Questioning How Limits Are Set

Andrew Becker [1,2], Paul Marsden [1], Sylvester Abanteriba [2] and David Forrester [1]

[1] *Air Vehicles Division, Defence Science and Technology Organisation, PO Box 4331, Melbourne, Victoria, 3001, Australia*
[2] *RMIT University, GPO Box 2476V, Melbourne, Victoria, 3001, Australia*

Abstract

Measuring a machinery health parameter in a condition monitoring program is only half of the equation; at some point a limit must be set that initiates either further intensive monitoring or maintenance action. Limits are a fundamental and critical component of an effective condition monitoring program (often an integral part of a HUMS) and they can directly impact aircraft availability (e.g. rotor track and balance checks), or result in substantial maintenance burden (e.g. removal of machinery for overhaul based on wear debris). Limits therefore need to have a solid basis and often this is unknown, unclear or is questionable resulting in sub-optimal performance of the condition monitoring program.

The traditional statistical-based philosophy behind setting limits dates back to early Statistical Process Control (SPC) work credited to Shewart [1] and Deming [2] in the 1930's. Whilst the familiar Gaussian-based statistical approach (based on the average plus or minus some number of standard deviations) and the Cumulative Distribution Function (CDF) approach have been used extensively, the authors argue that such a fundamental component of a condition monitoring program now deserves more attention and rigour. Interestingly there is scant information in the literature about methodologies for setting in-service limits where the requisite minimum data sample is unavailable, unreliable or simply where the traditional approaches are not optimal or lack physical meaning. Conceptually, methodologies can be researched and developed that have a reasoned and validated foundation in the physics or chemistry of the machine element.

This paper will provide some examples of areas where alternative approaches may have provided better outcomes. The authors argue that a purely statistical approach may not necessarily serve condition monitoring programs well in all instances and can in fact result in sub-optimal limits being set. Similarly, limits based on small data sets derived during accelerated production testing of aviation machinery may not be representative of the true in-service behaviour of the equipment. Questioning the pedigree of limits and reviewing limits in-service is essential for ensuring optimal performance of condition monitoring programs, however experience shows that this is often neglected.

Keywords: Condition monitoring, limits, setting.

Introduction

Limits are typically either numerical points or simple rule sets applied to measured parameters that trigger an action based on the context of the particular machine.

They are contextual decision points, not decisions per se. The required action can be anything from increased observation or sampling frequency of the particular parameter to physical removal of a component from the airframe for overhaul. Limits are a fundamental and critical element of any condition monitoring program, however there would appear to be some uncertainty surrounding how they are set or justified. Ideally limits should be set to ensure the useful life of a component is maximized whilst minimising the risk of catastrophic failure. Often however the justification for a a particular limit is either tenuous or unknown. Clearly there is a significant benefit for operators to be fully cognisant of the pedigree and justification of limits applied to their machinery. Additionally, fully justified limits contribute to instilling confidence in the condition monitoring program.

In the aviation context it is essential that only limits endorsed by the applicable engineering authority or manufacturer be applied to the measured parameters. For example some oil analysis laboratories offer to apply generic limits in the absence of any specific limits, however this is can result in sub-optimal maintenance by not reflecting the true importance of an exceeded limit. Generic limits are often generated from the vast database of machine types serviced by the laboratory and may have little if any relevance to a particular machine. Even the ubiquitous application of statistically derived limits to condition monitoring programs may not necessarily result in optimum performance. There would appear to be scope for further research into methods for setting meaningful and justified limits rather than blindly applying statistically determined limits.

Condition Monitoring Limits and their Meaning

Defining Limit Terminology

Before defining a suite of limits for a particular process or machine parameter, a discussion on what constitutes a failure is essential. Whilst it might seem superficially obvious what defines a failed component (such as a bearing), the typical reliability definition of failure is the point at which the component, system or process no longer provides the desired output in that operating context [3]. Nowlan and Heap [4] further define the concept of a functional failure being where the component is incapable of performing the desired task or where the degraded performance of the component is unacceptable. British Standard 4778 defines failure as “the termination of the ability of an item to perform a required function”[5].

Clearly where the consequence of failure is low a component such as a bearing can be run to functional failure. However where large, complex and expensive machinery is involved, or where the consequences of failure are significant (i.e. aircraft machinery) then what constitutes a failed bearing becomes less clear. Reliability engineering principles often lead users to define what a failure is, however without adequate experience or knowledge this can potentially lead to over conservatism (bearings removed from service prematurely) or an unanticipated functional failure of the bearing (i.e. the bearing ceases to rotate and carry load). Widner [6] suggests that where contact surfaces have changed sufficiently to adversely affect the performance of the contact surface, then the bearing is considered damaged. Further, he suggests that when the damage progresses to a point where the function of the bearing is threatened, then the bearing is considered to have failed. ISO

15243 [7] defines failure as a “defect or damage that prevents the bearing meeting the intended design performance”. SKF define the life of a bearing as “the number of revolutions the bearing can perform before incipient flaking occurs”; this definition suggests that *any* visible flaking of the bearing surface is considered a failure. ISO 281:2007 provides some insight into the definition of a failure relating to bearing life prediction [8]. Under the definition of bearing life, this document states that the bearing life has expired when the first evidence of fatigue in the material is evident on any of the bearing sub-components (inner race, outer race or rolling elements).

Hoeprich [9] states that the Timken bearing company define the end of bearing life as a spall area of 6.5 mm^2 for laboratory applications, however no clear justification is provided for this. Hoeprich also suggests that depending on the context, bearings could continue to operate well beyond this laboratory fatigue life limit. Li et al [10] state that the majority of condition monitoring techniques are aligned to detect any abnormality at the earliest possible stage and do not typically account for the substantial residual bearing life. Li et al also quote an industry standard of 6.25mm^2 as the fatal flaw size in a rolling element bearing however the origin or justification are not provided.

