Estimation of Structural Component Loads in Helicopters: A Review of Current Methodologies

Frank G. Polanco

Airframes and Engines Division
Aeronautical and Maritime Research Laboratory

DSTO-TN-0239

ABSTRACT

A review of the literature for models that use fixed-component loads and flight parameters to determine loads in a dynamic component is presented. The reviewed papers naturally divide into one of three categories depending on the information they use to determine the load in the dynamic component. An initial section on load variability demonstrates that even for the same aircraft under the same flight condition, the loading can vary dramatically due to pilot technique, altitude, and weight to name a few variables. Neural networks, regression, and statistical indicators prove invaluable in developing load models. The review also demonstrated a lack of solutions to fundamental questions concerning loads modelling.

RELEASE LIMITATION

Approved for public release
Estimation of Structural Component Loads in Helicopters: A Review of Current Methodologies

Executive Summary

Helicopter components whose fatigue lives are limited have prescribed Component Retirement Times (CRTs) that define the safe number of flying hours that they can remain in service. At the time a fleet of helicopters first enters service with a particular operator the CRTs are set by the manufacturer based on an assumed usage spectrum that has been defined with the agreement of the operator. The relationship between component loads and the various flight regimes that make up the usage spectrum is usually obtained via a flight load survey on the prototype helicopter. The Australian Defence Force discards components when they have reached their CRT. However, most components are, in effect, retired prematurely, due to the conservative nature of an assumed usage spectrum. At the other extreme, it is possible that some components could exceed their safe fatigue limit before reaching their CRT. The accurate determination of a loading history for dynamic components would have the potential to meet the two major objectives of reducing costs and improving safety.

We review the literature for models that use fixed-component loads and flight parameters to model the loads in dynamics components. The reviewed papers naturally belong in one of three categories depending on model inputs: fixed-component loads, flight parameters, or a combination of fixed-component loads and flight parameters.

A review of load variability reveals that even for the same aircraft under the same flight condition, the loading can vary dramatically due to pilot technique, altitude, and weight to name a few variables. For this reason, although the use of regime recognition models fleet-wide to obtain an estimate of the life fractions expended during flight on individual helicopters can produce results that are superior to the adoption of fixed CRTs, simply recognising the regime and its duration is insufficient to determine the true fraction of life expended for a component.

A technique that develops a calibration matrix in the frequency domain, termed holometrics, appears promising. The holometrics technique was initially developed using only fixed-component loads, but was later extended to include flight parameters. Statistical methods such as significance tests and confidence intervals allow different modelling options to be prioritised. Neural network (NN) and multiple regression models were also extensively utilised by researchers. NNs were more computationally intensive than the regression models, but the NNs had better generalisation capabilities. Most studies focused on high load manoeuvres (since these produce the majority of the fatigue damage), and thus used filtering or data weighting to produce a high load bias. Adding the product of certain input parameters, such as swashplate servo position and accelerations also improved results significantly. Expert NN systems appear to be a promising area for further research.

Fundamental loads modelling effects such as noise, rank deficiency, stability, and generalisation received no attention by most researchers. These essential questions need to be addressed if robust and accurate load estimation models are to be developed and implemented.
1. Introduction

Helicopter components whose fatigue lives are limited are assigned Component Retirement Times (CRTs) that define the safe number of flying hours that they can remain in service. At the time a fleet of helicopters first enters service with a particular operator the CRTs are set by the manufacturer based on an assumed usage spectrum that has been defined with the agreement of the operator. The Australian Defence Force discards components when they have reached their CRT. However, the CRTs are based on a worst-case scenario in terms of fatigue damage and do not take into account perturbations from this spectrum. Consequently, most components are, in effect, retired prematurely. The development of loading histories for critical components has two beneficial possibilities, increasing safety and reducing operator costs. The perception that these benefits are achievable explains the recent increase of work in this area [1–9].

The rotor systems and their attachments contain most of the fatigue-critical structural components in helicopters. To monitor loads on these components would require slip rings or rotating telemetry systems that currently lack the reliability and maintainability needed for everyday fleet usage. Hence, there is a great impetus to develop a load transfer function, which can predict the loads in rotor components using information from flight parameters and loads from static components.

In this report, we review published work on the estimation of loads necessary for calculating fatigue damage to helicopter components. Initially we demonstrate the vast variability in component loads and hence fatigue damage due to pilot technique. Subsequently we provide an overview of various methodologies for estimating loads in structural component in helicopters.

The parameters to be monitored for estimating component loads for helicopter component fatigue life estimation purposes fall into one of three broad categories: loads on fixed components, flight parameters, or both loads on fixed components and flight parameters.

Only one paper was found that implemented the estimation of dynamic component load based solely on fixed component load estimation. This approach, which was undertaken in the frequency domain and termed holometrics, develops a linear relation between the strain in fixed components and the strain in dynamic components. In contrast, several papers estimate dynamic component loads based on flight parameters. These methods based on flight parameters were further subdivided into four categories: regression, neural networks, flight condition recognition, and combined regression and neural networks. Using the loading history, determined using the above procedures, several authors went on to estimate component fatigue damage using strength data presented as S-N curves. In Section 5 of this report we review an extension of the holometric method, from just one paper found on this topic.
2. Load Variability

Schaefer [10] investigated the effects of aerial combat on helicopter structural integrity. Component loads were sensitive to a number of air combat variables including pilot technique, adversary type, control sensitivity, and aircraft response. The two test helicopters, the CH-53E and the AH-1W, were flown at two representative gross weights and piloted by at least three flight test pilots from both Marine Aviation Weapons and Tactical Squadron instructors. This pilot variation ensured that the effects of aircraft loading, pilot technique, pilot experience, and control variability would be recognisable during subsequent data acquisition and reduction. The evasive manoeuvring flight loads surveys were flown within the existing flight envelope limits.

For manoeuvres flown against a helicopter adversary, the pitch links, blades, and stationary controls of the main rotor experienced high loads. While against fixed wing aggressors, higher loads were more apparent for the tail rotor blade, yoke (hub), and control system. For example, during free engagement against an OH-58, the main rotor pitch links and yoke peak loads increased to 80% of the static “do-not-exceed” limit. Generally, fatigue loads were about 7–10% higher than those measured during the contractor’s flight stress survey and structural demonstration. Dynamic component loads were very dependent on control displacement rates, control magnitude, and control phasing. For example, in a “pitch-back-attack” manoeuvre, peak and vibratory loads varied by as much as 730% and 193% respectively for different pilots. Schaefer concludes that due to these elevated component loads, CRTs might be reduced by as much as 75%.

Boorla and Rotenberger [11] reported on the load variability of a two-bladed helicopter. Some effects that determine fatigue life include mechanical assembly tolerances, wear and tear, pilot technique, manoeuvre execution severity, gross weight, altitude, rotor speed, and ambient aerostatic conditions. The flight program, carried out on a Bell Model OH-58C helicopter, consisted of 33 manoeuvres each repeated 30 times using six pilots (three from Bell Helicopter Textron and three from the U.S. Army). Each pilot flew the helicopter five times, with the flights lasting 40–45 minutes. Using a three-parameter Weibull distribution the resulting data set was statistically analysed for distribution and variability of peak loads within manoeuvres, cyclic counted loads for the total manoeuvre, cyclic counted loads for the total flight, and fatigue damage rates. The Weibull distribution may be thought of as a generalised normal distribution, its probability density function having the form \( y^{b-1} \exp(y^{-b}) \), where \( y \) involves two parameters (analogous to the mean and the standard deviation) and \( b \) is the Weibull slope. The Weibull distribution was used because the majority of manoeuvres were heavily right skewed. A normal distribution proved sufficient for the peak loads analysis. During the flights 18 discrete load parameters were measured (an additional five parameters were derived from these measurements).

Analysing only peak loads within each manoeuvre was found to exhibit larger variability than analysing cycle-counted data from either each manoeuvre or an entire flight. Boorla
and Rotenberger concluded that ‘...fatigue damage analysis indicates that the load variability needs to be accounted for in fatigue analysis because of its powerful effect on life'. They also found that the Weibull distribution model produced correlation coefficients in excess of 90–95% (and in some cases above 99%). Finally, they suggest that the use of peak values when characterising load variability may result in more scatter than is truly representative of the inherent load variability.

