Proceedings of the ARPA/AFML Review of Progress in Quantitative NDE, July 1975–September 1976

Interdisciplinary Program for Quantitative Flaw Definition Annual Reports

9-1977

Discussion of Factors Controlling Army Helicopter Reliability

K. Rummel
Boeing Vertol Company

T. House
United States Army

Recommended Citation

Rummel, K. and House, T., "Discussion of Factors Controlling Army Helicopter Reliability" (1977). *Proceedings of the ARPA/AFML Review of Progress in Quantitative NDE, July 1975–September 1976*. 39.
http://lib.dr.iastate.edu/cnde_yellowjackets_1976/39

Discussion of Factors Controlling Army Helicopter Reliability

Abstract

The Army's interest in significant aviation improvements actually started, or at least gained steam, in the 1960's during the extended combat we had in southeast Asia where a lot of problems in the utilization of the rotary wing aircraft came to a head, and I'll talk about those in more detail. What I would like to do is go through how the Army has responded to those problems that were revealed, through R&D and even more specifically in the diagnostics area. When I say diagnostics, what I'm talking about is the second category of NDI that I see, the first category being the one I think that most people here have a prime interest in and that is the one time, static inspection during manufacture, a verification of quality control. The other type is the repetitive in-service inspection.

Disciplines

Materials Science and Engineering

DISCUSSION OF FACTORS CONTROLLING ARMY HELICOPTER RELIABILITY

K. Rummel
Boeing Vertol Company
Philadelphia, Pennsylvania
and
T. House
Eustis Directorate
U.S. Army Materials Research and Development Laboratory

I'm Kirk Rummel from Boeing Vertol. Tom House, from the Eustis Directorate of USAAMROL, had intended to present this paper and was looking forward to this very fine group. He apologizes for not coming.

What we would like to do is try to show the Army's program in the diagnostics area; specifically, I hope I can respond to some of the challenges that Charlie Smith threw out about having a logical approach to research planning.

The Army's interest in significant aviation improvements actually started, or at least gained steam, in the 1960's during the extended combat we had in southeast Asia where a lot of problems in the utilization of the rotary wing aircraft came to a head, and I'll talk about those in more detail. What I would like to do is go through how the Army has responded to those problems that were revealed, through R&D and even more specifically in the diagnostics area. When I say diagnostics, what I'm talking about is the second category of NDI that I see, the first category being the one I think that most people here have a prime interest in and that is the one time, static inspection during manufacture, a verification of quality control. The other type is the repetitive in-service inspection.

I think there's a lot of differences in those two categories. The first is very obvious: there are high levels of funding, and secondly, there appears to be a good deal of technology interchange which, quite seriously, draws my envy. I think that's one of the messages that you have given me from this symposium; that we need to do a much better job of technology transfer in this latter area of in-service diagnostics.

I do want to diverge just for a moment to express my appreciation of what I observe is the high caliber of the work being performed. The scientific progress that you're making is obviously quite considerable. Perhaps this paper addresses more of the engineering rather than the scientific aspects of NDI.

I would like to become more specific and talk about the approach that some like to call "deficiency oriented." It starts with problems rather than solutions, and on Fig. 1 is a simple outline of what I think is the generic approach to problem solution. I don't want to insult your intelligence but rather remind you of this fundamental aspect of good engineering application. I want to show it because I think all too frequently we start at step 6 and go to 8 or at least try to get to 8 and have to return eventually to 1 through 5. It's very important that we move on down that order in a very sequential manner, attempting to be as quantitative as possible, particularly in steps 2 and 3. I think quantitative assignments there can be either in a delta expression or in an absolute value. The material I'm going to show you, I hope, will adhere to that requirement that all of our goals are quantitative.

1. DEFINE REQUIREMENTS.
2. ESTABLISH WORKABLE GOALS.
3. ASSESS CURRENT POSITION.
4. REVIEW ALL POSSIBLE SOLUTIONS.
5. PERFORM PRELIMINARY EVALUATIONS.
6. DETAILED RESEARCH.
7. EVALUATE BEST CANDIDATES.
8. INCORPORATE.

Figure 1. Process for problem solution.

Figure 2 is a distribution of maintenance costs. This does not include fuel and crew. It happens to be on our CH-47. A thousand dollars an hour for any helicopter and I think there's where we start with the problem.

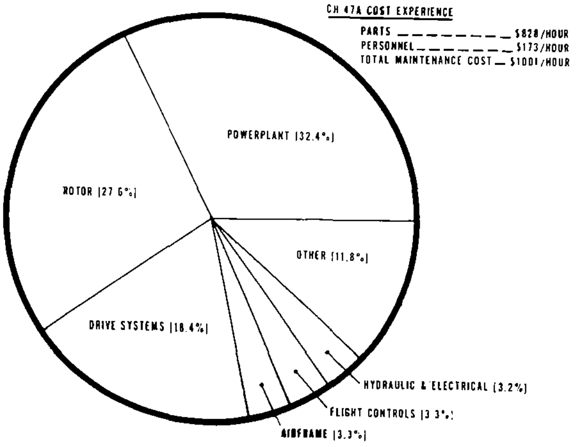


Figure 2. Distribution of CH-47A parts cost by subsystem.