Moubray [3] discusses the concept of a P-F curve which describes how a condition changes once a failure has commenced. Figure 1 illustrates the P-F interval concept. The point “P” on the curve represents a Potential failure that is physically detectable with current technology. The point “F” represents the functional failure of the machine or component. At some point between “P” and “F” the operators of the machine require notification that the failure is progressing (i.e. a limit).

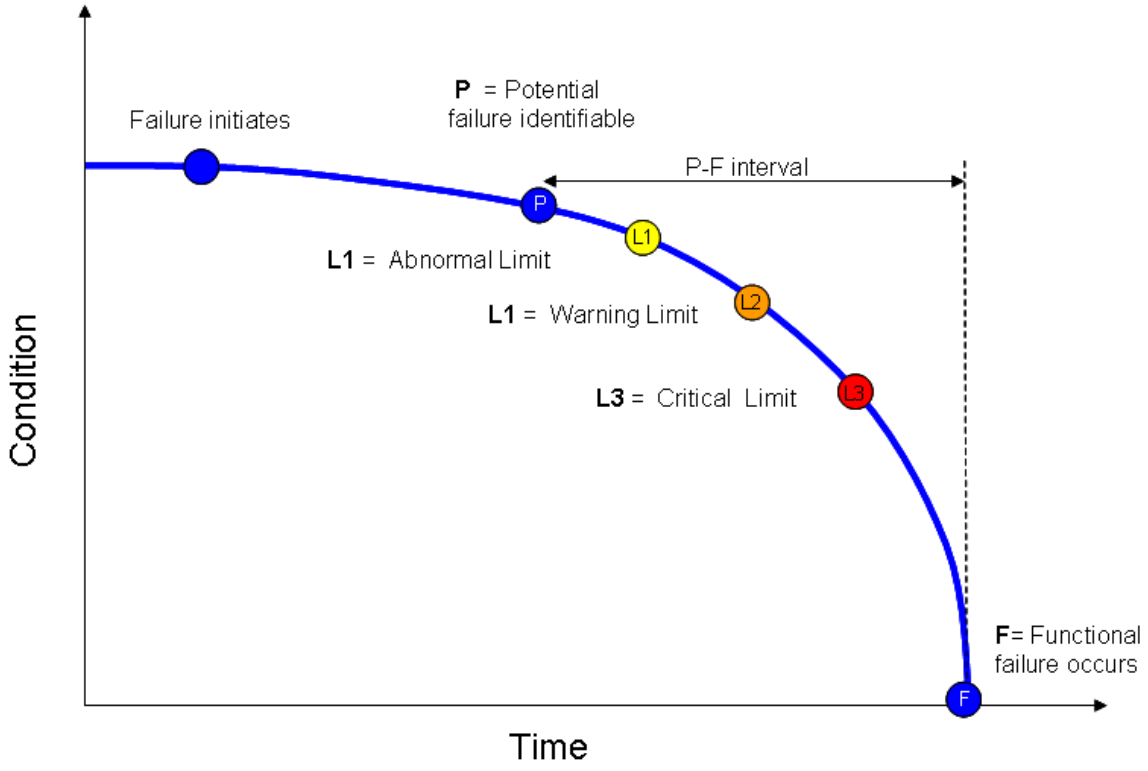

Fig. 1: P-F Interval

Traditional Limit Setting Methods

There are a number of traditional methods used to determine the value of a limit, however the most commonly used is the statistical analysis of a sample of historical data. When a statistical model is effective in describing the properties of a sample, so called parametric statistics can be used to make inferences about the population. The Gaussian distribution is the most obvious example. Where limits are set based on the average plus some number of standard deviations, there is an underlying assumption that the sample data follow a Gaussian distribution. If this assumption is flawed, the calculated limits may not truly reflect the intentions of the engineer.

Data can be said to be non-parametric when it is not normally distributed and it is not possible to apply a known distribution shape to the data. Rank statistics are commonly used in this situation. The cumulative distribution function (CDF) is the basis of a technique that may be used to determine limits from a data set without making any assumptions about its statistical profile. A CDF describes the fraction of the distribution that is at or below any given value [11]. Given a sufficient sized sample of historical data a CDF can be produced that allows limit levels to be defined in terms of their percentile. A judgment as to the reasonableness of the limit is still recommended according to ASTM D7720-11 [12]. An example of this type of judgment is given in the section on the T56 wear debris limit.

Garvey[13] provides a concise description of four traditional methodologies used to determine limits for oil analysis applications. In fact, these methodologies are used in many fields beyond oil analysis. The four methods described by Garvey are:

1. No predefined Limits: this technique is essentially judgment-based and relies on the availability of experienced analysts who are familiar with the type of machinery being monitored.
2. Industry Standards: These are typically generic limits based on groupings of machines based on working pressure or type (e.g. gearbox, hydraulic system, Diesel engine etc). Whilst this method might be satisfactory for some industrial applications, it is generally considered a starting point only for further development of system-specific limits.
3. Statistical Alarms: The two statistical methods described are the cumulative distribution function and Gaussian distribution. Both of these are widely used to set limits.
4. Trend-based or rate of change limits: These methods are essentially looking for an unacceptable departure from a usual level. Garvey quotes Poley [14] as suggesting there are three ways of achieving trend or rate of change alarms:
    a. Relative Magnitudes: This technique simply looks for a significant change in magnitude of the measurand.
    b. Rolling Average: This technique compares the current value to an average of several historical measurements.
    c. Weighted Delta Settings: This technique uses a weighting method that requires very large changes to occur before an alarm is tripped for small measurand values. As the measurand values increase, the required percentage change decreases.

    Whilst rate alarms can be useful for indicating trends prior to a fixed limit being exceeded, they can be susceptible to erratic indications where a short time base is used; in this case multiple consecutive excursions beyond the limit are required before action is taken.