Helicopter blades experience a constantly changing load resulting from aerodynamic and inertial forces during rotation. Ory and Lindert [12] developed a reconstruction method (RM) to determine airloads using a combination of eight spanwise blade strain gauge signals, the flapping angle, the azimuth position, and two accelerometer signals. They neglect coupling between flap, lag, and pitch motion. The blade is modelled as a slender, linear, elastic, hinged beam under centrifugal loading, while the mass distribution is modelled as lumped masses. The elasticity matrix was obtained from static deformation measurements. The RM requires modal parameters in the form of eigenmodes and frequencies of the rotating blades (which are stiffer than the corresponding stationary blades). These modal parameters were obtained from the stiffness matrix (which was stiffened numerically) taking into account the boundary conditions implied by the flap hinge. The RM essentially solves the equations of blade motion, and hence the blade deformation (derived from measured blade response data) allows the computation of blade forces.

The blade instrumentation signals were transmitted from the rotor to a receiving unit on the ground using a telemetry system. Very good receiving capacity is reported for distances of up to 150 m from the transmitter. Flight tests included hovering at altitudes of 2, 10, and 20 m and forward flight at speeds of 20, 40, 60, 80, 100, 120, and 140 km/h with two Hungarian Air Service helicopters. These helicopters were the Kamov-26 (which has co-axial counter-rotating rotors 1.2 m apart powered by two nine-cylinder radial engines) and the Hughes 500E (which has a five bladed rotor powered by a turbine engine). They suspect that blade-vortex interactions (vortices from the blade tip interacting with the pursuing blade) occur at low flight speeds and some hovering flights. Ory and Lindert recommend the RM as an easy and cost-effective alternative to complex pressure measurements on rotating blades.

Barndt and Moon [13] produced eight bar charts clearly depicting the differences between the assumed usage spectrum and actual usage monitored, based on 2100 flight hours from 35 aircraft. These eight bar charts are summarised in Figure 2.1, which show the ratio of the assumed usage to actual monitored usage based on data gathered for eight different manoeuvres from 27 different AH-1W helicopters. Both these usage figures were expressed as a percentage of time in flight condition. A box-whiskers diagram presents this usage variation information. (For more details on box-whisker diagrams see, for example, Goldman and Weinberg [14].) The upper and lower whiskers (vertical lines) show outliers in the data. The shaded box displays the range from the first quartile ($Q_1$) to the second quartile ($Q_2$, which is equal to the median), while the white box displays the range from
the second quartile ($Q_2$) to the third quartile ($Q_3$). Finally, the asterisk denotes the position of the data's mean value.

All flight conditions exhibit at least an order of magnitude difference in usage within the 27 helicopters monitored. The data in Figure 2.1 are correct to only two significant figures, since they were derived from the bar charts in Reference [13]. The actual proportion of time spent in severe fatigue-damaging flight conditions, such as gunnery turns and high-G gunnery turns, was found to be conservative compared to the assumed spectrum. Barndt and Moon caution that a few of ‘…these aircraft have only a small amount of flight time and may not be representative of long term use.’ They also state that ‘[w]ith such variability in accrued fatigue damage…it’s difficult to avoid substantially over penalizing many aircraft to protect the worst case usage.’

Fraser, King, and Lombardo [15] reported on the life substantiation program, for the U.S. Army’s UH-60 series Black Hawk helicopters, undertaken by Sikorsky Aircraft. The Royal Australian Air Force (RAAF) asked the Aeronautical Research Laboratory (ARL) to

---

1 The Aeronautical Research Laboratory was subsequently combined with the Maritime Research Laboratory to form the Aeronautical and Maritime Research Laboratory (AMRL).
comment on Sikorsky’s report [16] of the Structural Usage Monitor (SUM) program. (Component loads and aircraft parameters such as airspeed, altitude, and control stick position were measured in the SUM program.)

Fraser et al. state that personal communication (with a U.S. Army representative) pointed out the motivation behind the SUM program was to flight-test a usage monitoring system, and thus investigate the suitability of individual aircraft component life monitoring. Two other elements of this SUM program were to measure the usage spectrum and assess the impact on component retirement times, and to provide assessment of the degree of pilot-dependency of loads developed for a specified flight condition.

Loads were by far the most difficult parameters to measure; especially those on rotating components, with the conclusion that the SUM measured loads were not reliable. Although the flight condition recognition worked well, the SUM used was too complex, expensive, and difficult to maintain fleet-wide. The ARL analysis called into question Sikorsky’s findings that component lifing impact was nil for the U.S. Army’s Black Hawk fleet. The sensitivity of component retirement times to the usage spectrum implied that there is a need to obtain usage data for the Australian owned Black Hawks and Seahawks, and that some form of quantitative measurement of this usage is needed.

Using a three-parameter Weibull distribution Moon, Menon, and Barndt [17] modelled usage variation and variability in the alternating component loads associated with damaging flight conditions. (In the following paragraphs we use the word “usage” in the sense of Moon et al., namely, the amount of time spent in a particular flight condition.) The fifty AH-1W helicopters that were monitored (yielding 3400 valid flight hours of usage data) revealed significant variations in the usage between helicopters and a significant difference between the measured and the design mission spectrum.

The cumulative distribution of the Weibull function was linearised and the Weibull function’s parameters computed for 77 flight conditions at three gross weights using linear regression analysis. However, the resulting fits gave poor results. The data set was then partitioned into flight conditions and the Weibull distribution parameters recalculated. The mean variation-of-usage ranged from 0.11% to 51% for different flight conditions. For example, the fraction of time spent in forward flight was 45% according to the usage spectrum, but the measured fraction of time varied from 5% to 75% (with a mean of 51% variation) across the fleet. For forward flight, 20% of the fleet had usage that produced a damage rate for critical components lower than that according to the design usage spectrum, while 57% had a higher rate. Comparisons between the usage models of Weibull and normal distributions indicate that the Weibull model performs marginally better, especially towards the tail ends of the distribution.

The database consisted of rotor cycle alternating loads for all flight conditions in the original usage spectrum. The 1389 flight conditions (most flight conditions were flown several times) are identified by gross weight (low, medium, and high), centre of gravity (forward, mid, and aft), and altitude (0–3000 ft, 3000–6000 ft, and above 6000 ft). The loads
for the same manoeuvre flown under different conditions were combined to obtain the minimum, maximum, mean, standard deviation, and coefficient of variation (CV, defined as the standard deviation divided by the mean). The three-parameter Weibull distribution, fitted to the loads of 231 (77×3) flight conditions at three gross weights, passed the goodness-of-fit Kolmogorov-Smirnov test but failed the Chi-Squared test. (According to Press et al. [18] the Kolmogorov-Smirnov test is the most generally accepted test for continuous unbinned data as a function of a single variable.) The CV ranged from 15–31% for steady manoeuvres, and increased with the severity of the transient manoeuvre to a value ranging from 31–57% (CV examples include normal turns 30%, air combat 33%, gunnery turns 35%, and pullups 50%). Weibull distributions were also fitted to the results of fatigue specimen tests to determine the coefficient of variation. A value of 11%, which was consistent with the results reported by other researchers, was obtained.

Reliability computations were carried out using the joint (or multivariate) probability of the binned usage, load, and component strength Weibull distributions. All significant combinations of usage, load and strength were compared to obtain an ordering of fatigue life (based on Miner’s rule) and associated failure probability. There is an almost linear log-log relationship between probability of failure and hours to failure. For example, failure probabilities of $10^{-5}$ and $10^{-6}$ gave 2160 hr and 800 hr failure lives respectively. Given the effects of usage spectrum variation, usage monitoring with component lives based on a “six nines” (0.999999) reliability has the potential to increase the mean life of the pitch link by threefold. Finally, an increase in load severity from the 95th to 99.99th percentile resulted in a component life reduction from 5540 hrs to 110 hrs.
3. Deducing Loads on Dynamic Components from Loads Measured on Fixed Components

Gunsallus et al. [19] developed a linear relationship between fixed airframe loads and rotor loads (termed ‘holometrics’), using data from strain gauge readings on both sets of structures. In particular they develop a technique, based on vector orthogonalisation, which reduces linear dependence of the input data set, thereby rendering the resulting system of equations less ill-conditioned.

