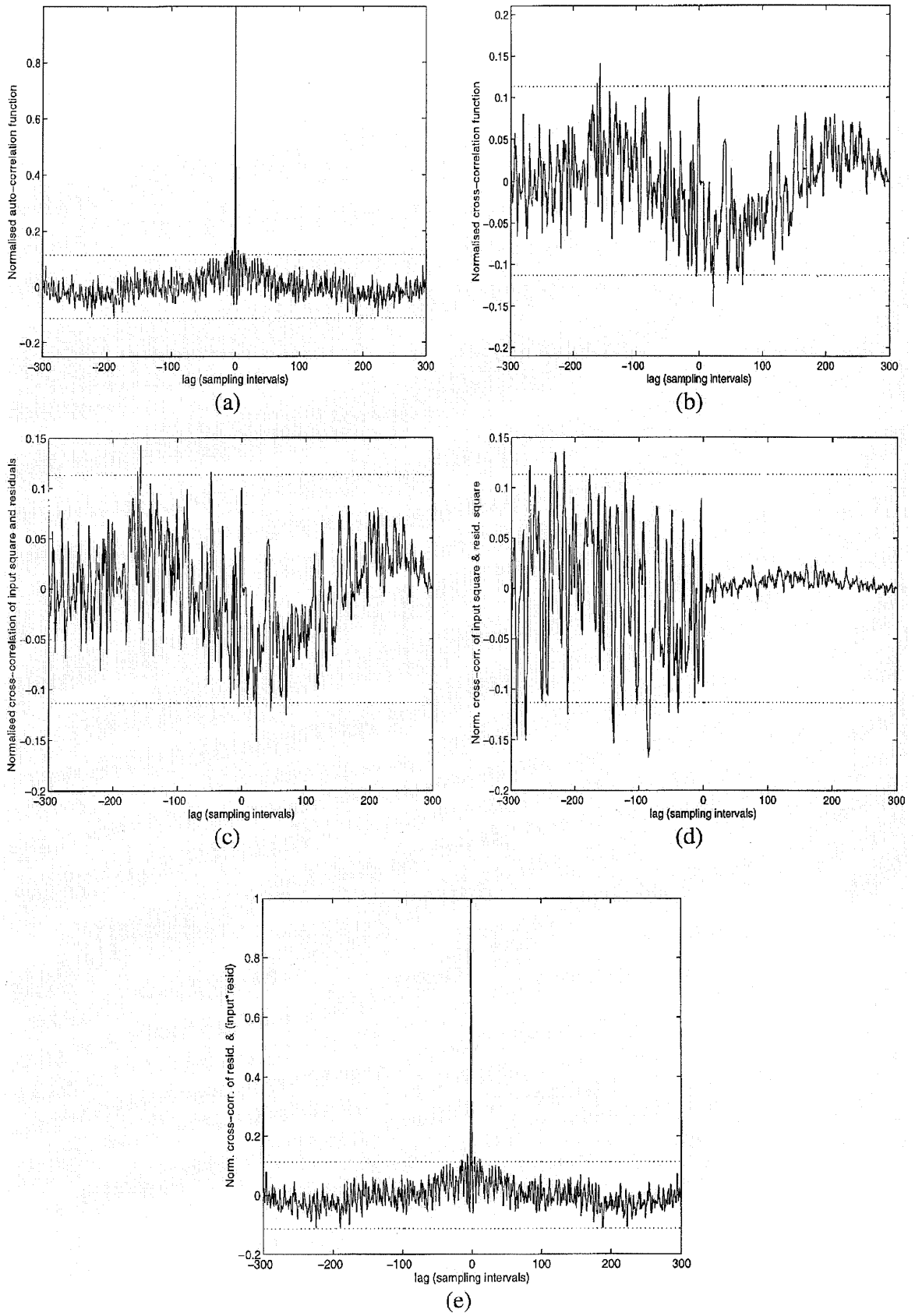
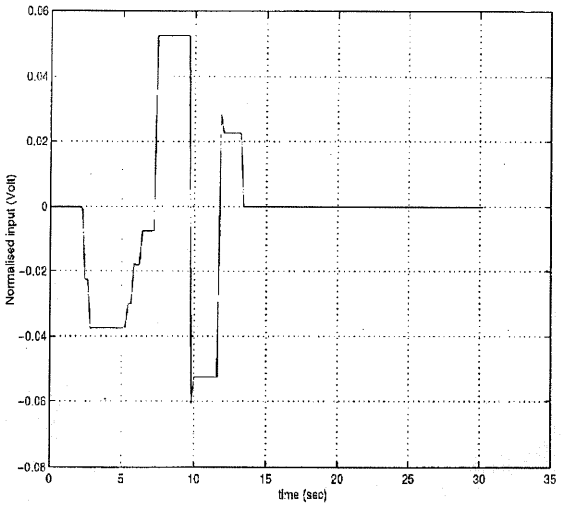
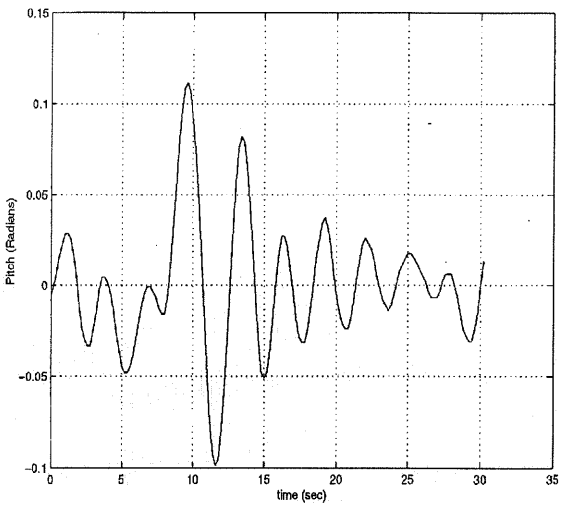


