A PERSPECTIVE ON THE USE OF N.D.T. IN THE INSPECTION OF AGING AIRCRAFT

by

D. C. HOLLAMBY
AUSTRALIAN CIVIL AVIATION AUTHORITY

The past thirty years have seen a great change in the perception of the place of nondestructive testing (NDT) in the field of aircraft maintenance. Thirty years ago the use of NDT was largely confined to providing inspection solutions in situations where a known type of damage was suspected and where more conventional inspection methods were ineffective. Now it is seen as an essential ingredient of continuing airworthiness. Much of this change has resulted from the introduction of damage tolerant design criteria with a concomitant increased reliance upon the inspection process. It must be stressed that this increased use of and reliance upon NDT methods has been imposed upon inspection personnel by the need created by the new design concepts.

Although some work has been done to justify the degree of reliance that can be put on NDT methods, aircraft designed to the new standard have not yet reached the stage where the resultant assumptions made in their design have been tested. Nevertheless, extensive use of NDT methods for the substantiation of the continuing airworthiness of aircraft, which were not designed totally to damage tolerant criteria but which have now exceeded their originally proposed design life, is now being proposed.

This paper is an attempt to survey the developments over the last 30 years which have led to the present position and to question whether they are sufficient justification for the direction taken. It is written from the point of view of an NDT specialist.

NDT methods of inspection have been used in aircraft maintenance inspections very successfully for many years and it is not the present intention to decry this in any way. When the type of damage to be sought, e.g. a crack, can be closely characterised with respect to its nature, expected location and direction of propagation, then it is usually possible to devise an NDT inspection procedure which will detect the presence of the damage with a high degree of sensitivity and reliability. What is in question is the reliability and practicability of NDT methods for the detection of damage, the position and nature of which cannot be exactly defined, in large areas of aircraft structure. One such structural area, which has been highlighted by recent events, is the fuselage of the typical commercial airline aircraft: more particularly the skin lap joints of the pressure cabin. It is problems relating to the inspection of these joints which will be used in this paper to illustrate the more general problems associated with the inspection of aging aircraft.

Changes in the Methods of Nondestructive Crack Detection Used in Airframe Maintenance.

In 1961, four methods of NDT were used, for the inspection of non-ferromagnetic materials, in airframe maintenance in the civil industry. Namely, the dye penetrant, ultrasonic, radiographic and eddy current methods. In 1991 the same four methods are used. During the intervening years, attempts have been made to introduce other methods, such as thermography, holography, acoustic emission etc., without success. Given that the methods used have not changed, to what extent have the methods themselves been improved? This may be summarised as shown in Table 1.

Although the improvements in the capability of the eddy current method have been substantial, it can be seen that the NDT methods used for crack detection in airframe maintenance have essentially changed little over a period during which its use has not only increased but also changed in nature.

METHOD	1961	1991
RADIOGRAPHIC	FILM TECHNIQUES IN THE 20 - 200kv ENERGY RANGE	NO SIGNIFICANT CHANGE
LIQUID PENETRANT	AEROSOL APPLIED RED AND FLUORESCENT DYE TECHNIQUES	NO SIGNIFICANT CHANGE
ULTRASONIC	PULSE-ECHO CONTACT TECHNIQUES	NO BASIC CHANGE BUT IMPROVED SENSITIVITY AND RESOLUTION
EDDY CURRENT	SURFACE CRACK DETECTION ONLY	IMPROVED SURFACE PLUS SUB-SURFACE CRACK DETECTION

Table 1. Nondestructive test methods used in aircraft maintenance over the preceding 30 years.

Changes in the use of NDT Methods

Thirty years ago, some of us who were involved with NDT in major manufacturing organisations were busy preparing Non-Destructive Testing Manuals. There were at that time no precedents for the form these Manuals should take but, we were told, they had to be written as a result of new requirements relating to fail-safe justification prior to certification of the aircraft. It soon became clear that what we were required to do was to nominate inspection procedures which could be applied to components and structural parts which were "difficult" to inspect visually. The procedure appeared to be that the Design Department nominated the components and structures to be inspected, the Service Department then surveyed the aircraft and determined the practicable extent of visual inspection and, finally, the left-over items were handed to the NDT section. It soon became clear that the engineers involved in this process had an exaggerated conception of the capabilities of the N.D.T. methods available to us. This was particularly so in the case of radiography.