The fixed data and rotor data are placed in matrices that are linearly related by a transfer coefficient matrix. That is, if $T$ is the fixed data matrix, $B$ is the rotor data matrix, and $C$ is the transfer coefficient matrix (or calibration matrix), then a linear approximation exists and is given by $TC = B$. The input matrices $T$ and $B$, which are complex, contain the strain and phase angle of rotor RPM harmonic for different flight cases. The columns represent a particular strain gauge, and the rows are constructed from the coefficients of the Fourier series approximations of the loads over one main rotor cycle.

A Moore-Penrose generalised inverse (which is required due to the non-square nature of the matrix to be inverted) is used in determining the transfer coefficient matrix. (For a more detailed discussion of the Moore-Penrose generalised inverse refer to any text dealing with singular value decomposition and least squares, for example, Golub and van Loan [20] or Press et al. [18].) Each row of data input is passed through a Jones Euclidean Limit Length Orthogonalisation (JELLO) procedure (which is reminiscent of a conjugate gradient minimisation technique). This JELLO procedure allows the choice of successive columns on their degree of informational independence to the thus-far-calculated calibration matrix. Finally, they apply a sensitivity analysis to the solution process, taking into consideration the case of random error in the measurements.

In a later paper [21] Gunsallus reported on the application and test results of the holometrics system on a SH-2F helicopter. Gunsallus and Robeson [22] then applied the holometric technique to an Apache helicopter (AH-64A) to synthesise four components using eleven, six, or three input parameters. The four synthesised components were the damper force and twisting moment at station 104.5 of rotor blade lead-lag, lateral control actuator force, and vertical bending moment at fuselage stringer station 416. The 11 input parameters were: the axial forces in six main transmission support struts; actuator forces from the longitudinal, collective, and directional controls; and lower right and upper right axial forces in the aft engine mount. After initialisation of the calibration matrices, an independent data set was used to compare the synthesised and actual cycle counts (that is, the comparison used a blind test). They report excellent correlation, with the exact cycle counts slightly underestimated by the synthesis in most cases. The majority of the uncounted cycles occurred in the range where the amplitude of the load was low (that is, the least critical range).
4. Deducing Loads on Dynamic Components from Flight Parameter Measurements

Here we review work that estimates rotor loads using flight parameters (which are normally measured and sometimes recorded). This section is subdivided into four subsections: flight condition recognition, regression, neural networks, and combined regressions and neural networks.

4.1 Flight Condition Recognition

Tang and O’Brien [41] developed a pattern recognition method to correlate quasi-static flight parameters with load data from flight tests. (The equipment to monitor quasi-static flight parameters was already installed in the helicopter.) Time history inputs are used with pattern recognition to identify the type of load cycle spectrum for a particular manoeuvre and the load range of that cycle (normalised against the maximum load range for that manoeuvre). The data from flight tests covered 51 manoeuvres and included the pitch link axial load (sampled at 235 Hz), main rotor torque, airspeed, roll and pitch attitudes, and vertical acceleration (all sampled at 30 Hz). Forty-eight manoeuvres could be categorised into one of four load cycle spectra. These categories were harmonic steps (all level-flight flight conditions), mixed steps and exponential decay (most turn flight conditions), rapid drop (sudden-change manoeuvres, for example jump take-off), and exponential decay (the remaining manoeuvres). Normal start, shutdown, and partial power descent could not be categorised into any of these four load cycle spectra and so were not included in the study. A number of apparently similar manoeuvres (longitudinal, lateral, and pedal control reversals) were unexpectedly classified into different categories.

The pattern recognition is based on the recognition of waveform and spectral form features of the input. Due to the small number of manoeuvres in the rapid drop category and the interpretation subjectivity of the mixed steps and exponential decay category, only harmonic steps and exponential decay were used in the training. This reduced the training set to 30 manoeuvres. Three different classification functions were tested: Gaussian discriminate (a plane dividing two patterns), Fisher discriminate (a line dividing two patterns), and nearest neighbour criteria (using the closest points and a majority rule to determine the pattern). A comparison of erroneous classifications of these methods showed that for both single and multiple parameter inputs the 3-point nearest neighbour criterion performed the best.

Three methods were compared against the ‘exact’ solution, which uses rainflow counting to determine fatigue usage. These three methods were spectrum recognition, load range determination, and on-line monitoring (a combination of the previous two methods). The spectrum recognition method uses the load range obtained from rainflow counting instead of the quasi-static parameters. This partitioning of the comparison allowed independent analyses of the spectrum recognition method and the load range determination methods.
The spectrum recognition method worked well (resulting in a small amount of scatter) despite consistently over-estimating the pitch link fatigue life usage. (This over-estimation was expected since the spectra used in the analysis were the upper bounds of the classification spectra). The on-line method resulted in a larger scatter than the spectrum recognition method, but this time the fatigue damage was under-estimated for some components and over-estimated for others, the method performing worst for low loads. A cumulative fatigue damage comparison exhibited better results for both the spectrum recognition and on-line methods.

Barndt and Moon [13] developed a flight condition recognition monitoring system for the AH-1W helicopter. Flight data were post-processed to determine fatigue damage in components for each aircraft. The monitoring system enabled identification of all flight conditions, and determination of individual aircraft usage and component fatigue damage. Each input parameter was binned (for example the range of vertical accelerations was partitioned into six bins). Using bin combinations for defining flight conditions, a flight condition recognition system was developed for the AH-1W helicopter. An example of a speed-altitude bin combination is speed in the range 10–20 kn and altitude in the range 1000–1500 ft. Using bin combinations alone was insufficient to distinguish certain flight conditions, and so the timing and sequence in which bin combinations occurred were also taken into account. Taking into account subdivision of various parameters (twin or single-engine modes, airspeed, load factor, bank angle, rate of climb, gross weight and altitude) 310 different flight conditions were defined. Flight trials permitted the use of an iterative process to manually correct for any unrecognised flight conditions, thus improving future flight condition recognition.

Post-processing the duration and sequence of flight conditions allows the determination of component damage. As already discussed, Figure 2.1 shows the variation in usage across a fleet of AH-1W helicopters. The fatigue life impact on three components (main rotor pitch link, swashplate anti-drive bellcrank, and swashplate inner ring) was considered for each monitored aircraft. Differences in component retirement times exceeding an order of magnitude were detected across the fleet.

4.2 Regression

Using multiple regression Haas [23] determined rotating component loads (main rotor blade bending and pushrod load) from fixed system measurements on a U.S. Navy SH-60B helicopter. Fixed measurements included pedal, collective, lateral, and longitudinal stick positions, roll, yaw, pitch, and climb rates, load factor at CG, and velocity. The coefficients of the regression were obtained using a least squares fit of the data. The $F$ statistic was used to determine the significance of each input on the unknown load, and a tolerance limit was used to control the collinearity of different inputs. Loads data were filtered into a min/max format over each rotor revolution, from which steady (min+max)/2 and alternating (min-max)/2 components were derived. This study focused on high-G and high-load manoeuvres, which are responsible for a major proportion of the fatigue damage. These
manoeuvres comprised eleven symmetric pullouts (SPO) and five rolling pullouts (RPO). Identification of non-damaging (below fatigue threshold) loads (such as those developed in climbing turns and level flight) is less critical. The three autorotative pullouts recorded were considered insufficient to be included in the modelling. The SPO flights included various pilot techniques and were preformed at different velocities and at maximum load factors.

Haas warns of the importance of having data with adequate variation of the independent inputs. Filtering out the low-G pre-maneuouvre flight section improved the model by creating a high-G bias. The SPO manoeuvre correlation coefficients were 0.88 and 0.81 for the pushrod and blade bending loads respectively. Cross coupling of inputs was also considered in the RPO model yielding a total of 29 possible regression inputs. The RPO model resulted in correlation coefficients of 0.91 and 0.89 for the pushrod and blade bending loads respectively. To improve the SPO model, two RPO manoeuvres were included and coupling terms were allowed; this yielded correlation coefficients of 0.95 and 0.89 for the pushrod and blade bending loads respectively.