In the 1930's Shewart [1] first suggested the application of statistics to the quality control of industrial processes, known as Statistical Process Control (SPC). SPC is a technique often used to measure the quality variation of product or process. This technique requires the measurement of a characteristic or feature of a sample of the population that can then be used to draw conclusions about the trend of the process. The results of this technique are generally displayed on a SPC control chart that can indicate visually the trend of the system. Whilst Shewart focused on manufacturing tolerances and setting limits for production processes, the concept has been applied to machinery performance monitoring and machinery condition monitoring. Shewart discusses the concept of a quality characteristic lying somewhere within a tolerance interval. If the quality characteristic exceeds the tolerance interval then the item is said to have failed; this is the classic go, no-go concept. The issue Shewart raises is that often we need some advance notice that the process is going awry and he introduces the concept of some basic statistical measures to provide an indication of process control. This results in the development of action limits that indicate the process is heading out of control but has not failed yet.

Deming and Birge [2] also discussed at length the issue of determining error tolerances for physical observations in particular the production of physical components. Whilst the concept largely revolved around choosing a suitable sample from a population, the concepts were further developed and resulted in 3 standard deviations (3σ) being accepted as the tolerance that minimises economic loss for any given process. By applying a number of standard deviations to a sample, the limits obtained represent a probability that a certain percentage of future data will lie within the limits. Table 1 shows the percentages for one sided and two-sided analyses (see also Figures 2 and 3).

Table 1: Percentage of population bounded by x standard deviations

Number of standard deviations (σ)	Average + (upper limit only)	Average +/- (upper and lower limit)
1	84.14%	68.27%
2	97.73%	95.45%
3	99.87%	99.73%

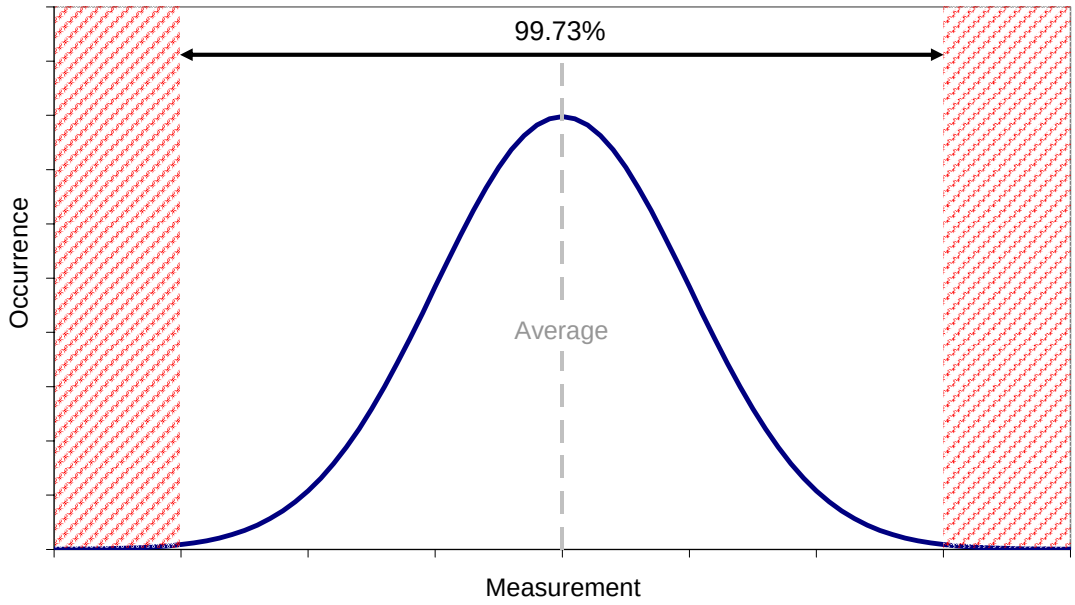

Fig. 2: Typical Gaussian distribution showing upper and lower limit at average +/- 3 standard deviations – hatched regions show where the limit would be exceeded

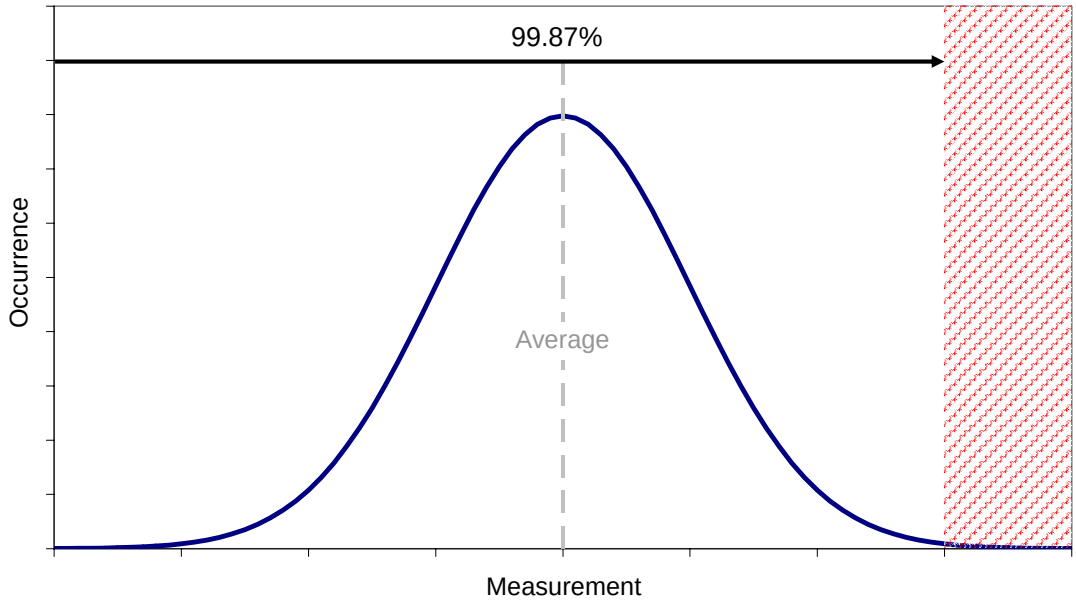

Fig. 3: Typical Gaussian distribution showing upper limit at average + 3 standard deviations – hatched regions show where the limit would be exceeded

