Flight parameter based fatigue analysis approach for a fighter aircraft

J. A. Tikka
Patria Aviation,
Halli, Finland

ABSTRACT

This paper describes a flight parameter based fatigue life analysis approach, which is developed for the Finnish Air Force F-18 fighters. It produces a flight specific fatigue life estimate for structural details using flight parameter data stored by each aircraft. Artificial neural networks are used to model structural response of analyzed details as a function of flight parameters. The analysis development is based on strain gauge data from 25 flights of an instrumented aircraft. The results show a satisfactory accuracy for the fatigue life estimates and prove the concept level analysis capability. The average difference between measured and modelled fatigue life is 21% for the fuselage bulkhead and 30% for the leading edge flap's hinge area. The total differences in the Finnish Air Force average usage are extremely small, being –2% for the bulkhead and +2% for the leading edge flap.

1.0 INTRODUCTION

The fighter aircraft structures, which are designed using safe-life principle, are well known for the potential to fatigue life problems during their actual service usage. Depending on the flight sortie profiles and the sortie distribution, significant differences in fatigue life consumption between individual aircraft and structural details exist. To account for this properly, operators should perform a structural detail specific individual aircraft tracking programme.

The detail and tail-number specific analysis requires, however, extensive instrumentation to the aircraft fleet or application of mathematical models between some more general variables and the stress states at critical locations. To avoid the expensive instrumentation of the whole fleet, the mathematical models combining stress states and significant flight parameters sound attractive – especially when the parameters are stored anyway for other purposes. This paper describes an approach, which uses artificial neural networks to combine the structural stress state at critical locations with the stored flight parameters of the Finnish Air Force (FiAF) F-18 fighter aircraft. Because the analysis results are tail-number and structural detail specific fatigue lives, they can be directly used in the definition of structural inspection intervals, evaluation of the need for preemptive modifications and aircraft rotation in the fleet.

2.0 RESEARCH AREA AND SCOPE

The current aim of the work was to prove a concept level capability for the flight parameter based fatigue life analysis of the FiAF F-18. In addition to the development of the analysis approach itself, the concept level work was supposed to provide limited evidence of the expected accuracy. As more data becomes available in the future, more sophisticated model development and validation can then be conducted.

Paper No. 3139. Manuscript received 16 September 2006, revised 29 August 2007, accepted 23 October 2007.

https://doi.org/10.1017/S0001924000002013 Published online by Cambridge University Press

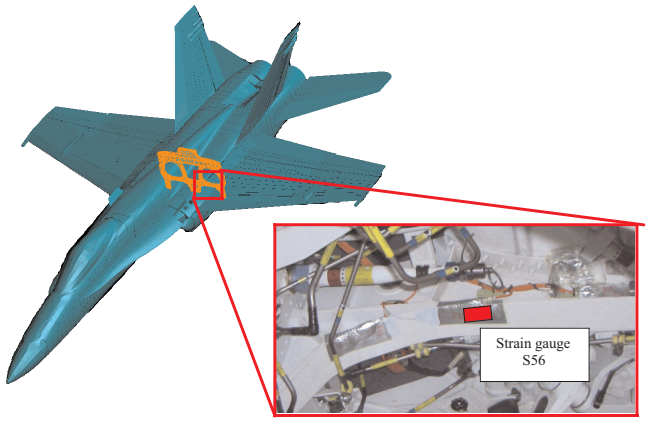


Figure 1. Analysed location in fuselage bulkhead FS488 and strain gauge S56 location.

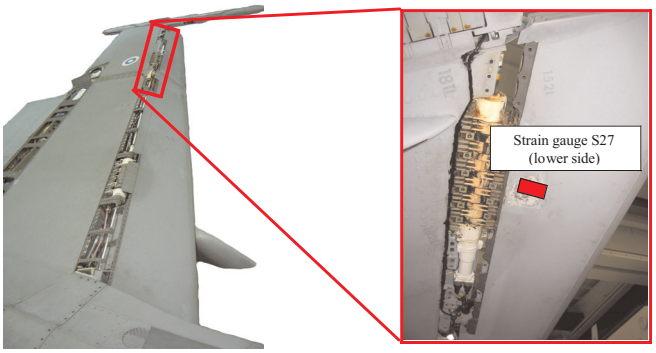


Figure 2. Analysed location in OLEF hinge and strain gauge S27 location.

To fullfil the concept requirements two structural details were selected to which the analysis approach was applied. These were the fuselage bulkhead and outboard leading edge flap (LEF) hinge area. The loads in the fuselage bulkhead (Fig. 1) are mainly caused by the wing root bending moment, being the less demanding of the two selected details. The loads of the LEF hinge area (Fig. 2) are more complicated depending on several flight parameters. For a successful analysis also the flight control system logic has to be taken into account on that area.

3.0 FLIGHT PARAMETER BASED APPROACH FOR FATIGUE LIFE ANALYSIS

A flight parameter based fatigue life analysis can be based on analytical reasoning, which models structural response as a function of usually very few parameters. This has been the case when using, for example, G-exceeding counters and reported sortie masses. After the mid 1990s more data intensive approaches are reported. These rely on measured responses, which are coupled to flight parameters by linear multivariate regression analysis or artificial neural network models[1, 2, 3]. Often a flight manoeuvre classification is applied before feeding flight parameter data to models.

The introduction of a specific flight manoeuvre classification algorithm for each detail, planned to be monitored in the future, would probably have required more work than the resources of this work allowed. Preliminary tests suggested, however, that the structural response inside the whole flight envelope could be modelled at

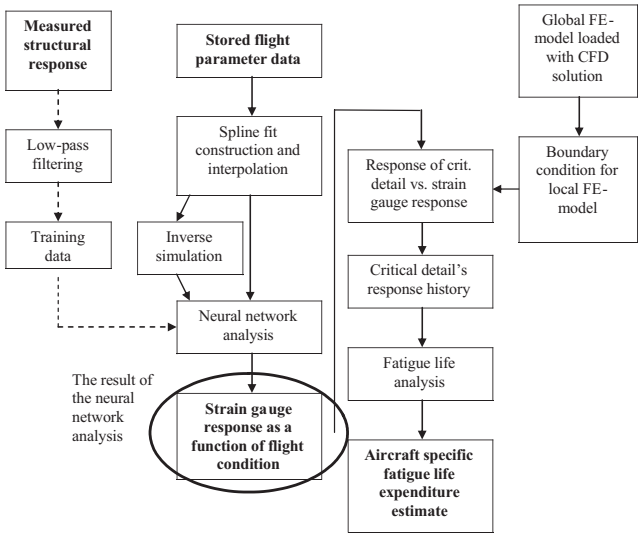


Figure 3. Analysis procedure.

a satisfactory engineering accuracy by a single neural network. For this reason a two-part analysis process was adopted: the first part models operational loads measurement (OLM) aircraft's strain gauge response as a function of flight parameters using preferably a single neural network for each analysed detail and for each flight configuration. The second part of analysis uses a finite element model to solve the stress at the critical detail as a function of strain gauge response, which is used as an input to fatigue life analysis software (Fig. 3). In this paper the latter part is left outside the scope, because it uses well-known standard tools.

3.1 Available input data

In the analysis of the whole F-18 fleet only parameters available from standard aircraft may be used, prohibiting desired but expensive modifications into the aircraft systems. This flight parameter information, which is stored to memory unit (called later as MU data), was enhanced in the analysis development phase by the more extensive instrumentation of an OLM aircraft. During the program its instrumentation included about 60 strain gauges and two bus monitors, which recorded the structural responses and associated avionics multiplexer bus (MUX) flight parameters.

The available OLM data consisted of 28 flights. Fifteen flights were representative of FiAF's typical operational usage. The rest of the flights were test flights specified in detail for structural analysis purposes consisting of separate flight test manoeuvres. The availability of the structural test flights was an advantage, because practically all flight envelope extreme points were present in the data. These extremes improved the network's generalisation capability. The disadvantage in the flight test program was that only a few typical operational flight profiles were repeated, restricting the validation possibilities. Almost every operational flight included flying in the flight envelope areas, which did not exist in other flights and thus, could not be changed from the training data set to the test set without having significant influence on the network training results.