Haas and Imber [24] extended previous work [23] by using a multiple regression analysis of fixed system parameters to predict rotating component loads. They report correlation coefficients above 95% for steady flight and 79–95% (described as ‘possible’) for manoeuvres contained in the data set used for regression. A blind data set was used in the verification of the regression model.

A regression model was developed for several different manoeuvres including symmetric pullout, rolling pullout, climbing turns, and level flight. In addition, measuring the position of the swashplate servo instead of the control stick meant that both the pilot and stability inputs were taken into account, instead of just pilot input. The main rotor swashplate collective, cyclic control inputs, and tail rotor collective, were calculated directly from the swashplate servo measurements. In addition, aircraft accelerations were used instead of angular rates. Other measured parameters included six accelerations, airspeed, aircraft mass, rotation rate, and rate-of-climb (in all 15 parameters were measured). Estimations of both the main rotor pushrod load and blade normal bending load (at the 15% radial station) were developed, with Haas and Imber stating that these loads should be representative of those applicable to the other critical rotating system components as well.

Moderate-to-high-speed level flight and several flight manoeuvres were analysed using stepwise multiple regression and least squares. In stepwise regression, variables are entered and removed from the regression in steps based on the significance of the \( F \) statistic. A tolerance limit controls multi-collinearity (where one of the regression parameters is expressible as a linear combination of other regression parameters). In addition, a non-linear velocity model was chosen to fit the data for steady level flight (an iterative technique was required to solve for the regression coefficients).

The helicopter used in the flight test program was the U.S. Navy SH-60B, which flew 35 individual manoeuvres lasting 10–90 seconds at the same altitude and nominal gross
The study aimed to predict alternating component load (the amplitude of the load over one main rotor revolution) during a given manoeuvre.

The measured data were filtered using deletion, smoothing (via averaging), and bin separation. As noted earlier by Haas [23], coupling of regression parameters (such as collective and longitudinal stick position) led to an improved model. A set of perturbation parameters (the absolute value of the difference between the control input and the steady flight value) was also found to give good results when used in the regression analysis. The high-load data points were given a higher weighting (and thus higher influence) in the regression model when compared to low-load data points.

Using 24 flight parameters, Tang and Moffat [25] used stepwise multiple regression to model the loads experienced by three dynamic helicopter components: pitch housing lead-lag lug, main rotor blade, and tail rotor fork. Nineteen flight test conditions were used to formulate the flight load models, with an additional four flights used to validate the model (in a blind test). Three load models were developed for each component: maximum mean load, minimum mean load, and maximum load range. In addition, 10 signal features were also extracted from each of the 24 flight parameters. These 10 features were the mean, mean square, maximum, minimum, amplitude, and standard deviation of the time signal, as well as normalised time duration above 25, 50, 75, and 90% of the peak.

The step multiple regression models began with a linear approximation and subsequently added higher order non-linear terms. These terms were either introduced, based on the $F$-to-enter statistic, or removed on the $F$-to-remove statistic. The authors cite Montgomery and Peck [26] for further details. Three flight load models were formulated for each component: maximum load range (RNG), maximum mean load (MAX), and minimum mean load (MIN). The pitch housing lead-lag lug models for the MAX, MIN, and RNG (MMR) were linear and had correlation coefficients of 0.80, 0.96, and 0.96 respectively. The main rotor blade models for the MMR had correlation coefficients and model types of 0.70 and linear, 0.89 and linear, and 0.95 and second order respectively. Finally the tail rotor fork models for the MMR had correlation coefficients and model types of 0.92 and second order, 0.92 and second order, and 0.96 and non-linear, respectively. The flight test parameter sampling rate varied (depending on the parameter) between 60–500 Hz. In contrast, the operational flight data recorder’s sampling rate varied only between 1–20 Hz. No degradation in performance was noted when the sampling rate was reduced from the flight test rate to the operational flight rate. The models developed were partitioned manually into time windows, varying between 0.5–22 seconds, to coincide with the duration of the manoeuvre. When the data were later partitioned into constant six-second duration windows, the results of all models were degraded. When the data reduction and a constant six-second time window were used on the blind data sets, the models were no longer adequate. The authors state that one of the four blind test data sets was spurious, and once removed the results improved.
In a later paper, Tang and Moffatt [27] used pattern recognition and multiple linear regression to develop a helicopter on-board fatigue processor, making use of Miner’s rule with a Goodman correction. The pattern recognition involves the following four steps:

1. Signal waveform measurement (termed pattern measurement).
2. Information analysis and combination, to optimise classification (termed feature extraction).
3. Compilation and analysis of features with known outcomes to obtain optimum features (termed training).
4. The use of training to decide outcome of future inputs (termed classification).

Their processor performs the following tasks in sequence: data features calculations, load-cycle spectrum classification, maximum load range determination, and fatigue usage calculations. Tang and Moffatt state that the vast amount of load-cycle information can be reduced to load-cycle spectra and maximum load ranges. The load-cycle spectra give the number of cycles at various load ranges for a given time window. The maximum load ranges are the largest difference between the highest and lowest stresses for a given time window. The flight of the helicopter is divided into contiguous constant time intervals (time windows), which have two conflicting length requirements. The interval has to be small enough to retain response characteristics to accommodate different manoeuvres, but long enough for the processor to compute fatigue damage within a defined time interval.

Three non-airframe components were chosen to validate the processor: pitch housing lead lag lug, main rotor blade, and tail rotor fork. Nineteen flight test conditions were used for fatigue processor training, and four conditions were set aside for verification. The duration of flight tests was between 8 and 73 seconds. Four flight load measurements were directly related to the three selected components. The 29 input flight controls included: airspeed, engine parameters, rates and attitudes of both pitch and roll, accelerations, control and actuator positions, chordwise bending moment, and pitch link and leading edge damper loads. Although the sampling rate in this study was high, analysis at a lower sampling rate (which is more typical) showed no performance degradation.

Ten waveform features were defined for each flight control data set. These were mean value, mean square value, maximum value, minimum value, amplitude range, and standard deviation of the time signal, and normalised time duration above 25, 50, 75, and 90% of the peak. The identification of a manoeuvre’s start and finish was a difficult task. The selection of the training-set time windows, which ranged from 0.5 to 22 seconds, was based on engineering judgement of a significant change in response characteristics. Tang and Moffatt use 90 time windows to train the pitch house lead lag lug response, and due to questionable data sets, only 57 time windows were used for the main rotor blade. Since high frequency fluctuations were superimposed on a low frequency fluctuation, 145 time windows were used to characterise the response of the tail rotor fork.

Each component had its own multiple linear regression model, independent of the load-cycle spectrum types. A stepwise $F$ statistic was chosen to obtain a subset of input features for each model. An additional model that incorporated mean loads was developed for each component.
Validation involved a comparison of the processor under two simulated conditions and the processor under cycle counting of the measure loads for each of the 19 test conditions. A uniform six-second time window, being the mean duration of all time windows used during the training stage, was used in the processor implementation. The long simulated condition, of 960 seconds duration, was a combination of back-to-back sequences of all the 19 test conditions. The short simulation condition, of 550 seconds duration, was a level flight condition constructed from a random selection of only the first window of each test condition. Flight test conditions that accumulated no fatigue damage were excluded. No comparison figures are given in Tang and Moffatt’s paper, but fatigue usage log-log plots of the processor versus the cycle counting method are shown. The pitch housing lead-lag lug had the best comparison; however, one test condition showed fatigue by the processor but none from cycle counting. Most of the test conditions accumulated no fatigue usage for the main rotor blade, the remaining conditions exhibiting a reasonable comparison. Three flight test conditions showed fatigue usage by the processor but none by cycle counting, while one test condition showed usage when cycle counted but not by the processor. The tail rotor also showed reasonable comparisons. However, one flight test condition had a much lower fatigue usage from the processor as compared to cycle counting. All three components had good processor-to-cycle counting comparisons for the total fatigue usage (that is, all flight test conditions incorporated). Similar results were obtained for fatigue usage that included the effects of mean stress. An excellent correlation was obtained between the processor and cycle counting for the long flight test simulation, both with and without the consideration of mean stress. The authors note that the use of a time window does not account for the effect of slow varying, large amplitude cycles on fatigue usage (such as ground-air-ground cycles).