The distribution of subsystems may be of some surprise to some of you. I have seen a great deal of emphasis on the airframe structures here; it's a small segment of that pie: the engine, the rotors, the drive system are the big contributors. Another display of the problem is Fig. 3 which is of inflight aborts on the UH-1. Here we see a slightly different subsystem distribution, the hydraulics there is a little bigger because of the single hydraulic boost system that the UH-1 has which makes the operator a bit more sensitive to hydraulic indications. Those are both false and real indications, I might remind you.

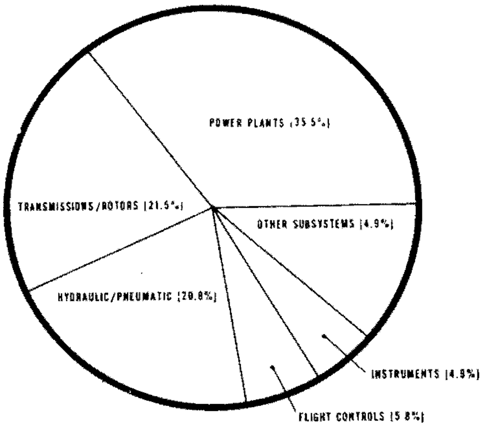


Figure 3. Subsystem causes of UH-1H Mission Aborts.

Figure 4 displays maintenance man hours per flight hour. Here I think you begin to realize some of the impact of scheduled inspections. Now, remember, that's organizational level maintenance man hours per flight hour. We're doing an awful lot of inspection.

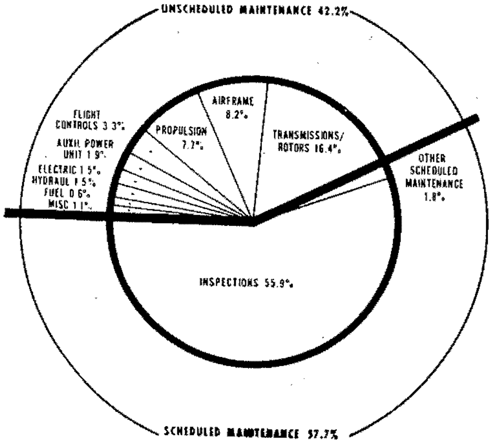
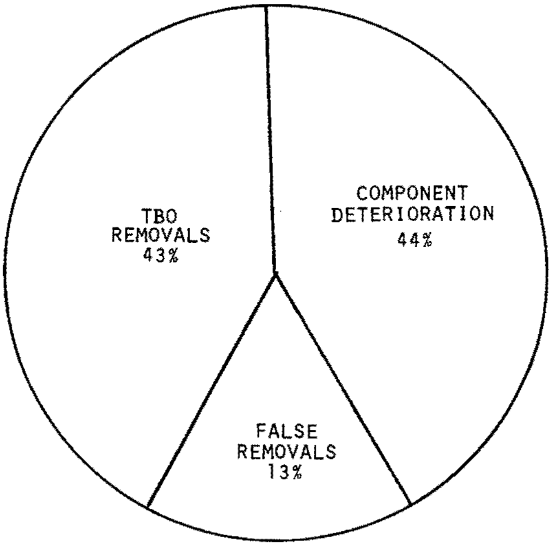


Figure 4. Distribution of MMH/FH for CH-47A by subsystem.

I don't have a chart, unfortunately, of accidents. The chart I had in mind to present had the simple message that rotary wing aircraft are considerably higher than fixed wing aircraft. Our goal is to bring it down to the fixed wing level and we think we can, at least on a realistic basis of accidents per landings.

In Army aviation, we're generally talking about goals, showing a factor of two improvement. In other words, we want to cut that experience by half. Now, that's rather optimistic, but I think you'll see in some of the later presentations that it's achievable.

Moving on from aircraft level display, let's follow the thread down through the transmission that we saw as a rather large contribution to aborts and maintenance man hours. In Fig. 5, you see a distribution of TBO (Time Between Overhaul) or scheduled removals, component failure deterioration, and false removals due to inadequate diagnostics.



TRANSMISSION REMOVALS (CH-47 EXPERIENCE)

Figure 5. Drive system major constraints.

Now, Charlie Smith alluded a little bit in the previous presentation to some of the work the Army is trying to do, which I'm proud to say I was a part of, in applying more rigorous analytical criteria for determining when and if we should have TBO removals. I don't want to explore that now, but just to throw one, perhaps controversial, thought out; our analysis has concluded that, in general, we do not have to have any additional diagnostics to allow our transmissions to go on-condition. If you want to pursue that with me later, I'll be glad to. It's a rather fundamental concept and some have had a hard time accepting it.

The inherent reliability of component deterioration is being addressed by many Army programs both in terms of material capabilities and, in fact, nondestructive inspection techniques that would improve reliability.

