ABSTRACT
In this paper a feedforward neural network is used to model the fuel flow to shaft speed relationship of a Spey gas turbine engine. The performance of the estimated model is validated against a range of small and large signal engine tests. It is shown that the performance of the estimated models is superior to that of the estimated linear models.

Keywords: gas turbines, system identification, nonlinear models, feedforward neural networks.

NOMENCLATURE

\( u(t) \) System input
\( y(t) \) System output
\( \varphi(t) \) Regressor vector
\( n_u, n_y \) Input, output lags
\( W_f \) Demanded fuel flow
\( N_H \) Maximum High Pressure (HP) shaft speed
\( \theta \) Number of estimation parameters (weights)
\( \hat{y}(t) \) model one-step ahead prediction
\( V \) cost function, sum of square errors
\( N \) Number of input/output data samples

INTRODUCTION
The modelling of gas turbine dynamics has been the subject of considerable study since the early days of jet propulsion. Gas turbines are now extensively used in aero, marine and industrial applications. With such widespread and increasing applications, the accurate modelling of such engines is an important issue.

Modelling of gas turbines is required both in the development and operational stages of an engine’s life. Models can be used to predict the engine’s performance, design and test the performance of the engine control systems and predict normal engine behaviour for on-line fault detection purposes. Models are usually validated against real data and in some cases a physical interpretation of the parameters can be made. This allows initial assumptions about the engine characteristics to be checked.

This paper deals with the nonlinear relationship between the fuel flow and shaft speed dynamics of a Rolls Royce Spey Mk202 aircraft gas turbine. Although no longer in service, the Spey possess the same characteristics, for control purposes as a modern engine such that of the EJ200 fitted to the Eurofighter [1].

Work conducted by Jackson for Rolls Royce plc [2], showed that the higher-order nonlinear thermodynamic models derived from the engine physics could be reduced to linear models of the same order as the number of engine shafts. The models were first linearised at a series of operating points, using small perturbations, and then a model reduction procedure was employed. Work done by Hill [3] examined the application of time-domain methods to estimate discrete \( z \)-domain engine models. Even though linear discrete models with good input-output properties were estimated, problems with the application of time-domain techniques were reported [4].

More recently, teams at the universities of Glamorgan, Birmingham and Sheffield utilised the same data to investigate various identification techniques on a Rolls-Royce Spey engine. Data were gathered under sea-level static conditions at the Defence Evaluation & Research Agency (DERA) at Pyestock. Multisine and inverse repeat maximum length binary sequences (IRMLBS) were used at amplitudes of up to \( \pm 10\% \) of the steady state fuel flow \( (W_f) \). In addition higher-amplitude signals, such as triangular waves and three-level periodic signals with input amplitudes of up to \( \pm 40\% \ W_f \), were used. These caused the HP shaft speed to vary between 65% and 85% of its maximum value \( (\%N_H) \).

Evans et al. [5-7] from the University of Glamorgan, used frequency techniques on the small-amplitude multisine and IRMLBS data to estimate linear models at different operating points. The errors due to noise and nonlinearities were assessed and found to be small for these small-signal models.
The fact that the dc gains and the dynamics of these models change with operating point showed that the gas turbine is nonlinear, so the need is apparent for a more complete gas turbine description using nonlinear models.

In [8] an extended least-squares algorithm with optimal smoothing was used by Norton, to identify time-varying transfer function models to represent large transient and non-equilibrium effects and provide a more detailed insight into the slow thermal dynamics of the engine. In order to identify a model capable of representing the engine at all operating points, Rodriguez [9] used a multiobjective genetic programming approach on the same data and allocated weights to various objectives, to assess their significance in the structure selection of Nonlinear AutoRegressive Moving Average with eXogenous inputs (NARMAX) models of the engine. Chiras et al. [10-11] used nonparametric data analysis in both time- and frequency-domains and an orthogonal estimation algorithm to estimate NARMAX models of the engine. A simple NARX model was identified which was able to represent both the small and large signal dynamics of the engine.

In this paper the nonlinear relationship of the fuel flow to shaft speed dynamics of a Spey gas turbine engine is modeled using a feed-forward neural network. The performance of the estimated model is validated against a range of small and large signal engine tests and it is shown that the performance of the nonlinear model is superior to that of the estimated linear models.