4.3 Neural Networks

For an introduction to neural networks see, for example, either Lippmann [28] (an article) or De Wilde [29] (a monograph). We will assume an elementary understanding of common terminology and techniques in this section.

Cook et al. [30] have investigated the use of neural networks for the estimation of dynamic helicopter loads. They use a component load and flight parameter database to determine the weighting coefficients of the neural network. The load-parameter database corresponds to histories taken from an instrumented military helicopter undergoing standard manoeuvres. Eight standard flight variables are used as inputs for predicting the time-varying mean and oscillatory components of the tail boom bending load and the pitch link load for seven flight manoeuvres. They report 90.7–97.7% accuracy in the neural network identification of mean and oscillatory components.

The network used was a feed-forward model with backpropagation (Cook et al. state that this type of network minimises the mean square error). The linear output of a hidden unit is the weighted sum of the inputs. The non-linear output merely involves the application of a
sigmoidal function to this linear output. The output of all hidden units (on a particular layer) are weighted and summed to produce a single output. The weights of individual hidden units and the layer of hidden units are determined using a backpropagation training scheme. Essentially the backpropagation scheme involves incremental changes of the weightings to reduce the error between the estimated output and the desired output using error surface gradients. (The authors reference Nelson and Illingworth [31] or Lippmann [28] for further neural network details.) The network was configured with four hidden units and its training tolerance was set to 2% (if the error between the estimated and desired outputs is less than the training tolerance, no weighting changes are made).

The database was generated from a helicopter heavily instrumented (in a separate structural analysis of 17 safety-critical components and 21 fatigue-sensitive locations) to monitor flight parameters and loads. The data consisted of time histories of 24 flight parameters and 17 component loads measured during seven standard flight manoeuvres. Each manoeuvre consisted of a step change in control stick position, either 0.5 in or 1.0 in, from a steady forward flight position. Of the 24 flight parameters, eight were chosen as network inputs (based on intuitive understanding of motion and loads). The chosen parameters were pitch, roll, and yaw rates, vertical, lateral, and longitudinal accelerations, and longitudinal and lateral control positions.

Only a small segment of data from each manoeuvre (two seconds centred around the stick position change) was used as network input, since each manoeuvre contained a substantial amount of steady-state data where neither flight parameters nor loads changed. A frequency analysis revealed two principal cyclic components, 4.75 Hz (main rotor frequency) and 19 Hz (blade passing frequency).

The network was trained to either estimate time-varying mean or time-varying oscillation of the tailboom bending load and pitch link load. This separation into oscillatory and mean components was necessary since none of the measured flight parameters varied in a sinusoidal fashion that was directly correlated to the main rotor blade motion. Training included data from six manoeuvres, while testing was conducted on data from the seventh manoeuvre. This procedure was repeated on each of the seven manoeuvres to test the estimation accuracy. Cook et al. state that randomising the time-history input speeds up the evaluation of network coefficients (they reference Nelson and Illingworth [31]), but warn that this procedure reduces the network’s ability to estimate temporally varying characteristic information.

The mean scaled percentage errors of the 0.5 in and the 1.0 in manoeuvres were respectively: 6.9% and 9.3% for the mean tailboom bending load, 2.3% and 5.2% for the oscillatory component of the same load, 4.3% and 6.5% for the mean pitch link load, and 3.2% and 6.8% for the oscillatory component of the same load. The higher error in the 1.0 in manoeuvres is explained as a result of estimation extrapolation (instead of interpolation) of the training data set.
Azzam [32] used a neural network to predict fatigue damage from flight parameters. The model input parameters have been classified as direct parameters (relating to rotor and engine systems, flight controls, etc.) and indirect parameters (speeds, gross weight, gearbox sensors, etc.). The parameter inputs were collective, cyclic, and lateral stick positions, tail rotor pitch angle, temperature, pitch and roll attitudes, main rotor speed, airspeed, roll, pitch, and yaw rates, normal and lateral accelerations, and both engine torques. The control inputs are transformed through the rotor system into high frequency varying loads that have components related to multiples of the rotor frequency.

The network model for the blade lug section performed poorly under some conditions (for example, high forward speed in turbulent weather with a 15 kn wind, and quartering flight at 10–30 kn with a 30°–90° starboard wind). Combined wind speed and quartering flight angle can move the main rotor wake close to the tail rotor, inducing appreciable high frequency vibration. Overall however, the network modelled the fatigue damage to within 5% of that obtained with strain gauge measurement, with similar results for a rotating pitch link. The damaging events which could not be estimated were quartering flight at 10–30 kn with a 30°–60° starboard wind, and 20°–45° bank turns to port and starboard at 125–145 kn. During the data acquisition period the stick positions were adjusted three times, but the stick position calibration factors were not available to Azzam. In addition, the blind test data included manoeuvres that were not present in the training data set. Both these facts led Azzam to believe that the network model possesses good generalisation capabilities.

Using data from a CH-46 helicopter, Hoffman [33] presented the effect of training data selection on the load estimation of a neural network model. Several modifications were made to the data sets including deleting redundant data, deleting variables that are highly correlated, determining regions of interest (explained below), and the removal of ‘contradictions’ in the data set. The removal of contradictions involved iteratively noting what effect removing a data record had on load estimations.

The data from a CH-46 extended stub wing flight test consisted of 99 variables of which nine were chosen as potential input variables and three as output variables. The potential input variables were vertical acceleration, collective lever, longitudinal stick, lateral stick, and directional pedal positions, pitch, roll, and yaw rates, and airspeed. The output variables were forward and aft blade flap bending loads and forward blade damper load. Taking out ‘extraneous’ spikes and interpolating 50 evenly spaced data points per main rotor blade revolution produced 81,900 data points. Ten manoeuvres were used for calibration and nine manoeuvres were used for verification. A universal model (one that modelled all manoeuvres) was found to produce excessive smoothing in load estimations. This smoothing did not account for peaks and valleys within the data, and so the model was altered to predict specific manoeuvres. Using a variance comparison the nine input variables were reduced to six by choosing those with greatest variance (the remaining three variables did not vary enough to be included in the neural network model). These six variables (vertical acceleration, lateral and longitudinal stick positions, pitch, roll, and yaw rates) had low correlation coefficients. Regions of interest (the range of space that data occupy in a dimension) were identified using Fuzzy C-Means method cluster analysis.
Hoffman concludes that identification of regions of interest was computationally intensive, with results no better than a simple random sample approach. Furthermore estimations did not capture extreme peaks and valleys, and so the cluster analysis was discarded.

Hoffman modelled a single cycle of output (at 30 kn level flight) using a neural network trained on 38 data records. The network had six inputs, five hidden neurons (on a single layer), a log-sigmoid transfer function from the input to the hidden layer, and a linear function from the hidden to the output layer. The sum-squared error for the network was $9.96 \times 10^{-3}$ after 66 epochs. Hoffman suggests using a log-sigmoid function from the hidden layer to the output layer to ensure values do not swing widely from the anticipated range. A larger training set (200 randomly chosen records out of 346) was then tested on a network with 10 hidden neurons (all other parameters remained the same), resulting in a sum-squared error of 2.657 after 100 epochs. Hoffman states that the ‘...results are not great, but they seem to be conservative.’ In discussing the blind test results for a level flight at 105 kn, Hoffman states that ‘[i]t looks like the objective to predict raw peaks and valleys is hopeless with these sets of input data and variable selection. The same conclusions have been drawn by Berens who used regression on this data and Tsai who used the holometric approach...’. Hoffman suggests that rainflow counting the data first should improve the model by optimising the network coefficients with the most important ranges.

Cabell, Fuller, and O’Brien [34] used a feed-forward neural network to predict the load in a pitch link, using information measured by the AH-64A helicopter’s flight control computer. Only the manoeuvring portions of flights were used to train the network. Flight loads and fixed system data were measured on 13 flights, with eight of these flights involving fatigue damage (as estimated by cycle counting) to the pitch link. Using frequency decomposition and a bandpass filter centred at 4.75 Hz with a 1.5 Hz bandwidth, the load data set was reduced to a load range over each cycle. The manoeuvre data sets were randomly split into seven training sets and six verification sets.