I want to concentrate on our efforts to reduce false removals through improved diagnostics. Before I do that, let me turn to a tabular list of accidents caused by transmissions. Figure 6 represents all Army rotary wing aircraft. It's a composite of many different aircraft. The top level causes are the lubrication starvation induced failures which we have addressed through many design improvements-redundant jets and last chance screens, etc., in new aircraft.

Component	Failure Mechanism
	Lubrication starvation induced
Scavange Line	Box spewed oil; inhaled by engine causing turbine failure
Input Pinion Bearing	Seizure - Lube starvation (local)
Input Pinion Bearing	Seizure - Lube starvation (local) residual contamination
Input Pinion Gear	Locknut backed off - Lube starvation of bearing-blocked oil jet
Input Pinion Bearing	Seizure - Lube starvation; filter studs pulled out
Filter Leak	Improperly reused gaskets leak badly
Input Pinion Bearing	Seizure and subsequent oil-fed fire - Lube starvation from quick disconnect failure
Gear Tooth	Possibly gear failure due to low lube - Possibly stained sight glass led to maintenance misreading level
Gear Scuffing	Tooth failures due to lube problems (common)
Gear Pitting	First stage bevel gear pitting clogs pump
	Isolated causes
Quill Shaft	Cracking through improperly machined hole in shaft
Bevel Gear	Crack through mounting flange due to material inclusion
Bevel Gear Retent	Failures of bevel gear retention cap screw
Stud Failure	Stud fatigue failure due to bending on hard landings
Control Yoke	Gross rotor imbalance or blade loss caused yoke failures
Mounting Studs	Transmission left aircraft - Mounting studs failed
	Repetitive causes
Clutch	Slippage and resultant rpm loss

Figure 6. Transmission/gearbox accident history.

The isolated causes are the one time events that are experienced, and down at the bottom is one of the few repetitive causes we see that create accidents, that is, clutch slippage. It's a particular insidious mode that so far has resisted any diagnostic technique to determine when the clutch is going to slip. Any sudden loss of power in a rotary wing aircraft is disturbing, particularly if it occurs as you are coming in for a landing. One of the things I want you to notice are the cracks we have had in shafting and bevel gears, and these are of concern to us, although, as you can see, they are isolated cases. They have not heretofore been detectable through debris monitoring because they simply don't generate debris until it's too late.

The next consideration that we ran into in looking at accidents, and this is what Charlie Smith was alluding to when he talked about the impact of precautionary landings, are the accidents which are simply due to misexecuted precautionary landings. The potential for these accidents is shown on Fig. 7.

Landing Mode	(≈Rvn)	(≈Conus)
P(Accident/ loss of structural integrity)	1/3	1/3
P(Accident/ power loss)	1/10	1/100
P(Accident/ controlled precautionary landing)	1/2,000	1/4,000
P(Accident/ normal landing)	1/60,000	1/120,000

Figure 7. Accident probabilities.

The third category in Fig. 7 is the accidents due to the controlled precautionary landings which occur in one in 2,000 landings to one in 4,000 landings. Now, that simply says the world is not a pool table and you're going to land on a stump once in a while, but your criteria for an adequate diagnostic system depends on your recognition of this.

In Fig. 8 is illustrated the methodology we have been developing to determine how the probability of an accident, given a precautionary landing, could affect the requirements for a reliability objective on your diagnostic system. What we show is that the accident rate can go up as the reliability of your diagnostic system goes down, and indeed, in certain kinds of conditions, you can actually cause more accidents than you would have had if you simply didn't detect it and let the inherent failure progression of the mode take its natural course.

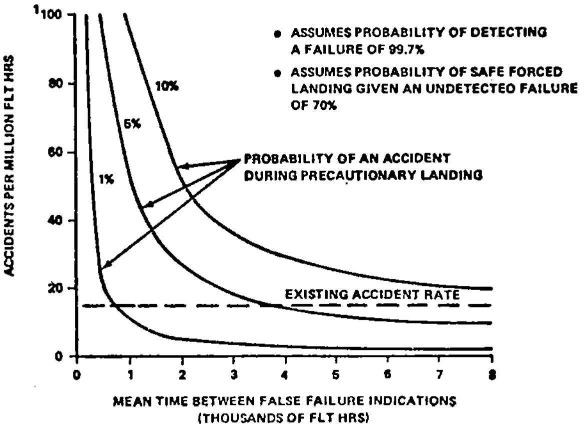


Figure 8. Accidents rate as a function of diagnostic system unreliability.

The important message here is that in trying to utilize a rigorous approach in looking at the problem, we end up defining some of the criteria that you then later apply to an adequate diagnostic system. In summary, after looking at all of the aircraft systems, the major needs that emerged were, 1) improved detection of the shaft and gear cracks in transmission for accident reduction 2) the reduction of erroneous removals of transmissions and 3) the reduction of in-flight aborts, both necessary and unnecessary.