3.1.1 Flight test data from avionics buses

Omitting small differences in data correction and engineering units, the flight parameter data recoded by flight test instrumentation directly from avionics buses is similar to the data stored to MU by standard aircraft. The difference is that MUX data includes several parameters, which are not present in MU data. It is also stored at a

https://doi.org/10.1017/S0001924000002013 Published online by Cambridge University Press

Table 1
Available parameters from standard aircraft instrumentation

Param. name	Used in anal.	Frequency	Resolution & accuracy	Explanation
Parameters directly available in memory unit data				
h_p	X	Satisfactory (1)	Satisfactory, includes systematic errors	Calculated pressure altitude
h		Unusable	Satisfactory	Altitude
V_i	X	Satisfactory	Low, includes systematic errors	Indicated air speed
V_T		Unusable (2)	Not analyzed	True air speed
α	X	Moderate	Moderate	True angle of attack
n_z	X	Moderate (1)	Moderate	Normal acceleration
n_y		Moderate	Unusable	Lateral acceleration
P	X	Low (1)	Satisfactory	Roll rate
Q	X	Low	Satisfactory	Pitch rate
R	X	Moderate	Moderate	Yaw rate
ϕ		Satisfactory for visualization, not for P,Q,R calculation		Pitch angle
θ				Roll angle
HDG				Heading
$F_{s\ lon}$		Low	Moderate	Longitudal stick force
$F_{s\ lat}$		Low	Moderate	Lateral stick force
F_r		Low	Moderate	Rudder pedal force
$PLA_{L\&R}$		Low	Satisfactory	Power level angle
$\delta h_{L\&R}$		Unusable	Not analyzed	Horizontal stabilator position
$\delta a_{L\&R}$		Unusable	Not analyzed	Aileron position
$\delta_{rL\&R}$		Unusable	Not analyzed	Rudder position
$\delta TEF_{L\&R}$		Unusable	Satisfactory	Trailing edge flap position
$\delta LEF_{L\&R}$		Unusable	Satisfactory	Leading edge flap position
Fuel	X	Satisfactory	Satisfactory, includes systematic errors	Fuel amount
LWR, …		Unusable (2)	Not analyzed	Left wing root strain [μ□] & 6 other S/G strains
WPN	X	Satisfactory (3)	Satisfactory	External stores information
LND	X	Satisfactory (3)	Moderate	Landing case description
Parameters calculated from memory unit data				
m	X	Satisfactory	Satisfactory	Aircraft mass
x_{cg}		Satisfactory	Moderate	Aircraft center of gravity
Ma	X	Satisfactory	Satisfactory	Mach number
		Unusable	Moderate	Roll acceleration
		Unusable	Moderate	Angle of attack rate
δLEF_{calc}	X	Satisfactory	Moderate	Leading edge flap position
$C_{h\ OLEF}$	X	Satisfactory	Moderate	Outboard leading edge flap hinge moment
AFU_{mode}	X	Moderate	Moderate	Flight configuration (clean, TO, approach)
(1) Enhanged by SAFE trigger parameters				
(2) Determined by SAFE trigger parameters				
(3) Written when change occurs				

considerably higher frequency than MU data. The extra flight parameters and the higher frequency of MUX data allowed the quality assessment of the MU data and of the interpolated fits required for the 10Hz analysis frequency of the neural networks. In theory the MUX data could be used directly for neural network training. This possibility was abandoned due to the later shown behaviour of the interpolating spline-fits, which were constructed for MU data.

3.1.2 Standard aircraft flight parameter data

Due to the F-18's low sampling rate, which is a result of the available memory capacity, continuous fits are required for the parameter value interpolation to the higher 10Hz frequency used in the rest of the analysis[1]. The MU data is stored at two different constant frequencies, and by one event-triggered process. Basic flight parameter data is stored at a higher, and control surface deflections at a lower frequency. The F-18 OEM fatigue monitoring

https://doi.org/10.1017/S0001924000002013 Published online by Cambridge University Press

system uses seven permanently installed strain gauges, which trigger the event dependent process storing not just the strain gauge responses but also a few flight parameters. Table 1 shows the available MU parameters, comments of data suitability for analysis and the later calculated flight parameters.

3.1.3 Flight test strain gauge data

The strain gauge data of the OLM aircraft was recorded at a high (at least 2500 Hz) frequency, covering all structurally important vibrations[4]. Because at this stage of work only the non-vibrating component of structural response was modelled, data filtering was required.

The separation of the vibrating response component was done by a digital third order Butterworth type low-pass filter run in forward and reverse directions. This led to zero-phase distortion and sixth order filter. Seven Hz was used as a cut-off frequency, defined as a frequency where the magnitude of the response after filtering equals to $\sqrt{1/2}$ of the original. The cut-off frequency selection was based on the control surface actuator performance and the frequency domain response of strain gauge data. The filtered signal included a small amount of lowest normal mode vibrations, but tests with lower cut-off frequencies tended to over-smooth signal extremes. After the filtering the strain gauge data was coordinated in time to MU data and re-sampled at 10Hz for neural network training.

3.2 Limitations to the concept level analysis

Only clean flight configuration was studied at this time, which means that samples where the flight control system was in take-off or power-approach modes were removed from the data. These excluded flight modes were expected be to relatively simple to model due to their limited manoeuvring envelope when compared to the clean configuration. Also the structural stresses on these modes are low and do not consume fatigue life considerably.

The second exclusion was flights which included heavy external loads. From the set of 28 flights three were classified as externally heavy configuration, including at least inboard wing pylons with launch rails for double AMRAAMs, a centre line external fuel tank and wing tip Sidewinders. The data from these three flights was not enough to construct a reasonable size training data set.

The third limitation was the omission of structural vibrations. In the F-18 case vibrations can be dominant for the fatigue lives of tail and outer wing areas but not for the studied details. The vibrating component was removed from the strain gauge signal by low-pass filtering.

3.3 Flight parameter data processing

The fatigue life of a fighter airframe can be assumed to depend on stress (ie. the response of flight parameters) history's turn points, making the calculation of response duration and path between the turn points unnecessary. A turn-point-only analysis could reduce the problem size considerably compared to the analysis of a continuous time history. The problem in turn point only analysis is that the peak responses in different areas of the airframe do not correlate in time. For this reason the analysis was run throughout the whole time history at a constant 10Hz sampling frequency. The analysis frequency was selected high enough for the fastest control surfaces, while simultaneously maintaining the amount of interpolation to an acceptable level for the MU flight parameters.

The Spline-fits, used for MU data interpolation, were in general satisfactory. The splines were constructed using Matlab's spaps function[5]. It minimises the second derivative of the fit within the required error tolerance. The spline f is chosen so that the cost W is minimised in equation:

$$W = \rho \cdot E(f) + F(D^m f) \quad \ldots (1)$$

The ρ in Equation (1) controls the allowable error between the fit and the MU data values and was selected empirically for each parameter by comparing the fit to the MUX data. The error E between data and fit is calculated by equation:

$$E(f) = \sum_{j=1}^{n} \left| y(j) - f(x(j)) \right|^2 \quad \ldots (2)$$

where $y(j)$ is a discrete data value and $f(x(j))$ value of the fit at point $x(j)$. The fit is smoothed by minimising the second derivative of the fit calculated by equation:

$$F(D^2 f) = \int_{x_{min}}^{x_{max}} \left| D^2 f(t) \right|^2 \mathrm{d}t \quad \ldots (3)$$

Three next studied parameters represent some typical characteristics of the interpolated MU data. The most important parameter for commonly analysed aircraft areas is normal acceleration (n_z). Its fit (Fig. 4) was better than expected during most of the studied flights, a major improvement being brought by the event-triggered samples. Some rapid and small changes in n_z remain undetected when they do not locate at the constant frequency sampling points and are too small to trigger the strain gauge's event based storage process. These points do not, however, consume fatigue life considerably.

Roll rate (P) was expected to be the most difficult parameter for spline-fit construction, because its changes can be fast and it does not trigger the event based storage process (Fig. 5). The spline-type behaviour is clearly visible in the beginning of abrupt manoeuvres, which led to the abandonment of the turn-point-only analysis. In spite of the spline-type behaviour, the quality of the fit proved to be satisfactory.

The changes in pressure altitude (h_p) are considerably slower than in roll rate and normal acceleration, enabling the basis for a good fit. A systematic error was, however, detected in high vertical rate manoeuvres at pressure altitude (Fig. 6). It was solved to some extent using a heavily smoothing fit, which did obviously filter real changes to some extent.

Because the MU data does not include control surface deflections at a useful frequency, an inverse simulation application was required to calculate them[6]. The inverse simulation takes into account the flight control laws adding model-based information to neural network input by supplementing MU data. In theory a neural network could adapt to the flight control system's logic, but its complex rules are extremely difficult to extract from interpolated flight parameters. For the two studied details only the deflection of LEF was needed as an input to neural networks.