The 11 measured inputs were airspeed, collective actuator position, and the rate, attitude, and actuator position for pitch, roll, and yaw. Dynamic information was included in the model by evaluating the changes in these 11 parameters over both one and two loading cycles. The inclusion of these 22 derived parameters brought the total number of network inputs to 33. The input variables were normalised to map them between zero and unity. In comparison to the loads, the input variables changed slowly and so were measured only at the beginning of each loading cycle. The network output gave both the mean and range of load over a single cycle. (The authors state that the estimation of load at each sample time was abandoned due to inaccuracies.)

The feed-forward neural network consisted of 33 inputs, 10 hidden layers, and 1 linear output. The network was trained several times using different randomly initialised weights, each time converging to one of two solutions (the two solutions performed almost identically on the validation data set).
Cabell et al. define $r^2$ as the square of the multiple correlation coefficient. The results return $r^2$ values between 0.86–0.96 and 0.60–0.90 for the training and validation data sets respectively, with a mean $r^2$ value of 0.72 for the validation data set. Comparisons are also made with a linear model that achieved $r^2$ values between 0.41–0.92 and 0.45–0.88 for the training and validation data sets respectively. The dispersion of estimation around small loads is cited as the reason for low $r^2$ values. The authors indicate that the $r^2$ values are ‘...not as high as might be needed for precise fatigue damage prediction...’.

To convert predicted loads (using the network model) and measured loads to fatigue damage using Miner’s rule, a comparison between the predicted and measured loads was carried out. However, due to the tiny fraction of fatigue life consumed, the results are particularly sensitive to the few high load cycles that occur in a given flight. The predicted fatigue damage was high on three flights and low on the other two when compared to the measured damage on the validation data set.

McCool, Flitter, and Haas [35] developed and evaluated a neural network model of vibratory loads on an SH-60B helicopter. In order to ensure that sparse data spaces are represented a self-organising feature map (SOFM) was used to develop a map of the input space. The SOFM transformed the 12-dimensional input space into a 2-dimensional space (a 50 by 50 node space) with 2500 data bins. The authors claim that this approach yields a data set that closely resembles the data population and includes sparse data regions. The load data set population was skewed towards low loads (since helicopters spend a large fraction of time in steady flight and hover). The importance of high loads was increased by using weightings on the data points.

A modular neural network (a collection of expert networks using supervised learning [36]) was used to predict load. Each of these sub-networks is an expert in some area of the problem domain, with a gating layer mechanism that learns how to partition the data space across the expert systems. The modular network consisted of a single gating network (with nine hidden nodes) and three expert networks (each having 65 hidden nodes). The 12 input parameters are airspeed, main rotor speed, engine torque, load factor, pitch and roll attitudes, lateral, longitudinal, and collective stick positions, pitch and climb rates, and gross weight. The authors use approximately 400 passes of the delta training rule. The baseline data were selected from 56 load survey flights yielding approximately nine flight hours (which when sampled at 8 Hz resulted in 255,401 data points). A 95th percentile analysis (binning flight conditions into 50 lb intervals) showed that turns and symmetric pullout manoeuvres contain high loads for both the pushrod and damper. Additionally, high damper loading occurs during low speed flight and power dives.

The authors argued that since the training data set (7500 points) was so small compared to the total baseline (255,401 points), the training data set inclusion in the validation data set (the entire baseline) would not significantly modify the results. The correlation coefficients between the measured and predicted loads for the main rotor pushrod and lead-lag damper were 0.94 and 0.78 respectively. The root-mean-square (RMS) errors were normalised as a percentage of the maximum measured loads (5000 lb and 7000 lb for the
pushrod and damper respectively), and were 3.65% and 7.62% for the pushrod and damper respectively. The pushrod load measurements and estimations were within 7.3% of the maximum load (that is within twice the RMS error) for 95% of the validation data. Separation of the damper loads into those corresponding to high (above 60 kn) and low (below 60 kn) speeds modified the results giving correlation coefficients (and normalised RMS errors) of 0.88 (and 4.14%) and 0.61 (and 13.51%) for high and low speeds respectively. The suspected causes of estimation variation were a change in characteristics (occurring between 60 and 70 kn) of the damper force-velocity curve and unreliable airspeed readings under 40 kn. The normalised RMS errors for each flight condition for both pushrod and damper are also given.

A flight test evaluation was conducted on an SH-60F, which had strain gauges on the main rotor pushrod, lead-lag damper, stationary scissors, and the lateral and longitudinal servo links. The network estimations were performed in real-time using an 8 Hz sample rate for network input parameters. The flight test accumulated approximately eight hours of data or 228,704 data points. The flight tests were within the normal operating flight envelope, unlike the baseline data that were accumulated under more severe operating conditions in order to train the neural network. The correlation coefficients (and normalised RMS errors) were 0.91 (and 2.11%) and 0.84 (and 7.24%) for the pushrod and damper respectively. As before, speed separation of the damper loads gave different results, correlation coefficients (and normalise RMS errors) of 0.83 (and 3.22%) and 0.49 (and 16.06%) for high and low speeds respectively. Only pushrod and damper loads were analysed by the authors, who state that the favourable results demonstrate the potential to train the network on one aircraft and use the results on a second aircraft. Correlation coefficient and normalised RMS errors are given for each manoeuvre; again they compare favourably with baseline results.

Flitter, McCool, and Haas [37] developed a neural network model to predict vibratory loads on an instrumented SH-60 helicopter (using a slip-ring and strain gauges on the rotating system). A self-organising feature map algorithm was used for training, and a modular network for load estimation. Loads were accurately modelled, and good correlation was obtained in predicting loads on a different SH-60 helicopter.

Three data sets were obtained, two from the first SH-60 helicopter (Aircraft 1) and one from the second SH-60 helicopter (Aircraft 2). The baseline data set, 2.8 flight hours from Aircraft 1, included high-G pullouts, turns, dives, autorotations, partial power descents, takeoffs, landings, level flight, and hover. The evasive manoeuvre (EVM) set, 2.2 flight hours from a separate test program on Aircraft 1, included manoeuvres that are performed while evading an aggressor aircraft. The final data set comprised one flight hour from Aircraft 2. It contained a spectrum quite similar to the baseline data set and was used to determine the transferability of the resulting network.

Less than 10% of the baseline data set was used for training. The remaining 90%, the EVM, and Aircraft 2 data sets, was used in the validation of the resulting network. A $50 \times 50$ SOFM was created with the baseline data set. Choosing an equal number of data points from each bin created a training data set that included sparse data areas (such as high loads, which are
important from a damage point of view). Previous work by Flitter et al. [38, 39] showed a
tendency to under predict high loads. In an attempt to correct for this under-estimation, a
1% bias was introduced by multiplying the actual loads by 1.01.

A modular neural network (MNN) resulted in lower average root mean square (RMS) error
and better mean correlation, when compared to several other networks including
backpropagation. The MNN network used was the ‘Adaptive Mixtures of Local Experts’,
which consists of several experts and a gating network (Flitter et al. cite [36]). The inputs
were airspeed, rotor speed, engine torque, load factor, pitch and roll attitudes, lateral,
longitudinal, and collective stick positions, aircraft weight, and pitch and climb rates. All 12
inputs are fed into all three expert systems (each with 65 hidden nodes) and the gating
network (with 9 hidden nodes). The gating network sets the weights of the outputs of all
the expert systems (the weights add up to unity) and sums the result. Network training
uses a delta rule. The network was trained for approximately 400 passes through the
training data set. The two measures used to assess performance were the correlation
coefficient and the RMS error (which was non-dimensionalised by dividing it by the load
range).

For the baseline validation set the correlation and RMS error were 0.98 and 2.2% for the
pushrod. The RMS errors were less than 4.3% for the damper and servo. The damper loads
were partitioned into those corresponding to low (less than 60 kn) and high (greater than 60
kn) speeds. This improved the RMS error for the low speed model to 3.3%, but resulted in a
deterioration of the high speed model to the extent of a 6.9% RMS error. The EVM data set
saw a decrease in modelling accuracy resulting in RMS errors ranging from 4.2% to 5.4%.
The estimation of the Aircraft 2 data set was better, with RMS errors ranging from 2.5% to
3.5%. This suggests that a network developed on one aircraft is probably transferable to
similar aircraft.