Now, improved detection can be accomplished in two basic approaches, as shown on Fig. 9. You can invent a new technique or you try to improve the old one. Improvement of old techniques can be done in three ways: you can change the threshold, the logic, or you can change the use intervals, inspect more frequently or less frequently as the case may be.

- INVENT NEW TECHNIQUE (DIFFERENT SENSING PROCESS)
- IMPROVE EXISTING SYSTEMS THRU:
  - DETECTION THRESHOLD
  - LOGIC
  - USE INTERVAL

Figure 9. Methods of improving detection systems.

What I would like to do is run through some of the things that the Army has tried to do in these kinds of categories. For instance, in the new technique area, in addressing the shaft and gear crack problems, some of the research we're doing is with a vibration detection technique using a very high frequency carrier, like 200 to 300 KHz. Figure 10 shows the result of a test for gear tooth cracking. What we did here is put a saw cut through the root of the tooth in a bevel gear, ran it about 270 percent load to try to get the crack to progress. What you see there are nice spikes that occurred as the crack progressed. So, there is some hope that we perhaps have a technique for detecting this heretofore undetectable mode. Of course, it still has to pass the test of some rigorous cost effectiveness analysis.

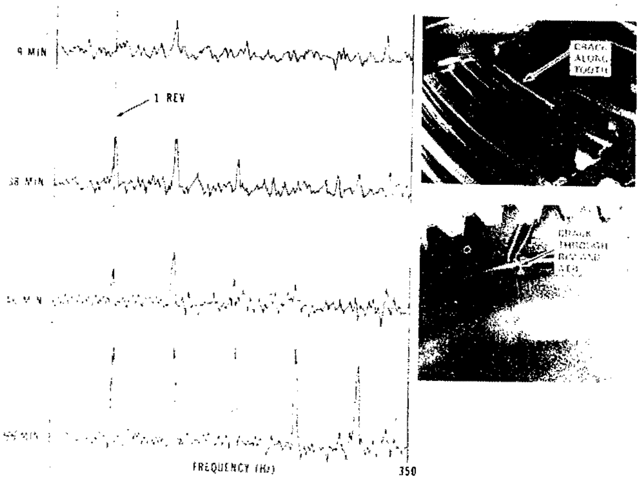


Figure 10. Spiral bevel gear crack progression test.

In terms of improving the existing techniques, the first thing we have to do is recognize that the primary diagnostic technique for helicopter transmissions is oil borne debris. Figure 11 indicates that 55 percent of the removals are caused by some sort of oil contamination, whether it be through a filter examination, soap analysis, or the notorious chip detector. Noise and vibration, that is, personally observed noise and vibration, is the next major cause. It is a large source of our false removals, just as in the oil contamination area. The miscellaneous visual observations such as oil pressure indicator fluctuations are the last group of symptoms.

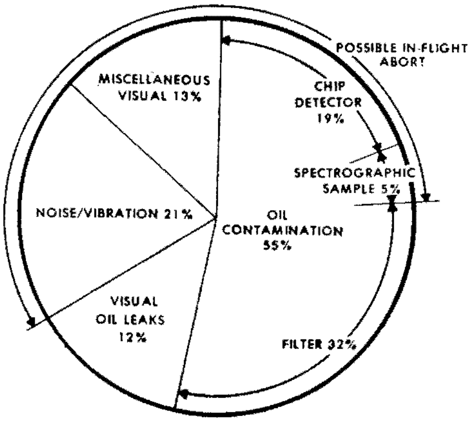


Figure 11. Debris in oil causes 55% of failure warnings. (Data from 193 CH-47 and 56 CH-46 transmission).

Before we can start to improve the false removals from debris, the first thing we have to do is basically quantify the relationship between the various levels of failure degradation and the detection signature, as illustrated in Fig. 12. It is so fundamental that we shouldn't have to put a chart up to remind anybody, and yet the amount of knowledge that we currently have in this area is disgraceful, woefully inadequate. We have optimized our diagnostic systems through gut feel for so many years that we no longer recognize that the poor engineers who are now trying to optimize it simply don't have this fundamental information available.

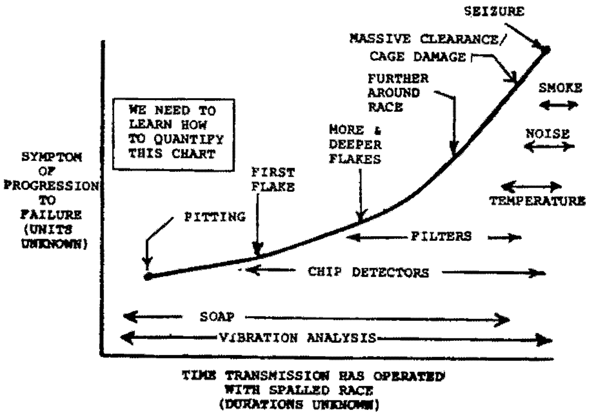


Figure 12. Crucial failure relationships.