In the case of the fuselage, these "left-overs" included large areas of structure concealed behind the cabin furnishings. It was clear that, if furnishings were not to be removed, the only possible inspection method for this structure was radiography. This led to a dispute. Presumably, the type of damage to be detected included cracks but, if this were so, it would be necessary to know exactly what to look for and in which direction, because to detect a crack radiographically it is necessary to direct the X-ray beam through the plane of separation. Naturally, there is a degree of latitude in this X-ray beam-to-crack plane angle but this latitude is very dependant upon crack tightness. Thus without considerable knowledge of the position, direction and nature of a crack, said the NDT people, a reliable procedure for its detection radiographically is not feasible. Nonsense, said the design people. The structure is fail-safe and can tolerate a degree of damage which will be obvious on any radiograph.

The arguments which ensued were insoluble since they were about hypothetical forms of damage which might, or might not, occur in the future. The inevitable result was that the argument was won by the protagonists with the greater prestige within the organisation and, inevitably, NDT manuals started to appear with procedures in them which required inspections to be made for "structural integrity", with no clear guidance to the inspector as to the type or position of the damage to be expected.

NDI and Damage Tolerance

The possible applicability to aircraft structures of damage tolerant design based on fracture mechanics/NDT criteria was first postulated by Packman et al. in 1968. The concept was a very attractive one offering potential savings in airframe weight as compared with the existing fail-safe criterion. However, there was one major problem - the non-existence of numerical data regarding the efficacy of the various NDT methods as crack detectors. What was required was a series of curves (one for each method) showing the probability of detection of a fatigue crack as a function of its size. It was only possible to produce such curves empirically. Also, if such curves were to be used with confidence, it would be necessary to perform a large number of tests on a large number of samples to establish them.

In 1973, Rummel et al.[1] published the results of trials to assess the detectability of fatigue cracks in aluminium alloy samples using the radiographic, ultrasonic, liquid penetrant and eddy current methods. These trials were statistically very thorough. A total of 166 specimens, containing 427 fatigue cracks of different shapes and sizes were inspected using all four methods. Each was inspected nine times - giving a total of over 15,000 go/no-go assessments. The results were processed statistically and portrayed as 95% confidence, probability of detection plots. A typical plot is shown in Fig.1.

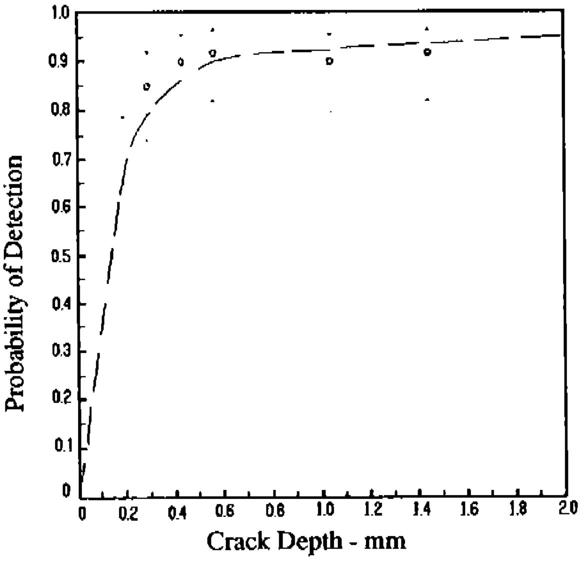

Fig.1 Typical probability of detection curve

Thorough as this work was, it is difficult to understand the rationale behind it. Since the subject of the trials was fatigue damage, it was presumably intended to provide information on the detectability of in-service cracking. However, the form of the cracks used, their environment and the way in which the test methods were applied were all quite different from what would pertain in airframe inspection. In fact, all these factors were such as to make the results very optimistic. Nevertheless, it is interesting to note that one of the conclusions drawn was that "X-radiography was demonstrated to be the least reliable of available non-destructive test methods for crack detection and should not be seriously considered as a sensitive, reliable technique for the detection of tight cracks". The liquid penetrant method was found to be potentially very sensitive but this was dependent upon surface condition. By contrast, the results obtained for the other methods were very encouraging. They suggested that there was at least a 90% probability with 95% confidence of detecting a crack 3 mm (0.12") long by 1 mm (0.04") deep using the ultrasonic and eddy current methods.