**4.4 Combined Regression and Neural Network Models**

Haas, Milano, and Flitter extended previous work on load estimations using regression [24]
to load estimations using neural networks [39]. Hass et al. developed a neural net model,
which is forward-chained and fully connected, and then combined it with their previous
regression model. The two learning models used were backpropagation and the extended
delta bar delta (EDBD) algorithm (they reference Zeidenberg [40] for further details). In the
backpropagation model the error between the network output and the desired output is
propagated back through the network via the delta rule (maximum gradient descent), and
used to adjust the connection weights so as to minimise the error function. The EDBD
model enhances the backpropagation model by using individual dynamic learning and
momentum terms for each connection.

The database was populated with component loads and fixed system parameters from an
SH-60B helicopter. The loads were initially sampled at a high rate, and the minimum and
maximum values over each revolution of the main rotor were extracted. Over each cycle
Hass et al. defined a steady load as (max+min)/2 and a vibratory load as (max-min)/2. The steady and vibratory loads can be thought of as mean and amplitude loads respectively. Over 1000 data points were contained in the training data set, which included airspeeds from 50 kn to 160 kn, load factors from 0.65 G to 3.3 G, six symmetric pullouts (with different pilot techniques), five rolling pullouts, and two climbing turn manoeuvres. The validation data set was chosen independently of the training data set, but within the parameter and load range limits. The validation data set contained nine symmetric pullouts, four rolling pullouts, and three climbing turn manoeuvres, which resulted in over 2600 data points.

Initially the 15 parameters chosen as inputs to the network model were the three main rotor control servo positions, one tail rotor control position, the horizontal stabilator position, six accelerations, airspeed, rate of climb, gross weight, and rotor rotation rate. The three component loads were pushrod load, blade normal bending moment (at the 15% radial position), and the rotor damper load.

Haas et al. compared the correlation coefficient and root mean square (RMS) error of a single-hidden-layer network with 5, 10, 15, and 20 nodes. (They later added 3, 4, 6, and 8 nodes to check for sensitivity around the 5-node network.) They found that the 5-node network gave comparable (usually superior) results to all other node combinations, when each network was trained for 10,000 iterations. Correlation coefficients of 96% for pushrod load (with RMS error of 239 lb), and 93% for the blade bending load (with RMS error of 2465 in lb), were achieved. A three-hidden-layer model confirms that a single-hidden-layer is sufficient for this problem. Furthermore, 3, 7, and 10-node EDBD network models were tested (as a check for global convergence) without yielding significantly different results to those obtained using the backpropagation model. When the number of training data points was reduced by 25% no appreciable difference in results was noticed, and when the data set was reduced by 50% only a slight degradation in results was reported. All the results quoted above are for a single output network (for example a model that predicts either pushrod load or blade bending load). The addition of eight parameters (three angular attitudes, three angular rates, and the torque of both engines) reduced the RMS error. However, the models perform worse under some manoeuvres with these additional parameters (Haas et al. label this a deterioration of the model’s ability to generalise). When compared to a multiple linear regression model, the RMS error of the 5-node network model was 35% lower for the pushrod load and 21% lower for the blade bending load. A similar analysis for the network model of the main rotor damper load produces a correlation coefficient of 90–93%; again the network model’s performance exceeds that of the regression model.

The inclusion of derived parameters (for example coupling control inputs and airspeed) increased the accuracy of the multiple regression model making it comparable to the network model (the network model maintained a slight accuracy advantage). However, an improvement in accuracy and generalisation capabilities was noted in the network model of pushrod load and blade bending load when these same derived parameters were
included in the network training. No such improvement was noticed in the network model of the damper load.

Finally Haas et al. develop a hybrid model that uses a multiple linear regression equation to predict the loads and a neural network to predict the remaining non-linear residual terms. The regression model is based on 12 derived parameters for pushrod and blade bending loads, and 9 derived parameters for the damper load. A neural network is then trained, using 15 basic parameters, on the residual of this linear regression model. Comparing just accuracy, the network model with derived parameters produced the best results.

As part of a continuing series of papers Haas, Flitter, and Milano [38] investigated the important features of load monitoring. These include the determination of the most significant fixed system parameters, important flight conditions, and the effect of a calibration data set on the load estimation.

The data set was composed of over 150 distinct flight manoeuvres from an SH-60 helicopter (each manoeuvre lasting between 10 and 60 seconds), which resulted in over 18,000 data points. The pushrod, blade normal bending, lag damper, and main rotor shaft bending loads were thought to be indicative of overall rotor behaviour, and as such these were the loads that were modelled. The 18 fixed parameters included six accelerations, velocity, aircraft mass, rate-of-climb, altitude, the torques from both engines, rotor rotation rate, four pilot control positions, and stabilator position. The manoeuvres were grouped as hover, level flight, turns, climbs, dives, partial power descent, sideslip, control reversals, pulls, and autos. Each group contained specific manoeuvres; for example the hover group contained approach to hover, hover control reversal, jump takeoff, and hover in and out of ground effect.

The neural network is the same as that described in Haas et al. [39], that is a single-hidden-layer model with five nodes and a single output. Both regression and neural network models were produced for the training data set as a whole, and also for flight condition partitioned training data sets. The results show that for the low speed hover flight condition the load estimations are generally poor. The correlation coefficients for the hover flight condition for the regression (and network) models were 44% (and 30%) for the pushrod, 14% (and 31%) for the blade bending, 59% (and 59%) for the damper, and 42% (and 43%) for the shaft bending loads. Interestingly, partitioning the training data set into flight conditions didn’t improve either model significantly, and in fact increased the error considerably for some of the loads! The hover correlation coefficients for the flight-condition-specific regression (and network) models were 56% (and 43%) for the pushrod, 15% (and 38%) for the blade bending, 60% (and 29%) for the damper, and 39% (and 10%) for the shaft bending loads. The particularly poor estimation of damper loads is thought to be due to the hydraulic nature of the damper, as compared to the constant stiffness of the pushrod, blade, and shaft. The damper exhibited non-linear behaviour for low loads but, due to the action of a pressure relief valve, the damper characteristics become less variable for high loads. To some extent, this explanation accounts for the large difference in the
damper load predictive capabilities for different flight conditions. The partial-power/autorotative flight condition was also difficult to predict.

A list of the most important parameters for the neural network model of each component follows. Important parameters for the pushrod load were load factor, velocity, first engine torque, and longitudinal cyclic control position. For the blade bending load, they were velocity, load factor, longitudinal cyclic control position, and first engine torque. For the damper load, they were velocity, load factor, first engine torque, altitude, and longitudinal cyclic control position. Finally, for the shaft bending load, they were load factor, second engine torque, velocity, longitudinal acceleration, aircraft mass, and longitudinal cyclic control position.
5. Deducing Loads on Dynamic Components from Flight Parameter and Fixed System Load Measurements

The method of holometrics \[19, 21, 22\] was further developed by Gustavson, Pellum, and Robeson \[42\]. They used existing flight load data to determine the most cost-effective combinations of fixed system parameters and data sample rates to be used to accurately synthesise rotating component loads for the AH-64A helicopter. These parameters included both currently monitored and newly added parameters.

The use of Fourier coefficients in the original holometrics model was modified to reduce the amount of computation. Gustavson et al. now develop a system of equations for a few parameter-azimuth combinations (thus variations in rotor speed need to be taken into account). The rows of the input matrices (fixed and rotor data matrices) now contain time history information instead of Fourier coefficients. The minimisation of a cost-function determined the best combination of parameters and associated sampling frequencies that would still achieve a required error tolerance in the rotating component load estimation. A high cost was assigned to both high frequency sampling and the need to measure extra parameters, while a low cost was assigned to low frequency sampling and the inclusion of currently sampled parameters only. Other costs assigned to the cost-function included engineering development, equipment acquisition, installation, field support, power required, weight, volume, and amount of processing required.

A parallel-orthogonal vector reduction algorithm was used. This algorithm selected the best potential starting parameter and then chose a minimal set of subsequently less effective parameters that would achieve the desired error tolerance. The first parameter was chosen based on parallelism (or similarity) to the rotating parameter. Since the dot product of two vectors is a measure of parallelism, the angle derived from the dot product was used as a measure of similarity.