Fig. 6 Correlation tests: (a) $\phi_{\varepsilon\varepsilon}(\tau)$, (b) $\phi_{u\varepsilon}(\tau)$, (c) $\phi_{u^2\varepsilon}(\tau)$, (d) $\phi_{u^2\varepsilon^2}(\tau)$, (e) $\phi_{\varepsilon(\varepsilon u)}(\tau)$. The dashed line shows the 95 per cent confidence interval

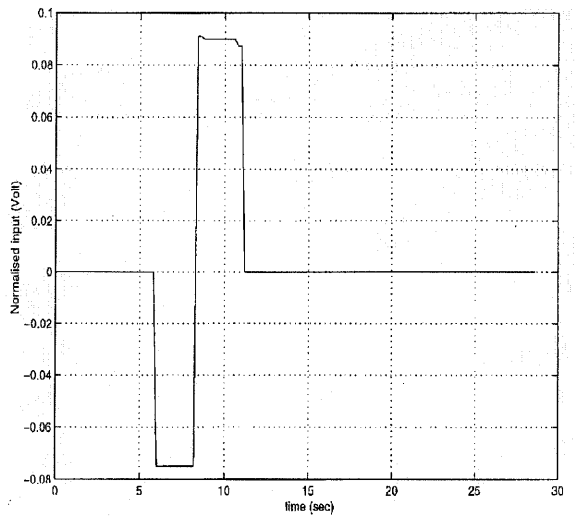


(a) Multi-step input (3211).

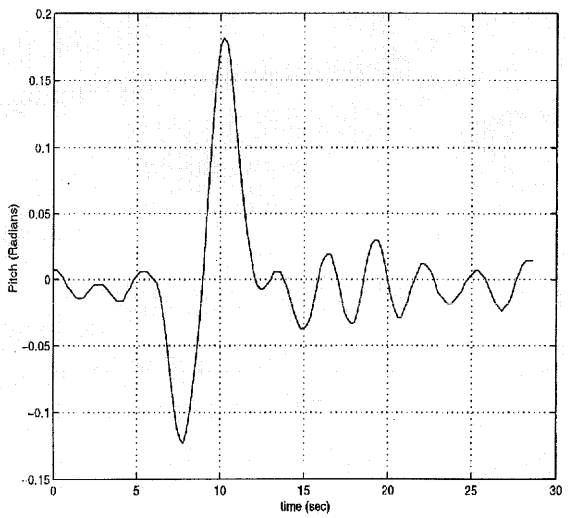


(b) System response to a multi-step input.

Fig. 7 Input and output signals used for model cross-validation (3211)



(a) Doublet input.



