Monograph:

Boaghe, O.M., Billings, S.A., Li, L.M. et al. (2 more authors) (2000) Time and Frequency Domain Identification and Analysis of a Gas Turbine Engine. Research Report. ACSE Research Report 778 . Department of Automatic Control and Systems Engineering

Time and Frequency Domain Identification and Analysis of a Gas Turbine Engine

O.M. Boaghe, S.A. Billings, L.M. Li, P.J. Fleming, J. Liu

Department of Automatic Control and Systems Engineering
University of Sheffield
Mappin Street, Sheffield S1 3JD
United Kingdom

Research Report No.778

September, 2000

University of Sheffield

Time and Frequency Domain Identification and Analysis of a Gas Turbine Engine

O.M. Boaghe, S.A. Billings, L.M. Li, P.J. Fleming, J. Liu

Department of Automatic Control and Systems Engineering, University of Sheffield, PO.Box-600, Mappin Street, S1 3JD, UK

Abstract

The NARMAX (Nonlinear AutoRegressive Moving Average model with eXogenous inputs) approach is used to analyse the dynamics of a gas turbine engine. The fuel flow - shaft speed relationship is analysed by identifying both time and frequency domain models of the system. The frequency domain analysis is studied by mapping the discrete-time NARMAX models into the Generalised Frequency Response Functions (GFRF's) to reveal the nonlinear coupling between the various input spectral components and the energy transfer mechanisms in the system. A continuous-time nonlinear differential equation model is also estimated using the Generalised Frequency Response Functions.

1 Introduction

Gas turbine engines are employed in aircraft and ship propulsion systems and as a consequence engine design is a critical issue in aircraft and ship performance. While designs based on linear models are adequate in certain circumstances, new techniques for the analysis and design of gas turbine engines have recently been employed [Godfrey and Moore, 1974] to accommodate nonlinear effects. As a consequence gas turbine engine modelling has received special attention in the last few years in the literature devoted to nonlinear system identification and analysis.

The first nonlinear models to be proposed were based on physical principles applied in the time domain. These models are also known as thermodynamic models. The continuous-time models, which are complex and nonlinear, were usually numerically linearised about a set of operating points [Jackson, 1988]. Following the increasing computer power and advances in systems theory, a new approach was considered more recently based on system identification procedures in the time and frequency domain.

The identification methods which were employed in the frequency domain were often based on the identification of linear s-domain models. Such models were extracted based on multisine testing, introduced and developed by Evans et al [1995, 1998, 1999] in which the linear Frequency Response Functions (FRF) were derived and used to compare and validate the models. The methods employed in the time domain were usually based on black-box discrete-time input-output identification methods. Time-varying models were estimated using extended least squares and optimal smoothing in Norton [1975] and Evans et al [2000]. Multi-objective genetic programming was also applied and discrete-time NARMAX models were identified in Rodriguez-Vasquez and Fleming [1998], Evans et al [2000] and Chiras et al, [2000].

While the previous approaches concentrate on either time-domain or frequency-domain techniques for identification, modelling and interpretation, the present paper is based on a combined approach and uses both the time and frequency domain. It is shown that new insights into the gas turbine dynamics can be obtained from studying both time and frequency domain properties of the system. While time-domain discrete NARMAX modelling has been adopted before in Rodriguez-Vasquez and Fleming [1998] and Chiras et al [2000], the frequency-domain approach based on Generalised Frequency Response Functions is entirely new in gas turbine engine analysis. The nonlinear Generalised Frequency Response Functions (GFRF), which are generalisations of the linear Frequency Response Functions (FRF) are shown to be crucial in system interpretation and modelling. This approach also enables the construction of a continuous time, nonlinear differential equation model.

The paper concentrates on the fuel flow - high pressure (HP) shaft speed relationship, measured on a Spey engine at DERA Pyestock. After a brief presentation of the NARMAX approach and frequency domain concepts, the paper presents the identification and analysis of gas turbine models in both the time and the frequency domain. Finally conclusions are drawn about the application of this approach.

2 Time and frequency domain identification techniques

NARMAX (Nonlinear AutoRegressive Moving Average model with eXogenous inputs) is a well known methodology used in nonlinear system modelling and identification. NARMAX procedures which have been developed over the past twenty years have a wide area of application, from real system identification to the analysis of nonlinear differential equations with strong nonlinearities. The NARMAX model, proposed by Leontaritis and Billings [1985a,b], is given as

$$y(k) = F[y(k-1), \ldots, y(k-n_y), u(k-1), \ldots, u(k-n_u), \xi(k-1), \ldots, \xi(k-n_\xi)] + \xi(k) \tag{1}$$

where $F[\cdot]$ denotes a nonlinear function, u and y are the discrete-time input and output signals with corresponding maximum lagged values of n_u and n_y. The quantity $\xi(k)$ accounts for possible noise and uncertainties. n_ξ represents the maximum noise lag. The nonlinear function F can be a polynomial, rational function, radial basis functions, wavelet decomposition or any other function subject to some mild constraints.