Gustavson et al. state that ‘[s]ynthesis using these parameters did not yield satisfactory results.’ Instead, they implemented a modified method where the first (most similar) parameter was chosen as described above, and then they chose the next seven parameters closest to being orthogonal to the first parameter. This led to a dramatic improvement and thus this variation to the original algorithm was adopted.

The loads that need to be identified for component fatigue life damage estimation purposes comprised chordwise bending moment, torsion, pitch link force, leading edge lead-lag damper force (all for the main rotor blade), and vertical bending moment at one of the fuselage stringers. The full set of parameters initially monitored comprised collective control position, lateral control actuator force, pitch and roll attitudes, lateral acceleration of the centre of gravity, yaw rate, longitudinal control actuator force, and longitudinal bending moment of the main rotor mast tube root. (Note that in the AH-64A helicopter the main rotor mast tube is a static member that takes the rotor bending moments, and acts like
a sleeve to the dynamic rotor shaft.) Of these, the parameters used in the final estimation were collective control position, the lateral and longitudinal control actuator forces, and longitudinal bending moment of the main rotor mast tube root.

Except for load cycle levels well below the endurance limit, ‘excellent’ correlation and damage fraction matching, between the actual and synthesised loads for the twisting moment and the lead-lag damper force, were reported. For the chordwise bending moment, there was a high correlation when the loading exceeded the endurance limit (the region of interest), but cycle range mismatches led to an under-estimation of the damage. The under-estimation could be corrected statistically due to its consistent nature, and this was done.

The U.S. Army applied a blind test with a statistical measure of correctness for the above procedure. Based on the correlation coefficient $r^2$, the first phase achieved 80–90% accuracy with six load inputs that are not normally monitored. The second phase involved the elimination of the Fourier decomposition and an optimisation of both the input parameters and sample rate. Using three load parameters and one control parameter with a sample rate less than 100 Hz, the mean accuracy is reported to be close to 90% (chordwise bending moment being the exception). In general, the load could be synthesised more accurately for high loads than for low loads.
6. Conclusion

Large variations in load were observed for the same helicopter type under the same flight condition. The load variations were caused by variations in pilot technique, adversary type (in aerial combat), control sensitivity, mechanical assembly tolerances, wear and tear, gross weight, altitude, rotor speed, ambient aerostatic conditions, and aircraft response to name a few. Dynamic component loads are very dependent on the magnitude, rate, and phasing of control movements. Furthermore, analysing only peak loads within each manoeuvre (these peak loads often incur high fatigue-damage) exhibited larger variability than analysing cycle-counted data from either each manoeuvre or an entire flight. Using current fatigue estimation techniques, it is difficult to avoid over-penalising component replacement times on most aircraft to protect against the worst possible usage for any given aircraft in the fleet. In fact, component replacement times might need to be reduced by as much as 75% if load variation is taken into account. Thus, there are good reasons for taking into account the powerful effect of load variability on life in fatigue life expenditure calculations.

In the frequency domain, the holometrics technique involves the development of a linear approximation (or a calibration matrix) of the loads in dynamic components from the loads in fixed components. The loads from fixed components are only included in the calibration matrix if they are sufficiently close to being orthogonal to the so-far-calculated calibration matrix. That is, the loads from fixed components are incorporated into the calibration technique one at a time, starting with the most influential load, and sequentially adding the next load that is closest to being orthogonal to the set of already collected loads. In the application of the holometric technique to an Apache helicopter, excellent identification of dynamic system loads was achieved. The only exceptions to this excellent identification result were the identification of loads that were well below the endurance limit. Both a parallel vector reduction algorithm and a cost-function (which determined optimum combination of inputs and hardware) provided an improvement relative to the original holometrics technique. It is unclear why this apparently promising technique wasn’t further pursued. It appears that using parameters other than just loads proved more productive for other researchers.

Several authors approximated the loads in dynamic components using regression of several key flight parameters. Most studies focused on high-G and high-load manoeuvres (which consume the majority of fatigue life). Results were improved by including terms in the approximation that involved the coupling of parameters and filtering out low-G data (to create a high-G bias). Alternatively, different weightings of data allowed the model to be biased in favour of higher loads. Measuring the swashplate servo position (which takes account of both pilot and stability inputs) instead of control stick position (which takes account of just the pilot input), and using aircraft accelerations instead of angular rates, also improved results.

One study developed a model using flight parameters and signal features (such as mean, maximum, and standard deviation). The main limitation of this work was the need to
manually partition the flight test data into time-windows. Partitioning the data automatically into constant time-windows rendered the procedure unusable due to inaccuracies, since manoeuvres were susceptible to partitioning over two different time windows. The more sophisticated procedures, such as combined pattern recognition and multiple linear regression, led to models that were more robust. A sophisticated pattern recognition procedure might allow variable time windows to be used. However, such an approach would be difficult to implement, and computational speed might be a severe restriction if on-line processing were required. In addition, the use of time windows does not account for the effects of slowly-varying large-amplitude cycles (such as ground-air-ground cycles). One study involved the testing of several classification criteria, with the 3-point nearest neighbour criterion performing best. A flight condition recognition system based on sequence and timing of bin combinations proved to be difficult to implement for one researcher.

Flight condition recognition offers an improvement over simply using the manufacturer’s design usage spectrum to determine retirement lives. This approach, however, is still susceptible to variability in load severity due to pilot techniques. Hence, using a flight condition recognition approach would only improve results if the fraction of time spent in the various flight conditions was different to the manufacturer’s design usage spectrum. Flight condition recognition, however, does not take into account the variability of loads experienced in any given flight condition. Given the extent of load variation within any particular flight condition, recognition of flight conditions appears to be of limited value.

Several researchers made extensive use of statistical methods (for example the $F$ statistic) to determine whether to include certain parameters. The use of significance tests and confidence intervals to determine the significance of the inclusion or omission of parameters on the performance of monitoring system seems an appropriate area for further investigation.

The neural network (NN) technique showed a significant improvement over that using linear regression. However, allowing non-linear terms (especially coupling between terms) reduced this difference significantly. The network models were computationally more intensive during calibration when compared to the regression procedures. However, the network models demonstrated better generalisation properties and slightly better results. Randomising the time history input for the NN sped up the determination of network coefficients, but reduced the NN’s ability to predict temporally varying characteristics. NNs that modelled all manoeuvres were found to produce excessive smoothing, and so specific manoeuvres were modelled by most researchers. Cluster analysis or the division of the data space into self-organising feature maps, which identify regions of interest, provided no appreciable improvement in results and were computationally intensive. On the contrary, partitioning the modelling between different expert systems appears promising.

Combining regression models with neural network models offered at least one main advantage. The major influence of measured parameters on load estimation is expected to be linear, and hence the regression component of the system could easily approximate the
linear component with a significantly reduced effort. The neural network could then concentrate on the remaining non-linear component (including elements arising from the cross coupling of terms), without having to waste any effort on the first order linear components. However, one of the papers indicated that the neural network alone was slightly superior to a combined linear regression and neural network model.

Very little work on the underlying problem of helicopter load modelling has been attempted. Questions such as the effects of noise, rank deficiency, and stability received no attention by most researchers. Furthermore, several of the models reported results for specific data sets relating to a single helicopter type. It is unclear whether these models would have generic applicability, or are restricted to a particular helicopter type. These fundamental questions need to be addressed if robust load models are to be developed and implemented.
7. References


12. H. Öry and H. W. Lindert, ‘Reconstruction of Spanwise Air Load Distribution on Rotorblades from Structural Flight Test Data’


A review of the literature for models that use fixed-component loads and flight parameters to determine loads in a dynamic component is presented. The reviewed papers naturally divide into one of three categories depending on the information they use to determine the load in the dynamic component. An initial section on load variability demonstrates that even for the same aircraft under the same flight condition, the loading can vary dramatically due to pilot technique, altitude, and weight to name a few variables. Neural networks, regression, and statistical indicators prove invaluable in developing load models. The review also demonstrated a lack of solutions to fundamental questions concerning loads modelling.