NONLINEAR SYSTEM MODELING USING NEURAL NETWORKS

The identification problem in the time-domain for either linear or nonlinear modeling is to infer relationships between past input-output data and future outputs. If a finite number of past inputs $u(t)$ and outputs $y(t)$ are collected into the vector $\varphi(t)$

$$\varphi(t) = [y(t-1) ... y(t-n_y) \ u(t-1) ... u(t-n_u)]^T$$

then the problem is to understand the relationship $f$ between the next output $y(t)$ and $\varphi(t)$

$$y(t) \rightarrow f(\varphi(t))$$

To obtain this understanding a set of observed data is required which consists of the input $u(t)$ and output $y(t)$, from which the vector $\varphi(t)$ can be built. The function $f$ can be any function and it is indeed this function which defines the model structure. In the case for which $f$ is a linear function several model structures are well documented such as ARX (Autoregressive with eXogenous inputs), ARMAX (AutoRegressive Moving Average with eXogenous inputs), OE (Output Error) and BJ (Box-Jenkins). A mature body of work exists for the estimation of linear models in the time-domain, which is considered a relatively simple task by practitioners and indeed, having to estimate a linear model for a system is considered to be a happy occasion [12].

The case for which $f$ is a nonlinear function presents the most challenging problem for practitioners due to the curse of dimensionality. It is important to note that the problem of obtaining a nonlinear model for a system for which no prior knowledge is available is sometimes a question of "curve-fitting" rather than "modeling", [13]. This justifies the large amount of model structures proposed by authors over the years.

Multilayer perceptron

Feedforward neural networks are proved to have excellent function approximation capabilities [14-16] thus justifying the enormous amount of research dedicated to the subject in recent years. The two most common types of feedforward neural networks are the multi-layer perceptrons (MLP) and the radial-basis functions (RBF) networks.

What is meant by neural networks depends from the author. Neural networks are so fashionable that even old types of models known by other names, have been converted to, or reinvented as neural networks. This makes it difficult to find a universal definition of what a neural network is and even impossible to cover all types of neural networks. A simple definition of neural networks can be formulated as in [17]: “A system of simple processing elements, neurons, that are connected into a network by a set of (synaptic) weights". The function of the network is determined by the structure of the network, the magnitude of the weights and the mode of operation of the processing elements.

The basic neural network element is a neuron shown in Figure 1. This is a processing element that takes a number of inputs, applies some weights and sums them up, and feeds the result to an activation function. The inputs to the unit can be external inputs or outputs of proceeding units. Bias inputs such as $b$ in Figure 1 can also be used to represent a displacement or a constant input. The activation function can be any kind of singular valued function, linear or nonlinear. The most popular activation functions used in system identification are the sigmoid function and the hyperbolic tangent function (tanh) shown in Figure 2. The sigmoid function is given by

$$f_{\text{sigmoid}}(x) = \frac{1}{1 + e^{-x}}$$

and the closely related hyperbolic tangent function is given by

$$f_{\text{tanh}}(x) = 2f_{\text{sigmoid}}(x) - 1$$
so that it makes no difference which of these two functions is used. These functions are nonlinear thus they define the nature of the particular neural network. A feedforward multilayer-perceptron network is constructed by ordering the units into layers, with each layer taking as input other external inputs or outputs of units in previous layers. An example of a two-layer feedforward network is shown in Figure 3 where it can be seen that the second layer produces the output, thus referred to as the output layer. The first layer is called a hidden layer since it is hidden between the external inputs and the output layer. Cybenco [14] proved that a neural network with one hidden layer of sigmoidal or hyperbolic tangent units and an output layer of linear units is capable of approximating any continuous function. The network is described by the magnitude of the weights and biases and should be determined by training the network on the estimation data. The estimation of the weights is usually a conventional estimation problem and several algorithms are available for this purpose.

The structure of a neural network is defined by the choice of regressors in equation (1). Consider the best known linear ARMAX model given by

\[ A(z)y(t) = B(z)u(t) + C(z)e(t) \]  

(5)

In this model \( \phi(t) \) is made up of past output, input and noise terms. Following the same regressor choice it is possible to define the following neural network dynamic models structures

- **NNFIR-models**, which use \( u(t-k) \) as regressors
- **NNARX-models**, which use \( u(t-k) \) and \( y(t-k) \) as regressors
- **NNOE-models**, which use \( u(t-k) \) and \( \hat{y}_u(t-k \mid \theta) \)
- **NNARMAX-models**, which use \( u(t-k), y(t-k) \) and \( \hat{y}(t-k \mid \theta) \)
- **NNBJ-models**, which use all four regressor types.