The identification of the NARMAX model structure and parameters can be performed using an orthogonal least-squares algorithm, where the candidate model terms are quantified according to the contribution that they make to the variance of the system output. The best model is then selected as the model which explains the total output variance and for which the number of terms is a minimum. This means that the model terms are selected according to their significance to the output variance. The model is then further validated both statistically and dynamically. The statistical validation involves higher order correlations which test if the residuals are unpredictable from all past values of the input and output [Billings and Zhu, 1994]. The dynamical validation is based on model predictions.

The time domain NARMAX model can also be mapped into the frequency-domain where Generalised Frequency Response Functions (GFRF), which are generalisations of the linear Frequency Response Function (FRF), are used to describe the system. The translation into the frequency domain is based on an algorithm developed by Peyton-Jones and Billings [1989]. The linear Frequency Response Function is defined as the Fourier transform of the impulse response function $h_1(\tau)$ in the convolution integral

$$y(t) = \int_0^\infty h_1(\tau)u(t-\tau)d\tau$$

Analogous to the linear case, the Generalised Frequency Response Functions are multi-dimensional Fourier transforms of the nonlinear impulse response functions $h_n(\tau_1, \ldots, \tau_n)$ in the nonlinear convolution integral, known as the Volterra series

$$y(t) = \sum_{i=0}^{\infty} \int_0^\infty h_i(\tau_1, \ldots, \tau_i)u(t-\tau_1)\ldots u(t-\tau_i)d\tau_1 \ldots d\tau_i \tag{2}$$

It is straightforward to note that the first order frequency response function

$H_1(\omega)$ explains the linear effects, while the nonlinear functions $H_n(\omega_1, \ldots, \omega_n)$ give a measure of the nonlinear coupling of the input spectral components and reveal energy transfer mechanisms to new spectral components in the output.

In the present work the discrete-time NARMAX approach was used to identify features of the nonlinear dynamics of a gas turbine engine. In the next section discrete-time NARMAX models are first identified for the fuel flow and the high-pressure (HP) shaft speed data sets. The discrete-time models are then mapped into the frequency domain, where the Generalised Frequency Response Functions reveal in a new way energy transfer phenomena. This approach also allows a continuous-time nonlinear differential equation model to be reconstructed from the frequency domain GFRF's, based on an algorithm developed by Li and Billings [1998].

3 Time domain analysis

The input data analysed in this paper, representing the fuel flow, is given as an IRMLBS (Inverse Repeat Maximum-Length Binary Sequence) signal imposed on a rectangular wave and sampled at 20 msec or 50Hz. The input signal varies within a 10% range of the steady-state fuel flow W_f. The output signal considered is the high pressure HP shaft speed, at three operating point values of 70%, 75% and 80% of the maximum value of the shaft speed NH. Both input and output signals are illustrated in Figure 1.

A delay of 2 samples was detected using an input-output cross-correlation test [Marmarelis and Marmarelis, 1978], corresponding to a 40 msec time delay for the original 20 msec sampling time. This is in good agreement with the time delays detected in Evans et al [1998] of 35 msec and in Evans et al [1999] of 21 msec. For sampling times higher than the one available of 20 msec an even more precise time delay would be expected. The data was decimated 4 times, to a sampling frequency of 12.5 Hz. Normally, the output bandwidth taken into account for the linear FRF is $[0; 1Hz]$ [Evans et al, 1998, 2000], therefore the 12.5 Hz frequency limit considered here should cover all possible nonlinear effects.

In practice nonlinear thermodynamic models of the engine are given as linearised versions around a series of single operating points. The time and frequency domain identification is also usually performed at single operating point modes in order to allow comparison with the linearised thermodynamic models [Evans et al, 1998, 1999] and also because it is usually considered that different operating points