Some of the work the Army is funding in this area, recognizing this problem, are research programs where we're examining the filters that have captured most of the debris in the oil and trying to quantify the various particle distributions of good and bad transmissions so we can begin to draw that distinction and understand where that threshold might be.

Figure 13 shows slides of some debris from failed transmissions. You can see the lower left one has some bronze from the cage and the upper left one has some rather large particles. There are obviously some 4,000 micron boulders floating around in the oil system. There is a certain reality here to the particle sizes floating around in a complex helicopter transmission that is not recognized by many people. Certainly engine lubrication systems run a little bit cleaner than these do, and this has got to be recognized in trying to optimize the diagnostic system.

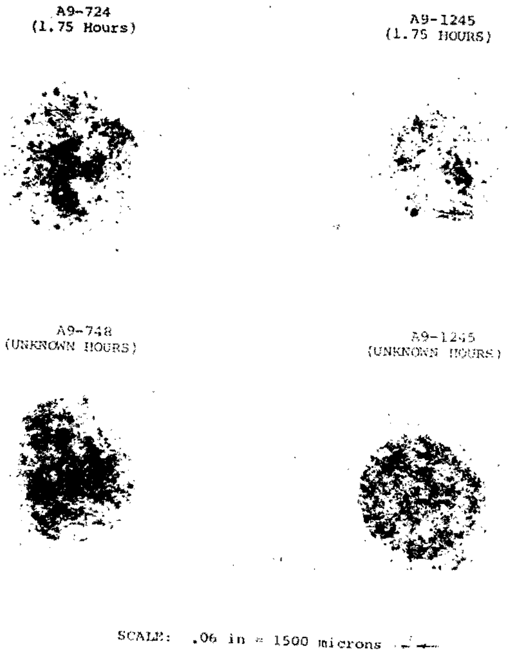


Figure 13. Examples of debris from failed transmissions.

We can convert these debris samples to some distributions. Figure 14 is an example of some work we just recently finished. This is a failed transmission particle size distribution. If you overlay the good transmissions on it you will begin to see that there is an awful tight overlap, and you conclude that you only begin to get some real distinction out at around 1500 microns. That's the kind of information that we need to change thresholds to improve our existing detection systems.

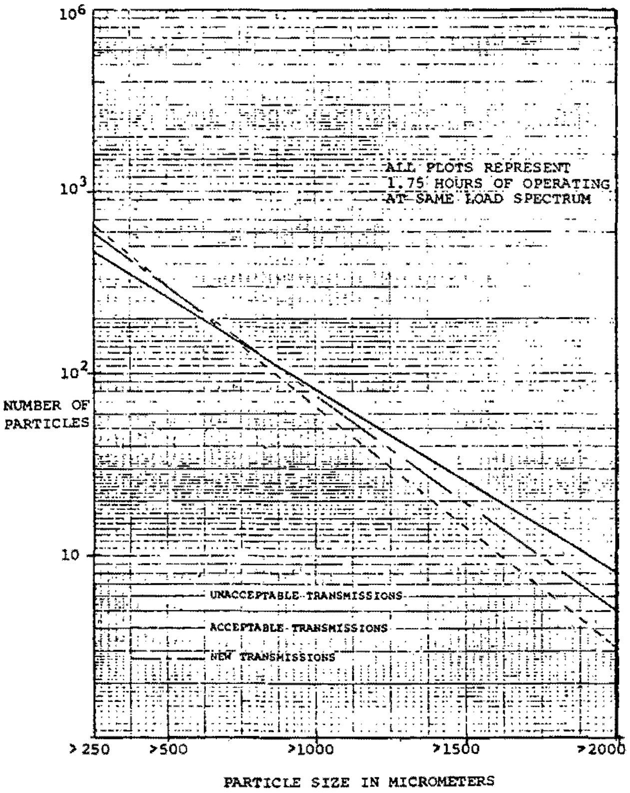
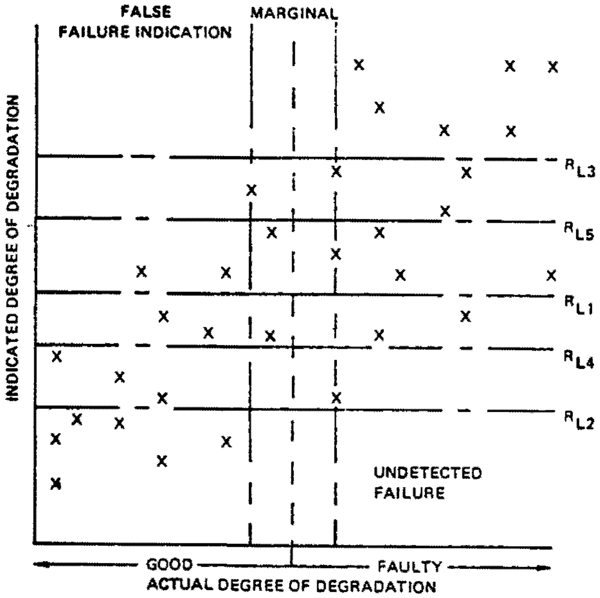


Figure 14. Large particle size distributions on CH-47 transmissions.