Some Common Flaws

There are some common errors that can result in the sub-optimal application of limits, including:

1. Too many limit levels: Some condition monitoring programs have as many as five limit levels, all with very similar recommendations. Typically a maximum of three limit levels should suffice:
    a. *Abnormal*: an indication that the measured parameter has exceeded the target value but is still within the in-service range.
    b. *Warning*: notification that the measured parameter has exceeded the desired in-service range and some maintenance action is required.
    c. *Critical*: should this limit be exceeded then the machine should be shut down immediately to avoid catastrophic failure.
2. Poorly defined actions at each limit level: This is extremely common and can be considered the "so what" factor. What is the significance of the limit being exceeded and what is the maintenance action required? If operators and maintainers don't have clearly articulated actions to carry out in the event of the limit being exceeded then serious conditions can be missed or conversely the consequence of benign conditions can be overestimated resulting in unnecessary expense.
3. Confusing set of parameters: Often the users of condition monitoring programs do not have a clear understanding of what measurements mean or what parameters are required to adequately assess machinery in their particular operating context. This can lead to analysis laboratories providing overly complicated lists of measurements to operators. This rarely enhances the machinery health and usually results in excess or redundant.
4. Unknown limit provenance: The reasoning, logic or justification for specific limits can be lost or indeed never fully understood by operators or manufacturers. Limits that were set at the beginning of a piece of equipment's life can become meaningless or irrelevant as measurement techniques improve or change or the operating context of the machine alters.

Some Curious Examples

Aircraft Wear Debris Limits

Helicopter wear debris limits usually refer to the number of particles allowed on a magnetic chip detector or recovered from a filter. In this context wear debris is defined as visible particles and not sub-visible particles detected by laboratory oil analysis techniques such as Spectrometric Oil Analysis (SOA). The wear debris limits can also be defined in terms of an allowable area of debris when the particles are spread out (see Figure 4).

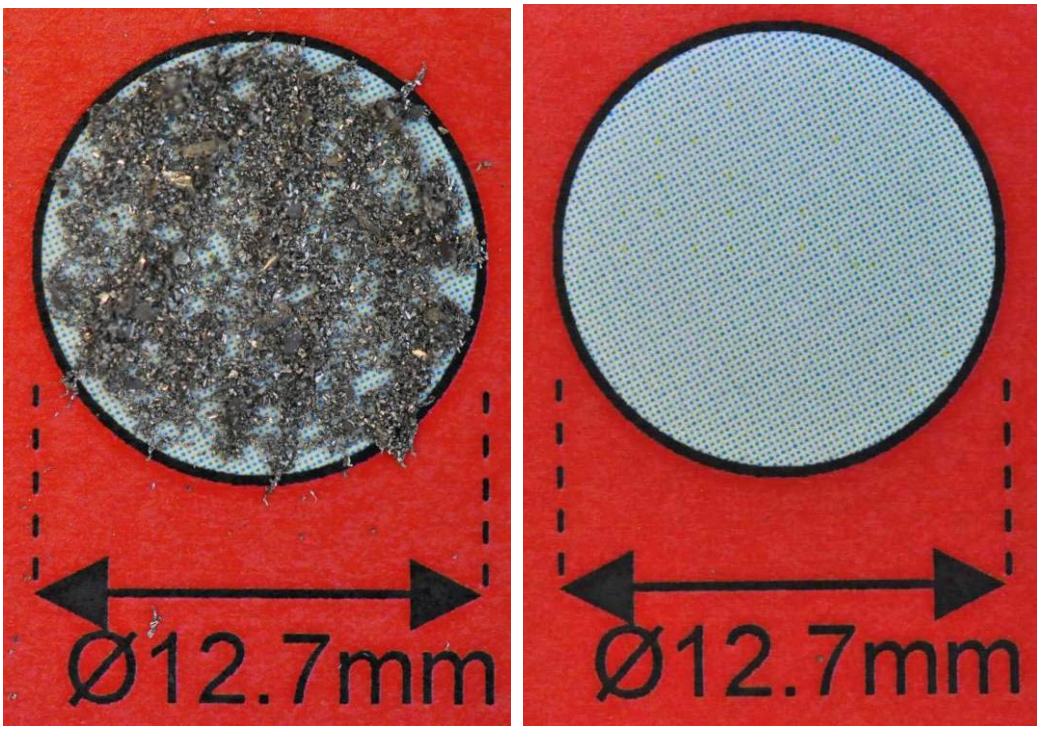

Fig. 4: Example of a wear debris area limit from a large aviation turbofan gas turbine

Whatever the measure (counts above a certain size, area) it is difficult to accurately assess in the field and the substantiation of these limits is usually lost to the archives of the OEM and not available to those who actually use the limits on a day to day basis. As an example, Table 2 shows the wear debris limits for a common military utility helicopter.

Table 2: Summary of military utility helicopter wear debris limits

Material	Allowable Wear Debris
Ferromagnetic debris	No more than10 particles > 1.27 mm or 0.81mm x 3.18mm
Aluminium	No more than10 particles > 1.27 mm or 0.81mm x 3.18mm
Bronze	No more than10 particles > 1.27 mm or 0.81mm x 3.18mm

Whilst these limits may be easy to convey to maintenance staff, it is difficult to see how the same number of particles can be justified for different material types given the vastly different quantities that exist in the particular gearbox. Indeed, the Australian Defence Force operators of this aircraft have religiously abided by the 10 particle limit, which has resulted in the removal of gearboxes for overhaul, however when confirmation of the 10 particle limit was sought from the OEM there was some ambiguity about whether it was a rigid limit or simply indicative.

Similarly, Table 3 contains in-service limits for a large cargo helicopter manufactured by a different OEM. Although the particle size is different, it can be seen that the quantity of ten particles persists for each gearbox type despite the significant difference in gearbox complexity. Whilst setting reliable in-service wear debris limits is not a trivial exercise, the substantiation behind these examples is not clear and perhaps a more sophisticated, robust and tailored approach that could be applied.