For the LEF also another additional parameter describing hinge moment was calculated. That parameter used OEM measurement data mainly from wind tunnel tests, being a function of kinetic pressure, angle-of-attack, Mach number and inverse simulated LEF deflection.

In the low sampling rate MU data the difference between the leading and trailing edge flap deflections was monitored to discriminate whether the flight control computer is in the auto-flaps up mode, e.g. clean configuration or not. This information was used to separate samples with different flight control laws and, at this stage of work, to remove samples of other than clean configuration.

https://doi.org/10.1017/S0001924000002013 Published online by Cambridge University Press

3.4 Neural network analysis

The purpose of neural network analysis is to model structural response at a strain gauge location as a function of selected flight parameters. In the analysis Matlab's neural network toolbox was applied. The developed networks are location specific, allowing the use of a different input parameter set for each detail. The outputs of neural networks are scalars, representing at this stage of work the low-pass filtered strain gauge response.

3.4.1 Input data pre-processing

An efficient data pre-processing is essential for a successful response modelling. All significant parameters affecting to the structural response shall be obviously included into the input data while simultaneously minimising the number of inputs to limit the problem size. The relationship between input parameters and output response should be as smooth as possible to allow a good generalisation capability for the developed model. This is evident especially in the transonic flight regime, where small changes in Mach number can change flight control laws and structural response considerably. Also the physical relationships of the modelled response should be as simple as possible for the neural network. In the studied case this led to separate networks for each analysed location and flight mode (clean *vs* power approach configurations).

One input parameter that was obvious in advance for further processing is normal acceleration. Both analysed details have relatively a small internal mass, leading to a stronger correlation to wing bending moment. As a first assumption, aircraft mass multiplied by normal acceleration was used instead of the acceleration alone.

The relative importance of the selected input parameters was studied using multi-linear-regression (MLR) analysis[7]. It linearly correlates normalised response Y to each normalised input parameter X_i by equation:

$$F_{MLR} = \left(X^T X\right)^1 X^T Y \qquad \ldots \quad (4)$$

The normalisation included data offsetting to zero mean and scaling to unit variance.

For the LEF hinge the weights for input parameters are shown in Table 2. All nine parameters have a moderate influence on the result, suggesting that all parameters can add information to the problem. Another method to examine the influence of input parameters is a principal component analysis (PCA), which solves an orthogonal input-dimensional space by maximising the input parameter variances. The new input space alone can enhance network performance, but more advantage can be gained if the least significant inputs are removed. The parameter variances of the PCA for the LEF problem are shown in Fig. 7, which suggests that the two least significant components could be unnecessary. Conducted tests did not, however, show noticeable difference in achieved performance and the PCA was not used in further analysis. After the tests of different input parameter sets (starting from a larger than mutually required set) and pre-processing techniques (Table 3) parameters were selected for the studied details.

3.4.2 Network performance

When evaluating different inputs and networks the stochastic nature of the network training process, which usually starts from a random initial state, has to be taken into account. When repeated, it leads to a distribution of achieved performances as displayed in Fig. 8, where ten similar networks are trained with the same training method, input parameters, and response data of OLEF. Experience has shown that in this case correlation coefficients $R \geq 0{\cdot}99$ would give good, and

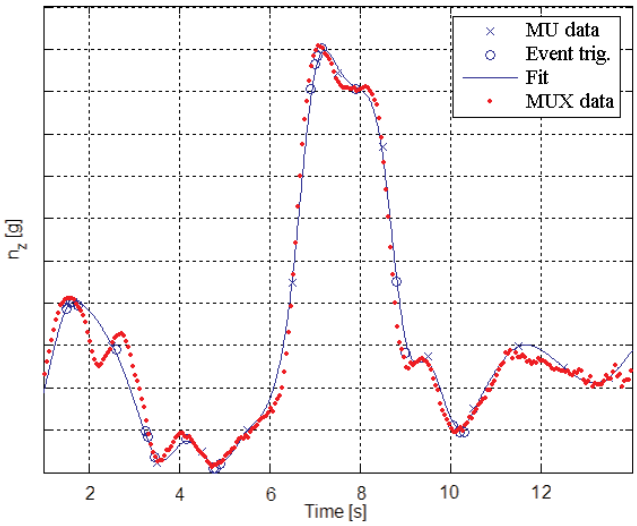


Figure 4. Example of interpolating fit for normal acceleration.

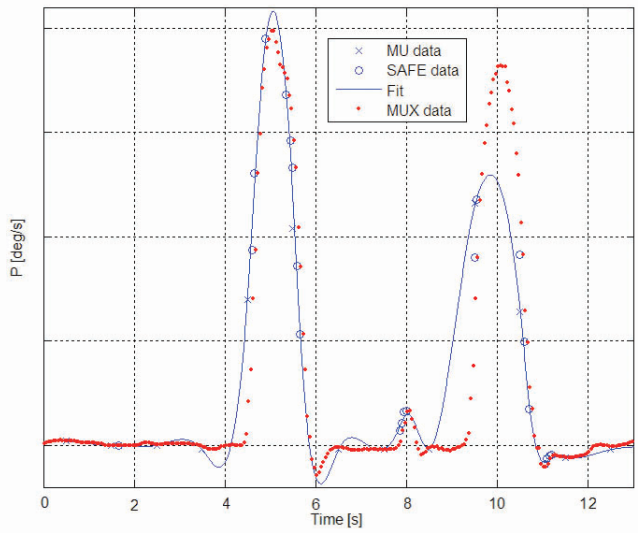


Figure 5. Example of interpolating fit for roll rate.

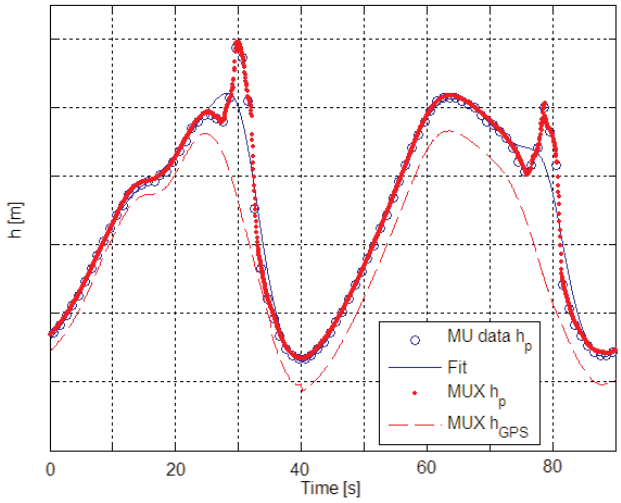


Figure 6. Example of systematic error in pressure altitude measurement and smoothing spline used for MU data interpolation.

https://doi.org/10.1017/S0001924000002013 Published online by Cambridge University Press

Table 2
MLR weights for the input parameters of OLEF

Parameter	Weight
Ma	–0·148
α	0.227
P	–0·756
P	1·460
R	–0·140
q_c	–0·563
$\delta_{LEF\ calc}$	–0·373
$C_{h\ OLEF}$	0·986
n_z x m	2·466

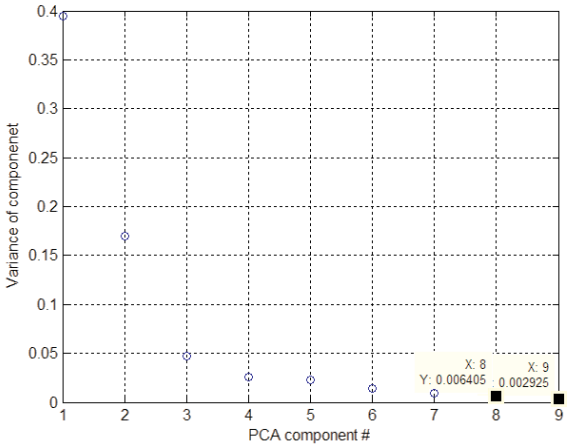


Figure 7. Variances of PCA component for OLEF.

$R \geq 0{\cdot}98$ satisfactory results, implying that the stochastic nature of network training has a significant influence on achieved performance. For this reason at least five networks were trained for each test scenario and the best performing one was selected.

The true indicator of network performance for the studied problem is its modelling capability measured in fatigue life. As the goal for the acceptable difference between measured and modelled results in individual flights, a tolerance interval –50% to +100% was adopted. The comparison of modelled fatigue lives is laborious, requiring the whole analysis chain of Fig. 3. Correlation coefficient and mean squared error between measured and modelled data are poor indicators alone, depending heavily on data distribution. They are useful, however, when comparing the network's relative performance and assuming that the modelled data distribution remains similar. A good, but labour intensive indicator of network's modelling capability is a visual comparison (Figs 9 and 10) between measured and modelled data.