In 1978, Lewis et al.[2] published the results of extensive crack detectability trials carried out by the Lockheed-Georgia Company. These trials, carried out under the terms of a U.S. Air Force contract, were aimed at assessing the detectability of fatigue cracks in actual aircraft structural samples. Reasonably large samples containing fatigue damage were transported to 21 different Air Force bases and depots for examination using NDT methods. This was the programme which became better known as the "have cracks will travel" programme. Again, these trials were statistically very thorough, with nearly 300 technicians completing tasks comprising the cumulative inspection of over half a million potential crack sites. The results were considered to be very disappointing - to quote from the conclusions drawn in the final report "With the exception of fluorescent penetrant inspection, the NDI techniques employed in the program demonstrated considerable difficulty achieving a 50% probability of detection for a 1/2-inch crack size with a 95% confidence level." This result certainly contrasted very unfavourably with the results obtained by Rummel.

Even the above conclusion could be said to be optimistic, in the context of this paper, in that it excepts the liquid penetrant method from the generally poor result. The components used to assess this method contained edge-cracks which were positioned remote from fastener holes, faying surfaces etc. This would not be representative of a typical pressure cabin situation where there is a multitude of natural interfaces to bleed out penetrant materials and thus vitiate the inspection.

There are additional reasons why this programme, which was intended to represent conditions typical of those pertaining in the maintenance environment, were, in fact, not typical of the inspection situation which is the present subject of discussion. The structural samples used were not part of complete aircraft, with the result that access for inspection was not a problem and, in the case of radiography, no other structure or furnishings etc. was superimposed upon the area under test and, in some specimens, fasteners had been removed from the holes to be inspected. All of these factors would contribute to producing results which were optimistic.

Other factors which must be considered relate to the method of analysis of the data - and this applies to both the Martin Marietta and Lockheed Georgia programmes. In both cases the data were analysed on the assumption that probability of detection must be a continually improving function of increasing crack size. This is reflected in the form of the P.O.D. curves which resulted. In fact, for the types of cracks and samples used and the techniques of inspection used in the programmes, this was a reasonable assumption. However, if an inspection situation is such that there may be some cracks present which will not be detected for reasons not represented in the programmes, the P.O.D. curves are no longer valid. The problem of crack orientation in the case of radiography was one such reason. Another is the effect of contamination and/or residual or applied compressive stresses on crack detectability in the case of the liquid penetrant method and, to a lesser extent, the ultrasonic method (e.g. the presence of a corrosion inhibiting penetrant, such as LPS3, in a crack can appreciably reduce its ultrasonic reflectivity).

There is yet another reason for treating the results of these programmes with caution. This relates to "expectancy of detection". The personnel performing the inspections were told that cracks were present in the samples. That this was certainly a factor in the case of the Lockheed programme is shown by the number of "false calls " i.e. the number of times cracks were reported at uncracked sites. Taking one series of inspections carried out on a total of 240 holes, using the manual eddy current bolt-hole technique, the two top performers, who located 24 and 28 of the 42 detectable cracks present, also reported an additional 87 and 107 cracks to be present at uncracked holes! Clearly these performances should have been rejected as part of the survey. This was an extreme example of a general trend for increasing numbers of successful detections to be accompanied by increasing numbers of false-calls. Even in the case of simple eddy current surface scanning around holes, of a sample of 96 inspectors, only 4 nominated no false calls while 28 nominated more false-calls than successful detections. This is hardly in accord with the conclusion drawn in the Lockheed final report that: "Fortunately, the instances of extremely high false call levels were not numerous enough to make the total data picture suspect".

Although there is no doubt that the U.S. Air Force gained much useful information from the Lockheed programme, for the many reasons described here, any extrapolation of the results to types of crack damage, structures and inspection environments other than those of the programme must be very suspect indeed. Nevertheless, we find that such extrapolation does exist. Thus, Goransen[3] in describing the elements of damage tolerance verification used for the certification of certain models of Boeing Aircraft states: "A specific probability of detection was assigned to the detection standard based on data obtained from Reference 9 " - The latter being the Report on the Lockheed Georgia programme. Also, the same source has been used by Broek[4] in developing the software for his proposed computerised damage tolerance analysis system.