An example of a NNARMAX model is shown in Figure 4. It must be noted here that this is a recurrent network since the past prediction errors depend on the model output and consequently establish a feedback. The same applies to the NNOE and NBJ networks.

Assuming that the neural network structure is chosen, the next step in procedure is to apply the data set to select the “best” model among the candidates contained in the model structure. This stage is called training, and it involves the minimisation of a fit criterion, the most common choice being the sum of square errors given by

\[ V_N(\theta, Z_N) = \frac{1}{2N} \sum_{t=1}^{N} [y(t) - \hat{y}(t \mid \theta)]^2 = \frac{1}{2N} \sum_{t=1}^{N} e^2(t \mid \theta) \]  

(6)

where \( \theta \) are the model parameters (weights), \( y(t) \) the system output, \( \hat{y}(t) \) the model estimate, \( N \) the number of samples and \( Z_N \) a matrix which contains the system output and the regressor matrix \( \phi(t) \) of equation (2).

\[ Z_N = [y(t) \ \phi(t)] \]  

(7)

known as the estimation or training data. The parameter estimate \( \hat{\theta} \) is thus obtained by

\[ \hat{\theta} = \arg \min \ V_N(\theta, Z_N) \]  

(8)

![Figure 1. Neuron: \( a = f(Wp + b) \).](https://via.placeholder.com/150)

![Figure 2. Commonly used activation functions (a) sigmoid, (b) tanh.](https://via.placeholder.com/150)
An important issue in system identification is the compliance with the principle of parsimony. It would be expected that increasing the size of the neural network model i.e. increasing the number of hidden units and consequently the number of weights will decrease the criterion in (7) thus obtaining a better fit to the estimation data. This does not necessarily mean that the model is able to “generalise” but only that the model is able to adjust of that particular piece of data that it has been trained with. A remarkable solution is provided by Ljung [14] who notes that if the criterion in (7) is evaluated on a validation data set with the same properties as that of the estimation data set then, asymptotically in $N$, the expectation of $\mathbb{E}(N)$ is given by

$$E(V_\theta(Z^N)) = V_\theta(Z^N)(1 + \frac{\text{dim} \theta}{N})$$

(9)

where $Z^N$ is a matrix corresponding to $Z^N$ evaluated on a validation data set, and the notation $\text{dim} \theta$ means the number of estimated parameters. It is thus, clear that by using validation data to evaluate the sum of square errors in equation (7) it is possible to penalise each additional parameter that is added to the model. Consequently, in order to select the size of the neural network model, models of different sizes are trained on the estimation data and the model which minimises the fit for the validation data in equation (9) is selected as the final model.

All estimators and algorithms required for the estimation of a multilayer perceptron feedforward neural network are implemented in the Neural Network System Identification Toolbox for Matlab written by Nørgaard [18]. Additional information for the methodology used to estimate a MLP neural network for system identification can be found in [17-18].

**GAS TURBINE MODELLING**

The selection of an appropriate signal for use for nonlinear identification was considered as an issue of some importance. Schoukens and Dobrowiecki [19] showed that multisine signals with a user-defined amplitude distribution can be designed. The authors suggest that a nonlinear system should be tested with a signal whose amplitude distribution matches as close as possible that of a typical input to the system. This poses a challenge for gas turbine modelling, since a typical input is difficult to define due to the diversity of inputs to the engine.

To this end, a concatenated set of small-signal IRMLBS tests covering an operating range between 55-75% $N_H$ was used for the identification of the nonlinear model to model the dynamic relationship between the fuel flow and shaft speed. A single period of one of these signals, at an operating point of 75% $N_H$...
and input amplitude of ±10% \( W_f \) is shown in Figure 5. The spectra of these signals are shown in Figure 6. It can be seen that the signal-to-noise ratios are very good for this input amplitude, up to 0.6 Hz. The concatenated data used for nonlinear model estimation are shown in Figure 7.