In the developed analysis approach the relative network performance is evaluated using the correlation between measured and

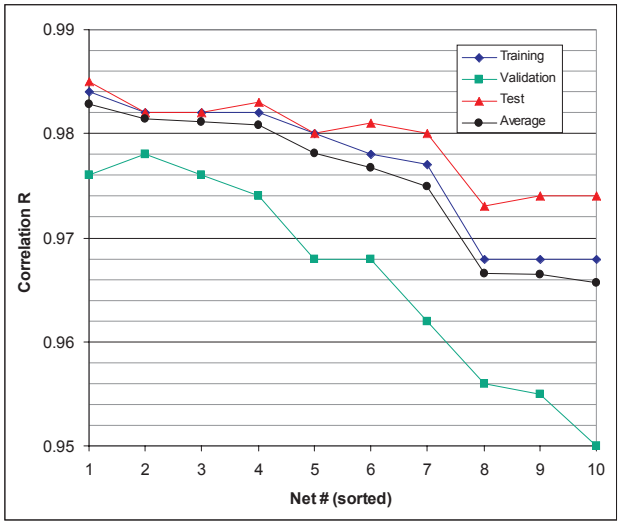


Figure 8. Achieved modelling capability for training, validation and test sets of ten similarly trained networks. Average values calculated according to set sample numbers.

modelled data supplemented by visual inspection of poorly modelled areas. The absolute performance is evaluated by comparing the measured and modelled fatigue life results.

3.4.3 Network size and structure

Network type, size and structure can vary considerably. Performances of radial basis and multilayer feed forward networks were compared in Salonen's study, which preferred multilayer feed forward networks (Fig. 11) for this paper's problem[8]. For the feed forward networks the output y of neuron i with inputs x_j from 1 to m can be calculated using equation:

$$y_i = f\left(\sum_{j=1}^{m} w_{ij} x_j + b_i\right) \quad \ldots (5)$$

The selection of transfer function f in Equation (5) between the ones provided in Matlab was done empirically resulting to the use of tansig function of equation:

$$y = \frac{2}{1+e^{-2n}} - 1 \quad \ldots (6)$$

where n is the weighted (w_{ij}) and biased (b_i) input to function f in Equation (5).

Based on tests with bulkhead and LEF hinge data, the number of hidden layers has no significant effect on the network performance, as long as the number of free parameters remains equal. The influence of the free parameters was tested between 111 parameters minimum in a 10 nodal network to 923 parameters maximum in a 30 + 17 + 5 nodal network. The number of free parameters did have an influence on the network performance but to a surprisingly small extent (Fig. 12). Already a relatively small network can produce an acceptable performance, which improves as the number of parameters increases, as also does the risk of over-fitting and required computing power during the training. The results suggest that there could be an optimum size for the network, where the possibility of poorly performing network is small. This would lead to considerably large networks (50 node network, 551 free parameters for OLEF) while an almost equal performance can be achieved with a much smaller 25 nodal network with 276 free parameters. The latter one was selected due to the smaller risk of over-fitting.

https://doi.org/10.1017/S0001924000002013 Published online by Cambridge University Press

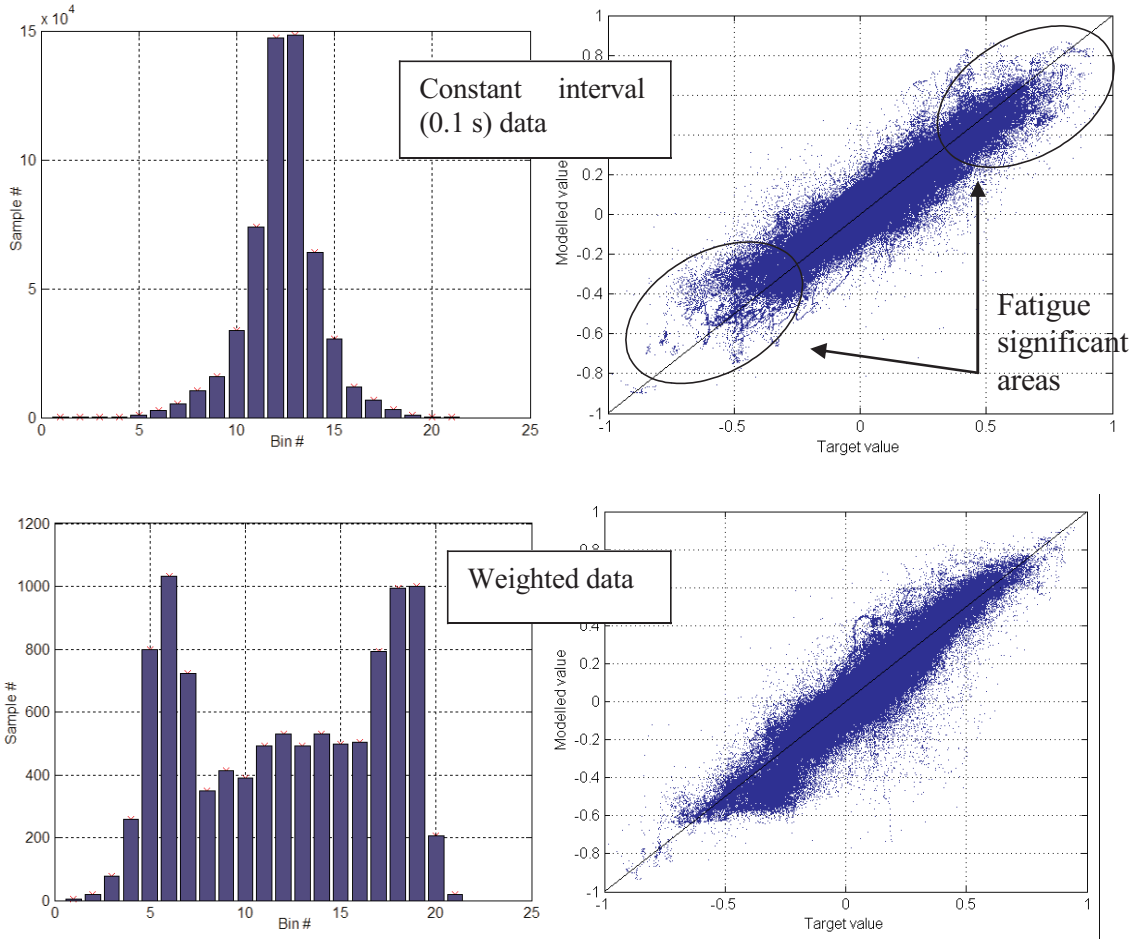


Figure 9. Response distribution for a network trained with constant frequency data (10Hz) and for a network trained with weighted data.

Table 3
Neural network input parameters

Parameter	FS488	OLEF
Ma	X	X
α	X	X
P	X	X
P	X	
R	X	X
q_c	X	X
$\delta_{LEF\ calc}$		X
$C_{h\ OLEF}$		X
n_z x m	X	X

3.4.4 Training method

Several training methods are presented in literature for feed-forward type networks. Within this project there was no intention to own development activities, resulting to the training method selection between the ones provided in Matlab's Neural Network toolbox[9]. In preliminary tests three functions seemed to be promising: traingdx (based on gradient momentum and adaptive learning rate), trainlm (based on Levenberg-Marquardt) and trainbr (based on Levenberg-Marquardt with Bayesian regulation). Although traingdx gave good results for small training data sets, its performance was lower than that of the two others as the size of the training set grew. Thus, only trainlm and trainbr functions were further tested. With 25 neuron nets the performance of trainlm and trainbr was comparable. With a 50 neuron net trainlm performed better, and was selected as the preferred training method.

To avoid over-fitting early stopping criteria was used in training. When using early stopping the available data is divided into two or three subsets. The largest subset is a training set being used for network adaptation. The second subset is a validation set that the network does not use in the adaptation phase. The training is stopped at the point, where the error of the validation set starts to increase while the error of the training set yet decreases. The third set is an optional test set, which can be used independently to compare competitive network performances, being not used during the training process at all.