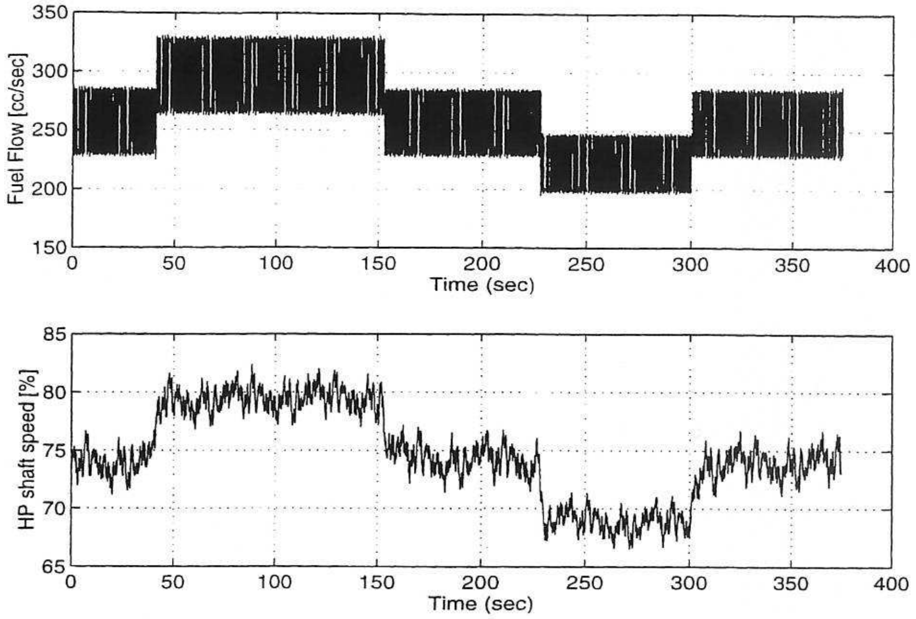

Figure 1: Input (10% range of the steady-state fuel flow W_f) and output (high pressure HP shaft speed) data sets

cannot be represented by a single model, especially for shaft speeds higher than 80% of the maximum value [Evans et al. 1999]. In order to test this final assertion the model identification in this paper was performed both on a single operating point at 75% NH, and also on multiple operating points at 70%-75%-80%.

It has been concluded in the literature, based on multisine testing and analysis that only weak second order nonlinearities are encountered in the gas turbine engine speed response under the influence of not more than 10% of steady-state fuel flow W_f [Evans et al, 1995]. Third order nonlinearities were identified in the case of 25% steady-state fuel flow W_f in Rodriguez-Vasquez and Fleming [1988]. For the case studied in this paper the fuel flow does not exceed 10% of the steady-state value, however the analysis is not restricted to second order nonlinearities, and higher order nonlinear model terms are considered.

A previous nonlinear NARMAX model which satisfied statistical and dynamical validation tests was identified using a multiobjective genetic algorithm by Evans et al [2000] and using an orthogonal least-squares algorithm by Chiras et al [2000], for the case of a small perturbation of 10% around the steady-state fuel flow W_f. The model terms reported by both Evans et al [2000] and Chiras et al [2000] included only $[y(k-1); y(k-2); y(k-1)^2; u(k-1); u(k-2)]$. This structure was tested in our analysis and was also compared with new model structures identified for higher order nonlinearities.

For the discrete-time varying model estimated by Evans et al [2000] using extended least-squares with optimal smoothing the sample means were removed to avoid the need for an offset term. The sample means were also removed for the identification procedure presented in Chiras et al [2000]. In order to test the effect of mean levels. in this paper the model identification was carried out with and without removing the means.

Table 1: Models identified for 10% steady-state fuel flow and 75% or 70%-75%-80% HP shaft speed

Model term	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
$const$	-0.0247	-0.0267	-0.0314	-1.4748	0	-0.0325
$y(k-1)$	+1.1145	+1.0310	+1.0246	+1.1895	+0.6335	+0.8728
$y(k-2)$	-0.1842	-0.1062	-0.0935	-0.1496	+0.3187	+0.0504
$u(k-1)$	+0.0008	+0.0009	+0.0008	+0.0007	0	+0.0007
$u(k-2)$	+0.0089	+0.0089	+0.0086	+0.0086	+0.0228	+0.0085
$y(k-1)^2$	-0.0016	0	-0.0008	-0.0007	0	-0.0014
$u(k-3)$	0	+0.0010	0	0	+0.0056	+0.0024
$y(k-2)^2$	0	-0.0015	0	0	0	0
$y(k-2)u(k-2)$	0	+0.0001	0	0	0	0
$y(k-1)^3$	0	0	0	0	0	-0.0002
$y(k-1)u(k-2)$	0	0	0	0	-0.0002	0
$u(k-1)u(k-2)$	0	0	0	0	4×10^{-6}	0

Six different models given in Table 1 were identified for different constraints relating to the number of terms, structure and identification data set. For all models the statistical correlation tests [Billings and Zhu, 1994] were inside the 95% confidence bands. The dynamical validation was based on model predictions. The model predicted output for these models over the entire data set, and on the test signals represented in Figure 2, was quantified with the Normalised Mean Square Error (NMSE) and the results are presented in Table 2.

$$NMSE = \sqrt{\frac{\sum(\hat{y}(k) - y(k))^2}{\sum(y(k) - \bar{y}(k))^2}} \qquad (3)$$

where $\hat{y}(k)$ is the model predicted shaft speed computed by simulating one of the models presented in Table 1, $y(k)$ is the measured shaft speed and $\bar{y}(k)$ is the mean value of the measured shaft speed.