The unfortunate truth seems to be that it is not economically possible to devise such programmes which will be relevant in a general way to structural inspections. Even if it were, it would be difficult to maintain their relevance. For instance, the example of eddy current surface scans and high incidence of false-calls described above, though to some extent understandable in the context of eddy current technology of the eariy '70s should not be relevant in 1991. The latest equipment with more easily minimised "lift-off" characteristics, used with screened probes to minimise effects from the proximity of fasteners, would be expected to have improved the resolution of real crack damage immensely.

If, as it would appear, the usefulness of POD trials is questionable, then any generalised approach to assessing crack detectability is impossible. It is then necessary to make individual assessments, for each inspection situation, based on past experience and considering all the factors which may be relevant. Not the least important of these factors is the character of the damage to be detected.

Characterisation of Geriatric Degradation

Much of the disagreement which exists regarding the proper way to deal with the problem of maintaining the airworthiness of aging aircraft by inspection revolves around the probability of multi-site fatigue damage propagating to a critical stage without detection. Those who once contended that this is not a problem, based their contention on the assumption that the normal scatter associated with fatigue initiation would ensure that there would always be a "lead crack" which would propagate to detectable dimensions before the total damage reached a critical state. Based on the experience of fatigue problems encountered during the normal life of an aircraft, this appeared to be a reasonable assumption. However, it is now apparent that this asumption cannot be confidently extrapolated into the largely unknown geriatric region.

Ideally, the aircraft designer tries to design a structure such that it will reach its design life without exhibiting any unacceptable fatigue damage. That fatigue damage of some sort will be present is inevitable, since aluminium alloys exhibit no fatigue stress limit and must start to accumulate damage from the first application of a fluctuating stress. The ideal design has, of course, not yet been achieved and, due to unpredicted local anomalies in stress distribution, fatigue cracks initiate early and are detected at various stages during the normal life of the aircraft. Once such an early incidence of fatigue has been established, there is usually no problem in monitoring an aircraft fleet for similar crack damage and, being able to characterise the nature of the damage, this is often the proper subject for NDT inspection.

Considering a structure as a whole, these early occurrences of fatigue cracking could be considered as "lead cracks" being indicative of its inevitable eventual demise. They would be expected to appear first where the anomalous stresses are highest and then where stresses are lower as cycles accumulate. The result of this mechanism is that the damage continually changes from a relatively high stress/low cycle phenomenon to a relatively lower stress/higher cycle phenomenon. This fact together with the increasing probability with expired time of the presence of potential fatigue initiators such as fretting, pitting and stress corrosion means that the probability of the presence of multi-site damage must increase with aircraft age. This can be exemplified by three recent occurrences in Australia of unexpected types of damage in high time aircraft - all related to fuselage lap joints.

First Example: Stress Corrosion Initiated Fatigue Damage.

During an external visual inspection of a high time (approx. 45,000 hours and 35,000 flights) airline aircraft, some slight skin distortion between the rivets in a longitudinal fuselage lap joint was detected. This indicated the possible presence of corrosion in the joint interface. Fasteners were removed and corrosion confirmed. The corrosion was not particularly gross and it was thought that it could be cleaned out within allowable limits. However, after cleaning the surface of the underlying skin, what appeared to be cracks were apparent on its faying surface. A section was removed for further examination and the presence of cracking confirmed. A section containing cracks was broken open to reveal the presence of what appeared to be stress corrosion cracks, which had initiated on the faying face. Closer examination in a scanning electron microscope revealed that, while the initial crack growth resulted from stress corrosion, subsequent cracking resulted from fatigue.