(b) System response to a doublet.

Fig. 8 Input and output signals used for model cross-validation (doublet)

used for linear system identification, it is therefore not suitable for the non-linear case. The model validity tests given in Sections 6.1 and 6.2 and the time-domain test illustrated in Figs 9 and 10 confirm that the RBF network obtained through the independent uniformly distributed signal is an adequate model of the TRMS.

7 CONCLUDING REMARKS

Radial basis function networks are shown to be suitable for modelling complex engineering systems in cases where the dynamics are not well understood or are not simple to establish from first principles, such as the next generation UAVs.

Careful selection of the excitation signal(s) is an important part of a system identification process. Without due consideration to this issue, the obtained model would not be able to capture the system dynamics, resulting in a poor model. Since no mathematical model is available, extensive model validation is imperative. This has been ensured by carrying out higher-order cross-correlation tests and MPO analysis. The extracted model has predicted the system behaviour well. Such a high-fidelity non-linear model is often

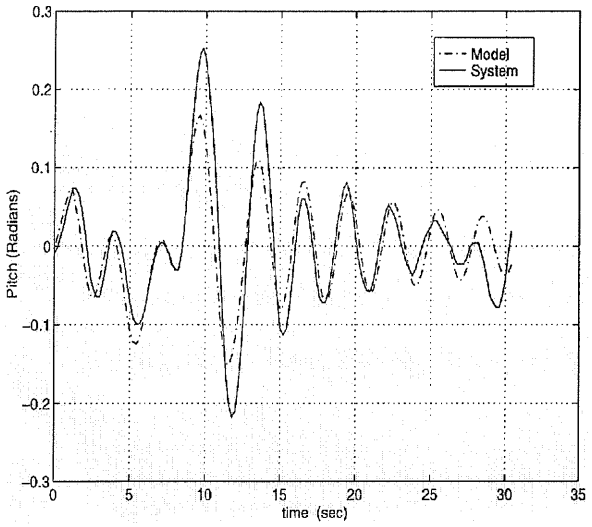


Fig. 9 The system and the non-linear model responses to a 3211 (noise signal used for estimation)

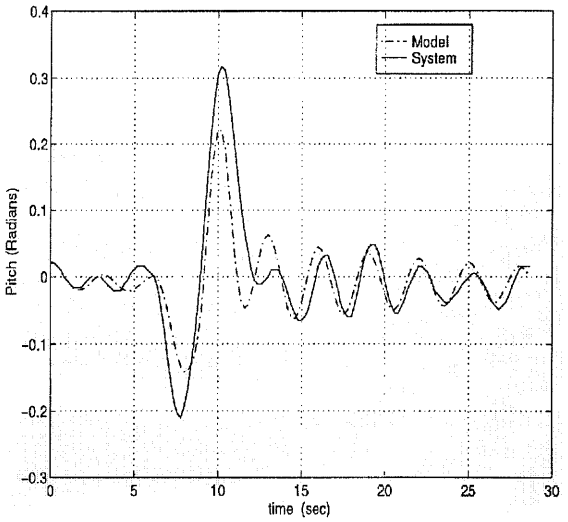


Fig. 10 The system and the non-linear model responses to a doublet (noise signal used for estimation)