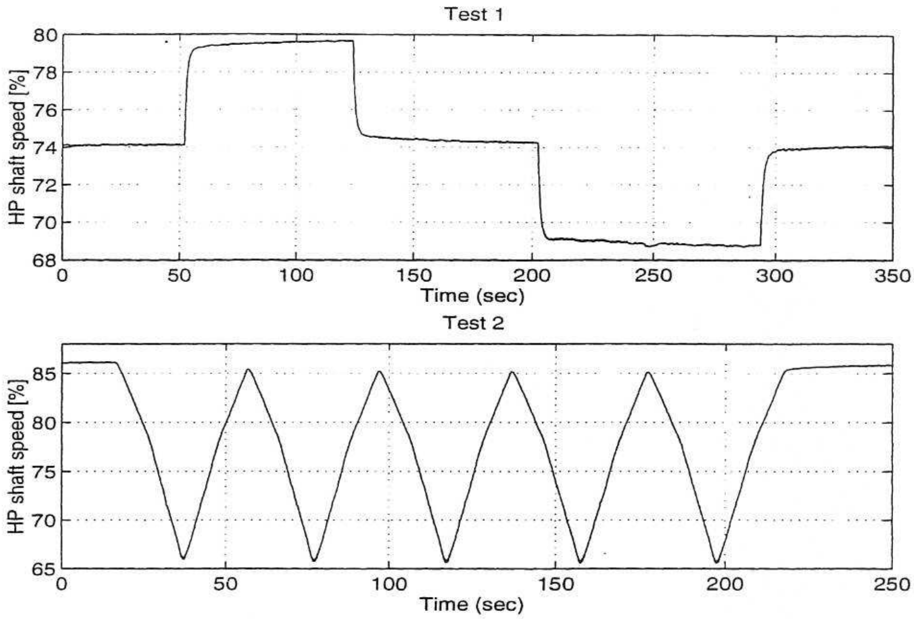

Figure 2: Test signals for the discrete-time models

The identification constraints applied to identify the six models are summarised in Table 2. The first two models Model 1 and Model 2 were identified for a single operating point at 75% HP shaft speed, after removing the means from both the input and the output data sets. In model 1 the structure was constrained to be identical with the model identified in Evans et al [2000]. The structure in Model 2 was not constrained and the term selection algorithm was used to determine the most significant model terms. The remaining models were all identified on the entire data set, for a 70%-75%-80% HP shaft speed. Models 3 and 4 were constrained to a structure identical with the one identified in Evans et al [2000]. For Model 3 the means were removed, while for Model 4 the means were preserved. A new structure was obtained for the same data set for Models 5 and 6 by searching for the significant model terms, identified for a data set without removing the means in the case of Model 5 and removing the means for Model 6.

If Model 1 and Model 2 are compared in terms of the NMSE values listed in Table 3, Model 2 which does not have a constrained structure performs better for all the tests considered. Note the presence in Model 2 of a few extra-terms added by the unconstrained selection $[u(k-3); y^2(k-2); y(k-2)u(k-2)]$. The NMSE values for the entire data set with 70%-75%-80% shaft speed are the highest for Model 1 and

Table 2: Constraints imposed on the identified models given in Table 1

Property	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
estimation set	75%	75%	entire	entire	entire	entire
mean level	removed	removed	removed	preserved	preserved	removed
structure type	imposed	free	imposed	imposed	free	free

Table 3: NMSE values for the models given in Table 1

Test Type	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
75% HP	0.0456	0.0447	0.1656	0.1772	0.1494	0.1825
70%-75%-80% HP	0.0977	0.0825	0.0645	0.0642	0.0670	0.0662
Test Signal 1	0.0960	0.0600	0.0483	0.0505	0.0451	0.0533
Test Signal 2	0.3768	0.2837	0.2356	0.2457	0.2375	0.1948

Model 2, identified on 75% shaft speed data set. The rest of the models have better NMSE values for the overall data set, however the performance of these models on the 75% shaft speed data set is poorer than in the case of the first two models. The test signals in Figure 2 cover the entire 70%-75%-80% data set, therefore better NMSE values are obtained by the global Models 3,4,5,6. One of the best models in terms of NMSE values is Model 3 which will be further analysed in the next section.

4 Frequency domain analysis

The Generalised Frequency Response Functions can be derived directly from the NARMAX models in Table 1 using the algorithm given in Peyton-Jones and Billings [1989]. The true first order generalised frequency response functions are illustrated in Figure 3 for all six models. The linear GFRF is almost identical for all models, showing a high amplification for low frequency, between -20 dB and -15 dB. Only Models 4 and 5 are both outside these limits, probably as a result of the data mean, which for these two cases was not removed from the estimation set. Both the linear magnitude and phase have been non-parametrically estimated in the literature, using multisine testing. The results from the models 1 to 6 presented here are in very good agreement with linear FRF's estimated for similar sets of data in Evans et al [1998, 1999, 2000] and Chiras et al [2000].