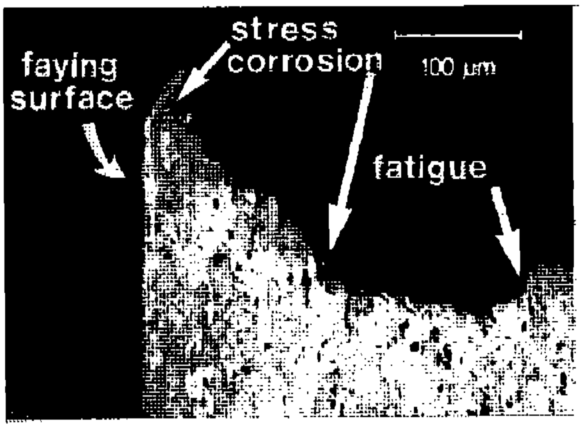

Fig.2 A section through the crack face.

Fig.2 shows a section transverse to the cracking showing that, as would be expected, the direction of propagation changed at the point where the crack mechanism altered. The stress corrosion cracking, being intergranular by nature, propagates at a comparatively shallow angle to the surface. When the fatigue mechanism takes over the crack turns to propagate across the thickness of the skin. Other sections revealed the presence of stress corrosion cracks which had not advanced far enough to change to fatigue. The significance of this example from an inspection point of view may be summarised as follows:

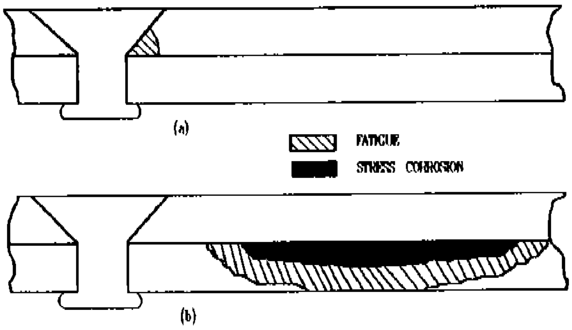

Fig.3 The expected nature of fatigue damage (a) compared with the cracking experienced (b).

(i) The aircraft manufacturer had postulated that the failure mechanism for this joint would result from fatigue damage initiating at the feather edge produced by the countersink in the overlying skin as shown in Fig.3 (a).

Consequently, the recommended inspection method was high frequency eddy current inspection of the top skin around the fasteners. Clearly this would have been completely inappropriate for the detection of the damage discussed above and shown schematically in Fig.3 (b). The size of the stress corrosion crack relative to the skin thickness as shown in Fig.3 (b) is typical of that present when the fatigue mechanism initiated, (i.e. the stress corrosion cracking penetrated to approx. 40% of the skin thickness at which time the lengths of the cracks were typically six times the skin thickness).

(ii) The only appropriate inspection method would have been low frequency eddy current inspection of the underlying skin through the top skin. However, the normal assumption that cracks would start at fastener holes might still have led to the inspection effort being misdirected.

(iii) In any case, if the whole area of the lap had been inspected, it is unlikely that, at the typical test sensitivity used for these joints, the presence of the stress corrosion cracks alone would have been positively detected.

(iv) The typical damage tolerance analysis for determining inspection intervals, would also not be appropriate. The assumption made for this type of structure would be that a fatigue crack would initiate as shown in Fig.3 (a). Crack propagation rate would be estimated for this case and taking no account of the presence of corrosion. However, in the case shown in Fig.3 (b), fatigue is initiating at the stress corrosion damage and propagating along a front in a direction across the sheet. This front, as shown in Fig. 3 would be approximately 6mm (0.25") long. Because of the pre-existence of the stress corrosion, the local stress level would be higher than for the predicted case and, in the presence of a corrosive environment, the propagation rate could be expected to be higher. The fatigue crack only has to propagate for approximately lmm to traverse the sheet, at which stage a 6mm long crack would be present. It can be seen that the analysis based on the predicted case would grossly underestimate the propagation rate and, therefore, overestimate appropriate inspection intervals for an eddy current inspection.

Example 2 - Multiple-site Fatigue Damage Initiating at Corrosion Pits

This second example of unusual crack damage also relates to longitudinal fuselage lap joints in a commercial airliner but concerns a different aircraft type produced by a different manufacturer. The aircraft in which the damage was found had all completed approximately 55,000 hours of operation and 55,000 flights.

The joints in question differed from those in the previous example in that countersinking was achieved by dimpling the skins rather than by machining.