Table 3: Wear Debris Limits for Tandem Rotor Helicopter Gearboxes

Gearbox	Ferromagnetic Debris Limit
Engine Gearbox	No more than 10 flakes up to 1.6mm diameter and 6mm long
Combining Gearbox	No more than 10 flakes up to 1.6mm diameter and 6mm long
Forward Gearbox	No more than 10 flakes up to 1.6mm diameter and 6mm long
Aft Gearbox	No more than 10 flakes up to 1.6mm diameter and 6mm long

Lubricant Viscosity Limits

Viscosity is arguably the most important lubricant characteristic, however the physical meaning of in-service limits is rarely clear. In aviation, viscosity typically has a serviceability window; in other words the viscosity is considered serviceable provided it remains within a defined range. The synthetic polyolesters typically used as aviation lubricants can become thicker (e.g. oxidise) or become thin (e.g. by the introduction of incorrect fluid) in service. Limits for aviation lubrications have been observed in the range of +/- 10% to +/- 50% of the fresh oil specification [15]. It is possible to determine the minimum film thickness that could sustain separation of the adjacent rotating dynamic components operating under elastohydrodynamic lubrication (i.e. gears, rolling element bearings etc., see Equation 1 [16]).

$$h_o = 3.63R'\left(\frac{U\eta_o}{E'R'}\right)^{0.68}\left(\alpha E'\right)^{0.49}\left(\frac{W}{E'R'^2}\right)^{-0.073}\left(1-e^{-0.68k}\right) \quad (1)$$

Where:

h_o = *minimum film thickness [m]*

U = *entraining surface velocity [m/s]*

η_o =*the viscosity at atmospheric pressure [Pa·s]*

E' = *reduced Young's modulus [Pa]*

R' = *reduced radius of curvature [m]*

α = *pressure viscosity coefficient [m^2/N]*

W = *contact load [N]*

k = *ellipticity parameter [non-dimensional]*

Whilst this is not a trivial calculation and would involve some design knowledge and assumptions, it would provide some physical basis of a viscosity limit. Applying a statistical approach may be an acceptable alternative to calculating the minimum oil film thickness in a complex machine, however the sample of data used to generate the statistics is critical. Carelessness in this regard can result in an overly conservative limit being applied for viscosity. As an example, the recommended in-service limits for a common synthetic ester lubricant (MIL-PRF-23699 compliant) are given in Table 4. These limits were derived from a known source of data (results from an extensive 12 month oil analysis trial [17]) Also shown in this table are the statistically derived limits from an unknown data sample incorrectly applied to an ADF aircraft.

Table 4: ADF viscosity limits for MIL-PRF-23699

Viscosity (cSt @ 40°C)	Fresh oil specification	ADF in-service limits	Statistically derived from unknown data set by oil analysis lab
Lower Limit	23	20	24.8
Upper Limit	*None specified*	30	30.4

Whilst the values in Table 4 might seem quite close, the subject aircraft was deployed remotely and was repeatedly (and incorrectly) flagged as having exceeded the viscosity limits and eventually the aircraft was grounded by maintenance crew awaiting further advice; after expert advice was sought, the aircraft continued in service with the same charge of oil. Another feature of this data is that the oil analysis laboratory's lower limit is above the minimum fresh oil specification; this means that fresh oil complying with the specification can immediately trip the unauthorized limits. Whilst no maintenance action was taken, the disruption to aircraft availability was significant and completely avoidable. Additionally, the fact that the aircraft continued flying for almost 200 operating hours suggests the oil analysis reports were not being considered credible.

Propeller Balancing Limits

The ADF uses trim balancing to reduce the amplitude of vibrations at propeller shaft frequency on the C-130J. In the past the procedure was performed periodically using a commercial vibration analyzer. A limit of 0.2 inches per second (IPS) peak [18] (measured on the gearbox casing) was used, and often this required several ground runs to achieve. The provenance of this limit was unable to be identified and there was no evidence that vibration amplitudes greater than 0.2 IPS peak represented a hazard to the subject bearings or other rotating components.

Additionally, this limit is also routinely applied to helicopter main rotor balance operations in the ADF with a similar absence of any solid technical justification. Subsequent, informal discussions with former employees of a major vibration measuring equipment manufacturer indicated that the origins of the limit related to the minimum resolution achievable by the electronics used when digital Fast Fourier Transform (FFT) analysers were first commercialised for aircraft dynamic balancing functions.

The limit was reviewed as a part of a program to implement a trim balance using in-flight data [19] from the on-board HUMS. It was found that there was no requirement imposed by the engine, propeller or airframe manufacturer to perform maintenance in response to propeller frequency vibration levels below 3.0 IPS peak; at which point a maintenance alert is triggered. The aircraft system will however record the vibration parameters required to perform a trim balance if the amplitude of the propeller shaft frequency vibration exceeds 0.5 IPS peak in flight. This is a more realistic and achievable limit that has been implemented with the new trim balance method and is performed on-condition using in-flight data.

T56 Wear Debris

A Condition Monitoring program using Wear Debris Analysis has been in place for RAAF T56-14-A engines since the mid 1980s. The results are used to inform

maintenance decisions that include more frequent WDA sampling, engine ground runs and engine removal where debris indicates impending failure.

Limits are set for the rate of ferrous debris produced, as measured by the Staveley Mark 3 debris tester, in units of Debris Quantity per hour (DQ/hr). The extant limits for this engine had no known provenance and no clear statistical origin. Some 1st stage (Abnormal) limits were set unreasonably low resulting in unnecessary maintenance actions, and the 2nd stage limits (Warning) were set so high that engine removals were sometimes performed before the limit was reached.