Once the signature to failure relationships are known, you can pick any threshold you want. We're trying to indicate on Fig. 15 that with any threshold of degradation level you end up getting different probability of false indications as well as the probability of missing a failure.



X TEST DATA
R_{L1} MINIMUM NUMBER OF FALSE INDICATIONS
R_{L2} MINIMUM NUMBER OF UNDETECTED FAILURES
R_{L3} MINIMUM NUMBER OF FALSE FAILURE INDICATIONS
P_3 PROBABILITY OF UNDETECTED FAILURE
P_2 PROBABILITY OF FALSE FAILURE INDICATION

	R_{L1}	R_{L2}	R_{L3}	R_{L4}	R_{L5}
P_3 =	0.19	0	0.63	0.06	0.44
P_2 =	0.25	0.63	0	0.44	0.06

Figure 15. Sensitivity of false indication to reject limits (R_L).

Other things we're doing is exploring with "and/or" logic. To illustrate, Fig. 16 displays how we could get different cumulative probabilities of false indications given some inherent reliabilities of an individual sensor.

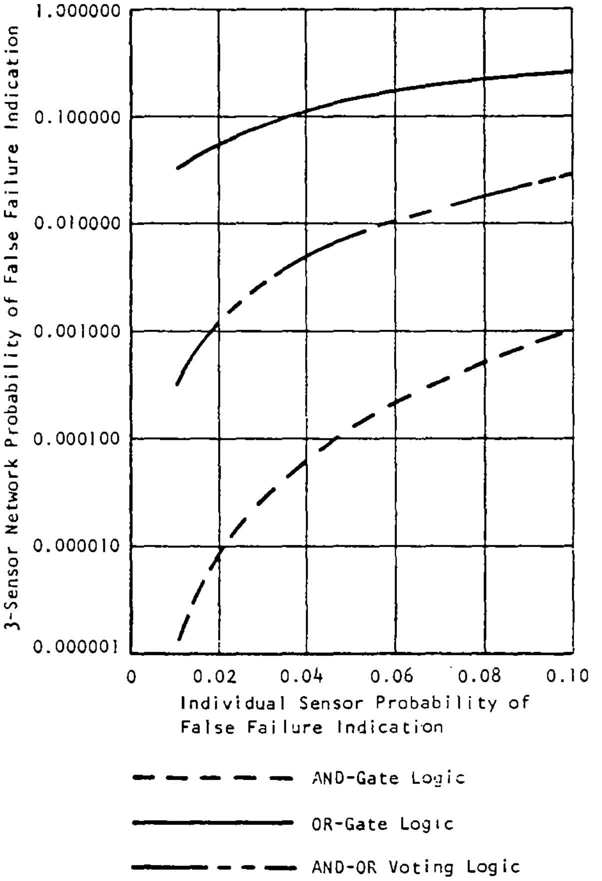


Figure 16. Three basic types of fault isolation logic on the probability of undetected failure.

The last area of improvement is inspection interval. Figure 17 is a plot of cumulative probability of detection versus the frequency at which you do an inspection with different progression intervals. I think what you see there is the problen we face of having to understand what is the failure progression interval of the failure modes we're trying to detect. Obviously, if it's 10 hours it's one thing; if it's 300 you would optimize it in an entirely different inspection interval. And again, here I must admit that our information is woefully inadequate. Utilizing in-service experience, information is simply truncated by whenever the diagnostic system took the failure out. Since we don't run our military systems for analysts, we don't let failures go all the way through to some catastrophic nature; we have to rely on R and D funding and limited test programs to fill the information gap in that area. They are obviously very critical to our understanding of failure progression levels.

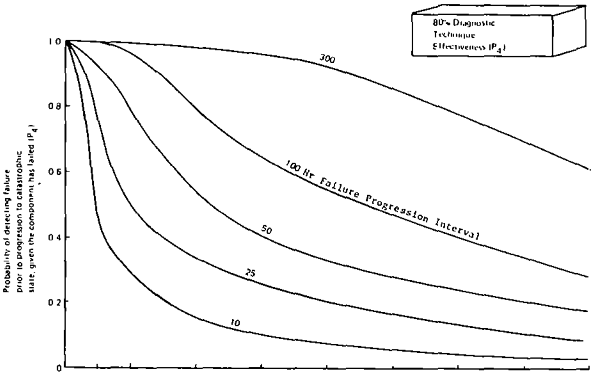
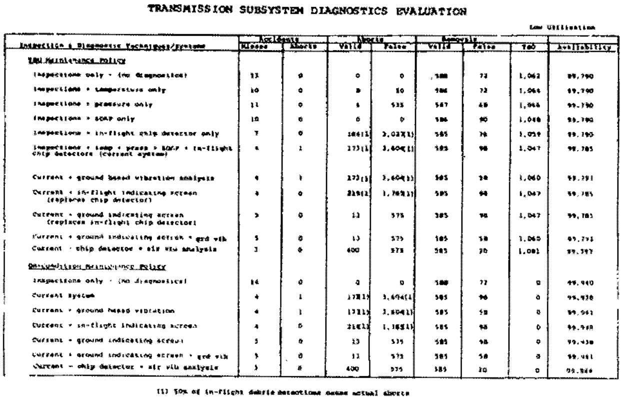


Figure 17. Detection probability as a function of progression interval and inspection interval.