The construction of training, validation and test sets depends on the data amount available. In an ideal case for a parameter based

https://doi.org/10.1017/S0001924000002013 Published online by Cambridge University Press

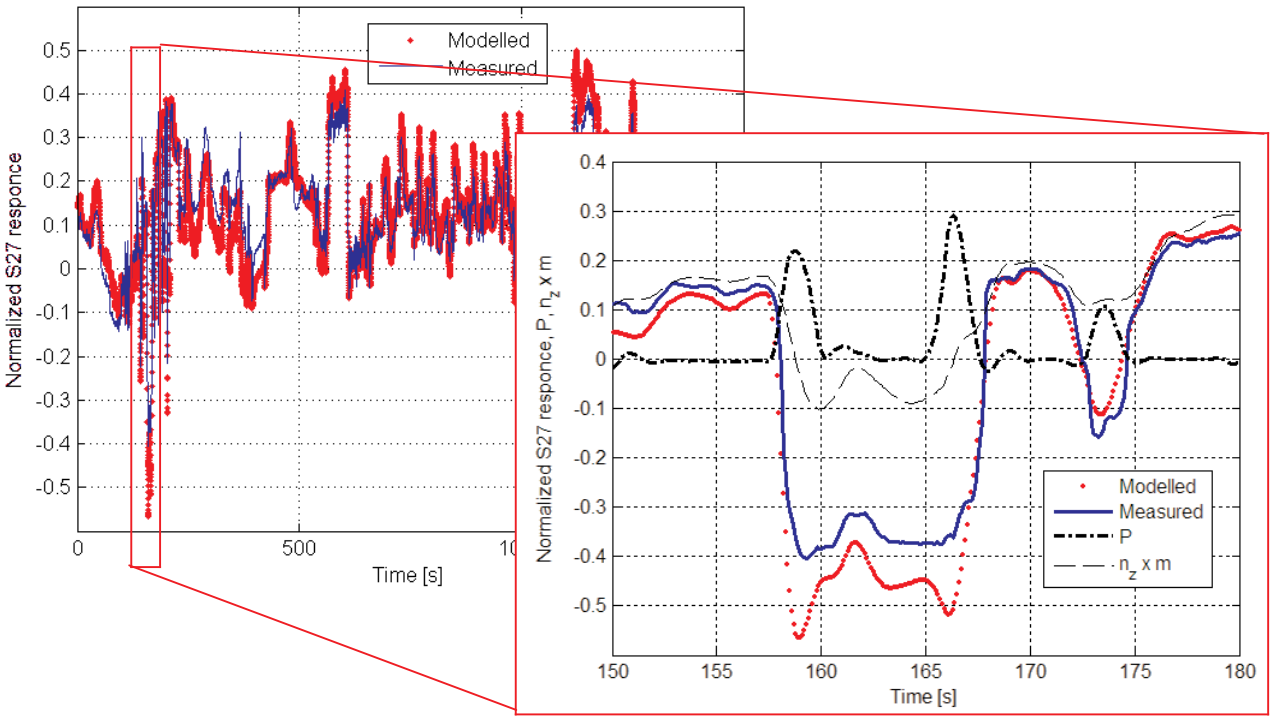


Figure 10. Cross validation test network's capability to predict flight #0028 response and essential flight parameters (normalised) for the modelling error.

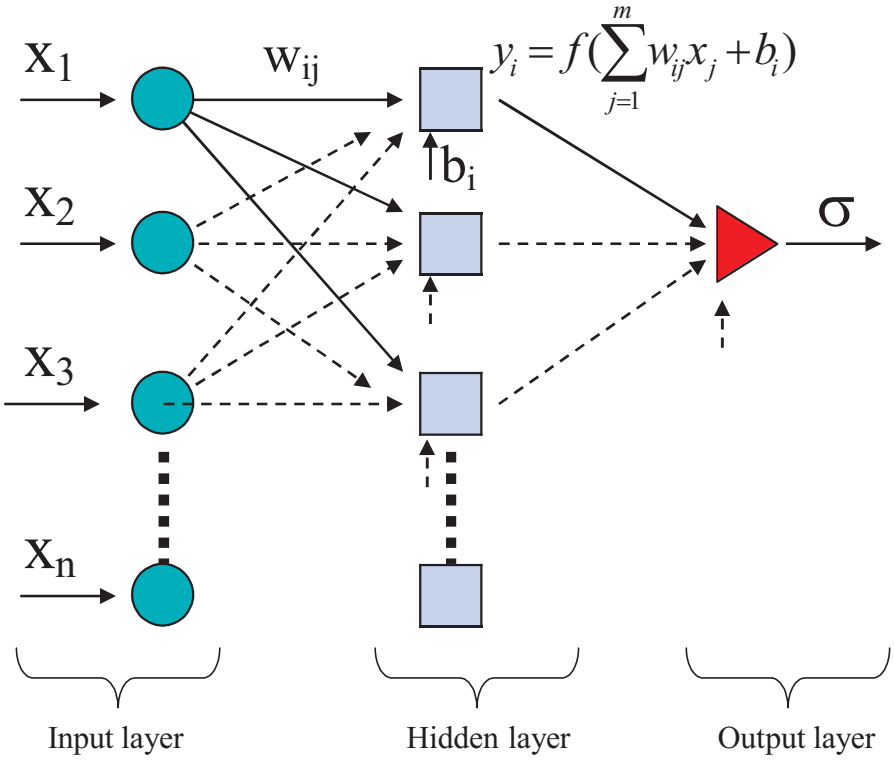


Figure 11. Principle of feed-forward type network with one hidden layer and scalar output.

https://doi.org/10.1017/S0001924000002013 Published online by Cambridge University Press

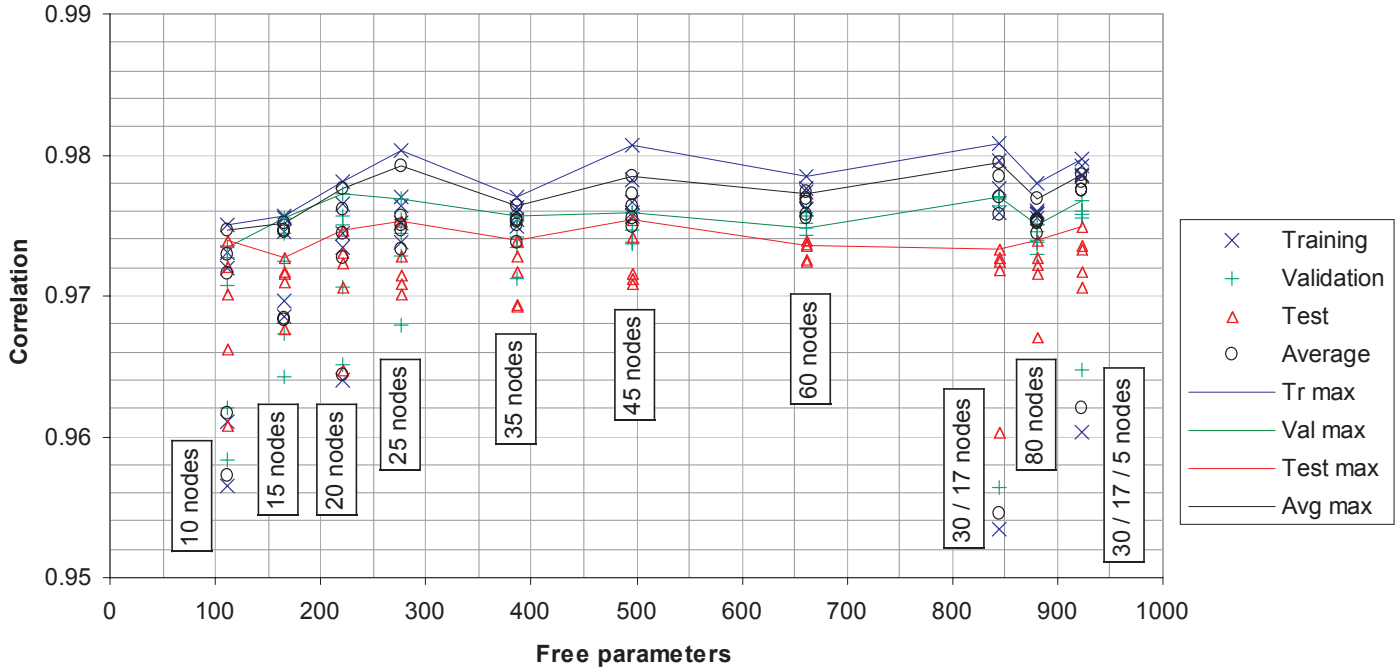


Figure 12. Effect of free parameters to the OLEF network performance
(19,000 samples divided to the blocks of 1,000 training, 300 validation and 100 test samples).

fatigue life analysis problem all sets would consist of several complete flights not interfering with each other. At the current stage of the work 25 flights were available, which did not yet allow several complete flights being left to validation and test sets. The second best option was to divide the data into such a long blocks that prevented in most cases the division of the same flight manoeuvre into training and validation sets. This led to the selection of at least 30 second long data blocks repeatedly, being further divided to the training, validation and optional test sets.