A vast historical archive of wear debris measurements on the Australian T56-14-A fleet exists. These data have been collected over 20 years for more than 100 engines. All of these data are generally from some time after failure initiation as material has been liberated. The ability to detect this material puts the machine condition somewhere in the P-F interval as shown in Figure 1. The probability distribution was highly right skewed; the long tail is clearly seen in Figure 5. The data are clearly not normally distributed, and are also not accurately described by other parametric distributions. For such a large data set, a cumulative distribution function is a useful tool for determining limit levels.

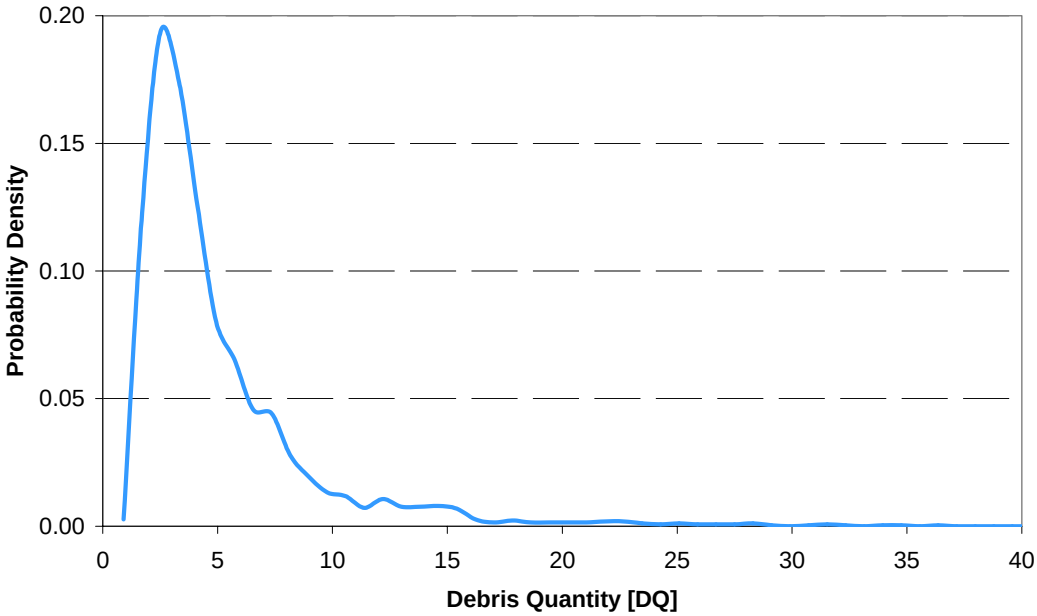

Fig. 5: Probability density function for wear debris from the RGB of RAAF T56-14-A engines

A percentile rank of 97.7% was chosen for the 1st stage alarm level, which would correspond to two standard deviations above the mean if the data were normally distributed. According to the standard ASTM D7720-11 [12] there is no requirement to set alarm levels that correspond to standard deviations when using the cumulative distribution technique. This value was chosen because it passes the assessment for “reasonableness” recommended by the standard. In this case the historical data were used to determine if there would have been any false negatives; component failures where the alarm threshold was not reached. Using similar reasoning, the 2nd stage alarm level was set at 1.5 times the level of the 1st stage alarm.

Oil Acidity

Total Acid Number (TAN) is a very important characteristic of oils and is particularly important for aviation synthetic ester oils. Synthetic ester oils are made by the reaction of alcohol and acid which produces the ester plus water as a by-product. If the ester is exposed to sustained high levels of water when in service, the reverse reaction can occur and result in the formation of acids. These acids can cause corrosion of the internal surfaces of the system and ultimately degrade dynamic components. The DSTO-recommended in-service limit of 2 mgKOH/g is and example of a limit that has a physical basis. It was derived experimentally using non-ferrous alloy samples immersed in oils of varying acidity. Figure 6 shows the results of the test. Interestingly, a limit published for TAN (not by the ADF) recommends an in-service limit of 1 mgKOH/g; this value is also the maximum allowable for fresh MIL-PRF-23699 oil [20] and hence does not allow for any degradation of the oil in service.