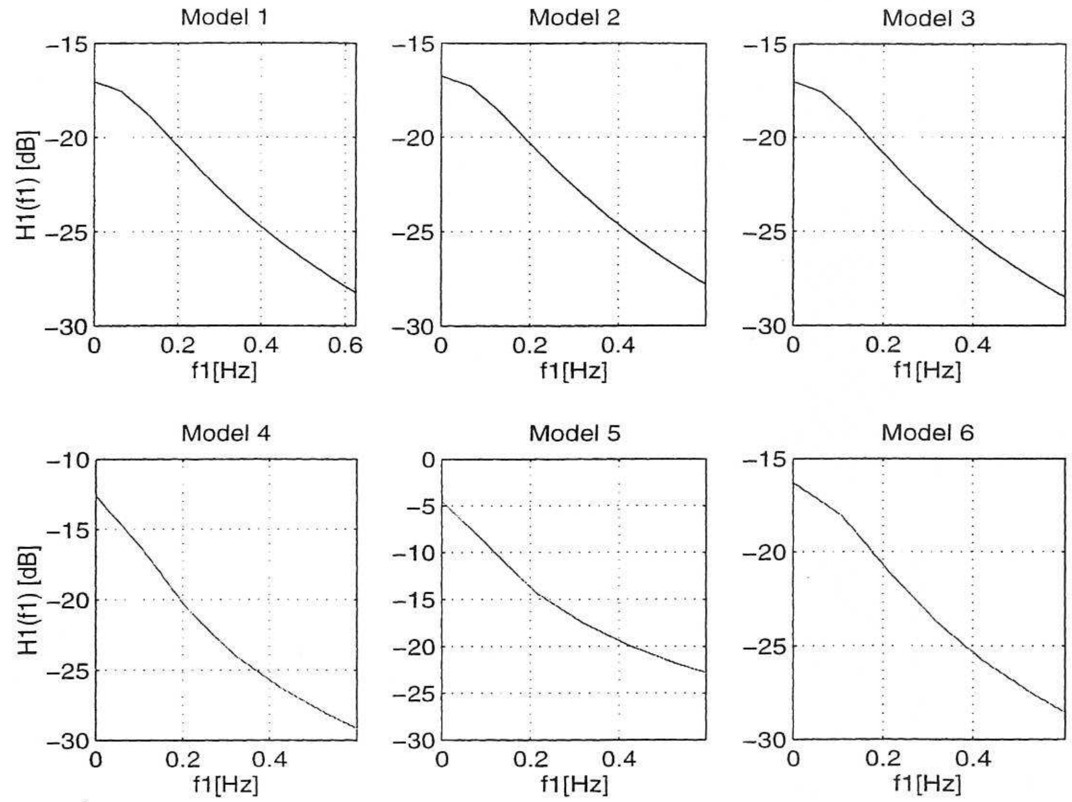

Figure 3: Linear GFRF $H_1(\omega_1)$ for all six models given in Table 1 (note the difference in scale for Model 4 and 5)

All six models also have nonlinear GFRF's, which are generated by the nonlinear terms in the discrete-time models. The second order GFRF's are derived here for the first time and represented in Figure 4. These functions show the same type of low-pass filter characteristic seen for the first order GFRF in Figure 3. The magnitude of the second order $H_2(\omega_1, \omega_2)$ also shows a high magnitude along two lines of frequency, defined by $f_1 = 0$, $f_2 = 0$, $f_1 + f_2 = 0$. This shows that among all possible frequency combinations caused by second order nonlinearities, this particular system will only generate new frequency components for input frequency components satisfying the conditions listed above.

The frequency line $f_1 + f_2 = 0$ can be seen as a storage of energy phenomenon, where energy at the input frequencies f_1 and f_2, which satisfy $f_1 + f_2 = 0$ are transferred to very low, close to zero frequencies. In other words, short time events are transfered to long time events. In contrast, the lines $f_1 = 0$ and $f_2 = 0$ can be