Charlie Smith has previously alluded to our caution in moving the chip detectors out of the cockpit. It arises from our lack of understanding of failure progression intervals from all the common failure modes that we might see in the transmission, and research only can fill this gap. Certainly, the traditional bearing spalling due to subservice fatigue is definitely 200 to 300 hours in our low speed transmission application, but there may be other failure modes which we must detect that may be considerably less than that.

I want to move on now to give you an example of some of the system evaluations we have been attempting to do recently. Figure 18 illustrates the results from a recent study where various diagnostic systems were examined. We tried to evaluate these different configurations in a given situation This happens to be a low utilization situation of only 10 hours a month (we did others at higher utilizations, such as 60 or 80 hours a month which would represent a combat situation) and tried to quantify the accidents, for instance, due to both missed failures as well as those aborts, those accidents caused by precautionary landings that I noted before. The number of in-flight aborts (in their own right as a mission abort parameter), the removals, both valid, false and the scheduled removals and the availability were all quantified. The next table (Fig. 19) simply takes all those parameters and converts them to a standard measure of life cycle costs.



TRANSMISSION SUBSYSTEM DIAGNOSTICS EVALUATION

Low Utilization

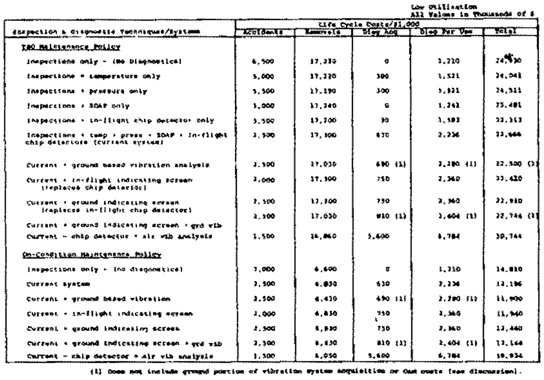
Inspection & Diagnostic Techniques/Systems	Accidents		Aborts		Removals			
	Losses	Aborts	Valid	False	Valid	False	TBO	Availability
TBO Maintenance Policy								
Inspections only - (no diagnostics)	11	0	0	0	588	72	1,062	99.790
Inspections + temperature only	10	0	8	50	586	72	1,064	99.790
Inspections + pressure only	11	0	6	535	587	48	1,066	99.790
Inspections + SOAP only	10	0	0	0	586	90	1,048	99.790
Inspections + in-flight chip detector only	7	0	166(1)	2,027(1)	585	76	1,059	99.790
Inspections + temp + press + SOAP + in-flight chip detectors (current system)	6	1	173(1)	1,604(1)	585	98	1,047	99.785
Current + ground based vibration analysis	4	1	173(1)	1,604(1)	585	58	1,060	99.791
Current + in-flight indicating screen (replaces chip detector)	4	0	219(1)	1,765(1)	585	98	1,047	99.785
Current + ground indicating screen (replaces in-flight chip detector)	5	0	13	575	585	98	1,047	99.785
Current + ground indicating screen + grd vib	5	0	13	575	585	58	1,060	99.791
Current - chip detector + air vib analysis	3	0	400	575	585	30	1,081	99.797
On-Condition Maintenance Policy								
Inspections only - (no diagnostics)	14	0	0	0	588	72	0	99.940
Current system	4	1	173(1)	1,604(1)	585	96	0	99.938
Current + ground based vibration	4	1	173(1)	1,604(1)	585	58	0	99.941
Current + in-flight indicating screen	4	0	219(1)	1,765(1)	585	98	0	99.938
Current + ground indicating screen	5	0	13	575	585	98	0	99.938
Current + ground indicating screen + grd vib	5	0	13	575	585	58	0	99.941
Current - chip detector + air vib analysis	3	0	400	575	585	30	0	99.946

(1) 50% of in-flight debris detections cause actual aborts

Figure 18. Transmission subsystem diagnostics evaluation.

Rather startling results emerged. It is a fact of life that in the current usage of helicopter systems, basically, the wholesale delivery of goods, it is extremely difficult to justify high cost airborne diagnostic systems. There are simply not enough benefits to be achieved to amortize, if you will, both the acquisition and the development cost, let alone the O and M cost of the diagnostic system.

Those are some examples of the Army's R and D approach in the diagnostics area. I think they reflect the methods I have earlier defined and produced a rigorous evaluation process. They first identify the problem and then try to fit some solutions to the problem. I would really encourage this group to look at the Army's problems and see if they don't have some solutions to these problems. I would encourage you to maintain a vigilance on the problems that you are trying to address. Consider the cost effectiveness of them.