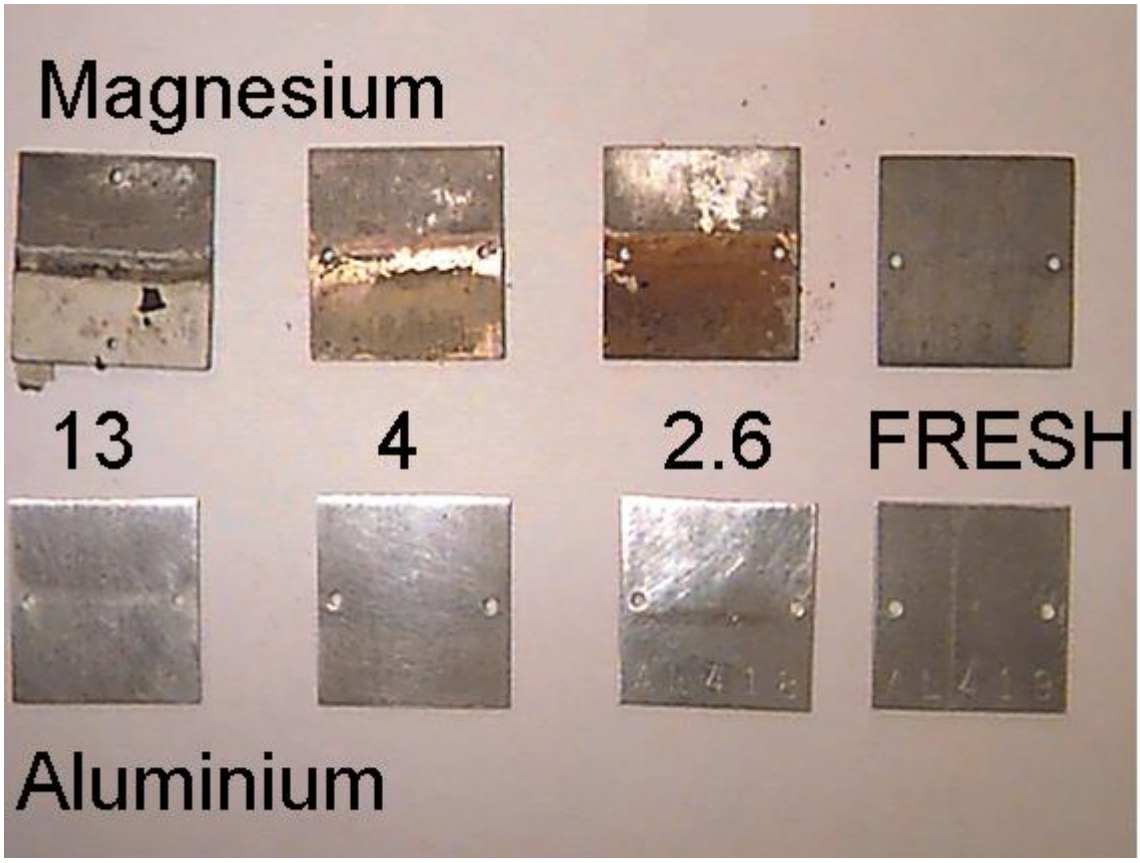

Fig. 6: Results of non-ferrous coupons immersed in MIL-PRF-23699 oil of varying acidity (numerical values shown are TAN in mgKOH/g)

Hydraulic Fluid Solid Particle Counting

Hydraulic systems are typically tested for solid particle contamination at some periodicity and reported in accordance with either ISO 4406 [21] or SAE AS 4059 [22] methods. The purpose of this test can sometimes be lost in the often confusing world of solid particle reporting for hydraulic fluids. The purpose of measuring the quantity of fine solid particulate is to ensure:

1. the filter is functioning correctly,
2. abnormal fine wear is not occurring, and
3. that significant solid particle contamination is not entering the system from the environment (i.e. via seals or breathers).

Setting an in-service limit that is not physically achievable by the extant filtration is pointless. This is a case where statistical analysis of sufficient in-service data can yield a reasonable limit provided extreme outliers are removed. The filter in an aviation hydraulic system would typically be rated as Beta 5 µm = 200 (i.e. removes 99.9% of 5 µm particles). This physical system component cannot be easily changed, so historical data will indicate the steady state condition that the filter can achieve. This average condition can then be relaxed by one or two standard deviations provided the data exhibits a normal distribution and extreme outliers are removed to avoid distortion.

When applying Gaussian statistics to solid particulate data care should be taken. An analysis of ISO 4406 codes using statistics may be valid, however only the initial limit (Abnormal) should be set by statistical means. Subsequent limits (Warning and Critical) should be determined by adding one ISO 4406 code number (Table 5), since an increase of one code number essentially doubles the physical number of particles as shown in Table 6.

Table 5: Suggested limit levels for hydraulic fluid solid particle contamination

Limit	Suggested Level
Abnormal	Average + 2 Standard Deviations
Warning	Abnormal + 1 ISO 4406 code number
Critical	Warning + 1 ISO 4406 code number

Table 6: Impact of increasing ISO 4406 cleanliness code by one code number

ISO 4406 Code	Number of particles > 4µmc per mL	Number of particles > 6µmc per mL	Number of particles > 14µmc per mL
19/16/12	2500 to 5000	320 to 640	20 to 40
20/17/13 (one ISO 4406 code higher than above)	5000 to 10000	640 to 1300	40 to 80

Inductive Wear Debris Sensors

Inductive wear debris sensors are a relatively recent development in the detection of metallic wear debris in oil-wetted systems. The sensors are typically fitted to the scavenge oil system and can detect individual particles as they transit the sensor. The particle detection range for this type of sensor is typically in the 100µm to 1000µm range. Having a sensor with this resolution in the operating system enables common failure modes such as roiling contact fatigue to be observed as they occur.

Figure 7 shows data from an inductive wear debris sensor fitted to a full scale helicopter main rotor gearbox. This gearbox was operated in the DSTO Helicopter Transmission Test Facility over approximately 80 operating hours (distributed over 6 weeks) primarily to assess the value of the sensor. The gearbox was operated with a periodic 150% torque overload in order to accelerate the test (i.e. not a seeded fault test). The cumulative ferromagnetic debris count and clearly shows the initiation and progression of the spall; in this instance the rolling contact fatigue occurred only in the planet gear support bearing. Whilst the effectiveness of this type of sensor was demonstrated (and confirmed in another similar test), there is virtually no guidance for

setting a meaningful set of limits for these sensors; this would appear to be an opportunity for further research.