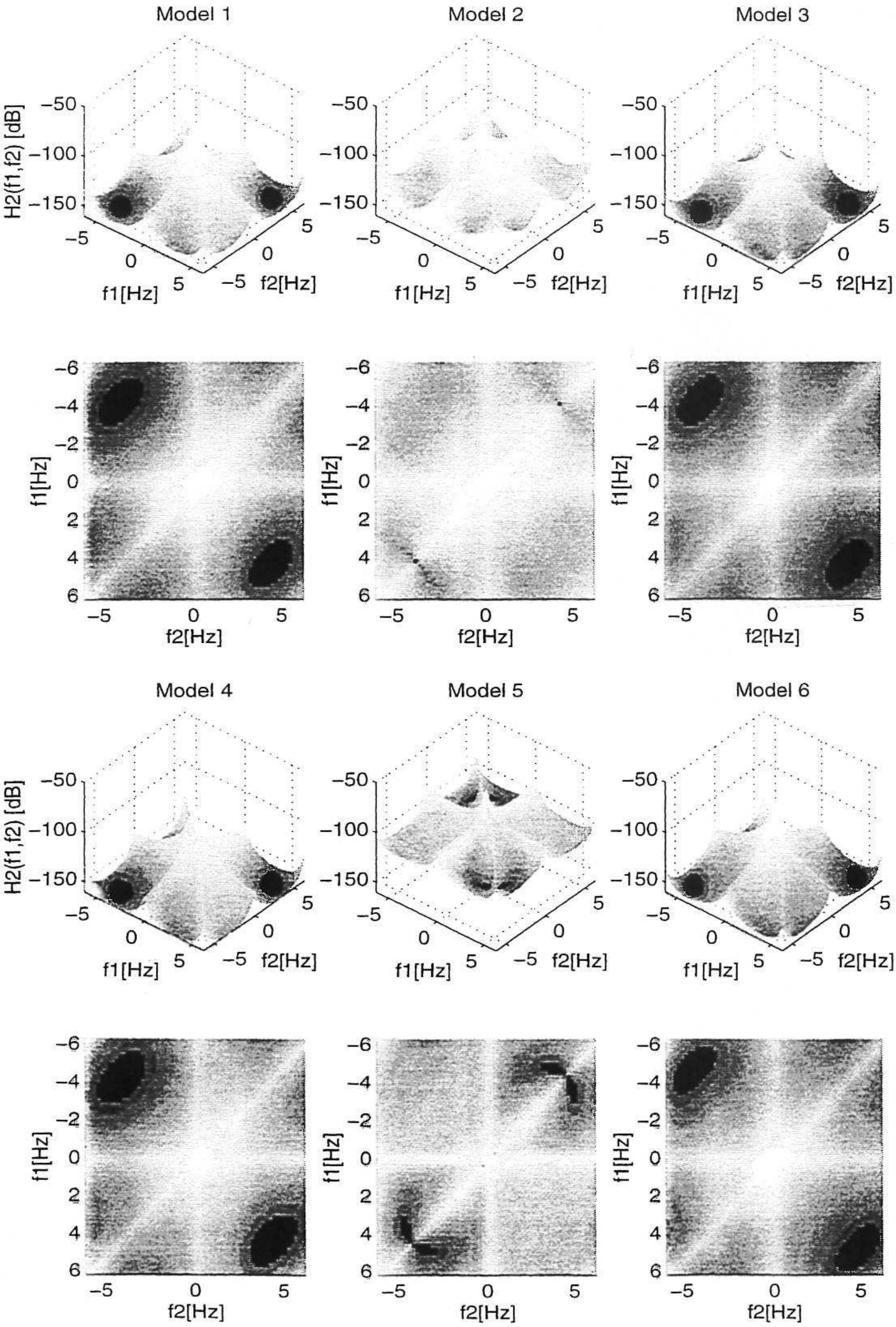

Figure 4: Second order GFRF $H_2(\omega_1, \omega_2)$ for all six models given in Table 1

seen as a release of energy phenomena, in which input frequency components close to zero are amplified by the system and transfered to frequency components in the output other than zero. The low frequency components correspond in this case to long time events, which are transferred in this case by nonlinear effects to higher frequency or longer time events.

A similar phenomenon including the storage and release of energy has been recently observed in the GFRF's corresponding to the geomagnetic activity in the earth atmosphere [Boaghe et al, 1999], where the effects of solar wind are related to the geomagnetic field dynamics and space weather predictions. The nonlinear GFRF's are excellent tools of analysis, they reveal the energy transfer mechanisms of nonlinear systems, and aid the understanding of phenomena evolving in the dynamics of nonlinear systems.

The third order transfer functions $H_3(\omega_1, \omega_2, \omega_3)$ illustrated in Figure 5 show similar types of amplification. The low frequencies are amplified, particularly on the lines of frequency $f_1 = 0$, $f_2 = 0$, $f_3 = 0$, $f_1 + f_2 + f_3 = 0$. All the models have almost identical $H_3(\omega_1, \omega_2, \omega_3)$, only Model 4 and 5 seem to have small differences, again due to the data mean. The interpretation of $H_3(\omega_1, \omega_2, \omega_3)$ is similar to that for $H_2(\omega_1, \omega_2)$, because the frequency lines listed above correspond to the same type of dynamic effects.

It is important to note that even though the six models analysed are different in terms of model structure (especially Model 1 and Model 2), all the models generated almost identical GFRF's. This observation confirms previous studies carried out in relation to GFRF's, in which the frequency domain response functions are seen as invariants of a nonlinear system. Based on this observation, algorithms for extracting the continuous-time model from the GFRF's were developed by Li and Billings [1998]. The continuous-time model corresponding to discrete Model 3 above was identified by using the method of Li and Billings [1998] to give:

$$y(t) = -0.5 + 0.1355u(t) - 2.4117\dot{y}(t) - 0.4661\ddot{y}(t) - 0.0209\dot{u}(t) - 0.0109y^2(t) \quad (4)$$

This model was simulated and the NMSE values recorded. The difference over the measured shaft speed for the discrete-time model at 70%-75%-80% was NMSE =0.0890. The test signals Test 1 and 2 from Figure 2 were reproduced with NMSE values of 0.0874 and 0.2325 respectively, showing an overall good performance. The advantage of the continuous-time model is that this can be easily compared with the thermodynamic models, and the differences can be investigated.