Low Utilization
All Values in Thousands of $

Inspection & Diagnostic Technique/Systems	Life Cycle Costs/$1,000				
	Accidents	Removals	Diag Acq	Diag Per Use	Total
TBO Maintenance Policy					
Inspections only - (no diagnostics)	6,500	17,330	0	1,210	[illegible]
Inspections + temperature only	5,600	17,220	300	1,521	24,041
Inspections + pressure only	5,500	17,390	300	1,321	24,511
Inspections + SOAP only	5,000	17,240	0	1,241	23,481
Inspections + in-flight chip detector only	3,500	17,200	30	1,383	22,113
Inspections + temp + press + SOAP + in-flight chip detectors (current system)	2,500	17,300	470	2,236	22,466
Current + ground based vibration analysis	2,500	17,030	690 (1)	2,280 (1)	22,500 (1)
Current + in-flight indicating screen (replaces chip detector)	2,000	17,300	750	2,360	22,420
Current + ground indicating screen (replaces in-flight chip detector)	2,500	17,300	750	2,360	22,910
Current + ground indicating screen + grd vib	2,300	17,030	810 (1)	2,404 (1)	22,746 (1)
Current - chip detector + air vib analysis	1,500	16,860	5,600	6,784	30,744
On-Condition Maintenance Policy					
Inspections only - (no diagnostics)	7,000	6,600	0	1,210	14,810
Current system	2,500	6,830	630	2,236	12,196
Current + ground based vibration	2,500	6,430	690 (1)	2,280 (1)	11,900
Current + in-flight indicating screen	2,000	6,830	750	2,360	11,940
Current + ground indicating screen	2,500	6,830	750	2,360	12,440
Current + ground indicating screen + grd vib	2,300	6,430	810 (1)	2,404 (1)	11,544
Current - chip detector + air vib analysis	1,500	6,050	5,600	6,784	19,934

(1) Does not include ground portion of vibration system acquisition or O&M costs (see discussion).

Figure 19. Transmission subsystem life cycle costs.

DISCUSSION

DR. MOW: We can take one or two questions.

DR. JOSEPH JOHN (IRT): I want to focus attention on the failed transmissions. I have heard a similar remark made several times now on the EPRI analysis.

DR. RUMMEL: What analysis?

DR. JOHN: EPRI analysis of the oil. Are there statistically meaningful data that exist now after analyzing some failed transmissions as to what you find in the oil, the particle size and elements?

DR. RUMMEL: The data you saw here, I must say with some regret, is about the best particle size distribution work that I've seen in complex transmissions. We're not very satisfied with it; it's a limited number of samples. There are a lot of problems in doing the data analysis whether you want to do metal or all particles, whether you're looking at just ferrous metals, whether you want to talk about shapes, whether you want to talk about particles that have been machined or pressed together going through gears. I'm not sure I'm answering your question. Yes, there is some quantitative display and distribution of particle sizes that are in the oil all the way from the smallest one micron soap type sample up to the largest 5,000 micron size.

DR. JOHN: Is that available on other transmissions or only for-----

DR. RUMMEL: I really don't know. I don't know how generic those distributions are. We just recently, honestly, started to do that work. We've done it on our new UTTAS, and on some CH-47 transmissions. I don't know whether the differences are due to the complexity of the transmission, the filtration levels which are different in those two transmissions, orwhat.

DR. JOSEPH HEYMAN (NASA, Langley): Kirk, would you care to mention something about the fact that a number of people have set different criteria on particle size as to relative predictability of failures, especially in oil lubricated systems?

DR. RUMMEL: I'm not sure I understand your question, Joe. Certainly the lower particle sizes are represented by the traditional spectrographic oil analysis program that is in use in the Air Force, Navy and Army programs, and that is a very viable tool, an important tool, for detecting many failure modes, particularly those failure modes which generate only the small debris. I'm thinking of spline wear and gear fretting. We're very concerned that perhaps we're going to infringe on the soap reliability with the fine filtrations that we're now going to.

Larger particle sizes, in our mind, seem to be where the action is. We can pick up later stages of failure and we feel we can let the transmissions run into those stages of failure and eliminate a lot of false removal that comes through wear and tear and the high quantities of particles that are generated at lower particle sizes. Did I answer your question, Joe?

I'm not sure how much difference of opinion there is in the technical community. I think most people accept the importance of soap. I think there are a lot of people who feel that its false removal rate is a bit too high and, indeed, there are programs going, particularly in the Navy; the Navy is doing some nice work in the soap area trying to improve their thresholds for soap detection. The Army has done some fine work in trying to evaluate how to take the sample, where to take it, how long to let the oil settle, etc. I don't think there's any disagreement on that. I think the disagreement may lie in how big the particle size is where the area of discrimination is. If this is what you're alluding to, and I do think there's a lot of open questions, what can I say, send money. Then we might be able to answer that question.

See you next year.