3.4.5 *Training data selection*

The selection of training data has a major influence on network performance, especially on generalisation capability. The need for the training data selection originates from the fact that the majority of recorded data represents steady level flight. The first alternative is to use only response turn points in training, which increases the relative amount of the interesting high response samples. For normal acceleration dependent areas the turn points based training would ideally leave one level flight sample for each 'high-g' sample. In tests this selection criteria produced, however, too much low response samples and networks did not learn the high response flight envelope at an acceptable level. Further tests with flight parameters n_z, P and Q as turn point criteria were conducted, which solved a majority of the problem. Yet some areas of the flight envelope remained underweighted to produce an acceptable network generalisation capability like, for example, high speed level flight conditions forcing the search for other training data selection criteria.

The total number of data points from 25 flights was about 556,000 when interpolated to 10Hz frequency. Available computer memory allowed only some 25,000 to 50,000 data samples depending on the network size, forcing the removal of 90–95% of the data before the neural network training. A better performance, compared to the turn points based training tests, was achieved by the removal of low response points on a statistical basis (Fig. 9). Yet similar generalization problems existed in the high speed low response flight.

The finally applied data removal was based on K-means clustered data, where samples with similar input parameters are associated to the same cluster. The K-means clustering process starts from a desired number, in this case 45, of arbitrarily selected cluster centres. Each sample is associated to the nearest one based on Euclidean distance of the normalized input parameters. The next step is to calculate new cluster centres according to the associated samples and to re-associate the samples using the new cluster centres. The iteration is continued until no changes in the association occur.[7]

In the data removal process the amount of samples within each cluster was randomly decreased until the desired number of samples in each cluster was achieved. This clusters based data weighting ensures more even training data distribution over the whole flight envelope. During the clusters based weighting, an additional check was done, which forced all samples with, even a single, input parameter outside 0·01 … 0·99 cumulative probability interval to remain in the training data set. This check increased the network performance especially in the supersonic low-altitude regime, where the number of recorded manoeuvres was extremely limited.[10]

3.5 Fatigue life analysis

Neural network analysis produces the modelled strain gauge response as a function of the selected flight parameters. This stress was linearly correlated with the stress at fatigue critical location using the results of structural finite element models. This resulted in the response history required in fatigue life analysis. The fatigue life analysis was done using aircraft manufacturers Life-Works software. Its analysis is based on strain-life theory, producing the calculated damages of the analysed flight for each detail.

4.0 RESULTS

The emphasis on the results was to prove to the fleet operator the analysis capability in typical FiAF usage. For this reason the results are mainly presented for a set of nine flights, which were classified as most representative to typical FiAF flying.

https://doi.org/10.1017/S0001924000002013 Published online by Cambridge University Press

Table 4
Trained networks in cross validation test (selected ones in bold)

Test fllight number	net 1			net 2			net 3			net 4			net 5				
	R_{train}	R_{valid}	R_{avg}	R_{train}	R_{valid}	R_{avg}	R_{train}	R_{valid}	R_{avg}	R_{train}	R_{valid}	R_{avg}	R_{train}	R_{valid}	R_{avg}	net_{test}	R_{test}
#0022	**0.970**	**0.960**	**0.968**	0.964	0.955	0.962	0.966	0.961	0.965	0.963	0.962	0.963	0.964	0.963	0.964	1	0.946
#0024	0.967	0.981	0.970	0.967	0.980	0.970	0.969	0.980	0.972	**0.970**	**0.980**	**0.972**	0.968	0.979	0.971	4	0.839
#0026	0.968	0.974	0.969	0.978	0.980	0.978	0.982	0.985	0.983	0.979	0.982	0.980	**0.982**	**0.986**	**0.983**	5	0.941
#0028	0.970	0.978	0.972	0.964	0.975	0.967	0.973	0.980	0.975	0.968	0.978	0.970	**0.974**	**0.980**	**0.975**	5	0.935
#0030	**0.980**	**0.985**	**0.981**	0.965	0.978	0.968	0.976	0.980	0.977	0.977	0.982	0.978	0.977	0.984	0.979	1	0.925
#0032	**0.981**	**0.976**	**0.980**	0.981	0.967	0.978	0.971	0.968	0.970	0.972	0.963	0.970	0.970	0.966	0.969	1	0.946
#0033	0.973	0.976	0.974	0.973	0.974	0.973	0.976	0.973	0.975	**0.980**	**0.977**	**0.979**	0.978	0.973	0.977	4	0.930
#0034	0.981	0.982	0.981	**0.984**	**0.984**	**0.984**	0.984	0.984	0.984	0.984	0.984	0.984	0.984	0.984	0.984	2	0.868
#0035	0.979	0.983	0.980	**0.983**	**0.985**	**0.983**	0.983	0.985	0.983	0.981	0.984	0.982	0.982	0.984	0.982	2	0.919

Table 5
Results of cross validation test

Test fllight number	Fatigue life [repeats]			Difference [%]	
	meas (raw data)	meas (filtered & GAG cycle removed)	modeled	Filtered / Raw	Modeled / Filtered
#0022	13862	17943	19764	29	10
#0024	20907	27133	50870	30	**87**
#0026	17722	23171	24574	31	6
#0028	66890	102163	37189	53	**-64**
#0030	92753	127153	79669	37	-37
#0032	1953	2430	3085	24	27
#0033	40219	53107	36133	32	-32
#0034	123280	219509	45323	78	**-79**
#0035	4986	6886	10608	38	54

4.1 LEF strain gauge S27 modelling

At first a larger than mutually required set of input parameters was selected and tested to ensure that all significant ones are included. The finally selected input parameters for LEF strain gauge S27 analysis were M*a*, α, *P*, *Q*, *R*, q_c, δ_{LEF}, $C_{h\ OLEF}$ and n_z *x m* (see Tables 1-3). The analysis used a feed-forward type network with one hidden layer including 25 tansig neurons and an output layer including one tansig neuron. The network was trained with Matlab's trainlm (Levenberg-Marquardt) function. Early stopping method was used to stop the training. After weighting, the selected training data was partitioned typically into the blocks of 1,000 training and 300 validation samples.

First a cross validation test was conducted to gather understanding of the over-fitting risk. In the test nine FiAF's operational use representative flights were one-by-one dropped out from the training and validation data sets and used as separate test data. The rest of the measurement data was weighted and divided into the blocks of 1,000 training and 300 validation samples. For each test case five networks were trained, the best performing one of which (measured in the achieved correlation) was selected. The excluded flight's fatigue life consumption was calculated using the measured and modelled S27 response histories (Tables 4 and 5). In addition to the fatigue lives, Table 5 shows the influence of response low-pass filtering. In the LEF case it increased the raw data fatigue life by about 30%. Its influence appears to be quite constant, allowing the use of an additional knock-down factor for the modelled fatigue life. The more appealing solution would be to model also the vibrating part of response, which is planned in the future.

The average difference between the measured and modelled fatigue life (Table 5), was 44% for the nine test flights. Using the table five fatigue lives, FiAF's average flying was modelled by weighting the flights according their share of the FiAF total flying hours. For the average flying the difference between the measured and modelled fatigue life was +27 %, being well inside the scatter factors usually required for analytical results. Although the modelling difference of +27 % in FiAF's average usage was small, there are much larger modelling errors in individual flights. When the tolerance band of –50% to +100% is used as the performance criteria, flights #0024, #0028 and #0034 fall outside it.

Flight #0028 consisted of low speed air combat manoeuvring. The largest errors were in Fig. 10 type manoeuvres that were rolls at high roll rate and transonic Mach number. The Mach number in these manoeuvres was about 0·9, being also the speed of aileron reversal. In this flight regime the aero-elasticity effects change rapidly, as do the aircraft flight control laws. As a possible solution for neural network adaptation, this kind of manoeuvres should be emphasised in the training data – if available. Flight #0034 was a supersonic air combat manoeuvring flight, including flight envelope areas not visited during any other flight. This highlights the challenge stated earlier: the relatively small data amount does not allow whole flights being left out of training data. The last poorly modelled flight #0024 represents beyond visual range air combat manoeuvring. No single manoeuvre with a major error was found from that flight. The flight included about ten quite similar high subsonic manoeuvres, whose response was modelled similarly lower than measured. In conjunction with the systematic temperature drift error this caused the large cumulative error in the modelled fatigue life.