A NNARMAX model was then trained using the data in Figure 7. In order to determine the optimum number of hidden units required to model the engine, the network was first trained using one hidden unit and the model structure was increased gradually with one hidden unit at a time. The performance of the model after each training session was evaluated using the long-term predictions on different validation data sets and calculating the sum of the squares of the prediction errors defined as in equation (6). It is thus possible to select the most appropriate model structure based on the performance of the model on validation data, different to the one used for estimation. Figure 8 shows the variation of this cost function with the number of hidden units. It is clear that a neural network with 8 hidden units provides the best performance on the existing validation data.

![Figure 6. Spectra of the IRMLBS test (a) measured fuel flow (b) HP shaft response.](image)

The quality of the estimated model can be assessed using higher order correlation functions [20], but no definite conclusion can be drawn unless cross-validation is employed. A range of engine tests is available for this purpose which consists of low-amplitude IRMLBS and multisine tests and high-amplitude signal tests such as triangular waves and three-level sequences. It is also possible in this way to compare the performance of the estimated neural network model and the performance of the previously estimated \( s \)-domain linear models [5-7] on the gas turbine data. Figure 9 shows a time-domain comparison between the linear and neural network model outputs and the measured gas turbine output for a small signal multisine test. It can be seen that both the linear and neural network models are capable of modelling the low-amplitude dynamics of the engine. Close inspection shows that the linear model performs slightly better. This can be attributed to the inclusion of the time delay into the model. Similar results are obtained when the performance of the neural network model is compared with the different \( s \)-domain models at different operating points, suggesting that the neural network model is capable of modelling the small-signal engine dynamics throughout the operating range. This allows the single nonlinear model to be used in place of the family of linear models previously estimated.

![Figure 7. Concatenated data set used for estimation (a) measured fuel flow (b) HP shaft response.](image)

![Figure 8. Ten different NNARMAX models showing the sum of squares of the long-term prediction errors on validation data (+).](image)
Figure 9. Output of multisine HP shaft test. (a) complete signal (b) portion of the signal. Measured output (solid), linear s-domain model (dashed), nonlinear model output (dash-dot).

Figure 10 shows a comparison of the performance of a linear model (estimated at 75% $N_{th}$) and the neural network model, on high-amplitude data. It can be seen from this triangular test that the linear model does not follow the engine data at high shaft speeds, whereas the nonlinear model has some problems at the lower shaft speeds. The difference is even more apparent in the next test, shown in Figure 11, where a three-level sequence was used as the input. Here, the linear model is completely unable to capture the high-amplitude dynamics.

A better idea about the quality of the estimated neural network model and its ability to model the engine dynamics throughout the operating range can be obtained if the static and dynamic behaviours are investigated. The static behaviour of the neural network model was derived through simulation and compared with the static behaviour of the engine derived from measured data. As it can be seen from Figure 12 the neural network model is capable of modelling the static behaviour of the engine throughout its operating range. In order to study the dynamic behaviour of the neural network model at various operating points the model was simulated with multisine signals at the four operating points for which linear models were previously estimated. Table 1 shows a comparison of the dominant modes (poles) identified from engine data using small-signal tests and frequency domain techniques, and the dominant modes identified when simulation data were generated using the neural network model. It is clear from the table that the local linearisation of the neural network model is very close to that of the engine at the different operating points.

It can be concluded that the estimated neural network model is capable of modelling both high and low amplitude engine dynamics. The linear models obtained perform very well with small-signal data at different operating points but cannot model the engine response to all large-signal tests. The results can also be extended to the LP shaft case, where a neural network model of similar structure (but different weights) was estimated.
CONCLUSIONS

The nonlinear modelling of a gas turbine using a feedforward neural network was discussed in this paper. A multilayer-perceptron neural network was trained and its performance was illustrated using a range of small and large signal tests and by examining the model’s static and dynamic behaviour. The estimated model was also compared with a previously estimated s-domain model. It was shown that even though linear models are capable of representing the small-signal dynamics of the engine with high accuracy, a nonlinear model is essential to model the high-amplitude dynamics.

This paper illustrates how it is possible to train a neural network to model the fuel flow to shaft speed dynamics of an aircraft gas turbine. The estimation of neural network models avoids the difficult problem of structure selection encountered in NARMAX modelling [8-11] and can simplify difficult mathematical analysis of the engine control system for the purpose of developing real-time predictive control strategies.

ACKNOWLEDGEMENTS

This work was conducted on data gathered at the Defence Evaluation & Research Agency at Pyestock with the support of Rolls Royce plc. The authors would like to thank all the staff involved.

REFERENCES