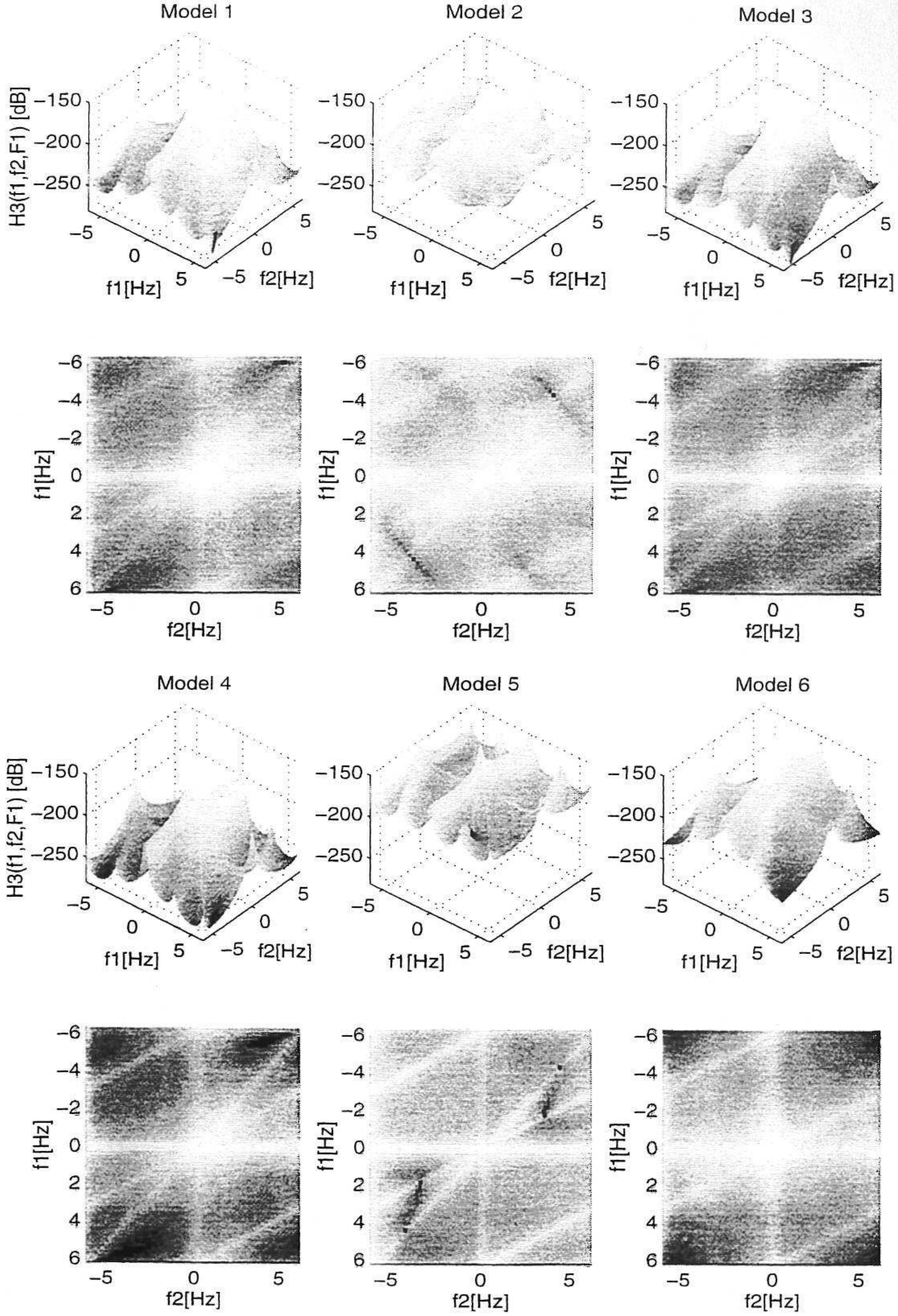

Figure 5: Third order GFRF $H_3(\omega_1, \omega_2, \omega_3)$ for all six models given in Table 1

5 Discussion and Conclusions

A combined time and frequency domain identification approach was considered in this paper to analyse data from a gas turbine engine. The NARMAX methodology has been applied to identify the system with fuel flow as input and HP shaft velocity as output. It has been shown that this system, which is known to be nonlinear for the operating points and fuel level considered, can be modelled with a series of discrete-time NARMAX models, derived for various conditions and constraints.

Although applying the NARMAX approach to this type of system is not new, the novelty of the present results relates to the nonlinear frequency domain analysis performed using the Generalised Frequency Response Functions. The Generalised Frequency Response Functions reveal, for the gas turbine engine application, nonlinear couplings between input harmonic components taking place at low frequency and also on particular lines of frequency. Energy release and energy storage phenomena were detected from the Generalised Frequency Response Functions plots.