In view of the poor modelled flight conditions, one can suppose that better training data selection will lead to better generalised results. Unfortunately the available data from the problematic flight envelope areas was extremely limited, and can be helped only by the new measurement data collected by the OLM aircraft in the future. On the other hand one can expect that the increasing information from these areas will improve the achievable modelling capability, and the current cross validation results can be considered as the lower boundary estimate of the analysis performance.

After studying the lower performance boundary of S27 modelling capability, two other tests were conducted to evaluate the expected and optimum performances. The expected performance was evaluated by training five networks with all available flights divided into the blocks of 1,000 training and 300 validation samples. The best performing network was used to produce the fatigue life estimate for the set of nine flights, which were also used in the cross validation test. Results of the analysis (Table 6) promise a better prediction capability than was achieved in the cross validation test. The average difference between measured and modelled data was 30%, being about two thirds of the difference in the cross validation test. For FiAF's average usage a modelling difference of only +2% was achieved. Yet two of the three problematic flights (#0028 and #0034) fell outside the –50% … +100% tolerance band and require in the production level analysis more attention. When comparing the results visually to the cross validation test data, no signs of over-fitting were found supporting the decision to use all flights in training.

https://doi.org/10.1017/S0001924000002013 Published online by Cambridge University Press

Table 6
S27 modelling capability using all flights for training (early stopping used)

Test fllight number	R_{test}	Fatigue life [repeats]		Difference [%]
		meas (filtered & GAG cycle removed)	modeled	Modeled / Filtered
#0022	0.955	17943	14073	-22
#0024	0.874	27133	31785	17
#0026	0.946	23171	27291	18
#0028	0.947	102163	49121	**-52**
#0030	0.934	127153	68457	-46
#0032	0.958	2430	2734	13
#0033	0.933	53107	40469	-24
#0034	0.915	219509	74336	**-66**
#0035	0.915	6886	7741	12

The upper boundary of S27 modelling capability was evaluated with all data samples in the training set. This test can be considered as the upper performance boundary, because all new data includes new flight conditions and more generalisation is required, thus lowering the achieved accuracy. If all the data is used in training, no indication of over-fitting is available and one has to rely on the fact that the amount of training data is overwhelming compared to the number of free parameters. The achieved performance increase shall be compared against the results, which are produced using separate validation and test data. The typical size of weighed data sets in this test was about 15,000 samples. When using, for example, 25 nodes in hidden layer the risk of over fitting for the 276 free parameters remains low and the probability of achieving a properly generalised solution is good. The Table 7 results indicate a small increase in achieved performance. The average difference between measured and modelled data was 31%; practically identical to the earlier test with early stopping. The difference for FiAF's average flying was also in this test extremely small, being –1%. The modelling accuracy of flights #0028 and #0034 increased slightly, but the same manoeuvres as earlier were modelled with an unsatisfying accuracy.

4.2 Fuselage bulkhead FS 488 analysis

The fuselage bulkhead FS 488 was considered as the easier one of the selected details. The selected input parameters were Ma, α, P, Q, R, $\dot{P}$, q_c and n_z x m. Response data from the strain gauge S56 was filtered by a 7Hz low-pass filter. Data from the 25 flights was used in training consisting of the clean configuration samples. Training data was weighted using the K-M clustering, resulting to the reduction of training data from 556,000 to about 10,000 samples. The training data was partitioned similarly to the S27 expected performance test: 1,050 and 300 samples per block for training and validation sets.

The S56 network was a feed-forward type with 15 tansig neurons in the hidden and one in the output layer. Matlab's trainlm function was used as the training function. Before early stopping activated about 100 training iterations were completed for the best performing network. The achieved performance (Fig. 13) provided correlation coefficients of 0·990 and of 0·985 for training and validation data respectively.

A Life-Works fatigue life analysis was conducted for both measured and modelled data. For the majority of the flights the results of Table 8 are inside a ±30% tolerance band. The difference between measured and modelled results remained for all nine flights inside the –50% to +100% band. The difference in FiAF average usage was extremely small, being only 2%. Similar reasons as in S27 case were found to be the origin of the flight #0034 poor modelling capability. The flight included rare supersonic manoeuvres, of which an especially problematic example is shown in Fig. 14.

Table 7
S27 modelling capability using all flights for training (early stopping not used)

Test fllight number	R_{test}	Fatigue life [repeats]		Difference [%]
		meas (filtered & GAG cycle removed)	modeled	Modeled / Filtered
#0022	0.951	17943	12317	-31
#0024	0.87	27133	33164	22
#0026	0.947	23171	25544	10
#0028	0.946	102163	60783	-41
#0030	0.938	127153	66749	-48
#0032	0.957	2430	2828	16
#0033	0.938	53107	38140	-28
#0034	0.921	219509	84439	**-62**
#0035	0.913	6886	8116	18

Table 8
Results of S56 fatigue life analysis

Test fllight number	R_{test}	Fatigue life [repeats]		Difference [%]
		meas (filtered & GAG cycle removed)	modeled	Modeled / Filtered
#0022	0.990	5302	5853	10
#0024	0.970	14453	15200	5
#0026	0.993	6703	5935	-11
#0028	0.991	13154	11250	-14
#0030	0.988	38100	33060	-13
#0032	0.989	3832	2889	-25
#0033	0.986	22832	16099	-29
#0034	0.943	45666	24864	**-46**
#0035	0.942	78371	49292	**-37**

5.0 CONCLUSIONS AND DISCUSSION

For FiAF F-18 fuselage bulkhead and leading-edge flap hinge area a flight parameter based fatigue life analysis capability is created. For the bulkhead the average difference between measured and modelled fatigue life during 9 FiAF's operational flights was 21% and for the LEF hinge area 30%. Both these numbers are at an acceptable level when compared to typical scatter factors used to analytical results. For FiAF's average usage the total differences were extremely small being –2% for the bulkhead and +2% for the LEF.

When using as the modelling accuracy criteria –50% to +100% for individual flights, the bulkhead modelling capability was at an acceptable level with results between –46% and +10%. For the LEF hinge area the differences of individual flights were between –66% and +18%. Two of the nine flights fell outside the accepted tolerance band requiring further work.

Although the available flight-test data consisted of 25 flights, the number of independent manoeuvres was limited, especially at high transonic and supersonic speeds. This led to modelling problems, which can be solved, to some extent, as more data from OLM instrumented aircraft becomes available. This additional data is expected to increase modelling accuracy, especially for the currently most problematic supersonic flight conditions.

During the neural network training and selection process two anomalies, worth noting were detected. The first results of modelled strain gauge S27 response revealed that in some cases (Fig. 15) the network modelled constant response, while the measured response experienced slow steady change. It was evident from flight parameters that the aircraft flew several minutes at steady level flight. Also some other strain gauge signals showed a similar behaviour, suggesting thermal expansion as the origin. When the

https://doi.org/10.1017/S0001924000002013 Published online by Cambridge University Press

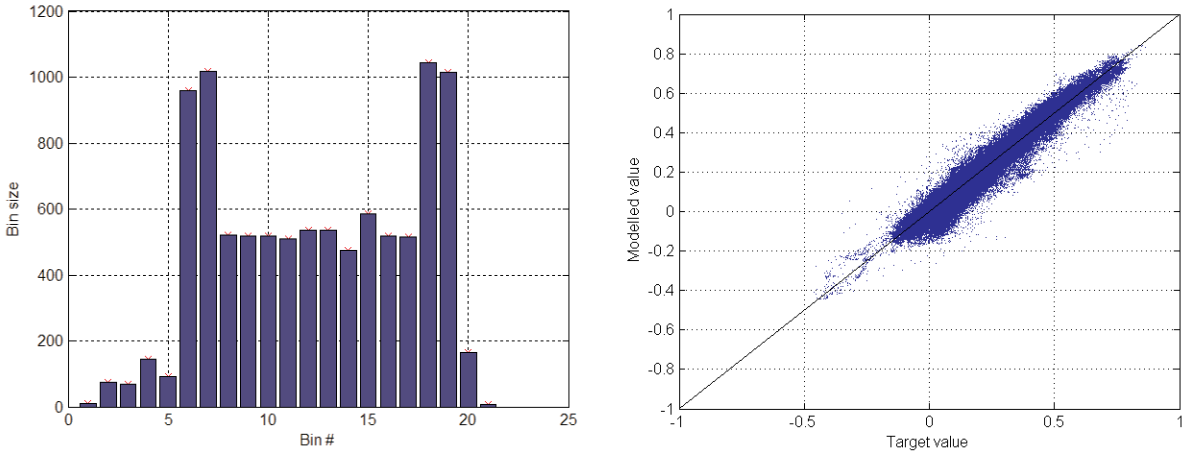


Figure 13. S56 training (including validation) data weighting and acquired performance.