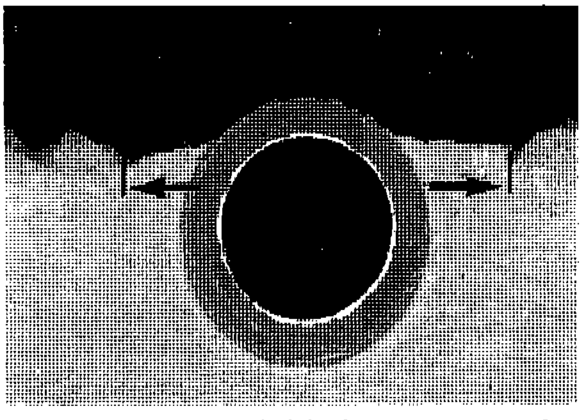

Fig.4 A fatigue crack initiating around a dimple.

Fatigue cracks were found to be initiating in the underside of the top skin in the dimpled area and propagating through the skin and circumferentially around the hole. Fig.4 shows such a crack having been broken open. It can be seen that the fatigue crack had propagated around the hole through an arc of approximately 45° and then turned to propagate towards the adjacent holes; the total extent of fatigue damage is shown by the black arrows. The cracks had initiated at corrosion pits occurring where the dimpling operation had apparently cracked the protective surface layers. Microscopic examination of the fracture face revealed very fine evenly-spaced striations indicating constant-amplitude loading from pressurisation cycles. Propagation was very slow there being approximately 16,000 striations present over a 2.8mm crack length.

Since the crack damage in this case was in the overlying skin, it was found that it could be detected using the eddy current method with a sliding probe technique. However, using the normal scanning technique, with the inspection coil traversed along the fastener line, those parts of the cracks which were concentric with the fastener would be completely undetectable and the sensitivity to the more advanced cracking between fasteners would be lower than would be the case for the presumed case of cracking along the fastener centre line.

An additional significant feature of this damage was the multiple-site character of the crack initiations. On one aircraft over a length of joint comprising 36 fastener holes, 28 were found to have cracks in the area of the dimple and of these 15 were at adjacent fasteners. None of these cracks were amenable to detection visually; low power microscopy indicating them to vary between 1mm and 2.5mm in length.

Example 3 - Unusual Fatigue Propagation Unrelated to Corrosion Damage.

This example relates to the same type of aircraft as the preceding one. During the investigation of the dimple cracking it was found that there were also cracks present in the underlying skin. There were many cracks present initiating at the faying surface and propagating transversely through the skin. The initiations were not at holes or in the dimpled area but in the skin and approximately in line with the edge of the fasteners. In some areas, these multiple transverse cracks had joined to form a continuous crack as shown in Fig.5 (note that the crack skirts around one hole in a manner which, without closer examination, may have led to the conclusion that it was of a similar nature to the previous example).

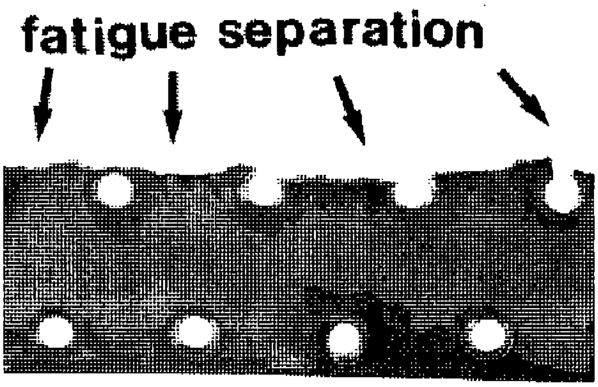

Fig. 5 An example of off-centre fatigue craking

The significance of the nature of these cracks from an inspection point of view is similar to the previous examples. However, whereas in those cases it could be argued that the unexpected position of the crack initiation points could be explained by the presence of corrosion, this was not so in this case.