As expected the removal of the mean from estimation data set affected the shape of the GFRF's, especially for the linear and the third order GFRF. In order to compare these models with other models derived in the literature the same data pre-processing is required. Indeed it was possible to obtain similar absolute values for the linear FRF's to those published in the literature using the models derived with the mean removed in this paper. Note that only the absolute value of the GFRF magnitude was considered in this paper, and not the phase information.

It can be argued that the level of the fuel flow was too low to excite the nonlinearities. Indeed for a level of 10% of the steady-state fuel flow W_f, multisine testing revealed only weak second order nonlinearities [Evans et al, 2000]. The second order nonlinearity quantified in Figure 4 using GFRF's shows a very small contribution to the system output, almost 100 dB lower than the linear effects. Moreover, the type of input used in the form of an IRMLBS signal was found to reduce the influence of the nonlinearity on the estimated model [Godfrey and Moore, 1974]. However linear models do not satisfy the statistical correlation tests [Rodriguez-Vasquez and Fleming, 1988] on this type of turbine engine, and in this paper a nonlinear model was the only one found to satisfy both the statistical and the dynamical tests. For an increased level of fuel flow, the effect of the nonlinearities on the system would be expected to be stronger [Rodriguez-Vasquez and Fleming, 1988].

The advantage of the continuous-time model (4) is that this can be compared with the continuous-time thermodynamic models. While the thermodynamic models are based on physical principles the discrete-time models are derived from the real system measurements. Therefore a comparison between these models could be very revealing and may lead to a final more representative and more accurate model of the gas turbine engine. Such an analysis is intended in a future work.

6 Acknowledgements

The authors would like to acknowledge the support of DERA Pyestock. O.M. Boaghe gratefully acknowledges financial support from the University of Sheffield and S.A. Billings gratefully acknowledges that part of this work was supported by EPSRC.

References

[1] Billings, S.A. and Zhu. Q.M., 1994, Nonlinear model validation using correlation tests, International Journal of Control, 60, 1107-1120.

[2] Boaghe, O.M., Balikhin, M., Billings, S.A., Alleyne, H.. 1999, Identification of Nonlinear Processes in the Magnetospheric Dynamics and Forecasting of Dst Index, accepted for publication by Journal of Geophysical Research.

[3] Chiras. N., Evans. C., Rees, D., 2000, "Nonlinear gas turbine modelling using NARMAX structure", Control 2000, Univ. of Cambridge, 4-7 September 2000.

[4] Evans, C., Rees, D.. Jones, L., 1995, "Identifying linear models of systems suffering nonlinear distortions, with a gas turbine application". IEE Proc. Control Theory and Applications, Vol.142, No.3, pp.229-240.

[5] Evans, C., Rees. D., Hill, D., 1998, "Frequency domain identification of gas turbine dynamics". IEEE Transactions on Control Systems Technology, Vol.6, No.5, pp.651-662.

[6] Evans. C., Borrell. A., Rees, D., 1999, "Testing and modelling gas turbines using multisine signals and frequency-domain techniques", Journal of Engineering for Gas Turbines and Power, Vol.121, July, pp.451-457.

[7] Evans,C., Fleming,P.J., Hill D.C., Norton,J.P., Pratt,I. Rees,D., Rodriguez-Vasquez,K., 2000. "System applied strategies applied to aircraft gas turbine engines", to appear in Control Engineering Practice.

[8] Godfrey, K.R., Moore, D.J., "Identification of processes having direction dependent responses. with gas-turbine engine applications", Automatica, Vol.10, pp.469-481.

[9] Jackson, D., 1988. "Investigation of state space architectures for engine models", Rolls Royce PLC, REPORT TDR 9331.

[10] Leontaritis, I.J., and Billings. S.A., 1985a, Input-output parametric models for nonlinear systems, Part I: Deterministic nonlinear systems, Int. Journal of Control, 41, 309-328.

[11] Leontaritis, I.J., Billings, S.A.. 1985b, Input-output parametric models for nonlinear systems, Part II: Stochastic nonlinear systems. Int. Journal of Control, 41, 329-344.

[12] Li, L.M., Billings. S.A., 1998. "Continuous time linear and nonlinear system identification in the frequency domain",submitted for publication.

[13] Marmarelis, P.Z.. Marmarelis, V.Z., 1978, "Analysis of physiological systems - The white noise approach", Plenum Press, New York.

[14] Norton. J.P., 1975. "Optimal smoothing in the identification of linear time-varying systems". Proc.IEE Vol.122, No.6, pp.663-668.

[15] Peyton-Jones, J.C., and S.A. Billings, 1989, Recursive algorithm for computing the frequency response of a class of non-linear difference equation models. Int. Journal of Control. 50, 5.

[16] Rodriguez-Vasquez. K., and Fleming, P.J., 1988, "Multi-objective genetic programming for gas turbine engine model identification". UKACC International Conference on Control, Vol.2, No.455, pp.1385-1390.