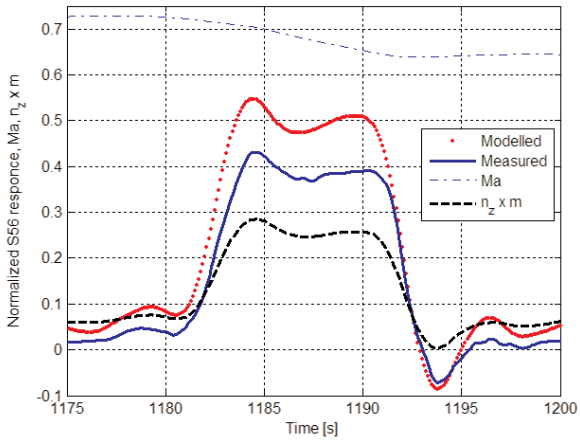


Figure 14. Example of poor correlation between measured and modelled data in a supersonic manoeuvre.

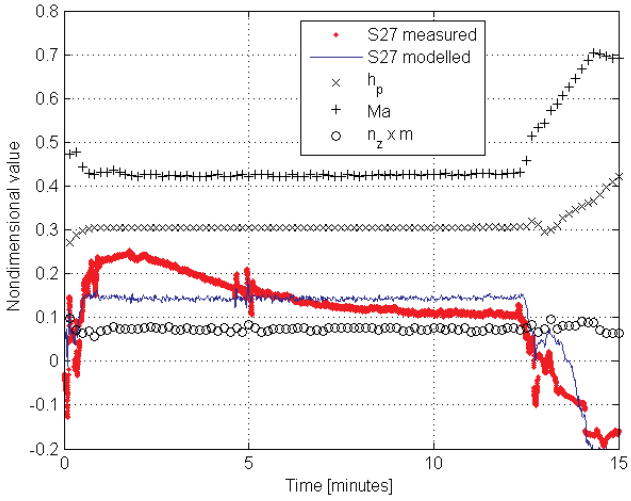


Figure 15. Example of temperature gradient caused drift in S27 response and relating flight parameters.

aircraft climbs rapidly to altitude and transits to steady level flight, the massive hinge of LEF experiences an internal temperature gradient, introducing stress at the surface. Tests to account for this stress were conducted using empirical fit to the heat expansion Equation (11):

$$\sigma_{surf} = E\alpha\left(T_{\text{int}} - T_{surf}\right) \quad \ldots (7)$$

where the hinge surface and internal temperatures (T_{surf} and T_{int}) were calculated for each time step i using equations:

$$T_{surf}(i) = T_{surf}(i-1) + C_{surf}\left(V_T^{0.8}\rho^{0.8}\right)\left(T_{amb}(i-1) - T_{surf}(i-1)\right)\Delta t \quad \ldots (8a)$$

$$T_{\text{int}}(i) = T_{\text{int}}(i-1) + C_{\text{int}}\left(T_{surf}(i-1) - T_{\text{int}}(i-1)\right)\Delta t \quad \ldots (8b)$$

The stress response was empirically fitted by selecting the coefficients α, C_{surf} and C_{int} to match the measured response. In Equation (7) the E equals to aluminium's Young's modulus. In Equations (8) the V_T equals to true air speed, ρ to the density of air at ambient temperature (T_{amb}) and Δt to sample time difference.

The empirical correction worked properly in Fig. 15 type cases but introduced new errors on other flight envelope areas, which included typically high vertical rate manoeuvres. The latter ones are more important for fatigue life prediction, thus the correction of thermal strain was not introduced into the analysis at this stage of work.

Another anomaly in the strain gauge data was discontinuities in S27 signal near zero level (Fig. 16). That was most likely originating from a backlash in LEF hinge. To reduce its influence on neural network training, the near zero response samples were classified less reliable in the training data selection and their amount was further decreased in the training data.

The results described in this paper proved the usability of the parameter based fatigue life analysis. The achieved accuracy fulfilled partially the requirements of FiAF and origins of the few poorly modeled samples could be identified. This supported the decision to continue the work towards a production level analysis environment. The analysis method will be developed to include also other than clean flight conditions, ground-air-ground cycle, and influence of external loads. Also efforts to enhanced training data selection and vibration content modelling are planned.

https://doi.org/10.1017/S0001924000002013 Published online by Cambridge University Press

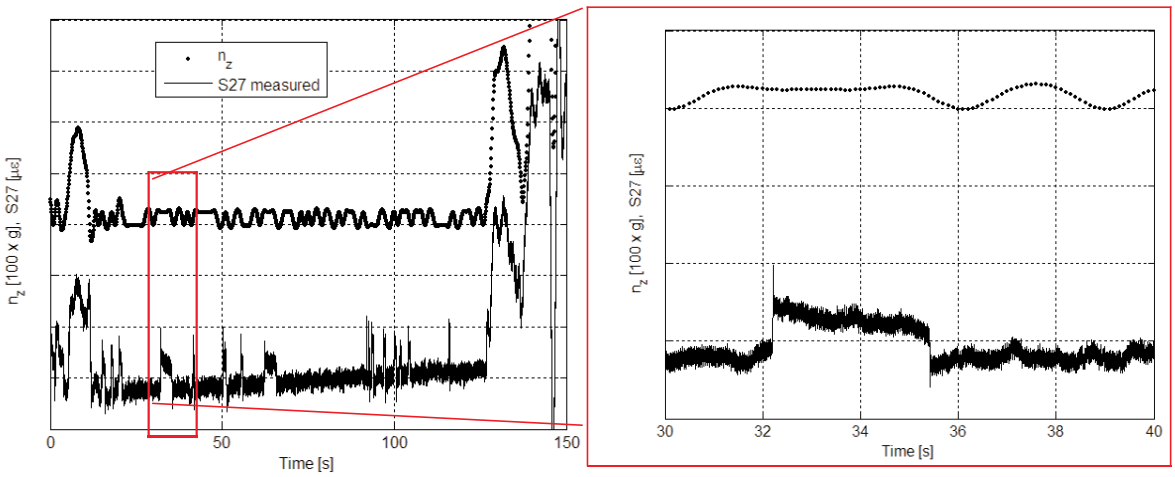


Figure 16. Example of response discontinuity in S27 signal.

ACKNOWLEDGEMENTS

The author acknowledges the FiAF for the possibility to realise this work. The support of the Technical Research Centre of Finland and Helsinki University of Technology are also highly appreciated.

REFERENCES

1. MOLENT, L. *et al*, Development of analytical techniques for calibration of F/A-18 horizontal stabilator and wing fold strain sensors, report DSTO-TR-0205, Melbourne, Australia, 1995.
2. UHL, T. Applications of neural networks for identifications of loads in mechanical structures, SMART conference, Warsaw, Poland, 2001.
3. REED, S. Development of a parametric aircraft fatigue monitoring system using artificial neural networks, at the 22nd ICAF conference, Lucerne, Switzerland, 2003.
4. SILJANDER, A. A review of aeronautical fatigue investigations in Finland during the period April 2003 to April 2005, 29th Conference of the International Committee on Aeronautical Fatigue (ICAF), Hamburg, Germany, June 2005.
5. *Spline toolbox for use with Matlab*, The MathWorks., Natick, MA, USA, 2004.
6. ÖSTRÖM, J. Inverse flight simulation for a fatigue life management system, AIAA Paper 2005-6212, in AIAA Modeling and Simulation Technologies Conference, San Francisco, California, 15-18 August 2005.
7. HYÖTYNIEMI, H. *Multivariate regression models,* Helsinki University of Tech., ISSN 0356-0872, Picaset Oy, Espoo, 2000.
8. SALONEN, T. An assessment of the feasibility of a parameter-based aircraft strain monitoring method, M.Sc. dissertation, Helsinki Univ of Tech, Espoo 2005 (in Finnish).
9. *Neural network toolbox for use with Matlab,* The MathWorks, Natick, MA, USA, 2001.
10. TIKKA, J. Parameter based fatigue life analysis, Influence of Training Data Selection to Analysis Performance, AIAA Paper 2006-6473, in AIAA Modeling and Simulation Technologies Conference, Keystone, CO, 21-24 August, 2006.
11. MCCRACKEN, K. Clarification on Reynolds Number and the correlator thermal design, SMA Technical Memorandum # 118, http://sma-www.cfa.harvard.edu/private/memos/118.pdf, 21.11.2005

https://doi.org/10.1017/S0001924000002013 Published online by Cambridge University Press