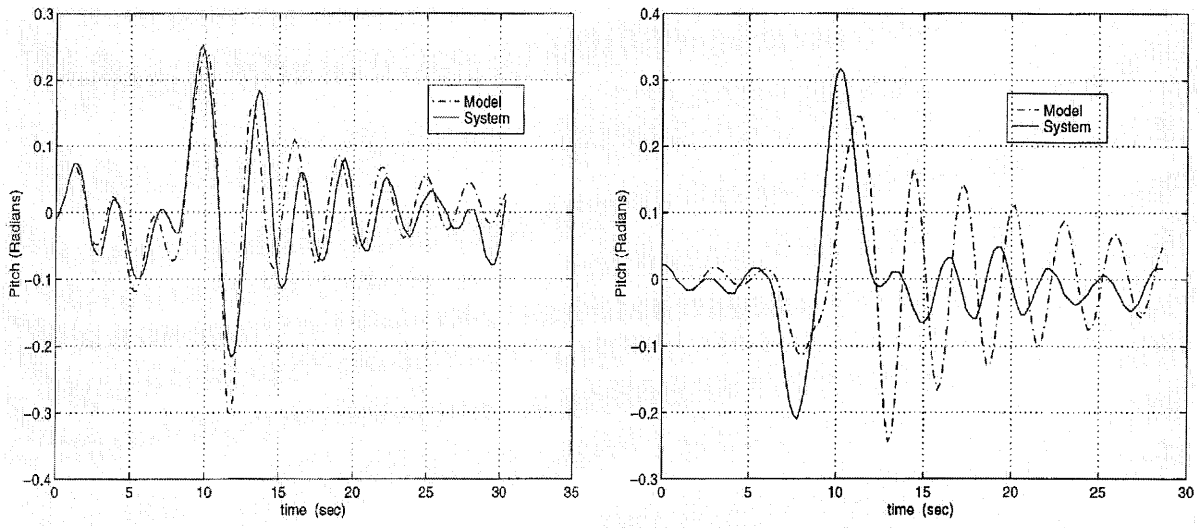


Fig. 11 The system and the non-linear model responses (PRBS used for estimation)

required for gauging the performance of control design and system analysis.

The technique discussed here thus provides a fast interim solution to model next generation air vehicles such as UAVs, X-wing, tilt body and delta-wing, whose flight dynamics are not well understood or are not easy to establish from first principles. If more rigorous analytical models become available, they can then be used to fine-tune the general solutions, if they prove to be more accurate.

ACKNOWLEDGEMENT

The authors are grateful to Professor S. A. Billings for his valuable comments and suggestions throughout this work.

REFERENCES

1 *Twin Rotor MIMO System, Manual 33-007-0*, 1996 (Feedback Instruments Limited, Sussex).
2 **Blythe, P. W.** and **Chamitoff, G.** Estimation of aircraft's aerodynamic coefficients using recurrent neural networks. In Proceedings of Second Pacific International Conference on *Aerospace Science and Technology*, Australia, 1995.
3 **Kim, B. S.** and **Calise, A. J.** Nonlinear flight control using neural networks. *J. Guidance, Control, and Dynamics*, 1998, **20**(1), 26–33.
4 **Lyshevski, S. E.** Identification of nonlinear flight dynamics: theory and practice. *IEEE Trans. on Aerospace and Electronics Systems*, 2000, **36**(2), 383–392.
5 **Bruce, P. D.** and **Kellet, M. G.** Modelling and identification of non-linear aerodynamic functions using B-splines. *Proc. Instn Mech. Engrs, Part G: J. Aerospace Engineering*, 2000, **214**(G1), 27–40.
6 **Ahmad, S. M., Shaheed, M. H., Chipperfield, A. J.** and **Tokhi, M. O.** Nonlinear modelling of a twin rotor MIMO system using radial basis function networks. In Proceedings of IEEE National Aerospace and Electronics Conference (*NAECON 2000*), Dayton, Ohio, 10–12 October 2000, pp. 313–320.
7 **Billings, S. A.** and **Voon, W. S. F.** Correlation based validity tests for nonlinear models. *Int. J. Control*, 1986, **44**(1), 235–244.
8 **Leontaritis, I. J.** and **Billings, S. A.** Input–output parametric models for nonlinear systems. Part 1: deterministic nonlinear systems. *Int. J. Control*, 1985, **41**(2), 303–328.
9 **Chen, S., Cowan, C. F. N.** and **Grant, P. M.** Orthogonal least squares learning algorithm for radial basis function networks. *IEEE Trans on Neural Networks*, 1991, **2**(2), 302–309.
10 **Bishop, C. M.** *Neural Networks for Pattern Recognition*, 1995 (Clarendon Press, Oxford).
11 **Soderstrom, T.** and **Stoica, P.** *System Identification*, 1989 (Prentice-Hall, Englewood Cliffs, New Jersey).
12 **Billings, S. A.** and **Chen, S.** Extended model set, global data and threshold model identification of severely non-linear systems. *Int. J. Control*, 1989, **50**(5), 1897–1923.
13 **Hagan, M. T., Demuth, H. B.** and **Beale, M.** *The Neural Networks Toolbox*, 1993 (MathWorks Inc., Natick, Massachusetts).