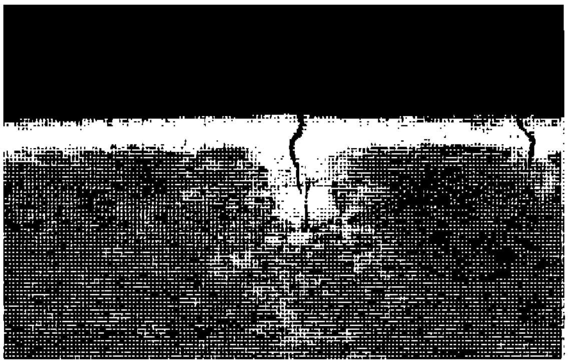

Fig.6 Small fatigue cracks penetrating the cladding.

Fig.6 shows a section taken through fatigue cracks which were at a very early stage of propagation and it can be seen that there is no corrosion present at the point of initiation. Presumably, the position of these cracks and the transverse nature of their propagation resulted from the fact that the primary stresses causing them were of a bending rather than tensile nature.

Summary and Conclusions

This paper has attempted to show that the use of NDT methods as a tool for the maintenance of the structural integrity of aircraft in a general sense, i.e. for the detection of damage the exact nature and position of which cannot be predicted, should be the subject of great caution and, in particular:-

(i) The trials which have been conducted to establish the probability-of-detection of cracks as a function of size have been too limited to be used in the establishment of realistic data to be used in the general case, because:

(a) The samples used were not representative of all the problems associated with a complete aircraft structure.

(b) The types of cracks used did not represent the great variety of crack types and conditions which can be encountered in practice.

(c) The trials were flawed by the artificial nature of their circumstances which must have had an effect on performance as a result of "expectancy of crack presence".

(d) In the analysis of the data obtained, no account was taken of "false calls" - a matter of great concern to a commercial airline operation.

(ii) Notwithstanding the suspect nature of these trials, there was one conclusion in which they were all unanimous, namely, that the reliability of radiography as a method for structural crack detection was so poor that its use should not be contemplated.

(iii) The use of NDT as a crack detection method in the substantiation of structural integrity into the geriatric regime is even more questionable because:

(a) The nature of the cracks sought can not be characterised with certainty.

(b) The location of the cracks cannot be precisely predicted.

(c) The possibility of the presence of multi-site damage only marginally below the threshold of detectability cannot be ignored.

Because of the above uncertainties, proposed nondestructive inspection procedures may be totally ineffective or significantly less efficacious than predicted. In addition, predicted crack propagation rates, based on the assumed nature and position of fatigue damage, could seriously underestimate the propagation rates encountered in practice.

The examples of unexpected types of fatigue damage which have been described all relate to skin lap joints in the pressure cabin. Although recent events have highlighted these joints as potential problem areas, it must be pointed out that they could be expected to represent a relatively simple problem as far as damage tolerant analysis is concerned. The predominant stresses in these joints, being due to pressurisation cycles, should be predictable and constant. Moreover, one side of the joints is completely accessible to inspection and accessibility to the other side is limited only by economic considerations. The possible relevance of these examples to structural elements of the wing and empennage, which are subjected to more complex stress fields and to which access is severely limited by physical constraints, should not be ignored.

It should be noted that, in all of the cited examples, the presence of a problem was first detected by the traditional method of careful visual inspection. In forty years of association with the inspection of aircraft structures, the author knows of no example of a fatigue problem which was first detected using NDT methods. In view of the present lack of experience of the types of damage which may be present in aging aircraft structures and of the use of NDT methods for their detection, it is considered that the substitution of the latter for the traditional inspection process, purely for reasons of convenience or economy, should not be permitted.

Thirty years ago the NDT fraternity understood that NDT methods could only be used reliably to detect fatigue cracks in aircraft structures when the nature and position of the cracks sought could be predicted with a high degree of certainty. Since then, although new design criteria have been introduced, the nature of NDT has not changed.

References

1 - Rummel W.D. et al - "The detection of Fatigue Cracks by Nondestructive Testing Methods" NASA Contractor Report CR-2369, Feb. 1974.

2 -Lewis W.H. et al. - "Reliability of Non-Destructive Inspections", SA - ALC / MME 76-6-38-1.

3 - Goransen U. - "Elements of Damage Tolerance Verification". Paper presented at 1983 ICAF Meeting, Toulouse.

4 - Broek D - "Fracture Control by Periodic Inspections with Fixed Cumulative Probability of Crack Detection". - FractuResearch TN 8604.