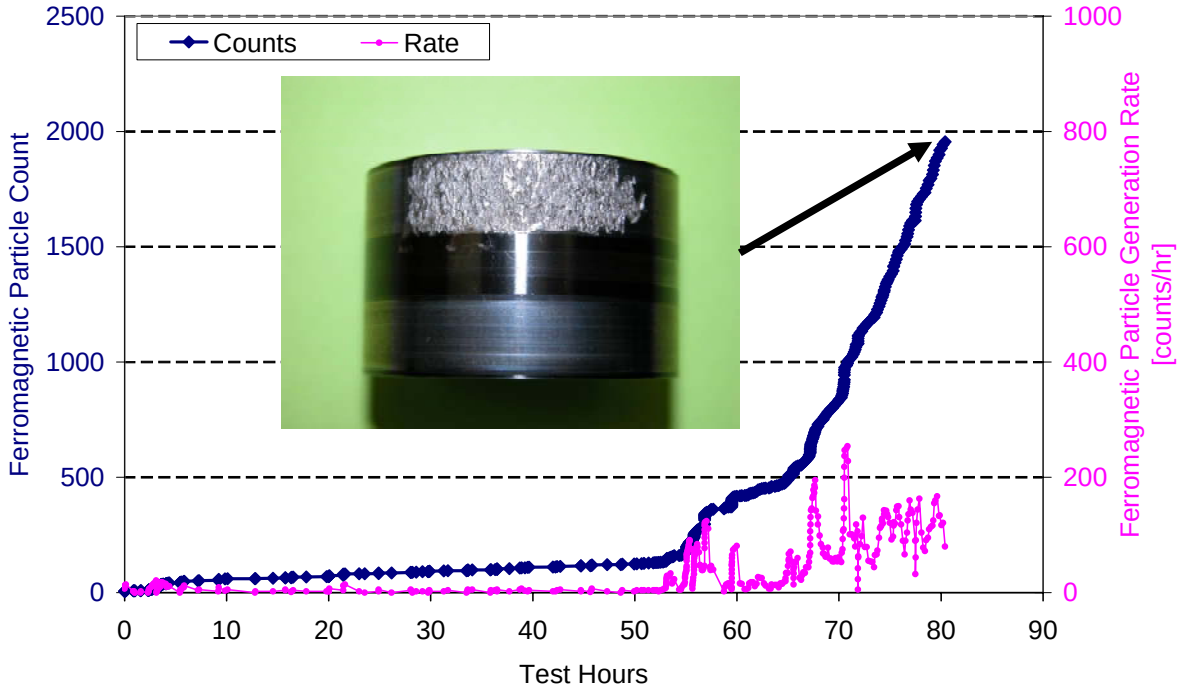

Fig. 7: Helicopter main rotor gearbox inductive wear debris sensor data (bearing raceway condition at conclusion of test also shown)

What to do?

Apart from questioning and reviewing extant limits applied to machinery health monitoring, one of the main purposes of this paper is to highlight the importance of well substantiated limits. Additionally, further research is recommended to identify alternative methods that could be used to justify limits in specific applications. Two areas of immediate interest would be setting limits for shaft unbalance that are related to the impact unbalance vibration has on bearing life and developing meaningful wear debris limits for inductive wear debris sensors. Certainly the effect of vibration on bearing life does appear in the literature, however the transition to a methodology for setting unbalance limits is not yet available.

Conclusion

This paper has discussed the importance of limits as applied to condition monitoring programs. Traditional methods have been explained and the apparent lack of physical meaning examined. Whilst the application of traditional statistical methods may still be valid in many cases, the authors contend that there are opportunities to research methods for determining limits that have physical meaning and can be appropriately justified. Several examples were discussed. Considerably more effort and rigour should be afforded the development, implementation and review of limits to ensure a condition monitoring program doesn't become simply a tedious data collection novelty.

References

1. Shewhart, W. A. (1938) Application of statistical methods to manufacturing problems. *Journal of the Franklin Institute* 226 (2) 163-186

2. Deming, W. E. and Birge, R. T. (1934) On the statistical theory of errors. *Reviews of Modern Physics* 6 (3) 119-161

3. Moubray, J. (1997) *Reliability Centred Maintenance*. Second ed. New York, Industrial Press Inc.

4. Nowlan, F. S. and Heap, H. F. (1978) RELIABILITY-CENTERED MAINTENANCE. *Proceedings of the Annual Reliability and Maintainability Symposium* 17 January 1978 through 19 January 1978

5. BSI (1991) *BS 4778-3.1 Quality Vocabulary - Part 3: Reliability, Availability and Maintainability Terms*. London. British Standards Institution.

6. Widner, R. L. and Littman, W. E. (1974) BEARING DAMAGE ANALYSIS. *Natl Bur Stand Spec Publ* (423) 8 May 1974 through 10 May 1974 67-84

7. ISO (2004) *15243 Rolling Bearings - Damage and Failures - Terms, Characteristics and Causes*.

8. ISO (2007) *281:2007 Rolling Bearings - Dynamic Load Ratings and Rating Life*.

9. Hoeprich, M. R. (1992) Rolling element bearing fatigue damage propagation. *Journal of Tribology* 114 (2) 328-333

10. Li, Y., et al. (1999) Dynamic prognostic prediction of defect Propagation on rolling element bearings. *Tribology Transactions* 42 (2) 385-392

11. Smith, P. J. (1993) *Into Statistics: A Guide to understanding statistical concepts in engineering and the sciences*, Thomas Nelson Australia

12. D7720-11 (2011) ASTM Standard Guide for Statistically Evaluating Measurand Alarm Limits when Using Oil Analysis to Monitor Equipment and Oil for Fitness and Contamination.

13. Garvey, R. (2002) Four Ways to Determine Alarm Settings. *Practicing Oil Analysis* (NOV./DEC.) 8, 10-12

14. Poley, J. (2005) *Proceedings of the World Tribology Congress III - 2005*, Washington, D.C.

15. Lynch, C. W. and Cooper, R. B. (1971) Development of a three- micron absolute main oil filter for the T53 gas turbine. *J Lubric Technol Trans ASME* 93 Ser F (3) 430-436

16. Stachowiak, G. W. and Batchelor, A. W. (2005) *Engineering Tribology*. 3rd ed, Elsevier Butterworth Heineman

17. Becker, A. J., Kappas, J. and Rawson, P. (2007) *An Analysis of Black Hawk Oil Health and Evaluation of the OilView 5100 In-field Oil Analyser*. Defence Science and Technology Organisation

18. RAAF (2006) *Technical Manual - Balancing Propeller AAP7211.031-2-61JG-00-1*.

19. Rebbechi, B., Marsden, P. R. and Church, S. B. J. (2011) C-130J In-Flight Propeller Balancing. In: *7th DSTO International Conference on Health & Usage Monitoring (Australian International Aerospace Congress),* Melbourne

20. Specification, M.-P.-F. P. (1997) *Lubricating Oil, Aircraft Turbine Engine, Synthetic Base, NATO Code Number O-156*. U.S. Department of Defense.

21. ISO (1999) *4406:1999(E) Hydraulic fluid power - Fluids - Method for coding the level of contamination by solid particles*.

22. SAE (2001) *AS4059 Aerospace Fluid Power - Cleanliness Classification for Hydraulic Fluids*.